**Markus Rabe & Uwe Clausen (eds.)**

# Simulation in Production and Logistics 2015

Fraunhofer Verlag, Stuttgart

Kontaktadressen:

Prof. Dr.-Ing. Markus Rabe
TU Dortmund, ITPL
Leonhard-Euler-Str. 5
44227 Dortmund
markus.rabe@tu-dortmund.de

Prof. Dr.-Ing. Uwe Clausen
TU Dortmund, ITL
Leonhard-Euler-Str. 2
44227 Dortmund
clausen@itl.tu-dortmund.de

**Bibliografische Information der Deutschen Bibliothek**

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar

ISBN 978-3-8396-0936-1

Simulation in Production and Logistics 2015
Eds. Markus Rabe & Uwe Clausen
Stuttgart: Fraunhofer Verlag, 2015.
zgl. Tagungsband 16. ASIM-Fachtagung Simulation in Produktion und Logistik, Dortmund, 23.-25. September 2015
zgl. ASIM-Mitteilung Nr. 157

Alle Rechte vorbehalten

Dieses Werk ist einschließlich aller seiner Teile urheberrechtlich geschützt. Jede Verwertung, die über die engen Grenzen des Urheberrechtsgesetzes hinausgeht, ist ohne schriftliche Zustimmung des Verlages unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen sowie die Speicherung in elektronischen Systemen.

Die Wiedergabe von Warenbezeichnungen und Handelsnamen in diesem Buch berechtigt nicht zu der Annahme, dass solche Bezeichnungen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und deshalb von jedermann benutzt werden dürften.

Soweit in diesem Werk direkt oder indirekt auf Gesetze, Vorschriften oder Richtlinien (z.B. DIN, VDI) Bezug genommen oder aus ihnen zitiert worden ist, kann der Verlag keine Gewähr für Richtigkeit, Vollständigkeit oder Aktualität übernehmen.

Umschlaggestaltung: Formkonfekt | konzept & gestaltung | Kassel, Karen Marschinke
Druck und Weiterverarbeitung: Druckerei Djucic, Remshalden

© by Fraunhofer Verlag, 2015
Fraunhofer-Informationszentrum Raum und Bau IRB
Postfach 800469, 70504 Stuttgart
Nobelstraße 12, 70569 Stuttgart
Telefon 0711 970-2500
Telefax 0711 970-2508
E-Mail verlag@fraunhofer.de
URL http://verlag.fraunhofer.de

# Foreword

Simulation has proven to be an important technology in production and logistics since the 1980ies. There has been a kind of hype in the very beginning, when some managers claimed that everything would be simulated in every detail in the future before turning a single real screw. After recognising that simulation is a precious tool in the results, but also sometimes in the required efforts, there has been a short phase of disappointment when unrealistic expectations scaled down to the real merits of this great technology. Now, for many years already we experience a phase of stable growth, where simulation technology is conquering new application fields and many auxiliary tools make its application more smooth, more simple, and less expensive.

This development is perfectly mirrored by the ASIM Dedicated Conference on Simulation in Production and Logistics (ASIM'SPL) that started its life thirty years ago, in 1985. After some great events it calmed down and was near to disappear in the early 1990ies. There has been something like a rescue event in 1996, organised by today's speaker of the related ASIM working group Sigrid Wenzel, who collected Germany's research and application scene in Dortmund and prepared the mood for something more. This has been taken up in 1998 by one of this year's chairmen, Markus Rabe, with the first ASIM'SPL organised in the current conference format in Berlin. With this success, the conference has steadily grown, organised in Duisburg, Kassel, Karlsruhe, Paderborn, and two more times in Berlin. Now, in 2015, we are proud to continue this sequence and at the same time close a circle: After 30 years, the conference is successful in its new format, and back to the city where it has been reborn in 1996.

Of course, we can recognise some developments among the presented topics over this long time. Some years ago, the integration of simulation with real control systems (emulation) has been a very interesting topic. Since many years, also the integration of simulation in the daily operation, sometimes called assistance systems, leads to interesting methods and applications. More recently, the combination of simulation with optimisation gains more and more importance. This is especially true for heuristic algorithms, leading to the new term simheuristics. Many papers in this book address this field or related areas. Of course, logistics – from intralogistics to supply chains and transport – also covers a wide range of the total 73 contributions, and also production simulation is still not losing any relevance. In the contrary: especially for production, energy is becoming a new additional simulation topic, driven by energy cost as well as the social request for reducing the footprint of our production. We hope that all these contributions further stimulate the use of simulation technologies and wish the reader a pleasurable time browsing and studying these proceedings.

Markus Rabe and Uwe Clausen,
September 2015

Die ASIM-Fachtagung „Simulation in Produktion und Logistik“ ist die regelmäßige Tagung der Fachgruppe „Simulation in Produktion und Logistik" der Arbeitsgemeinschaft Simulation (ASIM). Die ASIM ist zugleich der Fachausschuss 4.5 der Gesellschaft für Informatik (GI).

**Dieses Buch ist registriert als ASIM-Mitteilung Nummer 157.**

## Program Committee

Chairmen: Markus Rabe, Uwe Clausen, TU Dortmund (Dortmund, Germany)

Simaan Abourizk, University of Alberta (Canada)
Klaus Altendorfer, FH Oberösterreich (Steyr, Österreich)
Hans-Peter Barbey, FH Bielefeld
Hans-Joachim Bargstädt, Bauhaus-Universität Weimar
Jochen Bernhard, SSI Schäfer Noell GmbH (Dortmund)
Mathias Bös, SDZ (Dortmund)
Uwe Bracht, TU Clausthal-Zellerfeld
Peter Buchholz, TU Dortmund
Thorsten Claus, IHI Zittau
Uwe Clausen, TU Dortmund
Wilhelm Dangelmaier, Heinz Nixdorf Institut (Paderborn)
Volkhard Franz, Universität Kassel
Axel Hahn, OFFIS Oldenburg
Tobias Hegmanns, Fraunhofer-Institut für Materialfluss und Logistik IML (Dortmund)
Matthias Heinicke, Siemens Industry Software GmbH & Co. KG (Stuttgart)
Frank Herrmann, Hochschule Regensburg für angewandte Wissenschaften
Hironori Hibino, JSPMI Technical Research Institute (Tokio, Japan)
Angel Juan, Open University of Catalonia (Barcelona, Spanien)
Mark Junge, Limón GmbH (Kassel)
Jan Kaffka, TU Dortmund
Angeliki Karagiannaki, Athens University of Economics and Business (Athen, Griechenland)
Alexander Klaas, Miebach Consulting GmbH (Frankfurt am Main)
Markus König, Ruhr Universität Bochum
Christoph Laroque, Westsächsische Hochschule Zwickau
Lothar März, STREMLER AG (Lindau)
Gottfried Mayer, BMW Group (München)
Masaru Nakano, Keio University (Tokio, Japan)
Stefan Nickel, Karlsruher Institut für Technologie
András Pfeiffer, SZTAKI (Budapest, Ungarn)
Holger Pitsch, INCONTROL Simulation Solutions (Wiesbaden)
Carsten Pöge, Volkswagen AG (Wolfsburg)
Markus Rabe, TU Dortmund
Andrea Emilio Rizzoli, IDSIA (Manno, Schweiz)
Oliver Rose, Universität der Bundeswehr (München)
Andreas Schlegel, Fraunhofer IWU Chemnitz
Monika Schneider, Eisenmann AG (Holzgerlingen)
Anders Skoogh, Chalmers University of Technology (Göteborg, Schweden)
Sven Spieckermann, SimPlan AG (Maintal)
Dirk Steinhauer, Flensburger Schiffbau-Gesellschaft mbH und Co KG (Flensburg)
Patricia Stock, REFA-Institut e.V. (Dortmund)
Andrzej Szarata, Politechnika Krakowska (Krakau, Polen)
Simon Taylor, Brunel University (Uxbridge, UK)
Juri Tolujew, Fraunhofer IFF (Magdeburg)
Markus Vorderwinkler, PROFACTOR GmbH (Steyr-Gleink, Österreich)
Gerald Weigert, TU Dresden
Sigrid Wenzel, Universität Kassel
Enver Yucesan, INSEAD (Fontainebleau, Frankreich)

# Contents

**SIMULATION UND OPTIMIERUNG**
SIMULATION AND OPTIMISATION

Simulationsgestützte Optimierung der Feinplanung von synchronisierten Schmelzerei- und Gießereiaufträgen
*März, Lothar; Angelidis, Evangelos; Ohmann, Lars; STREMLER* ..... *1*

Simulationsbasierte Optimierung der Reihenfolgeplanung am Beispiel eines Liniensorters in der Automobilindustrie
*Heger, Jens; Leuphana Universität Lüneburg;*
*Hildebrandt, Torsten; BIBA Universität Bremen* ..... *11*

Flexible Job-shop Scheduling with Dynamic Stochastic Machine Sets
*Zhang, Tao; Shufang, Xie; Rose, Oliver;*
*Universität der Bundeswehr München* ..... *21*

Lösung eines Hybrid-Flow-Shop-Maschinenbelegungsproblems mit simulationsbasierter Optimierung
*Reggelin, Tobias; Tolujew, Juri; Nahhas, Abdulrahman; Aurich, Paul;*
*Fraunhofer IFF* ..... *29*

Reduction of Simulation Effort for Simulation-based Optimisation
*Uhlig, Tobias; Rose, Oliver; Universität der Bundeswehr München* ..... *39*

A Simheuristic for the Waste Collection Problem with Stochastic Demands in Smart Cities
*Gruler, Aljoscha; Universitat Oberta de Catalunya;*
*Fikar, Christian; Universität für Bodenkultur Wien;*
*Juan, Angel A.; Universitat Oberta de Catalunya;*
*Hirsch, Patrick; Universität für Bodenkultur Wien;*
*Bolton, Carlos Contreras; Universidad Andrés Bello* ..... *49*

Development of a Simheuristic Approach for Solving Realistic Inventory Routing Problems
*Pfeilsticker, Lena; TU Dortmund;*
*Juan, Angel A.; Universitat Oberta de Catalunya;*
*Rabe, Markus; TU Dortmund* ..... *59*

Simulation-based Optimisation in Industry 4.0
*Dunke, Fabian; Nickel, Stefan; Karlsruher Institut für Technologie* ..... *69*

Beitrag zur Abfallreduktion mittels mathematischer Optimierung und simulationsbasierter Validierung
*Janssen, Larissa; TU Dresden / IHI Zittau;*
*Herrmann, Frank; Hochschule Regensburg;*
*Claus, Thorsten; TU Dresden / IHI Zittau* ..... *79*

Sensitivitätsanalyse eines Kostenmodells zur Ableitung von Kriterien für die simulationsbasierte Optimierung von Planungsprozessen
*Kühn, Mathias; TU Dresden;*
*Angelidis, Evangelos; Bohn, Daniel; Universität der Bundeswehr München;*
*Völker, Michael; TU Dresden;*
*Zhou, Zhugen; Universität der Bundeswehr München* ..... *89*

Ein für die Wissenschaft freies Werkzeug zur Modellierung, Simulation und Optimierung
*Schönherr, Oliver; Winter, Elias; Fabig, Christian; TU Dresden;*
*Rose, Oliver; Universität der Bundeswehr München;*
*Völker, Michael; TU Dresden* ..... *101*

Verknüpfung von Simulation und Optimierung: Kategorien und Beispiele - Ein Bericht über die VDI Richtlinie 3633 Blatt 12
*Zisgen, Horst; SWZ Clausthal-Göttingen;*
*Hanschke, Thomas; TU Clausthal* ..... *111*

**LOGISTIK UND TRANSPORTNETZWERKE**
LOGISTICS AND TRANSPORTATION NETWORKS

Inhouse-Logistikplanung unter Verwendung von Ansätzen aus der Verkehrsflusssimulation
*König, Franz-Joseph; ZIP;*
*Heinrich, Marek; DLR;*
*Roth, Stefan; Volkswagen* ..... *119*

Ganzheitliche Supply-Chain-Optimierung mittels Kombination von mathematischer Optimierung und Simulation
*Klaas, Alexander; Klibi, Kamel; Miebach Consulting* ..... *129*

Simulative Planungsassistenz für die Disposition in mehrstufigen Distributionsnetzwerken
*Witthaut, Markus; Kamphues, Josef; Hegmanns, Tobias; Fraunhofer IML* ..... *137*

Konzeptionierung eines integrierten modellbasierten Ansatzes zur Prognose von transportlogistischen und intralogistischen Ereignissen in Logistiknetzwerken
*Schröder, Michael; Fraunhofer ITWM;*
*Beißert, Ulrike; Fraunhofer IML;*
*Jami, Neil; Fraunhofer ITWM;*
*Motta, Marco; Fraunhofer IML* ..... *147*

Simulation-based Decision Support for Vertical Supply Chain Segmentation Scenarios
*Terlunen, Sebastian; Horstkemper, Dennis; Wölck, Martin;*
*Hellingrath, Bernd; Universität Münster* ..... *157*

A Procedure Model for the Credible Measurability of Data Warehouse Metrics on Discrete-event Simulation Models of Logistics Systems
*Rabe, Markus; Dross, Felix; Vennemann, Andrea; TU Dortmund* ..... *167*

A Simulation-based Supply Chain Risk Analysis Framework
*Heckmann, Iris; FZI Forschungszentrum Informatik;*
*Nickel, Stefan; Karlsruher Institut für Technologie* ..... *177*

Meta-Modellierung und Werkzeug-Integration in der Simulation von Seehafen-Container-Terminals
*Lange, Ann-Kathrin; Jahn, Carlos; TU Hamburg-Harburg;*
*Pirovano, Giovanni; Politecnico di Milano;*
*Rossi, Tommaso; Università Carlo Cattaneo* ..... *187*

**INTRALOGISTIC**
INTRALOGISTICS

Measuring the Performance Availability of a Logistics Application Controlled by a Wireless Sensor Network
*Roidl, Moritz; Schieweck, Steffen; ten Hompel, Michael; TU Dortmund*..... *197*

Performance Availability Analysis of Autonomous Intralogistic Systems: An Agent-based Simulation Approach
*Güller, Mustafa; Karakaya, Elif; TU Dortmund;*
*Hegmanns, Tobias; Fraunhofer IML*..... *207*

Ein generisches Standardmodell zur Simulation spurgebundener Transportsysteme in der Halbleiterindustrie
*Rank, Sebastian; Hammel, Christian; Schmidt, Thorsten; TU Dresden;*
*Schneider, Germar; Infineon Technologies*..... *219*

Simulation-based Reduction of Traffic Jam Effects in a Distributed Automated Material Handling System
*Schmaler, Robert; Hammel, Christian; TU Dresden;*
*Schöps, Matthias; GLOBALFOUNDRIES;*
*Schmidt, Thorsten; TU Dresden* ..... *229*

Konfiguration von Kanallagern auf der Basis dynamischer Kanaltiefen mit Hilfe der Simulation
*Dietz, Thomas; Böttcher, Lars-Boris; Bernhard, Jochen; SSI Schäfer* ..... *239*

Teilautomatisierte Softwaretests mithilfe einer mechanisch plausiblen Robotersimulation
*Schenke, Björn; Aichele, Fabian; Eckstein, Bernd; TruPhysics*..... *249*

Verteilte Simulation und Emulation von Materialflusssystemen mit dezentraler Steuerung
*Daniluk, Damian; ten Hompel, Michael; Fraunhofer IML* ..... *259*

VR-basierte Simulation zur Ermittlung der Leistungsverfügbarkeit zellularer intralogistischer Systeme
*Eilers, Kevin; RIF e.V.;*
*Roßmann, Jürgen; RWTH Aachen*..... *269*

Mathematische Optimierung und Simulation von Paketsortieranlagen – Bessere Lösungen durch Kopplung der beiden komplementären Methoden
*Clausen, Uwe; Diekmann, Daniel; Baudach, Jens; Pöting, Moritz; TU Dortmund* ..... *279*

**MATERIALMANAGEMENT UND PRODUKTIONSPLANUNG**
MATERIAL MANAGEMENT AND PRODUCTION PLANNING

Anwendungen von Simulation zur Verbesserung von Prozessabläufen in Industriewäschereien
*Brandau, Annegret; Weigert, David; Universität Magdeburg; Tolujew, Juri; Fraunhofer IFF* ..... *289*

Prototyp eines simulationsbasierten Assistenzsystems zur Entscheidungsunterstützung bei der Pflege von Planungsparametern eines ERP-Systems im laufenden Betrieb
*Stumvoll, Ulrike; Ertl, Willi; Hochschule Regensburg; Claus, Thorsten; TU Dresden / IHI Zittau* ..... *299*

Bessere Prioritätsregeln für komplexe Produktionssysteme mittels multi-kriterieller simulationsbasierter Optimierung
*Hildebrandt, Torsten; Freitag, Michael; BIBA Universität Bremen* ..... *309*

Simulation zur Entwicklung und Dimensionierung von Schichtmodellen
*Mutzke, Harald; Simulation Expert Services; Blauermel, Gregor; B416 GmbH* ..... *319*

Material Flow Analysis at BASF - A Huge Variety in Complexity
*Feist, Markus; Fernández Alcalde, Ana Maria; BASF SE* ..... *329*

Joint Management of Material and Cash Flows
*Yucesan, Enver; INSEAD; Bendavid, Illana; Ort Braude College of Engineering; Herrer, Yale; TECHNION* ..... *337*

Inventory Management and Markup Pricing in the Presence of Price Fluctuations
*Canyakmaz, Caner; Ozekici, Suleyman; Karaesmen, Fikri; Koc University Istanbul* ..... *347*

Cost and Environmental Trade-offs in Supply Chain Network Design and Planning: The Merit of a Simulation-based Approach
*Keramydas, Christos; Mallidis, Ioannis; Vlachos, Dimitrios; Iakovou, Eleftherios; University of Thessaloniki* ..... *357*

Comparing Different Simulation-based Optimisation Approaches for Simultaneous Optimisation of Production Planning Parameters
*Felberbauer, Thomas; Schnirzer, Rene Kurt; Altendorfer, Klaus; FH Oberösterreich* ..... *367*

Influence of Customer Information Uncertainties on Production Order Variance Comparing two Different Kinds of Customer Order Behaviours
*Felberbauer, Thomas; Peirleitner, Andreas Josef; Altendorfer, Klaus; FH Oberösterreich* ..... *377*

Discrete Event Simulation of Modular Production System Models using Petri Nets
*Rabe, Markus; Deininger, Maik; TU Dortmund* ..... *387*

Scheduling under Consideration of the Machine Capability
*Doleschal, Dirk; Weigert, Gerald; TU Dresden* ..... *397*

**PRODUKTION UND DIGITALE FABRIK**
PRODUCTION AND DIGITAL FACTORY

Konsistente Simulationsmodelle für die virtuelle Inbetriebnahme fertigungstechnischer Anlagen mit Hilfe regelbasierter Modellverbinder
*Puntel-Schmidt, Philipp; Fay, Alexander; Universität Hamburg* ..... *407*

Ein Beitrag zum Lebenszyklusmanagement von Simulationsmodellen in der Digitalen Fabrik
*Völker, Sven; Hochschule Ulm;*
*Fischer, Jörg W.; Hochschule Karlsruhe* ..... *417*

A Case of Demonstrating Digital Factory Potentials through a Modular Approach
*Liotta, Giacomo; Aalborg University;*
*Jørgensen, Bent Aksel; Xcelgo;*
*Møller, Charles; Aalborg University* ..... *427*

Optimal Maintenance Resources Allocation Using Automated Simulation-based Optimisation and Data Management
*Ng, H.C. Amos; University of Skövde;*
*Skoogh, Anders; Chalmers University of Technology;*
*Lämkull, Dan; Volvo Car Corporation* ..... *437*

A Comparison of Discrete-event Simulation Approaches for Complex Manufacturing Systems and Healthcare Systems
*Fowler, John; Arizona State University;*
*Mönch, Lars; FernUniversität Hagen* ..... *447*

Einsatzmöglichkeiten der Simulation in einem cyber-physischen Produktionssystem
*Reuter, Christina; Brambring, Felix; Müller, Christiane;*
*WZL/RWTH Aachen* ..... *459*

Simulationstool zur alters- und belastungsorientierten Entwicklung und Bewertung von Arbeitssystemen in der industriellen Fertigung
*Müller, Ulf; Fachhochschule Köln;*
*Gust, Peter; Bergische Universität Wuppertal;*
*Feller, Nico; Technische Hochschule Köln;*
*Schiffmann, Michael; Fachhochschule Köln* ..... *469*

Abbildung von Steuerungslogiken durch maschinelles Lernen für die Simulation von Produktionssystemen
*Bergmann, Sören; Feldkamp, Niclas; Hinze, Ulrich; Straßburger, Steffen; TU Ilmenau* ..... *481*

**BAUSTELLENLOGISTIK**
CONSTRUCTION SITE LOGISTICS

Integration of Ergonomic Analysis into Simulation Modelling of Manual Operations
*Golabchi, Alireza; Han, SangUk; Abourizk, Simaan; University of Alberta* ..... *491*

Performing Scenario Simulations in Construction by Using Standard Software
*Bargstädt, Hans-Joachim; Feine, Immo; Universität Weimar* ..... *501*

GIL-Baustellensimulation - Ein entscheidungsunterstützendes Planungswerkzeug für Baustellen zur Errichtung kundenindividueller Anlagen
*Heck, Aron; Technische Hochschule Nürnberg; Habenicht, Ilka; SimPlan* ..... *513*

Prozesssimulation zur Analyse von Wartungsstrategien im maschinellen Tunnelbau
*Conrads, Alena; Scheffer, Markus; König, Markus; Thewes, Markus; Universität Bochum* ..... *523*

**SIMULATION UND ENERGIE**
SIMULATION AND ENERGY

Simulationsgestützte Planung und Bewertung der Energieeffizienz für Produktionssysteme in der Automobilindustrie
*Peter, Tim; Wenzel, Sigrid; Universität Kassel* ..... *535*

Energieeffizienzsteigerung im Tagebaubetrieb durch simulationsgestützte Untersuchung konkurrierender Flexibilitäten
*Fanghänel, Christin; Fraunhofer IWU;*
*Lange, Hans Rüdiger; Vattenfall Europe Generation;*
*Schlegel, Andreas; Stoldt, Johannes; Fraunhofer IWU;*
*Woldt, Thomas; Vattenfall Europe Generation;*
*Putz, Matthias; Fraunhofer IWU* ..... *545*

Ein Simulationsmodell zur Abbildung und Prognose der elektrischen Leistungsaufnahme in Seehafen-Containerterminals
*Grundmeier, Nico; Ihle, Norman; Hahn, Axel; Universität Oldenburg* ..... *555*

Allokation von Emissionswerten auf Behälterebene in multimodalen Umschlagsanlagen mittels Simulation
*Kaffka, Jan; Miodrag, Zoran; Clausen, Uwe; TU Dortmund;*
*Pitsch, Holger; INCONTROL Simulation Solutions* ..... *565*

Potentiale in der Reduzierung des Gesamtenergieverbrauchs einer Werkstattfertigung in der Maschinenbelegungsplanung
*Selmair, Maximilian; TU Dresden;*
*Claus, Thorsten; TU Dresden / IHI Zittau;*
*Herrmann, Frank; Hochschule Regensburg;*
*Teich, Enrico; TU Dresden* ..... *575*

Offshore-Windenergie: Kostensenkung durch Logistiksimulation
*Münsterberg, Torsten; Fraunhofer CML;*
*Jahn, Carlos; TU Hamburg-Harburg* ..... *585*

Vorstudie zur energetischen Modellierung automatischer Parksysteme
*Sinha, Katrin; Künne, Bernd; TU Dortmund* ..... *595*

**METHODEN UND UNTERSTÜTZUNGSSYSTEME**
METHODS AND AUXILIARY SYSTEMS

Geführte Assistenz für Ablaufsimulationsprojekte
*Mayer, Gottfried; Mieschner, Marielouise; BMW AG* ..... *605*

Aufbereitung von unvollständigen Daten für die simulationsbasierte Produktionsplanung im Schiffbau
*Steinhauer, Dirk; Flensburger Schiffbau-Gesellschaft;*
*Lödding, Hermann; Friedewald, Axel; Sikorra, Jan Niklas;*
*TU Hamburg-Harburg;*
*Haux, Moritz Alexander; Technische Universität Hamburg-Harburg* ..... *613*

Picker Blocking in manuellen Kommissioniersystemen – Eine simulationsbasierte Analyse
*Elbert, Ralf; Franzke, Torsten; Glock, Christoph H.; Grosse, Eric H.;*
*TU Darmstadt* ..... *621*

Avoiding Equidistances when Routing by Shortest Paths
*Hammel, Christian; TU Dresden;*
*Schöps, Matthias; GLOBALFOUNDRIES;*
*Schmaler, Robert; Schmidt, Thorsten; TU Dresden* ..... *631*

Correlated Random Number Generation for Simulation Experiments
*Bause, Falko; Kriege, Jan; TU Dortmund* ..... *641*

Implementation of Dynamic Biological Process Models into a Reference Net Simulation Environment
*Pettigrew, Liam; Hubert, Stefan; Groß, Frauke; Delgado, Antonio;*
*Universität Erlangen-Nürnberg* ..... *651*

A Natural-language-based Simulation Modelling Approach
*Mayer, Thomas; Pappert, Falk Stefan; Rose, Oliver;*
*Universität der Bundeswehr München* ..... *661*

Farming for Mining - Entscheidungsunterstützung mittels Simulation im Supply Chain Management
*Rabe, Markus; Scheidler, Anne Antonia; TU Dortmund* ..... *671*

Domänenspezifische Sprache für ein simulationsunterstütztes Supply-Chain-Design
*Parlings, Matthias; Fraunhofer IML; Sprenger, Philipp; TU Dortmund; Motta, Marco; Fraunhofer IML* ..... *681*

Entwicklung einer menschenähnlichen Puppe als Eingabegerät für die Menschsimulation
*Wiegmann, David; Hochschule Ostfalia; Brüggemann, Holger; Hochschule Ostfalia* ..... *691*

Agentenbasierte Simulation zur Abbildung von Intralogistikprozessen in der Grobblechindustrie zur Kampagnenoptimierung
*Hübl, Alexander; FH Oberösterreich; Strasser, Georg; voestapline Grobblech* ..... *701*

Agentenbasierte Planung und simulationsunterstützte Bewertung präferenzorientierter Einsatzzeiten des Personals
*Gamber, Thilo; Continental; Zülch, Gert; Karlsruher Institut für Technologie* ..... *711*

Ressourcenbeschränkte Terminplanung mit einem System kollaborativer Agenten
*Wenzler, Florian; Günthner, Willibald A.; TU München* ..... *721*

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulationsgestützte Optimierung der Feinplanung von synchronisierten Schmelzerei- und Gießereiaufträgen

## *Simulation-based Optimisation of Detailed Planning of Synchronised Orders in Melting Plants and Foundries*

Lothar März, Evangelos Angelidis, Lars Ohmann, Stremler AG, Lindau (Germany),
lothar.maerz@stremler.de, evangelos.angelidis@stremler.de,
lars.ohmann@stremler.de

**Abstract:** This paper describes a practical application in production planning and control in the automotive supplier industry. The company provides castings for passengers and commercial vehicles. The production process is very complex and can be divided in melting, moulding and finishing. Especially the processes in the melting plant and the foundries are closely related. A number of restrictions and mutual dependencies, which will be presented, make it difficult to establish a solid production plan. Therefore, a simulation based optimisation is introduced to find robust order sequences for the production areas. The methodology and the results in terms of calculation time and improved setup times are presented.

## 1 Einführung

Der Beitrag stellt eine Praxisanwendung in der Automobilzulieferindustrie vor, bei der in der Planung und Steuerung der Produktion eines Gießereiunternehmens die Methoden von Simulation und mathematischer Optimierung dazu eingesetzt werden, die hochkomplexen Zusammenhänge mit den anspruchsvollen Anforderungen in der Erfüllung von Kundenabrufen in Einklang zu bekommen.

Im täglichen Betriebsablauf sind die Zielsetzungen der einzelnen Abteilungen eines Unternehmens oftmals divergierend. Gegenläufige Anforderungen, die sich aus der Unstetigkeit der Marktabrufe und der angestrebten Stetigkeit der Produktion ergeben, müssen harmonisiert und synchronisiert werden. Marktanforderungen und Reaktivität der Lieferkette über mehrere Stufen stellen ein Spannungsfeld dar, das gezielt über die Leistungssteuerung ausgeglichen werden muss. Die Planung und Steuerung ist von entscheidender Bedeutung, um die Leistung zu maximieren und das Gesamtsystem effizienter zu gestalten. Mit innovativen Systemen ist die Steuerung der Produktion sogar in Echtzeit möglich.

Der technische Prozess des Gießens ist in viele Einzelschritte mit Zwangsfolge unterteilt, die nicht unterbrochen werden sollten. Die Formplatte erzeugt im Formsand innerhalb des Formkastens die Außenkontur des Gussteils, die später ausgegossen wird. In die Form der Außenkontur werden Kerne eingelegt, die später für den Hohlraum des Gussteils sorgen. Der Formkasten wird anschließend mit dem gewünschten Werkstoff ausgegossen. Dabei werden sowohl die Gussteile als auch Zuläufe erzeugt. Am Abschlagband werden abschließend die Gussteile von den Zuläufen getrennt. Nach dem Trennen werden die Gussteile gestrahlt und einer Sichtkontrolle unterzogen. Bestehen die Gussteile die Kontrolle, können sie im Anschluss weiteren Bearbeitungsschritten unterzogen oder direkt zum Kunden versendet werden.

Nachfolgend ist der grobe Prozess zu erkennen, der abhängig von der Gießanlage individuell abweichen kann.

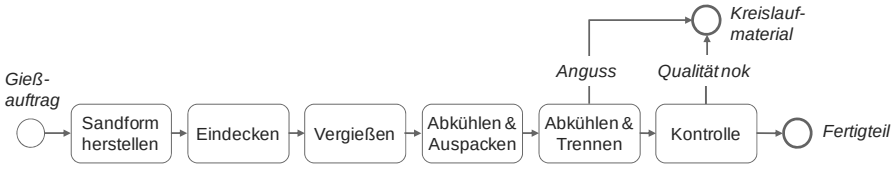



***Abbildung 1:*** *Technischer Prozess Gießen*

Im Zuge der Modernisierung der Gießerei ist der sukzessive Ersatz von drei Gießanlagen durch eine neue, weitaus modernere Gießanlage geplant. Diese geplante infrastrukturelle Maßnahme wurde zum Anlass genommen, die bestehende Organisation der Planung und Steuerung von Produktion und Logistik zu überdenken. Die Abstimmung der Planungs- und Steuerungsprozesse der Abteilung Schmelzen und Gießen wird von der eingesetzten Enterprise Resource Planning (ERP) Software nur rudimentär unterstützt. Die Detailplanung und –abstimmung erfolgt über Excel und unterliegt vielfältigen Einflussfaktoren, sodass aufgestellte Pläne bereits nach kurzer Zeit nicht mehr mit den realen Vorkommnissen in der Produktion übereinstimmen. Eine Beurteilung der getroffenen Planungsentscheidungen ist schwierig, da eine quantitative Bewertung der Auswirkungen auf Auslastung und Bestandsentwicklung einerseits und der Servicegrad-Erfüllung der Aufträge (Lieferfähigkeit und Termintreue) andererseits nicht transparent aufgezeigt werden.

Um die Potenziale in der effizienten Auftragsbearbeitung zu heben, hat das Gießereiunternehmen ein Teilprojekt des Infrastrukturprojektes gestartet, welches zum Ziel hat, einen Benchmark in Logistik und Feinplanung für die Konzerngruppe zu schaffen. Zur Sicherstellung der geplanten Effizienzsteigerung und Erreichung der strategischen Zielsetzungen strebt das Gießereiunternehmen an, die Prozesse und die Systemwelt zur Unterstützung der Supply Chain Management Funktionen im Hinblick auf die zukünftigen Herausforderungen weiterzuentwickeln. Diese Weiterentwicklung auf Basis einer optimierten Einplanung von synchronisierten Schmelz- und Gießereiaufträgen erfolgt in der Simulations- und Optimierungssoftware LOM.Cubes® und wird nachfolgend beschrieben.

# 2 Problem- und Zielbeschreibung einer Produktionsfeinplanung in einem Gießereiunternehmen

Die Gießerei ist der Bereich mit den meisten Restriktionen und komplexen dynamischen Wirkzusammenhängen. Dies ist der Hauptgrund, warum eine Planung und Steuerung der Gießerei nicht im ERP-System direkt erfolgt. Die Komplexität sprengt die Grundfähigkeiten jedes ERP-Systems und eine Erweiterung erscheint aufgrund der zu erwartenden Kosten nicht zweckmäßig. Im beschriebenen Beispiel werden eingehende Kundenabrufe zwar noch in Planaufträge umgewandelt, diese jedoch anschließend durch den Fertigungsplaner unter Zuhilfenahme von unterschiedlichen Excel-Dokumenten in Fertigungsaufträge umgewandelt und manuell in die Betriebsdatenerfassung des Unternehmens eingegeben.

Um die Komplexität der vom Fertigungsplaner zu berücksichtigenden Restriktionen und dynamischen Wirkzusammenhänge darzustellen, sollen an dieser Stelle sowohl die wichtigsten Restriktionen als auch Wirkzusammenhänge beschrieben werden.

Über die *Formplatte* müssen bereits eine Vielzahl von Restriktionen berücksichtigt werden. Ein überwiegender Anteil der Formplatten wird nur an einer Gießanlage eingesetzt, weshalb in diesen Fällen eine 1:1 Beziehung zwischen Gießanlage und Formplatte existiert. Eine kleine Menge an Formplatten kann jedoch mit Hilfe eines Adapters auch auf einer anderen Gießanlage laufen. Zusätzlich kann eine Formplatte auch eine Mischplatte sein. Das bedeutet, dass sich auf einer Platte unterschiedliche Modelle befinden. Ist es erforderlich ein Gussteil zu erzeugen, werden automatisch auch die anderen Gussteile der Platte mitgegossen. Dies kann im ungünstigsten Fall zu einer massiven Erhöhung der Bestände führen, sollten sich die abgerufenen Mengen der Artikel auf der Mischplatte unterscheiden.

Aus den Formplatten wiederum werden die *Formkästen* gebildet. Die Formplatten sind aufgrund von unterschiedlichen Größen jedoch nicht beliebig kombinierbar und müssen teilweise sogar vorgegebene Positionen innerhalb des Formkastens einnehmen.

Zur Erzeugung des Hohlraums im Gussteil werden die *Kerne* benötigt. Ungefähr 80 – 85 % der zu gießenden Teile benötigen Kerne. Somit stellen auch die Kerne eine Restriktion dar, die in der Planung berücksichtigt werden muss. Da es kein Lager für Kerne gibt, welches die benötigten Kerne für einen größeren Zeitraum (für mehr als eine Woche) aufnehmen könnte, werden Kerne zeitnah zum Gießen gefertigt. Außerdem unterliegt die Kernmacherei ebenfalls gewissen Restriktionen, sodass die Forderung an Kernen die Kapazitäten der Kernmacherei nicht überschreiten darf.

Die Kerne führen darüber hinaus zu einer weiteren Restriktion an einigen Gießanlagen, dem sogenannten *Eindeckwert*. Der Eindeckwert gibt die Anzahl der Mitarbeiter an, die für das Einsetzen der Kerne in die Formkästen vor dem Gießen benötigt werden. Dieser Eindeckwert sollte speziell bei der Kombination von mehreren Formplatten in einem Formkasten den Wert 7 nicht überschreiten.

Zu den Restriktionen, die bereits in der Optimierung der Fertigungsreihenfolge einfließen, müssen in der Simulation noch die dynamischen Wirkzusammenhänge berücksichtigt werden.

Als erster Wirkzusammenhang ist das *Gießgewicht* zu sehen. Das Gießgewicht eines einzelnen Formkastens stellt selbstverständlich keine Restriktion dar. Allerdings kann die Summation mehrerer Formkästen hintereinander und die Kombination mit

anderen Gießanlagen über einen spezifischen Zeitraum dazu führen, dass mehr Werkstoff benötigt wird, als durch die Schmelzerei nachproduziert werden kann. Zwar verfügt die Schmelzerei über eine Art Puffer, doch ist auch dieser nicht dauerhaft in der Lage einen hohen Werkstoffbedarf über einen längeren Zeitraum zu decken. Von daher gilt es, den Werkstoffbedarf der einzelnen Gießanlagen über die Zeit aufeinander abzustimmen.

Im Zusammenhang mit dem Gießgewicht ist auch der zu vergießende *Werkstoff* zu sehen. Dies hängt damit zusammen, dass auch die Menge eines spezifischen Werkstoffes über die Zeit limitiert ist. Zusätzlich müssen größere Sprünge in der Güte von zwei Werkstoffen nach Möglichkeit vermieden werden. Größere Sprünge führen ebenfalls zu Rüstzeiten, die es sonst nicht geben würde. Genannt sei an dieser Stelle der Anteil an Silizium und Kupfer im Werkstoff.

Die *Gießtemperatur* spielt bei allen Gießanlagen eine Rolle. Jedes Gussteil darf nämlich nur in einem spezifischen Temperaturfenster gegossen werden. Somit bietet es sich an, die Reihenfolge unterschiedlicher Gussteile so zu wählen, dass sich die Temperaturfenster überlappen. Ist dies nicht möglich, so müssen zusätzliche Zeiten für das Erwärmen/ Abkühlen eines Werkstoffes berücksichtigt werden, wodurch zusätzliche Rüstzeiten entstehen.

Als nächster dynamischer Wirkzusammenhang muss die *Abkühlzeit* betrachtet werden. Die meisten Gussteile kühlen während ihres Transportes durch die Halle ausreichend ab, um sie in den nachfolgenden Schritten direkt weiter bearbeiten zu können. Lediglich bei dickwandigen Gussteilen kann es vorkommen, dass diese während des verketteten Weitertransportes nicht ausreichend abkühlen. In diesem Fall muss der Takt, mit dem Formkästen gegossen werden, verlangsamt werden. Dies hat im Umkehrschluss Auswirkungen auf die Formkästen der vor- und nachgelagerten Artikel und dem Wirkungsgrad der Gießanlagen.

Zusätzlich zu den zuvor beschriebenen Restriktionen und dynamischen Wirkzusammenhängen, die für die gesamte Schmelzerei und Gießerei gelten, existieren an jeder einzelnen Gießanlage und an den Anlagen der Schmelzerei individuelle Restriktionen und dynamische Wirkzusammenhänge, auf die an dieser Stelle nicht näher eingegangen werden soll. Genannt sei hier zum Beispiel der Wirkungsgrad jeder Anlage, der einen ganz erheblichen Einfluss auf die in der Planung anzunehmende Fertigungsgeschwindigkeit hat.

Zur Lösung der Komplexität und damit der Ablösung der Planung in Excel soll die simulationsbasierte Optimierung eingesetzt werden. Dazu werden zukünftig die Planaufträge für die Schmelzerei und Gießerei von der ERP-Software automatisiert an das Simulations- und Optimierungssystem LOM.Cubes® übergeben. Dieses ermittelt unter Berücksichtigung aller Restriktionen und der dynamischen Wirkzusammenhänge einen optimierten Produktionsprogrammplan. Als Ergebnis werden die ermittelten Fertigungsaufträge mit Menge und Termin zur Ausführung an das ERP-System des Unternehmens zurückgespielt.

## 3 Beschreibung des Lösungskonzepts

Durch die Grobplanung im ERP-System erhält die Feinplanung Schmelzerei und Gießerei einen Produktionsprogrammvorschlag für die nächsten zwei Werktage sowie für die darauffolgenden vierzehn Kalendertage wie in Abbildung 2 dargestellt.

Die ersten zwei Werktage sind dabei eine Frozen Zone, die bereits in den Tagen zuvor fertig ausgeplant wurde und nur im Ausnahmefall angepasst werden sollte. Die anschließenden vierzehn Kalendertage sind frei disponierbar, wobei die Planungssicherheit mit abnehmendem Horizont zunehmen sollte.

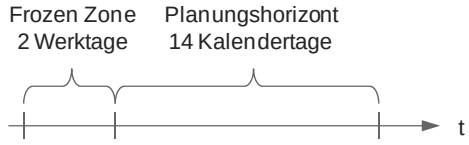



***Abbildung 2:*** *Planungshorizont Schmelzerei und Gießerei*

Da die Grobplanung nur eine zeitliche Einordnung aufgrund der hinterlegten Durchlaufzeit eines Artikels und einen einfachen Abgleich der Kapazitätsbedarfe der jeweiligen Gießanlage mit ihrer zur Verfügung stehenden Gesamtkapazität vornimmt (Dickersbach und Keller 2014), berücksichtigt die Feinplanung sämtliche Restriktionen und dynamischen Wirkzusammenhänge. Auch hier soll am Ende des Planungsprozess eine optimierte Produktionsprogrammplanung stehen, die in der Schmelzerei und Gießerei umgesetzt werden kann. Allerdings werden auch Engpässe an die Grobplanung zurückgemeldet, die mit den Kapazitäten der Schmelzerei und Gießerei nicht durchgeführt werden können.

Jeder Auftrag innerhalb der Produktionsprogrammplanung enthält dabei folgende Informationen:

- Artikelnummer
- Spätester interner Liefertermin
- Stückzahl

Anschließend werden über die Artikelnummer die für die Optimierung entscheidenden Stammdaten ausgelesen bzw. errechnet. Errechnet werden muss über die Stückzahl je Planauftrag und die Stückzahlen pro Modellplatte die erforderliche Anzahl von abzuformenden Kästen je Fertigungssauftrag. Die Einheit Kasten ist die der Optimierung und Simulation zu Grunde liegende Recheneinheit. Die zeitlichen Grenzen der Simulation erhält man über den Liefertermin des Auftrags, von dem die nachfolgenden Durchlaufzeiten subtrahiert und damit auf den spätesten Gießtermin rückgerechnet werden. Über die Artikelnummer des Rohteils, die Kastenzahl und das Gießgewicht kann der Werkstoffbedarf berechnet werden. Der grundlegende Ablauf des Datenmanagements für die Feinplanung in der Gießerei kann der Abbildung 3 entnommen werden.

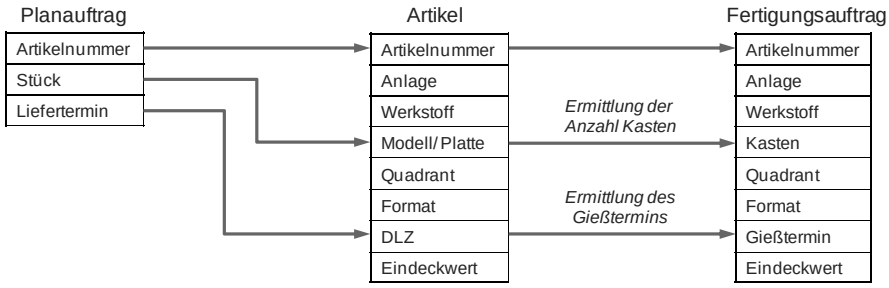



***Abbildung 3:*** *Logik zur Ermittlung der Arbeitsdaten*

In der Feinplanung muss jede Gießanlage individuell geplant werden. Dies geschieht dadurch, dass im Vorfeld über die Artikelnummer die Aufträge einer speziellen Gießanlage zugeordnet werden. Für Artikel, die mit Hilfe eines Adapters auf mehr als einer Gießanlage gelegt werden können, erfolgt die Auswahl durch den Planer. Die Aufträge werden je Gießanlage zusammengefasst. Erst nach Planung jeder einzelnen Gießanlage werden diese in Gesamtheit gegenüber den Restriktionen, die für alle Gießanlagen gelten, abgeglichen.

Bei der Vorgehensweise zur Feinplanung sind folgende Aspekte bei allen Gießanlagen gleichermaßen zu behandeln. Zuerst wird über den Liefertermin und die Durchlaufzeit der späteste Gießtermin für jeden Auftrag sowie die minimal erforderliche Kastenzahl errechnet. Eine Spezialität stellen bei dieser Berechnung die Mischplatten dar. Muss zur Erfüllung eines Auftrags eine Mischplatte herangezogen werden, gilt es im Gegenzug auch zu prüfen, ob ein Auftrag für den oder die zusätzlichen Artikel auf der Mischplatte existiert und mit verrechnet werden kann. Der Artikel, dessen Stückzahl laut Auftrag die größte Anzahl von Kästen erfordert, ist dabei maßgeblich. Existieren keine Abrufe zu dem oder den Gegenartikeln auf der Mischplatte, werden diese trotzdem mitgegossen. An dieser Stelle wird jedoch eine Warnmeldung an den Bediener ausgegeben, damit dieser die Möglichkeit zur Klärung dieser Problematik erhält.

Über die Artikelnummer des Auftrags werden je Gießanlage weitere Informationen aus den Stammdaten gezogen. Dies sind:

- Stück pro Kasten (Modell/ Platte)
- Werkstoff
- Mischplatte
- Gießgewicht

Über die Anzahl der Kästen gleichen Typs und das Gießgewicht wird der Werkstoffbedarf berechnet. Das benötigte Werkstoffgewicht darf nicht größer werden als die maximale Menge, die an die jeweilige Anlage geliefert werden kann. Sind diese Vorarbeiten geleistet, können die Aufträge und mit ihnen die zugehörigen Kästen in eine Reihenfolge gebracht werden. Dabei wird berücksichtigt, dass die Werkstoffe möglichst selten gewechselt werden und der Sprung von einem Werkstoff auf den anderen so gering wie möglich ausfällt. Es wird so geplant, dass größere, nicht zu vermeidende Sprünge über einen freien Kalendertag erfolgen. Ist dies nicht möglich, werden solche Sprünge innerhalb der ersten 6 Stunden einer Schicht durchgeführt, da in diesem Falle die Pause hierzu genutzt werden kann. Berücksichtigt wird auch,

dass die spätesten Gießtermine nicht verletzt werden. Die ersten beiden Werktage der Planung stellen dabei eine Frozen Zone dar, die nicht oder nur in Ausnahmefällen angepasst wird. Die nächsten 14 Kalendertage sind dann optimiert. Es entsteht ein Vorschlag für ein Produktionsprogramm der folgenden zwei Werktage und der anschließenden 14 Kalendertage.

## 4 Vergleich von Szenarien mit optimierten Sequenzen und dynamischen Kapazitätsanalysen

Der Ablauf zur Ermittlung bestmöglicher Auftragsreihenfolgen folgt der Logik von iterativen Näherungsprozessen. Zunächst wird für jeden Produktionsbereich eine Liste von möglichen Auftragsreihenfolgen unter Berücksichtigung der vorab erläuterten Restriktionen und Terminvorgaben ermittelt. Dazu wird ein heuristischer Optimierungsalgorithmus verwendet. In der Praxis zeigt sich, dass es nicht immer eine beste Lösung geben muss, sondern dass es eine Reihe von Alternativen geben kann, die ebenso gut die Zielvorgaben zu erreichen versprechen. Manchmal ist eine leicht schlechtere Lösung eine zu bevorzugende Alternative, da sie gegenüber Prozesszeitschwankungen oder sonstigen Fertigungseinflüssen weniger sensitiv und somit robuster gegenüber Störungen ist.

Nach Auswahl einer Auftragsreihenfolge erfolgt die Simulation, die die dynamischen Wirkzusammenhänge bei der Abbildung der Auftragsabwicklungsprozesse im Produktionsbereich berücksichtigt. Dazu greift das Modell auf die im ERP-System hinterlegten Vorgabezeiten (Rüst- und Bearbeitungszeiten) zurück. Das Simulationsmodell bildet das Ablaufverhalten der Material- und Informationsflüsse ab und ist so aufgebaut, dass alternative Konfigurationen der Produktionsfaktoren von beispielsweise Kapazitäten (Anzahl Mitarbeiter, Qualifikationsmatrizen, Schichtmodelle) oder Arbeitsplandaten (Leistungsdaten, alternative Arbeitspläne, Rüstmatrizen) änderbar sind. Das Ergebnis der dynamischen Kapazitätsanalyse durch Simulation wird anschließend bewertet.

Die Bewertung erfolgt hinsichtlich der logistischen Zielvorgaben, die sich aus den Zielgrößen von Termintreue, Durchlaufzeiten, Beständen und Auslastung der Anlagen zusammensetzen. Je nach Priorisierung von Service und Kosten kann Einfluss auf die Auftragsabwicklung genommen werden: im Falle maximaler Termintreue ist damit zu rechnen, dass höhere Rüstaufwände und eine niedrigere Auslastung die Folge ist. Demgegenüber ist bei der Verfolgung einer hohen Auslastung mit höheren Beständen und Kosten zu rechnen.

Eine gute Lösung kann durch den Vergleich unterschiedlicher Szenarien ermittelt werden. Dazu wird zunächst in einer ersten “Optimierungsschleife“ die möglichen Sequenzvorschläge der Optimierung durchsimuliert und miteinander verglichen. Die Konfiguration des Simulationsmodells entspricht der Produktionskonfiguration, die der Planer einsetzen möchte. Sollten die Ergebnisse dieser ersten Optimierungsrunde nicht zufriedenstellend sein, so kann der Planer an verschiedenen Stellschrauben drehen:

- Verschiebung der Gewichtung von Zielgrößen
- Anpassung der Kapazitäten, z. B. durch Verlängerung der Arbeitszeiten bzw. Einsatz zusätzlicher Schichten

- Manipulation der Terminvorgaben von Aufträgen, z. B. Verschieben eines Fertigstellungstermins um eine Woche nach hinten
- Verändern von Produkt – Ressourcenbeziehungen wie beispielsweise dem Einsatz externer Kapazitäten oder alternativer Arbeitspläne
- Usw.

Der Planer kann sein Wissen über mögliche Konfigurationsmöglichkeiten nutzen und dem Planungssystem durch einfache Parametrierung vermitteln. Dadurch verändert sich der Suchraum für die simulationsgestützte Optimierung und die Chance, ein für die gegebene Auftragslast optimales Planungsergebnis zu finden. Sofern der Planer einer der genannten Parameteränderungen vorgenommen hat, durchläuft der Planungsablauf wieder die erste Optimierungsschleife. In Abbildung 4 sind die Zusammenhänge der iterativen Planungsdurchläufe zueinander dargestellt.

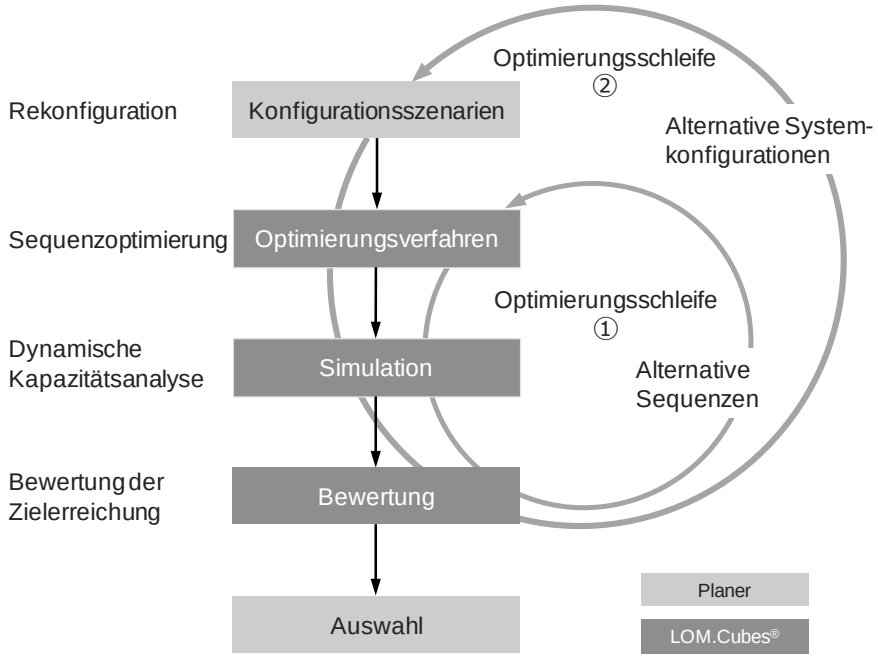



***Abbildung 4:*** *Iterativer Planungsablauf mit LOM.Cubes®*

Der Abgleich der Planungsergebnisse mit den vor- und nachgelagerten Produktionsbereichen erfolgt durch den Vergleich der Terminvorgaben der Aufträge mit den Terminvorschlägen aus den Einzeloptimierungen. Im Falle von Verletzungen sind produktionsübergreifend weitergehende Planungsschritte zu initiieren.

## 5 Ergebnisse

Die sehr komplexen und teilweise zeitlich abhängigen Bedingungen erfordern jahrelange Erfahrung, um Einplanungen von Aufträgen vornehmen zu können, die in der Praxis auch umsetzbar sind. Aber auch dann greifen erfahrene Planer auf bewährte Ablaufkonfigurationen zurück, da eine vollständige Durchrechnung der Auswirkun-

gen eines Planungsvorschlages weder zeitlich noch praktisch durchführbar ist. Somit bleibt unklar, welches Potential an möglichen Einsparungen bei gleichbleibender Leistungserfüllung gehoben werden könnte.

Die ersten Ergebnisse sind vielversprechend. In Abbildung 5 wird aufgezeigt, welche Verbesserungen durch eine systembasierte Planungsunterstützung mit LOM.Cubes® im Vergleich zu den Ergebnissen einer manuellen Planung möglich ist. Dargestellt sind die Rüstvorgänge zwischen 5 Rüstzuständen, deren Aufwände durch den Abstand auf der Ordinate in Relation zueinander veranschaulicht werden. Der Ausschnitt zeigt die Reihenfolge von Güten, die in der Schmelzerei für die nachfolgenden Gießanlagen zeitbezogen bereitgestellt werden müssen. Insgesamt konnten die Rüstaufwände deutlich reduziert werden ohne Verlust der Akzeptanz in der Umsetzbarkeit der vorgeschlagenen Lösung.

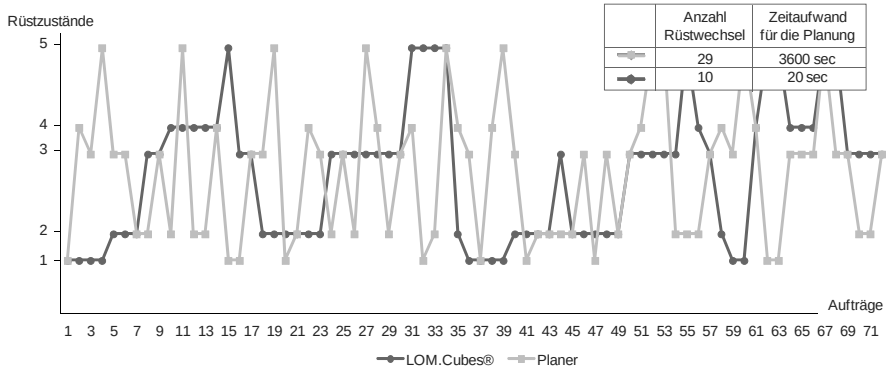



***Abbildung 5:*** *Rüst- und Planungsaufwände im Vergleich*

Gegenüber dem Planer, der für die Lösung dieser Planungsaufgabe eine Stunde benötigt, steht mit dem Einsatz von LOM.Cubes® mehr oder minder unmittelbar ein Ergebnis zur Verfügung. Dies ist insbesondere von Vorteil, wenn aufgrund von unvorhergesehenen Einflüssen wie beispielsweise Produktionsstörungen oder Eilaufträge eine Neuplanung erforderlich wird.

## 6 Zusammenfassung und Ausblick

Die Automobilindustrie erwartet von seinen Lieferanten termin- und mengengerechte Lieferungen. Dazu ist die Auftragsabwicklungskette der Zulieferer so auszurichten, dass flexibel und kurzfristig auf die Abrufe reagiert werden kann. Um den Forderungen nach Flexibilität und Reaktionsfähigkeit zu genügen und gleichzeitig mit wettbewerbsfähigen Preisen zu agieren, sind Mechanismen zur kontinuierlichen Anpassung so effizient wie möglich zu gestalten.

Die Synchronisation der Planungs- und Auftragsabwicklungsprozesse erfordert eine hohe Integration der Modelle. Eine integrierte, systemgestützte Planung erlaubt die effiziente Erfüllung von Kundenabrufen bei volatilem Marktverhalten und Störeinflüssen in Produktion und Logistik.

Dabei soll nicht der rein systembasierten Planung der Mund geredet werden, sondern im Gegenteil: der Planer steht im Mittelpunkt und soll sein Wissen über mögliche Konfigurationsalternativen einbringen können. Die Durchrechnung und Überprüfung der Auswirkungen dieser Konfigurationen erfolgt im Rechner. Damit kann sich der Planer auf den eigentlichen, kreativen Planungsprozess konzentrieren, während die repetitiven Berechnungen dem Computer überlassen werden können. Das effiziente Zusammenspiel von Frage und Antwort in der Mensch-Maschine-Schnittstelle wird durch ergonomische Bedienoberflächen und zeitnaher Rückmeldungen durch das System gefördert. Der in diesem Beitrag aufgezeigte Anwendungsfall zeigt auf, wie eine solche Unterstützung in der Feinplanung von Schmelz- und Gießereiaufträgen durch eine simulationsgestützte Optimierung nutzbringend erfolgen kann.

## Literatur

Dickersbach, J.T.; Keller, G. (2014) Produktionsplanung und –steuerung mit SAP ERP®. Galileo Press, Bonn

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulationsbasierte Optimierung der Reihenfolgeplanung am Beispiel eines Liniensorters in der Automobilindustrie

## *Simulation-based Scheduling for a FIFO Line Sorter in a Car Manufacturing Scenario*

Jens Heger, Leuphana Universität Lüneburg, Lüneburg (Germany), Jens.Heger@leuphana.de

Torsten Hildebrandt, BIBA – Bremer Institut für Produktion und Logistik GmbH an der Universität Bremen, Bremen (Germany), hil@biba.uni-bremen.de

**Abstract:** This study analyses different scheduling methods for a FIFO multiple line buffer in a scenario with sequence-dependent setup times. A simple selection rule working after the shortest-setup-time-next principle is compared to an approach with a genetic algorithm and a MILP-Solver combination. Although the scenario is np-complete the occurring small instances can be solved in acceptable time. The scenario is implemented in two different simulation systems (Tecnomatix Plant Simulation and jasima). Results show that the rule-based scheduling is outperformed by both other algorithms, where the MILP-Solver achieves the highest improvement in mean flowtime with over 26%.

## 1 Einleitung und Szenariobeschreibung

Die in dieser Studie betrachtete Problemstellung ist aus der Automobilindustrie abgeleitet und betrachtet einen Liniensorter, der aus mehreren parallelen FIFO-Warteschlangen besteht und einer Arbeitsstation (hier: Lackierstation) vorgeschaltet ist (Eley 2012). Der Liniensorter wird eingesetzt, da an der Arbeitsstation reihenfolgeabhängige Rüstzeiten anfallen und diese zur Verbesserung der mittleren Durchlaufzeit reduziert werden sollen. Es werden 20 verschiedene Rüstzustände (bspw. Farben) angenommen, die je nach Vorgänger eine andere Rüstzeit benötigen.

Es gibt zwei Entscheidungspunkte, an denen unterschiedliche Strategien angewendet werden können. Einerseits werden neu eintreffende Karosserien einer Linie zugeordnet und dort eingelagert und anderseits kann die Arbeitsstation aus den fünf Linien eine der jeweils ersten Karosserien auswählen. Eley (2012) hat dazu bereits einfache Greedy-Heuristiken vorgestellt und diese mit einem Genetischen Algorithmus in einer Simulationsstudie verglichen. Die Einlastungsstrategien sind so gewählt, dass einerseits zufällig eine Verteilung vorgenommen wird und andererseits

eine zweite Strategie angewendet wird, die die Karosserien wenn möglich in eine leere Warteschlange einlagert. Sind in allen Warteschlangen bereits Karosserien vorhanden, wird die Linie ausgewählt bei der zwischen der letzten Karosserie und der einzusortierenden die geringste Rüstzeit benötigt wird.

Als erste Auslagerungsstrategien wird eine einfache Regel angewendet, die jeweils die Karosserie auswählt, die zur zuletzt bearbeiteten die geringste Rüstzeit aufweist. Die zweite Auslagerungsstrategie betrachtet das Ausgangsproblem als kombinatorisches Optimierungsproblem, das mit einem Genetischen Algorithmus gelöst wird. Ziel ist es, die Rüstzeiten und damit auch die Durchlaufzeiten zu minimieren.

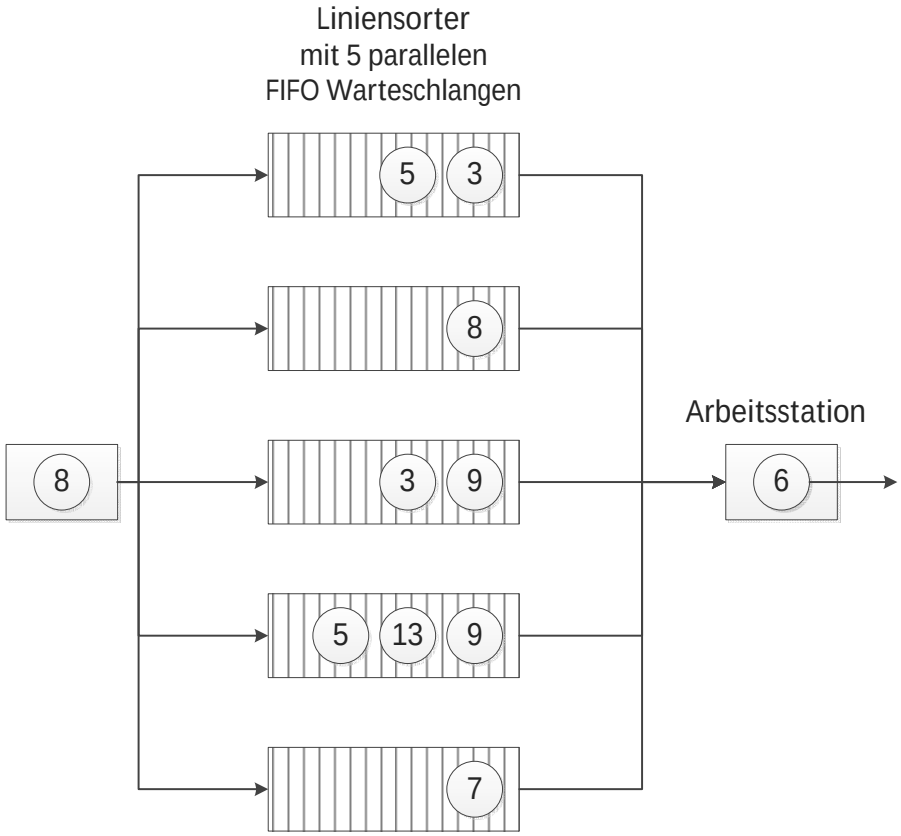



***Abbildung 1:*** *Paralleler Liniensorter bestehend aus FIFO-Warteschlangen; Zahlen stellen den Rüstzustand dar (Eley 2012)*

Das betrachtete Szenario ist in Abbildung 1 dargestellt. Beispielhaft sind einige Aufträge in die Warteschlagen eingereiht worden. Die Arbeitsstation kann, nachdem sie den Auftrag mit Rüstzustand 6 abgearbeitet hat, zwischen den jeweils vorderen Aufträgen einen als nächsten auswählen (3, 8, 9, 9, 7). Die dafür anfallenden Rüstkosten sind in der nachfolgenden Tabelle 1 aufgeführt.

***Tabelle 1:*** *Rüstzeiten in Sekunden*

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 0 | **237** | 353 | **324** | 91 | **274** | 434 | **311** | 102 | **214** | 228 | **64** | 253 | **103** | 170 | **162** | 196 | **108** | 167 | **371** |
| **2** | 126 | **0** | 181 | **234** | 125 | **64** | 355 | **146** | 227 | **187** | 99 | **98** | 143 | **228** | 205 | **268** | 282 | **142** | 173 | **134** |
| **3** | 259 | **197** | 0 | **53** | 94 | **235** | 395 | **316** | 62 | **369** | 296 | **178** | 95 | **59** | 284 | **122** | 205 | **222** | 128 | **317** |
| **4** | 346 | **284** | 185 | **0** | 151 | **322** | 470 | **396** | 149 | **275** | 383 | **124** | 42 | **244** | 231 | **210** | 335 | **168** | 163 | **418** |
| **5** | 169 | **307** | 151 | **204** | 0 | **217** | 334 | **298** | 213 | **165** | 245 | **233** | 174 | **210** | 339 | **143** | 184 | **277** | 148 | **441** |
| **6** | 61 | **298** | 159 | **112** | 152 | **0** | 372 | **82** | 163 | **276** | 289 | **126** | 79 | **164** | 232 | **223** | 257 | **85** | 200 | **432** |
| **7** | 294 | **303** | 204 | **19** | 170 | **325** | 0 | **328** | 168 | **253** | 321 | **143** | 60 | **170** | 160 | **228** | 316 | **187** | 182 | **200** |
| **8** | 420 | **210** | 219 | **118** | 270 | **275** | 363 | **0** | 217 | **393** | 309 | **243** | 85 | **275** | 349 | **100** | 70 | **3** | 106 | **345** |
| **9** | 197 | **135** | 259 | **285** | 260 | **135** | 252 | **217** | 0 | **83** | 234 | **233** | 192 | **299** | 340 | **60** | 286 | **220** | 66 | **269** |
| **10** | 65 | **95** | 163 | **72** | 156 | **4** | 376 | **85** | 167 | **0** | 193 | **129** | 38 | **168** | 236 | **186** | 155 | **89** | 160 | **229** |
| **11** | 371 | **308** | 101 | **155** | 195 | **150** | 329 | **124** | 163 | **247** | 0 | **276** | 134 | **96** | 380 | **224** | 193 | **127** | 75 | **329** |
| **12** | 195 | **245** | 178 | **222** | 27 | **154** | 361 | **236** | 215 | **150** | 272 | **0** | 189 | **116** | 106 | **141** | 110 | **44** | 146 | **379** |
| **13** | 379 | **317** | 144 | **33** | 185 | **197** | 376 | **279** | 183 | **266** | 416 | **158** | 0 | **143** | 264 | **243** | 252 | **202** | 122 | **452** |
| **14** | 239 | **176** | 251 | **221** | 205 | **176** | 293 | **258** | 42 | **125** | 275 | **178** | 134 | **0** | 284 | **102** | 146 | **214** | 107 | **311** |
| **15** | 249 | **186** | 261 | **231** | 215 | **186** | 303 | **268** | 51 | **135** | 285 | **188** | 144 | **10** | 0 | **112** | 156 | **224** | 117 | **321** |
| **16** | 88 | **118** | 32 | **65** | 126 | **27** | 191 | **108** | 94 | **23** | 216 | **152** | 32 | **26** | 258 | **0** | 136 | **112** | 5 | **252** |
| **17** | 193 | **123** | 217 | **233** | 232 | **187** | 320 | **115** | 68 | **151** | 222 | **205** | 160 | **27** | 310 | **129** | 0 | **118** | 134 | **257** |
| **18** | 185 | **189** | 129 | **162** | 223 | **124** | 288 | **181** | 135 | **120** | 288 | **249** | 129 | **93** | 355 | **97** | 66 | **0** | 102 | **324** |
| **19** | 137 | **198** | 27 | **80** | 121 | **75** | 254 | **157** | 63 | **146** | 296 | **199** | 121 | **21** | 305 | **123** | 130 | **160** | 0 | **332** |
| **20** | 118 | **148** | 47 | **100** | 141 | **57** | 221 | **138** | 109 | **53** | 120 | **182** | 71 | **106** | 288 | **169** | 208 | **142** | 175 | **0** |

Im Folgenden wird kurz der Stand der Technik dargestellt und auf Vorarbeiten speziell zu Szenarien mit Rüstzeiten und Prioritätsregelsteuerung eingegangen. Anschließend werden der hier untersuchte Ansatz und die Simulationsstudie beschrieben. Dabei wird auch auf die unterschiedliche Simulationssoftware eingegangen. Nachfolgend werden die Ergebnisse der Studie aufgezeigt und diskutiert. Der Beitrag schließt mit einer Zusammenfassung und einem Ausblick zu zukünftigen Forschungsfragen.

## 2 Bekannte Ansätze für die Reihenfolgeplanung mit reihenfolgeabhängigen Rüstzeiten

In diesem Abschnitt werden zunächst allgemeine Ansätze zur Reihenfolgeplanung mit reihenfolgeabhängigen Rüstzeiten dargestellt. Anschließend wird auf Vorarbeiten und Implementierungsdetails zu dem hier untersuchten Szenario eingegangen.

## 2.1 Allgemeine Ansätze zur Reihenfolgeplanung mit reihenfolgeabhängigen Rüstzeiten

Das Problem der Reihenfolgeplanung mit reihenfolgeabhängigen Rüstzeiten wird schon seit langer Zeit aufgrund seiner hohen Praxisrelevanz untersucht. Eine Übersicht über vorgestellte Ansätze auf Basis von Prioritätsregeln geben etwa (Pickardt und Branke 2011). Je nach Zielkriterium (bspw. mittlere Durchlaufzeit, Verspätung etc.) führen verschiedene Regel zum besten Ergebnis. Weiterhin spielt auch das Verhältnis der Rüstzeit zur Bearbeitungszeit des Szenarios eine entscheidende Rolle. Beste Ergebnisse lieferten die Regeln FE[MMS; SPT] (van der Zee 2010), SNSPT (Kochhar und Morris 1987), ACTS (Lee und Pinedo 1997) und ECR (Chiang und Fu 2009). Eine weitere Verbesserung kann durch die dynamische Anpassung der ACTS-Regel auf Basis vorgelagerter Simulationen und der Leistungsprognose mithilfe von Gaußscher Prozesse Regression erreicht werden (Heger et al. 2015; Heger 2014).

Im Gegensatz zu den vorgestellten dezentralen regelbasierten Steuerungsverfahren existieren zentrale Ansätze, die sämtliche Informationen berücksichtigen. Da es sich aber bereits beim Ein-Maschinen-Problem mit reihenfolgeabhängigen Rüstzeiten um ein NP-vollständiges Problem handelt (Monma und Potts 1989), können optimale Verfahren, wie beispielsweise MILPs beziehungsweise mathematische Solver, nur von kleinen Instanzen Lösungen berechnen. Daher muss genau untersucht werden, ob ihr Einsatz in einem solchen Szenario in der Praxis möglich ist. Alternativ werden diese Verfahren häufig eingesetzt, um die Leistungsfähigkeit von nicht optimalen Heuristiken zu beurteilen (Kurz und Askin 2004; Naderi und Salmasi 2012). Als Heuristik bietet sich beispielsweise das Simulated Annealing an (Tan und Narasimhan 1997).

Alle diese beschrieben Ansätze und Untersuchungen haben gemein, dass sie von nur einem Puffer ausgehen. Es werden zwar teilweise parallele Maschinen untersucht, aber keine parallelen Puffer. Weiterhin wird davon ausgegangen, dass sämtliche Aufträge innerhalb des Puffers ausgewählt werden können.

## 2.2 Untersuchung einer Lackieranlage in der Automobilindustrie

Das in dieser Studie ausgewählte Szenario wurde bereits von Eley (2012) untersucht. In seiner Simulationsstudie hat er dazu bereits zwei Einlagerungsstrategien und zwei Auslagerungsstrategien miteinander verglichen. Die Einlagerungsstrategie teilt die ankommenden Karosserien auf die Linienpuffer auf, wohingegen die Auslagerungsstrategien eine Karosserie für die Bearbeitung aus einem der Linienpuffer auswählt. Bei der Einlagerungsstrategie stellte sich heraus, dass es am günstigsten ist, die ankommende Karosserie in einen freien Puffer einzulagern, alternativ den auszuwählen, bei dem die geringste Rüstzeit zu der vorherigen Karosserie anfällt. Die zufällige Auswahl eines Linienpuffers führt zu schlechteren Ergebnissen bezüglich der Durchlaufzeit und der benötigten Rüstzeiten.

Weiterhin wurden zwei Auslagerungsstrategien untersucht. Die erste Regel wählt die Karosserie aus, die die geringste Rüstzeit zur zuletzt bearbeiteten Karosserie aufweist. Die zweite Auslagerungsstrategie reduziert das Entscheidungsproblem auf das des Handlungsreisenden (TSP: Travelling Salesman Problem), indem die Rüstzeiten als Entfernungen zwischen den verschiedenen Rüstzuständen interpretiert

werden. Das TSP wird mit einem genetischen Algorithmus (GA) gelöst, der bei jeder Auswahl einer Karosserie aufgerufen wird, also sobald die Lackieranlange wieder zur Verfügung steht.

Die Implementierung dieser Studie hat Eley (2012) in *Tecnomatix Plant Simulation* der Firma Siemens vorgenommen, und auch den dort integrierten Baustein des genetischen Algorithmus verwendet. Es zeigte sich, dass die Einlagerung nach Rüstzeitendifferenz und die Auslagerung mithilfe des genetischen Algorithmus die besten Ergebnisse lieferten. Für diese Studie haben wir die Einstellungen übernommen, die zu besten Ergebnissen geführt haben (Crossover-Parameter = Partially Matched Crossover (PMX); siehe Eley 2012, Experiment 4).

# 3 Simulationsstudie und untersuchter Ansatz

Ziel dieser Studie ist es einen verbesserten Ansatz zur Reihenfolgeplanung für das hier dargestellte Szenario zu finden. Dazu haben wir ein mathematisches Modell (MILP) entwickelt, das wir in Kombination mit einem Solver als Alternative zum genetischen Algorithmus einsetzen wollen, um so mit optimalen Lösungen des Teilproblems bessere Ergebnisse zu erzielen. Dies erfordert die Kopplung des Simulationsmodells beziehungsweise der Simulationssoftware mit dem Solver, da regelmäßig bei der Fertigstellung eines Auftrags ein neuer Reihenfolgeplan berechnet werden muss. Es gilt zu untersuchen, ob der Solver die auftretenden Instanzen in vertretbarerer Zeit, also innerhalb von Sekunden oder wenigen Minuten, lösen kann.

Da es in unserer Fragestellung weniger auf die visuelle Darstellung der Abläufe ankommt, sondern vielmehr um die Entscheidungsfindung und die Steuerung der Anlage geht, haben wir uns für eine alternative Simulationssoftware mit hoher Geschwindigkeit und geeigneten Schnittstellen entschieden. Die java-basierte Simulationssoftware *jasima*® (JAva SImulator for MAnufacturing and logistics ;jasima solutions UG) ist effektiv in der Ausführung und bietet eine leichte Integration insbesondere in Java-Projekte. Als Solver wurde *Gurobi* von Gurobi Optimization eingesetzt, der sich durch seine hohe Leistungsstärke auszeichnet und ebenfalls eine Java-Schnittstelle besitzt. So ist es leicht möglich das entwickelte Modell in der Praxis einzusetzen und in bereits bestehende Steuerungssoftware zu integrieren.

Das hier verwendete MILP stellt eine Erweiterung der Modellierung des TSP dar, indem zusätzliche Reihenfolgebedingungen ergänzt wurden, die die verwendeten FIFO Puffer-Bedingungen sicherstellen.

## 3.1 Analyse des Simulationsmodells und alternative Implementierung

Für das Simulationsmodell werden einige Annahmen getroffen: die Karosserien treffen negativ exponentialverteilt mit einem Erwartungswert von $1/\lambda = 300$ Sekunden ein, der Abstand beträgt dabei mindestens 60 und höchstens 12000 Sekunden. Es werden 20 verschiedene Rüstzustände, hier Farben, angenommen, die gleichverteilt auftreten. Die Rüsttabelle ist in Tabelle 1 dargestellt. Die Bearbeitungszeit auf der Lackierstation beträgt jeweils 5 Minuten. Eley (2012) hat in seiner Untersuchung einen Zeitraum von 8 Stunden betrachtet.

Bei der Implementierung des Szenarios in *jasima* stellte sich heraus, dass die in *Plant Simulation* berechneten Werte beispielsweise für die Durchlaufzeit nicht mit

denen des *jasima* Modells übereinstimmten, wenn man die Simulation über einen längeren Zeitraum laufen ließ. Erste Vermutungen betreffend unterschiedlicher Implementierungsdetails beispielsweise bei den Verteilungen konnten nicht der Grund für die deutlichen Abweichungen sein. Nach genauer Betrachtung stellte sich heraus, dass das System mit den gewählten Parametern überlastet ist und langsam vollläuft. Betrachtet man also unterschiedlich lange Zeiträume erhält man entsprechend unterschiedliche Durchlaufzeiten, da diese bei den gewählten Parametern ansteigen. Bei der gewählten Einlastung ist dies unabhängig von der gewählten Steuerungsmethode.

Nach diesen Ergebnissen wurde die Einlastung der Aufträge verringert, sodass der Erwartungswert der negativen Exponentialverteilung $1/\lambda = 360$ Sekunden entspricht. So wird sichergestellt, dass das System auch mit den einfachen Regeln zur Einbeziehungsweise Auslagerung nicht vollläuft. Diese Erkenntnis unterstreicht weiterhin die Bedeutung der Auswahl eines geeigneten Steuerungsalgorithmus möglichst bereits bei der Dimensionierung von Anlagen, da dieser mitunter nicht nur zu leicht verbesserten Durchlaufzeiten führt, sondern auch andernfalls nötige Kapazitätserweiterungen vermeiden kann.

## 3.2 Durchführung der Simulationsstudie

Mit beiden Simulationssystemen wurde das beschriebene Szenario implementiert und mit der verringerten Einlastung untersucht. Als Steuerungsverfahren kamen dazu als Auslagerungsstrategie insgesamt drei Verfahren zum Einsatz. In beiden Systemen wurde die einfache Regel implementiert, die den nächsten Auftrag anhand der kürzesten Rüstzeit zum aktuellen Rüstzustand auswählt. In *Plant Simulation* wurde zusätzlich das Verfahren basierend auf dem genetischen Algorithmus eingesetzt. In der *jasima* Implementierung wurde eine Kopplung zu *Gurobi* umgesetzt, der jedes Einzelproblem optimal löst. Aufgrund der regelmäßig nach der ausgewählten Verteilung neu in das System eintreffenden Aufträge kann nicht davon ausgegangen werden, dass eine insgesamt optimale Strategie berechnet wird. Da allerdings aufgrund der FIFO-Linienpuffer Eigenschaften jeweils nur der erste Auftrag ausgewählt werden kann, kann man von sehr guten Ergebnissen ausgehen, denn eingehende Aufträge können in der Regel ohnehin erst später berücksichtigt werden. Weiterhin wurde das *jasima* Modell in einer zweiten Untersuchung auf 10 Linienpuffer erweitert und es wurden ebenfalls beide Steuerungsverfahren miteinander verglichen. Die Simulation wurde für jeweils 1000 Aufträge durchgeführt und bildet damit je nach Verfahren etwas weniger als 5 Tage ab. Um statistisch signifikannte Aussagen treffen zu können, wurden etwa 100 bis 200 einzelne Experimente durchgeführt.

Es hat sich gezeigt, dass sich in der Regel weniger als 20 Aufträge im System befanden und diese Menge in wenigen Sekunden, in Ausnahmen in wenigen Minuten, vom Solver gelöst werden konnte. Damit ist der Einsatz eines Solver in diesem Szenario sehr gut möglich. Sollten schwierige Instanzen auftreten oder deutlich mehr Aufträge im System vorhanden sein, könnte man beispielsweise die Anzahl der berücksichtigen Aufträge reduzieren und nur eine gewisse Anzahl an Aufträgen pro Warteschlange betrachten. Weiterhin könnte man die maximale Zeit, die dem Solver zur Berechnung gegeben wird, begrenzen und sich mit der bis dahin besten Lösung zufrieden geben oder alternativ auf eine Regelauswahl zurückgreifen.

# 4 Ergebnisse

Die Ergebnisse der Studie zeigen, dass der Genetische Algorithmus genauso wie der Solver im Vergleich zu der einfachen Steuerungsregel signifikant bessere Ergebnisse in Bezug auf die Durchlaufzeit erzielen. Die mittlere Durchlaufzeit beträgt bei den Experimenten die mit *jasima* durchgeführt wurden 2872 Sekunden mit der Regelsteuerung und 2114 Sekunden, wenn der Solver die Reihenfolge berechnet hat. Die Unterschiede sind signifikant, wie der Standardfehler zeigt; ein paarweiser Vergleich (paired t-test) hat dies ebenfalls bestätigt. Die Verbesserung beträgt 26,39 % und liegt damit sehr hoch (Tabelle 2).

***Tabelle 2:*** *Simulationsergebnisse: jasima - Regel und MILP*

| Verfahren | Durchlaufzeit | Standardfehler |
|---|---|---|
| Regel | 2872 s | 79,3 s |
| MILP | 2114 s | 63,1 s |
| Differenz | 26,4 % | |

Der Vergleich zwischen der Entscheidungsregel und dem genetischen Algorithmus zeigt, dass die Regel eine mittlere Durchlaufzeit von 2829 Sekunden liefert, die damit signifikant über der des genetischen Algorithmus mit 2180 Sekunden liegt. Der Unterschied beträgt 22,9 % und ist signifikant (Tabelle 3). Die Ergebnisse bestätigen weiterhin, dass die unterschiedlichen Implementierungen, im Rahmen der auftretenden Schwankungen, die gleichen Ergebnisse liefern.

***Tabelle 3:*** *Simulationsergebnisse: Plant Simulation - Regel und GA*

| Verfahren | Durchlaufzeit | Standardfehler |
|---|---|---|
| Regel | 2829 s | 42,8 s |
| GA | 2181 s | 24,5 s |
| Differenz | 22,9 % | |

Weiterhin ist zwar anzunehmen, dass der Solver mit seinen optimalen Lösungen bessere Ergebnisse liefert als der genetische Algorithmus, da seine Verbesserung gegenüber der Regel etwa 3,5 % höher ausfällt als die des genetischen Algorithmus; allerdings lässt sich dies nicht signifikant zeigen. Der Standardfehler ist dafür zu groß und ein paarweiser Vergleich ist aufgrund der unterschiedlichen Simulationssysteme hier nicht möglich.

In Tabelle 4 sind die Ergebnisse der Simulationsstudie mit 10 statt 5 Linienpuffern aufgeführt, die ähnliche Vorteile für den Solver aufzeigen, allerdings fällt die Verkürzung der mittleren Durchlaufzeit mit 18,3 % kleiner aus. Dies geht einher mit insgesamt kürzeren Durchlaufzeiten (2476 Sekunden und 2023 Sekunden).

***Tabelle 4:*** *Simulationsergebnisse: jasima - Regel und MILP mit 10 Linienpuffern*

| Verfahren | Durchlaufzeit] | Standardfehler |
|---|---|---|
| Regel | 2476 s | 47,3 s |
| MILP | 2023 s | 25,6 s |
| Differenz | 18,3 % | |

Sämtliche Ergebnisse sind in Abbildung 2 mit dem zweifachen Standardfehler grafisch dargestellt.

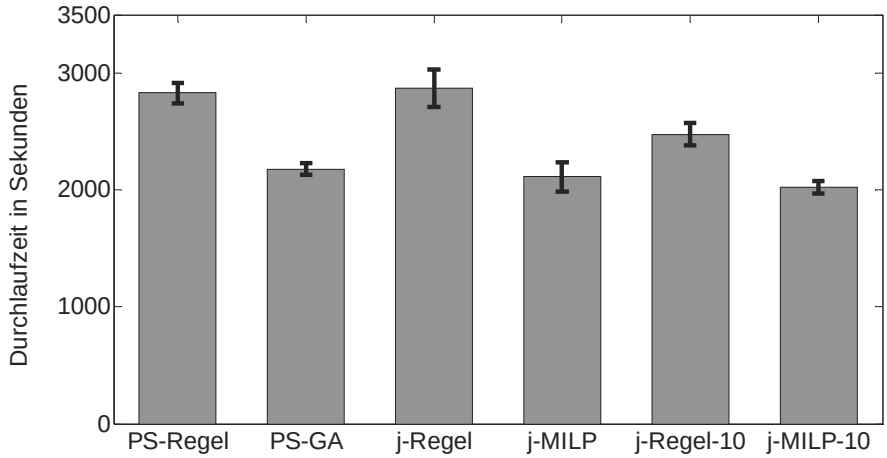



***Abbildung 2:*** *Simulationsergebnisse (mit zweifachem Standardfehler) in der Übersicht*

Ein Geschwindigkeitsvergleich zwischen den beiden Simulationssystemen bietet sich bei dieser Studie nur bedingt an, da die Länge der Simulationsläufe hauptsächlich von den verwendeten Steuerungsverfahren abhängt und deutlich weniger von der Simulationssoftware an sich. Auf einem aktuellen Intel i7 4510 System benötigen die Simulationsläufe mit hunderten Replikationen mehrere Stunden. Prinzipiell liegen die Stärken von *Plant Simulation* sicherlich in der Visualisierung und der relativ einfachen Bedienbarkeit. Bei der Geschwindigkeit und bei der Integration in weitere Systeme spielt *jasima* seine Vorteile aus. Speziell bei der simulationsbasierten Optimierung oder der Integration in Entscheidungsunterstützungssysteme bietet sich daher *jasima* an.

## 5 Zusammenfassung und Ausblick

In dieser Studie wurden verschiedene Steuerungsverfahren zur Reihenfolgeplanung in einem Szenario mit mehreren parallelen Linienpuffern und reihenfolgeabhängigen Rüstzeiten untersucht. Obwohl es sich dabei um ein NP-vollständiges Problem handelt, sind die auftretenden Problemgrößen in diesem Szenario in vertretbarer Zeit

lösbar. Ein Ausweichen auf eine nichtoptimale Heuristik kann in diesem Fall vermieden werden. Die Kopplung mit einem Simulationssystem bietet dabei vielversprechende Möglichkeiten, beispielsweise die Verknüpfung von dynamischen Abläufen mit (teil)-optimierten Lösungen.

Bei der Durchführung dieser Studie ist weiterhin bestätigt worden, dass es wichtig ist, Annahmen des Simulationsmodels, wie hier die Einlastung der Aufträge, zu überprüfen, um vergleichbare Ergebnisse zu erhalten.

In weiteren Studien wäre es interessant zu untersuchen, ob alternative Einlagerungsstrategien zu weiteren Verbesserungen führen können. Es könnte beispielsweise sinnvoll sein, einen Linienpuffer freizuhalten, um gut passende Aufträge direkt verfügbar zu machen. Dieses Vorgehen könnte je nach Anzahl der Linienpuffer mehr oder weniger vorteilhaft sein. Treten die Aufträge nicht gleichverteilt auf, da einige Farben in diesem Beispiel häufiger nachgefragt werden, könnte dies ebenfalls in der Einlagerungsstrategie berücksichtigt werden.

Eine interessante Fragestellung ist darüber hinaus, wieweit die Steuerungsentscheidungen verbessert werden können, wenn Informationen über zukünftige Aufträge bekannt sind. Gerade im hier betrachteten Beispiel der Automobilproduktion sind die in näherer Zukunft zu erwartenden Karosserien mit geringer Unsicherheit bekannt. Diese Information lässt sich wahrscheinlich für bessere Ergebnisse nutzen, insbesondere für verbesserte Einlagerungsstrategien in die Puffer.

Nicht zuletzt ist ein Vergleich der optimierenden Verfahren (seien es heuristische Optimierungsverfahren wie GAs oder MILPs) mit optimierten Prioritätsregeln sehr interessant (ein aktueller Überblick über deren Erstellung findet sich in (Branke et al. 2015). Diese Regeln konnten manuell entwickelte Regeln in verschiedenen Szenarien zum Teil deutlich schlagen (Hildebrandt et al. 2014). Sie haben gerade unter praktischen Gesichtspunkten den Vorteil, dass sie sich, einmal erstellt, über einen längeren Zeitraum einsetzen lassen und dabei nur sehr geringe Rechenzeitanforderungen stellen. Somit eignen sie sich sehr gut zum Treffen von Planungs- und –Steuerungsentscheidungen in Echtzeit. Für das in diesem Paper betrachtete Szenario ließe sich dieser Ansatz einsetzen, um sowohl optimierte Strategien für die Einlagerung als auch die Auslagerung zu entwickeln.

## Literaturverzeichnis

Branke, J.; Nguyen, S.; Pickardt, C.; Zhang, M.: Automated design of production scheduling heuristics: A review. IEEE Transactions on Evolutionary Computation, PP 99 (2015, early access).

Chiang, T.-C.; Fu, L.-C.: Using a family of critical ratio-based approaches to minimize the number of tardy jobs in the job shop with sequence dependent setup times. European Journal of Operational Research 196 (2009) 1, S. 78–92.

Eley, M.: Simulation in der Logistik: Einführung in die Erstellung ereignisdiskreter Modelle unter Verwendung des Werkzeuges Plant Simulation. Berlin: Springer Gabler 2012.

Heger, J.; Hildebrandt, T.; Scholz-Reiter, B.: Dynamic adjustment of dispatching rule parameters in flow shops with sequence dependent setup times. International Journal of Production Research 2015 (accepted with minor revisions)

Heger, J.: Dynamische Regelselektion in der Reihenfolgeplanung: Wiesbaden: Springer Fachmedien 2014.

Hildebrandt, T.; Goswami, D.; Freitag, M.: Large-scale simulation-based optimization of semiconductor dispatching rules. In: Winter Simulation Conference, Savannah 2014, S. 2580-2590.

jasima solutions UG: jasima. Online verfügbar unter http://code.google.com/p/jasima/.

Kochhar, S.; Morris, R. J. T.: Heuristic methods for flexible flow line scheduling. Journal of Manufacturing Systems 6 (1987) 4, S. 299–314.

Kurz, M. E.; Askin, R. G.: Scheduling flexible flow lines with sequence-dependent setup times. European Journal of Operational Research 159 (2004) 1, S. 66–82.

Lee, Y. H.; Pinedo, M.: Scheduling jobs on parallel machines with sequence-dependent setup times. European Journal of Operational Research 100 (1997) 3, S. 464–474.

Monma, C. L.; Potts, C. N.: On the complexity of scheduling with batch setup times. Operations Research 37 (1989) 5, S. 798–804.

Naderi, B.; Salmasi, N.: Permutation flowshops in group scheduling with sequence-dependent setup times. European Journal of Industrial Engineering 6 (2012) 2, S. 177–198.

Pickardt, C. W.; Branke, J.: Setup-oriented dispatching rules - a survey. In: International Journal of Production Research 50 (2011) 20, S. 1–20.

Tan, K. C.; Narasimhan, R.: Minimizing tardiness on a single processor with sequence-dependent setup times: a simulated annealing approach. Omega 25 (1997) 6, S. 619–634.

van der Zee, D.-J.: Non-exhaustive family based dispatching heuristics – exploiting variances of processing and set-up times. International Journal of Production Research 48 (2010) 13, S. 3783–3802.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Flexible Job-shop Scheduling with Dynamic Stochastic Machine Sets

## *Flexible Job-Shop-Planerstellung mit dynamischen stochastischen Maschinen-Mengen*

Tao Zhang, Shufang Xie, Oliver Rose, Universität der Bundeswehr München, Neubiberg (Germany), tao.zhang@unibw.de, shufang.xie@unibw.de, oliver.rose@unibw.de

**Abstract:** In the general flexible job-shop scheduling problem, each type of operation has a given static and deterministic machine set. However, in practice the machine set is mostly neither static nor deterministic. Approaches to the general FJSP will fail to solve the practical case. Our study focuses on this practical FJSP with both dynamic and stochastic machine sets. Because the problem involves stochastic factors, a simulation-based meta-heuristic procedure is introduced. To speed up the solution procedure the stochastic simulation is replaced with a deterministic simulation model and a stochastic process generator. The experimental results show that the computing time is dramatically reduced while the performance of the schedule decreases only a little.

## 1 Introduction

In the flexible job-shop scheduling problem (FJSP), an operation may be processed on several machines which form a machine set. Comparing to the job shop scheduling problem, the raised issue is the machine allocation, i.e., assigning each operation to a particular machine from the machine set. Usually, for the general FJSP the assumption is made that each type of operation has a given static and deterministic machine set. However, in practice the machine set is mostly neither static nor deterministic. As a consequence, approaches to the general FJSP will fail to solve the practical case.

The FJSP is one of the most important combinatorial optimisation problems. It has been proven to be a NP-hard problem. Usually FJSP is formularised as a mixed binary integer linear programming problem (MBILP) (Demir and Kürşat İşleyen, 2013) in which the machine allocation variables are binary and the starting time of the operations are continuous. There are some exact methods to solve the MBILP problem, such as branch and bound (B&B) (Brucker et al, 1994; Pan and Chen 2005), cutting-plane, branch and cut (Karimi-Nasab and Seyedhoseini 2013), and so

on. Because the size of the practical FJSP can become very large, these methods are not able to find an optimal solution in a reasonably short time. In order to shorten the computing time, researchers are trying to find an approximate optimal solution instead of the exact solution. Hence, a variety of heuristic procedures and meta-heuristic procedures have been used. Decision rules are typical heuristic procedures (Holthaus and Rajendran 1997). The meta-heuristics procedures include local search (Murovec and Šuhel 2004), tabu search (Bożejko and Makuchowski 2009), simulated annealing (Satake et al. 1999), genetic algorithms (Qing-dao-er-ji and Wang 2012), ant colony algorithms (Rossi and Dini 2007), particle swarm algorithms (Lin et al. 2010), and so on. In addition, considering the stochastic factors, simulation is often combined with these procedures to solve the problem. There are two ways to combine the optimisation approaches with the simulation method. One way is to embed the optimisation approaches in the simulation. The optimisation approaches make decisions during the simulation. An example can be found in Balin's work (Balin 2011). Another way is to embed the simulation in the optimisation approaches (Gholami and Zandieh 2009). Running the stochastic simulation many times can help to evaluate candidate solutions in the stochastic environment. The solutions will be improved based on the evaluation (Zhang and Rose 2013). However, in order to obtain good evaluation results, the simulation has to run lots of times for each candidate solution. The simulation runs are very time-consuming and the procedure becomes very slow.

Our study will focus on this practical FJSP with both dynamic and stochastic machine sets. The problem is denoted by FJSP* in this paper. Because the problem involves stochastic factors, we adopt a simulation-based meta-heuristic procedure to solve the problem. To speed up the solution procedure the stochastic simulation is replaced with a deterministic simulation model and a stochastic process generator. The paper is structured as follows. In Section 2, we address the dynamic and stochastic features of the machine sets in details. Section 3 introduces the improved simulation-based meta-heuristic procedures to solve the special FJSP. Experiments are carried out in Section 4 and the paper is concluded in Section 5.

## 2 Description of FJSP*

### 2.1 Dynamic Machine Sets

We use an example to illustrate the dynamic machine sets. Given that two machines A and B are able to process one operation and the location of the concerned semiproduct is close to machine A but far away from B, it is inappropriate to carry the semiproduct to machine B. In this case, the machine set for the operation at the decision point actually includes machine A only which is determined by the location of the semiproduct. The current location of the semiproduct depends on the machine allocation decisions made for previous operations. In other words, the previous decisions we made decide upon the machine set of the current operation. Thus, the machine set is changing dynamically.

A machine network is used to map the dynamic relationship between the operations and the machines for each type of product, as shown in figure 1. The nodes denote the machines. The machines are arranged in the order of the operations. The machines that can process the same operation are put in one column. The arrows

originating from one machine point to the alternative machines which can process the next operation if the current operation is processed on the machine.

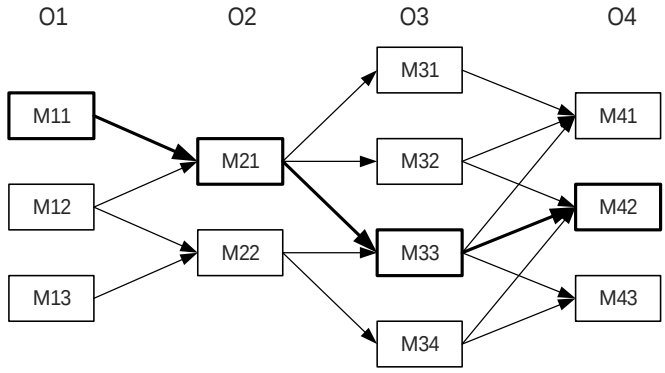



***Figure 1:*** *Machine network for one type of product*

## 2.2 Stochastic Machine Sets

Many reasons lead to the stochastic machine sets, such as breakdowns, maintenance, utilisation balancing strategies, and so on. We consider two stochastic events, breakdown and maintenance, for each machine. Because the maintenance activities are usually carried out according to the maintenance plan, we assume that the time between two successive maintenance activities follows the normal distribution; the time to maintain follows the normal distribution, too. The maintenance activities are carried out only if the machines are idle. For the machine down times, we assume their probability density functions are piecewise functions,

$$f_B(b;\xi,\lambda,k)=\begin{cases}(1-\xi)\dfrac{k}{\lambda}\left(\dfrac{b}{\lambda}\right)^{k-1}e^{-(b/\lambda)^k}, b>0\\ \xi \quad , b=0\\ 0 \quad , b<0\end{cases} \tag{1}$$

where $\xi$ is the probability that the machine does not break down; $k$ is the shape parameter; $\lambda$ is the scale parameter. The function is depicted in Figure 2. This function integrates the breakdown time with the breakdown probability. The breakdown time follows the Weibull distribution while the time is greater than 0. The breakdown time "0" means that the machine won't break down at sample points. If the machine breaks down while it is processing one job, we assume that the job will be blocked there and resumed after the machine is fixed.

Note that the probability that the machine breaks down is related to the time difference between current time and the last breakdown time. Usually the probability will go up as the time difference increases because the newly repaired machine has a

higher reliability than the machines which have been running for a long time. Thus $\xi$ can be computed as,

$$\xi = (1-\upsilon)(1-(\tau-\tau_l)/\delta) \tag{2}$$

where $\upsilon$ is the average probability of machine breakdowns, $\tau$ is current time, $\tau_l$ is the last break time, $\delta$ is the remaining life of the machine. Note that $\upsilon$ is not the times of breakdowns during certain period. It is computed from data which is sampled in the fixed time interval. If during the interval, the machine does not break down, the sample data is 0. If the machine is broken, the sample data is the repair time. The histogram of this data set can be found in figure 2.

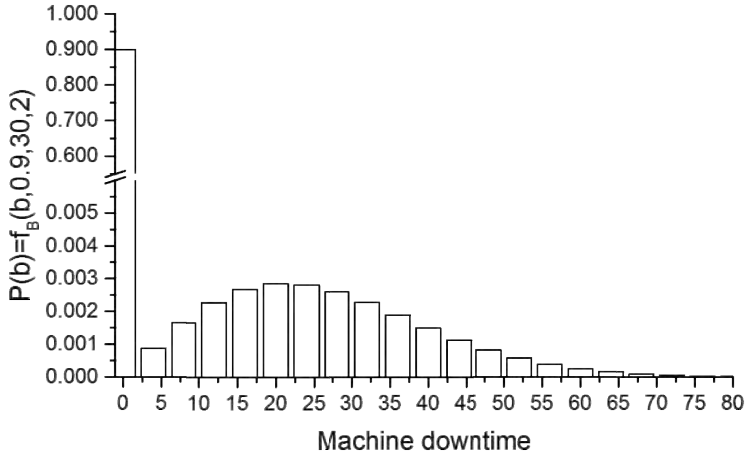



***Figure 2:*** *Probability density function of the machine down time*

# 3 Simulation-based Meta-heuristic for FJSP*

## 3.1 Meta-heuristic for FJSP*

A basic metaheuristic procedure contains one loop and three components. The components are initial solution generation, solution evaluation, and solution improvement, as shown in figure 3. The solution evaluation and the solution improvement are carried out repeatedly in the loop. Different meta-heuristics have different solution improvement components. In our study, we will adopt genetic algorithms and ant colony algorithms to solve the problem, respectively. Both the genetic algorithms and ant colony algorithms have population-based improvement components. Because the improvement components in these two algorithms (Qing-dao-er-ji and Wang 2012; Rossi and Dini 2007) are already very mature, we won't discuss them here. The solution evaluation will be addressed in the next section. So the only component left is the generation of the initial solution. The most important issue is to design the appropriate structure of the solution.

From figure 1, we know that each job has only one operation flow (i.e., a series of operations), but possibly has several optional processing paths (i.e., a series of machines). The task of assigning jobs to machines can be transformed into selecting a processing path from the optional paths for each job. Thus, the problem is equivalent to finding an optimal combination of all jobs' processing paths. The solution is a set of selected paths. The solution to the sequencing task can be easily depicted as a set of position indices of operations. Then, the structure of the solution to FJSP* is a combination of the selected path set and the operation position set.

## 3.2 Improved Simulation-based Solution Evaluation

As we mentioned before, because the problem involves stochastic factors, we adopt a simulation-based metaheuristic procedure to solve the problem. To speed up the procedure the stochastic simulation is replaced with a deterministic simulation model and a stochastic process generator, as shown in figure 3. In the deterministic simulation, the parameters, such as processing times, release times, setup time, and so on, are average values. Thus they are deterministic. The breakdown and the maintenance won't be considered. The deterministic simulation runs only once and the simulation results will be adjusted according to the outputs of the stochastic process generator. The adjusted results are returned to the evaluation module. The number of simulation runs is always one, thus the procedure is speeded up considerably.

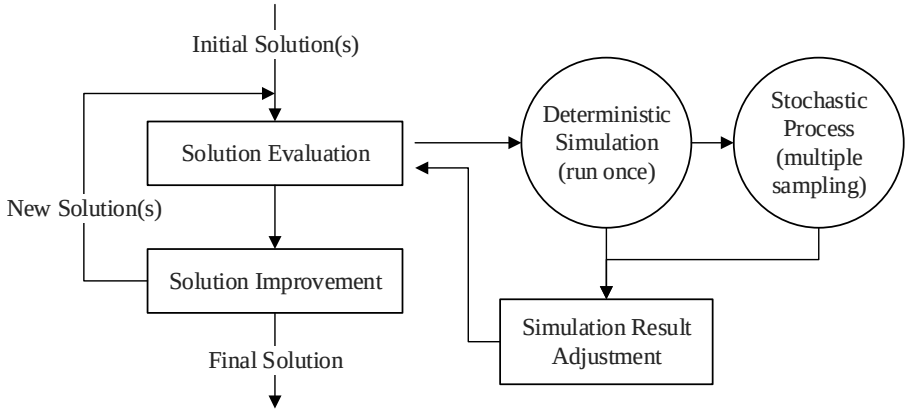



***Figure 3:*** *Improved simulation-based meta-heuristic procedure*

The stochastic process generator and the result adjustment are the kernel of our study. The deterministic simulation records each job's start/end times at each machine. The stochastic process generator generates and lists all stochastic events in the period of the simulation for all machines. We adjust the simulation results, i.e., start times of the jobs, to remove the conflicts caused by the stochastic events. For the same reason, the generator has to generate the event list many times and the adjustment procedure also has to be carried out the same number of times. Finally, we obtain several groups of adjusted simulation results. In our study, the objective is to minimise the makespan. The average makespan is returned to the metaheuristic.

Because the event sampling is very simple and the adjustment procedure is relatively fast, the whole metaheuristic procedure is speeded up considerably compared to a traditional multi-run simulation-based meta-heuristic.

### 3.3 Stochastic Process Generator and Simulation Result Adjustment

The stochastic process generator is responsible for creating the stochastic events and activities. The maintenance activities are generated in a sequential manner. First, we sample a time period from the stochastic variable of time to next maintenance (TTNM). Current time plus the time period will be the starting time of the first maintenance activity. Then, we sample another time period from the stochastic variable of time to maintenance. This time period is the duration of the first maintenance activity. Next, we sample another time period from the TTNM. The end time plus the time period will be the starting time of the second maintenance activity. The procedure will stop as soon as the starting time of a new maintenance activity is out of scope. The breakdown activities are generated in a fixed-interval manner. The sampling is carried out at some fixed time points. The interval between two successive sample points is fixed and equal to the sample interval used when obtaining the breakdown probability. The sample point is the starting time of the breakdown activity. The sampled time period is the duration of the breakdown activity. If the sampled time period is 0, the activity won't happen.

If we put these generated stochastic activities into the schedule built by the simulation, and replace the other parameters, such as processing times, release times, setup time, and so on, with randomly generated values, there may be some conflicts. For example, a broken machine is processing an operation. Thus we have to remove these conflicts. In our study, the conflict operations will be put off and the related successors are also adjusted. Then a non-conflict schedule is obtained. However, the schedule may not be optimal anymore. Thus a critical-operation adjustment (COA) procedure is introduced. The procedure (algorithm 1) will adjust the non-conflict schedule further and try to achieve a better performance. The COA always focuses on the critical operations, for instance in our study they are the latest finished operations, and try to adjust these operations. After adjustment, the critical operations may not be critical and other operations may become critical. The COA works in a recursive way. The procedure is as follows. If one critical operation cannot be adjusted, we will move to its' predecessor.

## 4 Experiments

We create a flexible job shop model to evaluate the proposed approach. The objective is to minimise the makespan. In the model, there are 8 machines and 2 types of jobs, type A and B. Type A contains 3 operations while type B contains 4 operations. Each operation has 1 to 4 alternative machines. The operations need different amounts of time if they are processed on different machines. The machine sets for the operations are heavily overlapping. The shop plans to produce 13 jobs of type A and 15 jobs of type B. So there are 99 operations in total. We assume that the processing times follow the normal distribution. The release times and the setup times are constant. Four experimental cases are designed according to the metaheuristics and the evaluation components. Two metaheuristics are involved

including genetic algorithms (GA) and ant colony algorithm (Ant). SS denotes the classical stochastic simulation. DS+SP denotes our improved evaluation. The results are shown in Table 1. From Table 1 we can see that no matter which metaheuristic we adopt the improved approach is always faster than the SS approach while the objective values we achieve are very close.

***Algorithm 1:***

```
void adjust() {
     while (true) {
          if (!adjust(getCriticalOperation())) {
               break;
          }
     }
}
boolean adjust(Operation operation) {
     if (isAdjustable(operation)) {
          adjustOperation(operation);
          return true;
     } else if (!operation.isFirstinJob()) {
          return adjust(operation.getPredecessor());
     } else {
          return false;
     }
}
```

***Table 1:*** *Comparisons among different meta-heuristics and evaluations*

| Meta-heuristic | Evaluation | Mean Objective (hours) | Mean Run Time (min) |
|---|---|---|---|
| GA | SS | 21.9 | 17.2 |
| | DS+SP | 23.5 | 6.7 |
| Ant | SS | 26.1 | 21.1 |
| | DS+SP | 28.7 | 8.3 |

## 5 Conclusion

The difficulties raised by the dynamic machine sets are relieved by converting the machine allocation problem to a path selection problem. The stochastic machine sets may cause unstable schedule performance. This problem is reformulated as a new optimisation problem which tries to minimise the expected value of the objective (makespan) in stochastic situations. Simulation is always a good way to analyse stochastic systems. But when using the simulation in meta-heuristic search algorithms to optimise the performance of stochastic systems, the simulation shows its weakness. The simulation in the meta-heuristic search usually plays the role of an evaluator in stochastic situations. Our study implements this function by means of a stochastic process generator and an adjustment procedure. From the experiment results, the computing time is dramatically reduced while the performance of the schedule decreases only a little.

## References

Balin, S.: Parallel machine scheduling with fuzzy processing times using a robust genetic algorithm and simulation. Information Sciences 181 (2011) 17, pp. 3551–3569.

Bożejko, W.; Makuchowski, M.: A fast hybrid tabu search algorithm for the no-wait job shop problem. Computers & Industrial Engineering 56 (2009) 4, pp. 1502–1509.

Brucker, P.; Jurisch, B.; Sievers, B.: A branch and bound algorithm for the job-shop scheduling problem. Discrete Applied Mathematics 49 (1994) 1–3, pp. 107–127.

Demir, Y.; Kürşat İşleyen, S.: Evaluation of mathematical models for flexible job-shop scheduling problems. Applied Mathematical Modelling 37 (2013) 3, pp. 977–988.

Gholami, M.; Zandieh, M.: Integrating simulation and genetic algorithm to schedule a dynamic flexible job shop. Journal of Intelligent Manufacturing 20 (2009) 4, pp. 481–498.

Holthaus, O.; Rajendran, C.: Efficient dispatching rules for scheduling in a job shop. International Journal of Production Economics 48 (1997) 1, pp. 87–105.

Karimi-Nasab, M.; Seyedhoseini, S.M.: Multi-level lot sizing and job shop scheduling with compressible process times. European Journal of Operational Research 231 (2013) 3, pp. 598–616.

Lin, T.-L.; Horng, S.-J.; Kao, T.-W.; Chen, Y.-H.; Run, R.-S.; Chen, R.-J.; Lai, J.-L.; Kuo, I.H.: An efficient job-shop scheduling algorithm based on particle swarm optimization. Expert Systems with Applications 37 (2010) 3, pp. 2629–2636.

Murovec, B.; Šuhel, P.: A repairing technique for the local search of the job-shop problem. European Journal of Operational Research 153 (2004) 1, pp. 220–238.

Pan, J.C.-H.; Chen, J.-S.: Mixed binary integer programming formulations for the reentrant job shop scheduling problem. Computers & Operations Research 32 (2005) 5, pp. 1197–1212.

Qing-dao-er-ji, R.; Wang, Y.: A new hybrid genetic algorithm for job shop scheduling problem. Computers & Operations Research 39 (2012) 10, pp. 2291–2299.

Rossi, A.; Dini, G.: Flexible job-shop scheduling with routing flexibility and separable setup times using ant colony optimisation method. Robotics and Computer-Integrated Manufacturing 23 (2007) 5, pp. 503–516.

Satake, T.; Morikawa, K.; Takahashi, K.; Nakamura, N.: Simulated annealing approach for minimizing the makespan of the general job-shop. International Journal of Production Economics 60–61 (1999), pp. 515–522.

Zhang, T.; Rose, O.: Simulation-based overhead-crane scheduling for a manufacturing plant. In: Pasupathy, R.; Kim, S.H. (Eds.): Proceedings of the 2013 Winter Simulation Conference. Washington D.C.: 2013, pp.2633-2642.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Lösung eines Hybrid-Flow-Shop-Maschinenbelegungsproblems mit simulationsbasierter Optimierung

## *Solving a Hybrid Flow Shop Scheduling Problem with Simulation-based Optimisation*

Abdulrahman Nahhas, Paul Aurich, Tobias Reggelin, Otto-von-Guericke-Universität, Magdeburg (Germany), abdulrahman.nahhas@hotmail.com, paul.aurich@mail.de, tobias.reggelin@ovgu.de

Juri Tolujew, Fraunhofer IFF, Magdeburg (Germany), juri.tolujew@iff.fraunhofer.de

**Abstract:** This paper describes and compares two approaches to solve a hybrid flow shop scheduling problem in the production of surface-mount devices (SMD) with the help of simulation-based optimisation. The hybrid flow shop consists of four surface-mount technology (SMT) placement machines on the first production step and five automated optical inspection (AOI) machines on the second production step. The first approach described is integrated simulation-based optimisation and the second approach described is simulated annealing. Both approaches lead to an improvement in terms of producing more orders on time compared to the decision rules used so far by this company. The implemented integrated simulation-based optimisation delivers results much faster than the implemented simulated annealing, whereas the simulated annealing leads to slightly better results.

## 1 Einleitung

Im Rahmen dieser Arbeit wird ein Maschinenbelegungsproblem (Scheduling-Problem) eines Hybrid Flow Shops in der Leiterplattenbestückung mit multiplen Zielfunktionen behandelt. Bei einem Hybrid Flow Shop handelt es sich um eine mehrstufige Fließfertigung mit mehreren parallel arbeitenden Maschinen je Produktionsschritt. Die Maschinenbelegungsplanung behandelt die Zuordnung von Produktionsaufträgen auf Maschinen und die Reihenfolgeplanung unter Beachtung von einzelnen Zielfunktionen wie Minimierung der Zykluszeiten, Durchlaufzeiten oder Minimierung von Abgabeverspätungen oder von kombinierten Zielfunktionen (Pinedo 2012, S.151).

Maschinenbelegungsprobleme sind Inhalt verschiedener wissenschaftlicher Betrachtungen (Graham et al. 1979; Wittrock 1990; Voß 1993). Eine Maschinenbele-

gungsplanung ist eine sehr komplexe Entscheidung, die täglich vom Management getroffen werden muss. Dazu werden in der Praxis häufig Prioritätsregeln wie Shortest Process Time (SPT), Longest Process Time (LPT) oder Earliest Due Date (EDD) verwendet, da diese einfach anwendbar sind und oft gute Ergebnisse liefern. Die Ergebnisse dieser Prioritätsregeln entsprechen aber nicht unbedingt dem möglichen Optimum (Pinedo 2012).

Im Gegensatz dazu können mit analytischen Modellen optimale Lösungen gefunden werden. Beispiel dafür ist der Ansatz von Johnson (1954), mit dem optimale Maschinenbelegungen zur Minimierung der Zykluszeit für einen zwei Stufen Flow Shop mit reihenfolgeabhängigen Rüstzeiten erstellt werden (Johnson 1954). Die Zykluszeit ist die Gesamtbearbeitungszeit aller Aufträge in einem bestimmten Zeitraum (Wittrock 1990). In einem Flow Shop gibt es m Maschinen in Serie. Jeder Auftrag muss auf allen der m Maschinen bearbeitet werden. Dabei folgen alle Aufträge demselben Weg durch das System (Schocke 2000). Analytische Modelle sind nur geeignet, kleine und sehr spezifische Probleme zu lösen.

Hybrid-Flow-Shop-Maschinenbelegungsprobleme wie das in diesem Paper behandelte Problem sind hingegen in der Regel NP-schwer (Wittrock 1990). Daher ist die Entwicklung eines Algorithmus, der eine optimale Lösung garantiert, unwahrscheinlich (Baker and Trietsch 2009). Deshalb wird oft die simulationsbasierte Optimierung genutzt, um solche Maschinenbelegungsprobleme zu lösen (Anderson 2008). Die Aufgabe der Simulation ist dabei die Bewertung der von der Optimierung entwickelten Lösung (März 2011).

## 2 Maschinenbelegungsproblem eines Hybrid Flow Shops in der Leiterplattenbestückung

Maschinenbelegungsprobleme lassen sich nach folgenden Kriterien charakterisieren (Graham et. al. 1979), nach denen nachfolgend das im Paper behandelte Maschinenbelegungsproblem eines Hybrid Flow Shops in der Leiterplattenbestückung beschrieben wird:

- Produktionsumgebung
- Aufträge
- Restriktionen
- Zielfunktionen.

*Produktionsumgebung*: Der betrachtete Ausschnitt der Leiterplattenbestückung besteht aus vier identischen parallelen Bestückungsautomaten (SMD) und fünf identischen parallelen automatisierten Sichtkontrollen (AOI) und kann als Hybrid Flow Shop klassifiziert werden (Abb. 1). Ein Hybrid Flow Shop beinhaltet c Produktionsstufen mit m parallelen Maschinen. Jeder Auftrag j muss auf jeder Stufe bearbeitet werden, allerdings nur auf einer Maschine je Stufe (Schocke 2000).

*Aufträge* werden durch den Abgabetermin, die Stückzahl, den Produkttyp, Bearbeitungszeiten und verwendete Rohmaterialen charakterisiert. Auf Basis der Rohmaterialien werden die Produkttypen zu Produktfamilien gruppiert. Produkttypen, Bearbeitungszeiten und Rohmaterialien sind im vorliegenden Fall sehr heterogen.

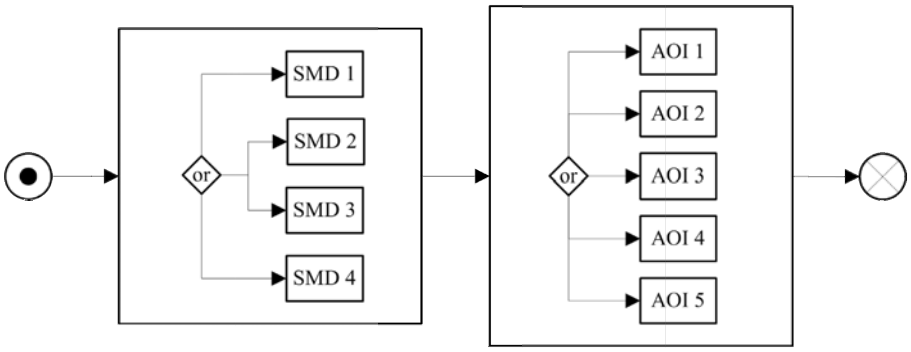



***Abbildung 1:*** *Leiterplattenbestückung als Hybrid Flow Shop*

*Restriktionen:* Maschinenstörungen werden indirekt über eine reduzierte verfügbare Produktionszeit abgebildet. Produktionsabbrüche und Auftragssplitting sind nicht zulässig. In der ersten Produktionsstufe gibt es reihenfolgeabhängige Rüstzeiten. Es gibt dabei Major- und Minor-Rüstzeiten. Major-Rüstzeiten entstehen, wenn der nachfolgende Produktionsauftrag nicht zur selben Produktfamilie wie der vorhergehende Produktionsauftrag gehört. Minor-Rüstzeiten entstehen, wenn der nachfolgende Produktionsauftrag zur selben Produktfamilie gehört, was als Familienrüststrategie bezeichnet wird (Wittrock 1990; So 1990). Die reihenfolgeabhängigen Rüstzeiten haben einen erheblichen Einfluss auf die Komplexität eines Maschinenbelegungsproblems (Gupta und Kyparisis 1986).

*Zielfunktion:* Hauptziel dieser Untersuchung ist die Minimierung der Zykluszeit. Darüber hinaus ist sollte der Abgabetermin möglichst nicht verletzt werden.

## 3 Vorgehensweise und Implementierung

Um die Komplexität der Maschinenbelegungsplanung des vorliegenden Hybrid Flow Shops zu verringern, wurden Allokation (Zuordnung) und Sequenzierung (Reihenfolgeplanung) voneinander getrennt (vgl. Baker und Trietsch 2009, S.221). Es wird zunächst das Allokationsproblem gelöst, in dem die Aufträge auf eine definierte Maschine je Stufe zugewiesen werden. Dadurch vereinfacht sich die Sequenzbildung für die erste Produktionsstufe (SMD) zu einem Einmaschinenproblem mit reihenfolgeabhängigen Rüstzeiten und für die zweite Produktionsstufe (AOI) zu einem Einmaschinenproblem mit reihenfolgeunabhängigen Rüstzeiten. Das Einmaschinenproblem ist eines der einfachsten Maschinenbelegungsprobleme und wird in der Literatur als das Einplanen von *n* Aufträgen auf eine einzige Maschine beschrieben (Baker und Trietsch 2009).

Zur Lösung des Maschinenbelegungsproblems wurden zwei Ansätze implementiert. Im ersten Ansatz (siehe Abschnitt 3.1) wurden definierte Entscheidungs- und Optimierungsregeln (Heuristik) in ein Simulationsmodell integriert, da hierdurch sehr schnell gute Ergebnisse erwartet wurden. Im zweiten Ansatz (siehe Abschnitt 3.2) wurde die Meta-Heuristik des Simulated Annealing gewählt, da dieser Ansatz bereits oft bei Maschinenbelegungsproblemen eingesetzt wurde und oft gute Ergebnisse liefert (Schocke 2000).

## 3.1 Integrierte Simulationsbasierte Optimierung

Bei der hier beschriebenen Vorgehensweise (siehe Abb. 2) handelt es sich nach März et al. (2012) um eine in ein Simulationsmodell integrierte Optimierung. Die Allokations- und Sequenzierungsalgorithmen sind direkt im Simulationsmodell implementiert. Das vorliegende zweistufige Produktionssystem wurde mit einem mesoskopischen Simulationsmodell (vgl. Tolujew und Reggelin 2011) abgebildet, um kurze Rechenzeiten zu erzielen. Zur Umsetzung des mesoskopischen Simulationsmodells wurden die discrete-rate und discrete-event Bibliotheken von ExtendSim genutzt. Die Prozesse Leiterplattenbestückung (SMD) und Qualitätskontrolle (AOI) wurden mit discrete-rate modelliert. Die Flussrate hängt hierbei von den Taktzeiten der verschiedenen Produkttypen ab. Für die Allokations- und Sequenzierungsentscheidungsprozesse wird die discrete-rate Abbildung der Aufträge in einzelne Flussobjekte aggregiert. Nach der Allokations- und Sequenzierungsentscheidung werden die aggregierten Flussobjekte (Aufträge) wieder aufgelöst und als discrete-rate dargestellt.

Vor dem Start der Simulation erfolgt eine Initialisierungs-Allokation von zu Produktfamilien zusammengefassten Aufträgen durch eine Zuordnung reihum auf die Maschinen der ersten Produktionsstufe (SMD) nach der Regel Shortest Process Time (SPT). Während der Simulation interagieren der Sequenzierungs- und Allokationsalgorithmen mit einander (Abb. 2). Durch diese Interaktion soll eine Balancierung der Auslastung der Maschinen und eine optimale Produktionsreihenfolge der Aufträge erreicht werden, wie nachfolgend beschrieben wird.

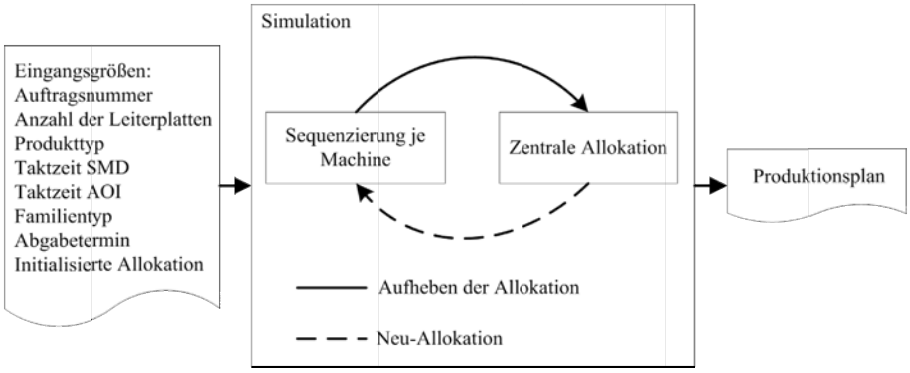



***Abbildung 2:*** *Integrierte simulationsbasierte Optimierung*

Zur Minimierung der gesamten Rüstzeiten, was der geforderten Minimierung der Zykluszeit entspricht, ist es erforderlich, alle Aufträge einer Familie in einer Sequenz nacheinander zu produzieren. Das kann dazu führen, dass nicht alle Aufträge zu ihrem Abgabetermin fertiggestellt werden können. Deshalb wird nach folgendem Sequenzierungsalgorithmus (Abb. 3) vorgegangen.

Es wird davon ausgegangen, dass zu Beginn der Betrachtung alle SMD-Maschinen für die Produktion eines Auftrags der höchsten Priorität vorbereitet sind, wobei die Priorität den Abgabetermin darstellt. Als erstes wird nun ein Auftrag der höchsten Priorität von der Produktfamilie mit der kleinsten Gesamtbearbeitungszeit gewählt.

Anschließend plant der Algorithmus solange Aufträge derselben Produktfamilie geordnet nach deren Abgabeterminen ein, bis keine Aufträge dieser Produktfamilie mehr vorhanden sind oder bis es nicht mehr möglich ist, Aufträge einzuplanen, ohne dadurch die Abgabetermine von Aufträgen anderer Produktfamilien zu verletzen (kritischer Punkt). Dabei kann es allerdings vorkommen, dass, wenn es nicht möglich ist, einen Auftrag mit einer hohen Priorität zu fertigen, ein Auftrag derselben Produktfamilie mit einer geringeren Priorität gefertigt wird, um eine eventuelle spätere Major-Rüstung zu vermeiden. Ist ein kritischer Zeitpunkt erreicht, wird wiederum die Produktfamilie mit der geringsten Gesamtbearbeitungszeit gewählt. Hiermit soll eine mögliche Major-Rüstung vermieden werden. Dies kann durch die höhere Wahrscheinlichkeit, dass eine kleinere Produktfamilie vollständig bearbeitet werden kann, bevor der nächste kritische Punkt erreicht wird, erklärt werden.

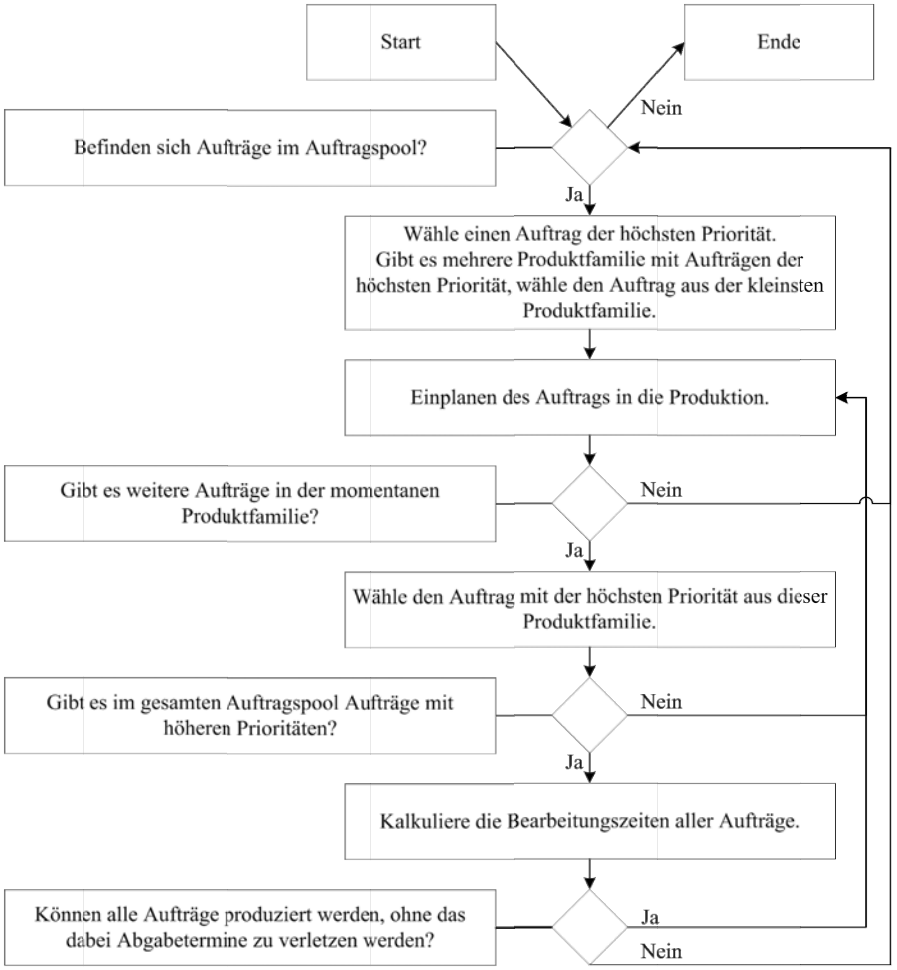



***Abbildung 3:*** *Sequenzierungsalgorithmus*

Mit der Sequenzierung ist ein ereignisbasierter Allokationsalgorithmus verbunden. Dieser wird genutzt, wenn ein kritischer Punkt erreicht ist, aber noch nicht alle Aufträge einer Produktfamilie bearbeitet wurden. Die noch vorhandenen Aufträge einer Produktfamilie werden zu einem zentralen Allokationsblock geschickt. Ebenso werden jeweils nach drei Tagen alle sich momentan nicht in der Produktion befindenden Produktfamilien zu diesem zentralen Allokationsblock geschickt, was als zeitlich definierte Allokation bezeichnet wird. Für beide Allokationsarten ist das Verfahren im Allokationsblock zunächst identisch. Die Produktfamilien werden nach ihrer Gesamtbearbeitungszeit aufsteigend sortiert, anschließend wird als erstes die Produktfamilie mit einer möglichst kleinen Gesamtbearbeitungszeit gewählt, wobei mindestens ein Auftrag in dieser Produktfamilie von der höchsten Priorität sein muss. Bei der ereignisbasierten Allokation wird diese Produktfamilie auf die momentan am geringsten belastete Maschine zugewiesen, um eine Auslastungsbalance zwischen den Maschinen zu erreichen. Das Verfahren der zeitlich definierten Allokation funktioniert ähnlich, jedoch werden bei der Zuweisung die Aufträge mit Abgabetermin innerhalb der nächsten vier Arbeitstage betrachtet. Das geschieht, um die Eintrittswahrscheinlichkeit einer eventuellen Major-Rüstung zu verkleinern, die aufgrund von vielen gleich hohen Prioritäten von Aufträgen unterschiedlicher Produktfamilien auf einer Maschine entstehen kann.

## 3.2 Simulated Annealing

Simulated Annealing (SA) wird häufig zur Lösung von Maschinenbelegungsproblemen eingesetzt (Schocke 2000, S.38). Für den hier vorgestellten Ansatz wird das Simulated Annealing mit einem Simulationsmodell gekoppelt. Das Simulationsmodell fungiert nach März et al. (2012) als Bewertungsfunktion der Optimierung (Abb.4).

Das Simulated Annealing gehört zu den von der Natur inspirierten Meta-Heuristiken. Die Methode wurde gleichzeitig von Kirkpatrick et al. (1983) und Cerny (1985) entwickelt. Es soll den physikalischen Prozess des Abkühlens von Materialien nahe dem Schmelzpunkt simulieren (Schocke 2000). Wird zum Beispiel Metall langsam abgekühlt, erstarrt es in einem energetisch günstigen Zustand. Dieser Sachverhalt wird genutzt, um kombinatorische Probleme zu lösen. Das Vorgehen beim Simulated Annealing kann wie folgt zusammengefasst werden. Zunächst wird eine Initialisierungslösung erstellt. Auf Basis dieser werden Nachbarschaften vorgeschlagen und deren Zielwerte errechnet. Sind die Zielwerte einer neuen Nachbarschaft besser als die der zugrundeliegenden Nachbarschaft, so wird diese überschrieben. Sind sie nicht besser, wird mittels der Boltzmann-Verteilung entschieden, ob die neue Nachbarschaft dennoch ausgesucht wird. Das Simulated Annealing dient in dieser Anwendung der Bestimmung der besten Produktfamilienallokation auf die SMD-Maschinen.

Die Nachbarschaftssuche wurde an das von Naderi et al. (2009) vorgeschlagene Single-Point-Operator-Verfahren (SPO) angelehnt und stellt jeweils neue Produktfamilienallokationen zur Verfügung, wobei jeweils eine zufällig ausgewählte Produktfamilie von einer zufällig bestimmten SMD-Maschine auf eine andere zufällig gewählte SMD-Maschine verschoben wird. Um auszuschließen, dass eine Maschine wesentlich mehr oder weniger belastet wird als die anderen, wird auf Basis der gesamten Produktfamilienanzahl eine Ober- und Untergrenze für die Produktfamilienmenge je Maschine definiert. Die Schrittweite (Anzahl der Produkt-

familien, die bei einer Nachbarschaftssuche neu allokiert werden) sollte ebenfalls behutsam gewählt werden, da sich die Meta-Heuristik wie eine Zufallssuche verhält, falls die Schrittweite zu groß ist.

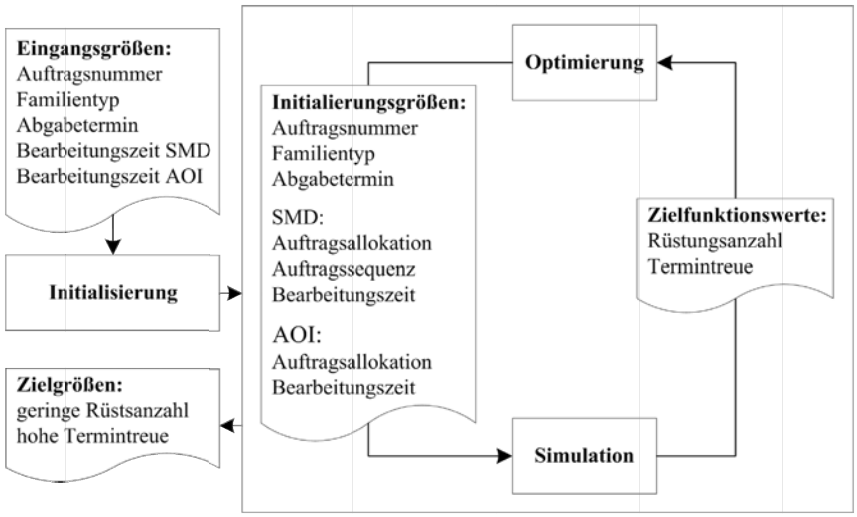



***Abbildung 4:*** *Simulation als Bewertung der Optimierung*

Nach der Auftragsallokation auf die SMD-Maschinen wird der in Abbildung 3 bereits beschriebene Sequenzierungsalgorithmus genutzt, um eine Auftragsreihenfolge für die SMD-Maschinen zu erstellen. Anschließend erfolgt die Allokation der Aufträge auf die AOI-Maschinen auf Basis der Reihfolgeplanung der SMD. Die Aufträge werden nach dem voraussichtlichen Fertigstellungszeitpunkt auf der ersten Stufe (SMD) geordnet und anschließend unter Beachtung einer gleichmäßigen Auslastung auf die AOI zugewiesen. Sind die Auftragsallokationen der SMD und AOI und die Reihenfolgeplanung der SMD abgeschlossen, wird ein ereignisdiskretes Simulationsmodell gestartet.

Um die hohen Rechenzeiten der ereignisdiskreten Simulation zu umgehen, wird ein Auftrag zu einem Flussobjekt aggregiert. Die Gesamtbearbeitungszeiten werden zuvor aus den bekannten Taktzeiten je Leiterplatte und der Auftragsmenge berechnet. Alle anderen Informationen müssen nicht zusammengefasst werden, da sie für alle Produkte eines Auftrags einheitlich sind.

Das Simulationsmodell bearbeitet die Aufträge nach ihren Arbeitsplänen, lediglich die Bearbeitungsabfolge der AOI-Maschinen wird im Modell entschieden. Kommt ein Auftrag von der SMD auf eine leere AOI, so wird dieser direkt produziert. Bei einer Warteschlangenbildung werden die Aufträge anhand ihres Abgabetermins eingeplant, wobei die EDD-Prioritätsregel gilt. Am Ende der Simulation erfolgt die Rückgabe der Zielwerte an das Simulated Annealing. Da mehr als ein Zielkriterium betrachtet werden soll, ist es notwendig, den Entscheidungsprozess genau zu definieren. Dominieren die Zielwerte der neuen Nachbarschaft die der alten, so wird diese überschrieben. Dominiert die alte Nachbarschaft die neue oder dominiert keine der beiden Lösungen die andere, so wird eine gewichtete Summe gebildet. Über die

Boltzmann-Verteilung wird anschließend entschieden, ob die neue Nachbarschaft gewählt wird (Ulungu et al. 1999).

Um einen sinnvollen Verlauf des Simulated Annealing sicherzustellen, ist die Wahl der Steuerparameter Temperatur, Iterationsanzahl und Abkühlungsrate entscheidend, da diese einen Einfluss sowohl auf die Dauer und Güte der Tiefensuche als auch auf den Entscheidungsprozess, ob eine Lösung gewählt wird, haben (Pirlot 1996). In der vorhandenen Anwendung wurde die Starttemperatur in einem Intervall zwischen 20 und 30 Grad gewählt. Die Iterationsanzahl liegt in einem Wertebereich von 10 bis 20 Iterationen je Temperatur. Die Abkühlungsrate hat einen linearen Verlauf mit einem Abkühlungswert zwischen 0,1 und 0,25 Grad. Die gewählten Steuerparameter wurden auf Basis von Experimenten ermittelt.

## 4 Ergebnisse

Für die Anwendung der beiden vorgestellten Lösungsansätze wurde ein reales Produktionsprogramm mit folgenden Parametern genutzt:

- Rüstzeiten: 65 Minuten (Major-Rüstung), 20 Minuten (Minor-Rüstung)
- Auftragsanzahl: 164
- Produktfamilienanzahl: 41
- Produktionsmenge: 40 - 109.920 Leiterplatten/Auftrag
- SMD Bearbeitungszeit: 4 - 3.142 Minuten
- AOI Bearbeitungszeit: 4 - 4.351 Minuten

Tabelle 1 vergleicht die Ergebnisse der beiden vorgestellten Lösungsansätze mit den Prioritätsregeln SPT und EDD und der Familienproduktion.

***Tabelle 1:*** *Vergleich verschiedener Lösungsansätze für das Hybrid-Flow-Shop-Maschinenbelegungsproblem*

| | Makespan (Minuten) | Major-Rüstungen (Anzahl) | Verspätungen (Anzahl) | Ø-Verspätung (Minuten) |
|---|---|---|---|---|
| Familienproduktion | 59.550 | 37 | 40 | 2.686 |
| Shortest Process Time | 63.690 | 129 | 23 | 3.760 |
| Earliest Due Date | 62.385 | 100 | 0 | 0 |
| Integrierte simulationsbasierte Optimierung | 59.780 | 43 | 5 | 694 |
| Simulated Annealing | 59.910 | 45 | 0 | 0 |

Der Eintrag „Familienproduktion“ steht für die Strategie, die das Unternehmen des Fallbeispiels nutzt und bei dem eine Familie komplett produziert wird, wenn sie auf eine Linie kommt. Sie stellt somit eine Lösung mit einer minimalen Major-Rüstungsanzahl und einer minimalen Zykluszeit dar. In der Ausgangssituation wird eine feste Allokation genutzt. Auf jede Linie wird ein fixes Portfolio an Produkt-

familien zugewiesen. Die Sequenzierungsentscheidung wird auf Basis von Erfahrungswerten getroffen.

Wie erwartet können die häufig verwendeten Prioritätsregeln SPT und EDD das vorhandene Problem nicht zufriedenstellend lösen. Die Ergebnisse zeigen, dass entweder vermehrt Liefertermine nicht gehalten werden können oder eine große Anzahl an Major-Rüstungen ausgeführt werden müsste.

Um die Systemauslastung zu verbessern, wurden in die integrierte simulationsbasierte Optimierung zusätzliche Effizienzkriterien eingeführt. Anhand derer können die Verspätungen bei dieser Methode erklärt werden. Zum einen wurde eine Regel eingeführt, durch die eine Produktionsentscheidung für Aufträge gefällt wird, die wegen eines kritischen Punktes eigentlich nicht mehr produziert werden dürften. Dies geschieht, wenn es sich um den letzten Auftrag einer Produktfamilie handelt. Ziel ist es, so eine Major-Rüstung für nur einen Auftrag einzusparen. Zum anderen entstehen Verspätungen durch die wenigen Aufträge mit einer Gesamtbearbeitungszeit auf SMD und AOI von über 5 Tagen. Diese könnten nach der ersten Produktionsstufe gesplittet werden, wodurch es möglich wäre die erste Auftragshälfte bereits auf einer AOI zu überprüfen, während sich die zweite Hälfte noch im Bestückungsprozess befindet. Es ist wahrscheinlich, dass durch das Splitting die Verspätungen bei diesen Aufträgen erheblich minimiert werden könnten.

Das Ergebnis der simulationsbasierten Optimierung wird bereits nach etwa 15 Sekunden Simulationszeit erreicht. Die etwas besseren Ergebnisse des Simulated Annealing, 45 Major Rüstungen ohne Verspätungen, werden erst nach etwa 1500 Schleifendurchläufen erreicht, was einer Rechenzeit von etwa 3 Stunden entspricht (System: CPU 4 x 2,6 GHz, RAM 8 GB, Windows 7). Anhand eines Laufes mit 4400 Iterationen, also einer wesentlich größeren Durchsuchung des Lösungsraumes, konnte gezeigt werden, dass das Ergebnis mit 1500 Iterationen bereits sehr gut ist, denn es wurden hierbei ebenfalls 45 Major-Rüstungen ohne Verspätungen erreicht.

Es wurde gezeigt, dass die beiden vorgestellten Lösungsansätze hinsichtlich der Termintreue der produzierten Aufträge ein wesentlich besseres Ergebnis erzielen, als die bisher im betrachteten Unternehmen genutzten Regeln zur Auftragseinplanung. Die etwas erhöhte Anzahl der Rüstvorgänge stellt hierbei kein Problem dar.

Um die Optimierungsmodelle noch besser an die Realität anzunähern, könnten Erweiterungen im Modell vorgenommen werden. Maschinenstörungen könnten direkt in den Simulationsmodellen abgebildet werden. Weiterhin gibt es im realen Produktionssystem für bestimmte Produkttypen noch weitere Fertigungsprozesse, die in die Betrachtung mit aufgenommen werden könnten.

## Literatur

Anderson, M.; Ng, A.H.C.; Grimm, H.: Simulation Optimization for Industrial Scheduling Using Hybrid Genetic Representation. In: Jefferson, T.; Fowler, J.; Mason, S.; Hill, R.; Moench, L.; Rose, O. (Hrsg.): Proceedings of the 2008 Winter Simulation Conference (WSC), Miami (USA), 07.-10. December 2008, S. 2004-2011.

Baker, K.R.; Trietsch, D.: Principles of Sequencing and Scheduling. Hoboken, New Jersey: John Wiley & Sons, Inc. 2009.

Cerny, V.: Thermodynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Algorithm. Journal of Optimization Theory and Applications 45 (1985) 1, S. 41 – 51.

Graham, R.L.; Lawler, E.L.; Lenstra, J.K.; Kan, A.H.: Optimization and Approximation in Deterministic Sequencing and Scheduling: a Survey. In: Hammer, P.L.; Johnson, E.L.; Korte, B.H. (Hrsg.): Annals of Discrete Mathematics 9 (1979). Amsterdam: Elsevier S. 287-326.

Johnson, S. M.: Optimal two- and three-stage production schedules with setup times included. Naval Research Logistics Quarterly 1 (1954) 1, S. 61-68.

Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P.: Optimization by Simulated Annealing. Science 220 (1983) 4598, S. 671 – 680.

März, L.; Krug, W.; Rose, O.: Simulation und Optimierung in Produktion und Logistik. Heidelberg: Springer 2011.

Naderi, B.; Zandieh, M.; Roshanaei, V.: Scheduling hybrid flowshops with sequence dependent setup times to minimize makespan and maximum tardiness. The International Journal of Advanced Manufacturing Technology 41 (2009) 11-12, S. 1186 – 1198.

Pinedo, M.L.: Scheduling: Theory, Algorithms, and Systems. New York: Springer 2012.

Pirlot, M.: General local search methods. European Journal of Operational Research 92 (1996), S. 493 – 511.

Schocke, K.-O.: Maschinenbelegungsplanung mehrstufiger Fließfertigungen: vom Modell zum Leitstand für die chemische Industrie. Wiesbaden: Deutscher Universität Verlag, Gabler 2000.

So, K.C.: Some Heuristics for Scheduling Jobs on Parallel Machines with Setups. Management Science 36 (1990) 4, S.467-475.

Tolujew, J.; Reggelin, T.: A mesoscopic approach to modeling and simulation of logistics processes. In: Jain, S.; Creasey, R.R.; Himmelspach, J.; White, K.P.; Fu, M. (Hrsg.): Proceedings of the 2011 Winter Simulation Conference (WSC), Phoenix (USA), 11.-14. December 2011, S. 1513-1523.

Ulungu, E.L.; Teghem, J.; Fortemps, P.H.; Tuyttens, D.: Mosa Method: A Tool for Solving Multiobjective Combinatorial Optimization Problems. Journal of Multi-Criteria Decision Analysis 8 (1999), S. 221 – 236.

Voß, S: The Two - Stage Hybrid - Flowshop Scheduling Problem with Sequence - Dependent Setup Times. In: Fandel, G; Gulledge, T.; Jones, A. (Hrsg): Operations Research in Production Planning and Control. Berlin, Heidelberg: Springer 1993, S. 336-352.

Wittrock, R.J.: Scheduling parallel machines with major and minor setup times. The International Journal of Flexible Manufacturing Systems, 2 (1990) 4, S. 329-341.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Reduction of Simulation Effort for Simulation-based Optimisation

## *Reduzierung des Simulationsaufwandes bei simulationsbasierter Optimierung*

Tobias Uhlig, Oliver Rose, Universität der Bundeswehr München, Neubiberg (Germany), tobias.uhlig@unibw.de, oliver.rose@unibw.de

**Abstract:** Simulation often represents a bottleneck with regard to the applicability of simulation-based optimisation. Generally, simulation-based optimisation requires a sufficient amount of simulation runs to operate reasonably well. Since every simulation run is a valuable resource, it is important to use it efficiently. This is especially true for stochastic simulation, where additional runs are needed to provide statistical significance for the evaluation of candidate solutions. This paper discusses three approaches that aim to reduce the amount of required simulation runs for simulation-based optimisation. The approaches are compared to the classic technique, which relies on a fixed number simulation runs to evaluate a candidate solution. Our experiments show a significant gain in efficiency for the new approaches.

## 1 Introduction

Simulation-based optimisation (SBO) is an established optimisation approach relying on simulation as a tool to evaluate candidate solutions. In recent years its popularity has increased significantly (Hagendorf 2010). This can at least partially be attributed to the increased availability of computing power. While the optimisation component usually is quite undemanding, the simulation oftentimes requires a lot of computing power. Unfortunately, an effective SBO approach usually needs a large amount of simulation calls to evaluate candidate solutions. This effect is amplified whenever stochastic simulation models are considered, since repeated simulation runs are required for one model to guarantee statistical significance. Generally, the number of simulation runs N reduces the corresponding standard deviation by a factor of square root of N (Law and Kelton 1999). A common practice is to use ten repetitions to evaluate a simulation model. A reduced simulation effort or faster simulation can be attained in many different ways, e.g., the use of more efficient simulation procedures, the use of more abstract simulation models, or the use of distributed parallel simulation. Our goal is to reduce the actual number of simulation runs, without causing a negative impact on the returned quality. In other words, we want to use the available simulation runs more

efficiently. We take general inspiration from the field of metaheuristics. Basically, SBO can be considered a (meta)heuristic that uses simulation to evaluate candidate solutions (see Fig. 1). In fact, simulation can be seen as the basis for a fitness function that asserts the quality of solutions generated by the optimiser. Previous research on metaheuristics has already considered the challenge of noisy fitness functions. On the whole, three classes of approaches were proposed to handle uncertain fitness functions, i.e., explicit averaging, implicit averaging, and modification of selection (Jin and Bratke 2005). With regard to SBO the work of Chen (2003) demonstrated an efficient sampling approach that lead to a significant reduction of simulation runs.

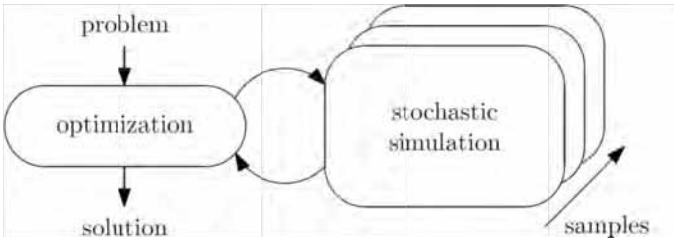



***Figure 1:*** *Basic concept of simulation based optimisation.*

For our paper, we rely on the idea that most selection processes in SBO can be broken down to a comparison of two candidate solutions. With this in mind, we want to compare four approaches to handle uncertainty in SBO.

This paper has the following structure. After this introduction, the next section discusses the optimisation approach used for our experiments. Subsections 2.1 to 2.4 introduce the techniques used to compare potential candidate solutions using simulation. Afterwards, we elaborate our experimental setup (see Section 3) and present the results of our experiments in Section 4. We close with a discussion (see Section 5) and summary (see Section 6) of the content provided by the paper.

## 2 The Optimisation Approach

This section discusses different optimisation approaches used in conjunction with simulation. In detail, we will focus on the decision processes used in optimisation to identify suitable solution candidates. Generally, we use a classic evolutionary algorithm for optimisation. For selection purposes, we evaluate four approaches to compare two candidate solutions:

- The classic comparison based on averaging the results of a fixed number of simulation runs,
- a phase-based approach that adapts the number of simulation runs used for a comparison as the optimisation progresses,
- a statistical approach, that uses Student's t-test to determine the number of simulation runs used for a comparison on a case by case basis,
- and a novel "Natural" approach.

These approaches rely on two or more simulation runs to assess the relative fitness of two candidates. For each candidate solution we use at least one simulation run to

evaluate it. Additional runs may be used to increase the statistical significance of the evaluation. The results of the evaluation are used by the evolutionary algorithm to select promising candidate solution for further optimisation. Considering, the evolutionary algorithm we recognise three aspects of the optimisation, which require simulation runs:

1. We already mentioned the necessity to use more than one simulation run to evaluate a candidate solution. This explicit *sampling* is required to provide statistical significance. With each additional simulation run we increase the reliability of the observed fitness of a candidate solution. However, eventually adding more simulation runs results in diminishing returns.
2. Generally, evolutionary algorithms operate by accumulating small changes to generate a good result over many *generations*. Accordingly, the number optimisation steps must be large enough to support this process. To this end, we need to allocate a large amount of simulation runs to the evaluation of successive candidate solutions. Once again, we observe diminishing returns when to many simulation runs are used for the generations, since the optimisation process eventually converges. At the point of convergence, the process has reached a local optimum and further optimisation will not result in an improvement.
3. Finally, we can assign simulation runs to consider more than one candidate solution in parallel – a *population* of potential solutions. Basically, a large population increases the robustness of an optimisation process. It can prevent premature convergence to a single local optimum by considering alternative solutions. Additionally, a larger population facilitates implicit sampling and thereby improves the performance of the optimisation in the context of uncertain evaluation functions like stochastic simulation (Arnold and Beyer 2003).

Fundamentally, all three aspects – sample size, generations, and population size – benefit from additional simulation runs. To realise an efficient optimisation it is necessary to balance their demands to match the number of available simulation runs. The following subsections discuss four strategies to distribute the available runs to the three aspects.

## 2.1 The Classic Approach Using Fixed Sample Sizes

The classic approach for SBO relies on fixed parameter assignments. To this end, the experimenter defines a parameter setup in advance to the optimisation. During the optimisation process these parameters remain unchanged. For our scenario, using SBO with an evolutionary algorithm, the parameters are: number of generations, population size, and the number of simulation runs used to sample one candidate solution. These values are often assigned based on intuition or experimentation. There are no exact guidelines but rather rules of thumb for their setup. Considering, for example the sample size used to provide statistical significance a value of ten is often used. Employing ten simulation runs for the evaluation is regarded as a reasonable compromise between calculation time and accuracy. Subsection 4.1 illustrates the relative influence of the three parameters based on experimental evidence.

## 2.2 Phase-based Approach

The phase based approach was proposed in Frank et al. (2013). It divides the optimisation into three phases: exploration, exploitation and accurate simulation. Based on the assumption that during the initial exploration phase candidates solutions are more varied it uses only a small amount of simulation runs in the beginning. As the optimisation progresses the algorithm switches to the exploitation phase and eventually to accurate simulation phase and uses more simulation runs. More simulation runs are required later on, since it is more difficult to distinguish candidates the closer we get to the optimum. The algorithm decides when to switch to the next phase based on the observation of the optimisation progress. It monitors the progress of the fitness and calculates the slope of the fitness curve. Whenever the algorithm reaches a predefined slope it switches to the next phase. Equation 1 defines the calculation of slope (m) for each generation (g) using an offset (o). The offset is used to avoid erratic switching caused by short-term stagnation of the optimisation. Using the offset the algorithm only reacts to longer-term changes over multiple generations.

$$m = \frac{fitness_{g-o} - fitness_g}{d} \tag{1}$$

## 2.3 Statistical Approach Using Students T-Test

The statistical approach is discussed in more detail in Uhlig and Rose (2015). The basic idea is to start with a small sample size of simulation runs to get an initial evaluation of two candidate solutions. Subsequently a paired Student's t-test is applied to the samples. If the test indicates a significant difference between the samples a comparison of the sample averages is returned. Whenever the test is inconclusive additional simulation runs are performed. Essentially, this approach reduces the number of simulation runs for cases were candidate solutions are sufficiently different and only uses more runs whenever the candidates have similar quality. Algorithm 1 summarises the employed technique.

***Algorithm 1:*** *The statistical approach using Student's t-test.*

```
1.  For every model execute 3 simulation runs and
    evaluate the objective value for each run.
2.  Use the t-test for the samples of objective values
    using the null hypothesis (H0) that the two samples
    have equal means. We calculate the minimal
    significance level (p) at which we can reject the H0.
3.  If p > 0.5 and the maximum number of simulation runs
    (20 per model) is not reached go to step 4 otherwise
    go to step 5.
4.  Execute one more simulation run and add the result to
    the appropriate sample. Return to step 3.
5.  Compare the two models using the sample average of the
    respective objective values.
```

### 2.4 "Natural" Approach

The final approach we will discuss is inspired by biological evolution. Accordingly, we refer to it as the "Natural" Approach. Conceptually, it is very simple and requires in its current implementation no additional parameters. It relies on implicit sampling, but in contrast to simply using a large population it does so over time. This is achieved by limiting the use of simulation results to just one generation. During each generation candidate solutions are evaluated based on only one simulation run. Subsequently, the simulation results are discarded at the end of each generation and for the next generation the candidates are re-evaluated. This mimics selection processes of natural evolution. In nature, an organism does not have an assigned fixed fitness value, but rather has to repeatedly prove to be fit enough for survival and procreation. Similarly, a candidate solution in the "Natural" approach is compared repeatedly with other candidates. Only candidates that are successful consistently over multiple generations will "survive" the optimisation process. In contrast, statistical anomalies in the simulation will only let a candidate survive a small amount of generations.

## 3 Experimental Setup

To test and compare the proposed approaches we perform a series of experiments. Our experimental setup relies on the SEREIN (Uhlig 2013) presented in Uhlig (2015). The framework provides implementations of various evolutionary algorithms and supports a comparison-based fitness evaluation. With regard to the implementation of our "Natural" approach, it is important that the framework has a distinct fitness concept. This concept models fitness as an observation by a selecting agent rather than being an inherent feature of a candidate solution. This approach lets the selecting agent handle all fitness decisions and accordingly it can easily reset the values for each generation.

The complete setup we use is illustrated in figure 2. It shows the four implementations of our comparison approaches depicted in blue and the interfaces and classes provided by the SEREIN framework in white. Depicted in orange are classes used to model a simulation component of SBO. We do not use an actual simulation for our experiments, since the amount of computation time required to perform all experiments is too large. Instead of an actual simulation we use a benchmark function with comparable behaviour and complexity. This technique has already been used successfully in Frank et al. (2013). We use the deterministic Rastrigin function (Mühlbein et al. 1991) and add normal distributed noise to it. In addition to the reduced computation time, this approach has two advantages. On the one hand, we know the theoretical optimum of the fitness function, which usually is not attainable for an actual simulation. On the other hand, the added error term in conjunction with a deterministic function provides fine grained control of the observed uncertainty.

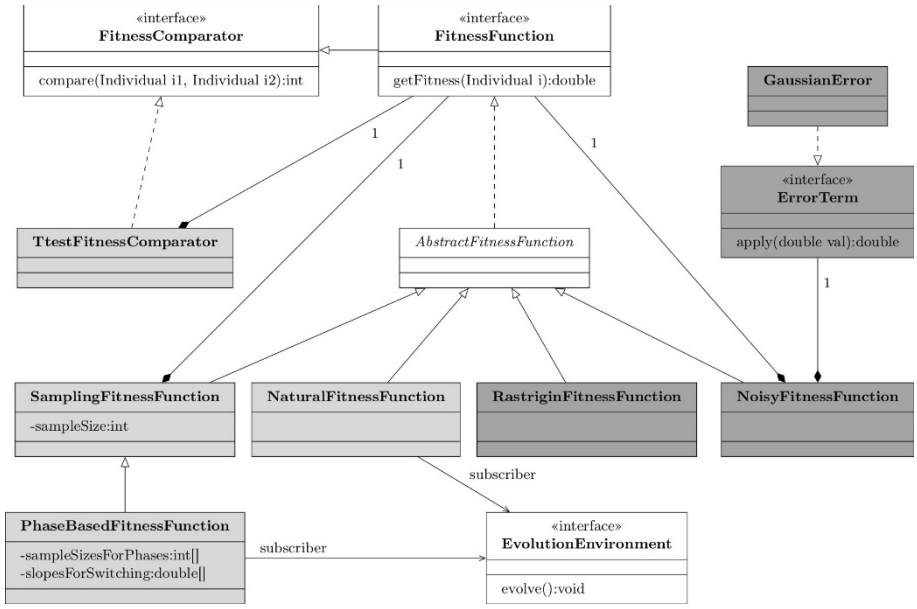



***Figure 2:*** *UML class diagram for our experiment setup. Depicted white, classes provided by the SEREIN framework; in light gray, our implementations for evaluation methods that can be used in SBO; and depicted in dark gray, classes that are used to model a benchmark problem that mimics a simulation.*

For our experimental design, we use a standard evolutionary algorithm. Most parameters for this algorithm were determined using a meta-optimisation, e.g., mutation and selection operators. For our experiments, we mainly consider the three previously discussed parameters: number of generations, population size, and number of sample simulation runs see. We considered the following settings:

- Population size – 1 to 100
- Sample size – 1 to 20 simulation runs (where applicable)
- Generations – Implicitly given, each optimisation terminated after 36,000 simulation runs
- Error term – $N(0, \sigma^2)$, $\sigma = 0...10$

Each experiment is repeated 100 times to ensure statistical significance. In the following section we discuss the results of our experiments.

## 4 Experimental Results

The results of our experiments are divided into two parts. For the first part, we analyse the general impact of the three parameters – generations, population, and sample size – on the optimisation process. We evaluate them given different magnitudes of uncertainty using the classical approach with fixed parameter settings. We use the results to illustrate influence of these parameters and to determine a

reasonable reference setup. This reference is used in the second part of the experiments to compare the other approaches to the classic setup.

## 4.1 Influence of Parameters

Figures 3 and 4 exemplify the results of our initial experiments. Generally, a higher amount of noise causes the performance of SBO to decrease. Increasing the population size (fig. 3, on the left) or the number of samples reduces the performance decline. However, both values have to be balanced reasonably since too large values for both of them subtracts from the number of available optimisation steps (fig. 4).

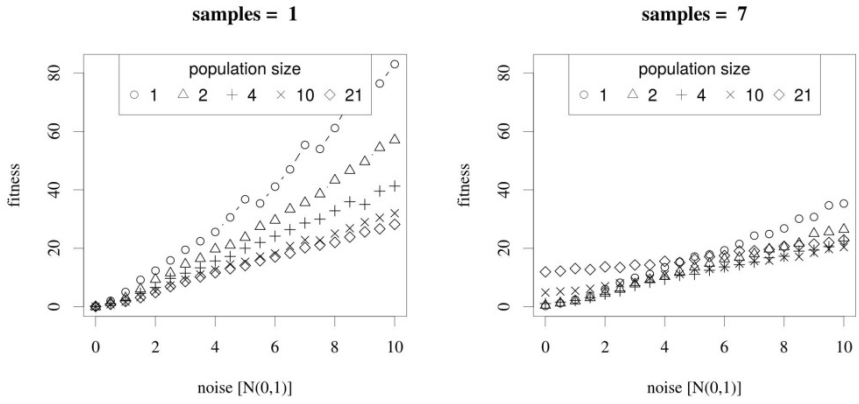



***Figure 3:*** *Influence of population size and the magnitude of the error on SBO*

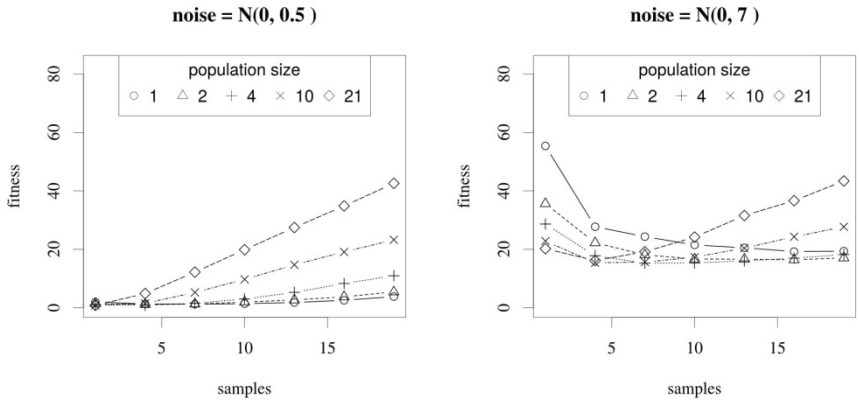



***Figure 4:*** *Influence of sample size and population size on SBO.*

## 4.2 Comparison of Approaches

For our final comparison, we use the optimal setups determined in the initial experiments. The four compared approaches are:

- The classic approach implemented by the SamplingFitnessFunction (SFF)
- The phase-based approach implemented in the PhaseBasedFitnessFunction(PBF)
- The "Natural" approach (Natural)
- The statistical approach implemented in the TtestFitnessComparsion (TFC)

The results of our experiments are shown in figure 5. All alternative approaches exceeded the performance of the reference. However, the phase based approach shows only minimal improvements. Considering the added amount of parameterisation required (sample sizes for each phase and switching parameters) we do not recommend its use over a well set up classic approach. The Natural approach returns good results and needs only minimal parameterisation. Although the discarding of simulation results seems to be counterintuitive the results are promising. Finally, the statistical approach using Student's t-test outperformed all other techniques significantly. Using additional simulation runs only when they are actually required lead to the most efficient use of available resources. In the next section, we will discuss the results of the paper and put them into the right context.

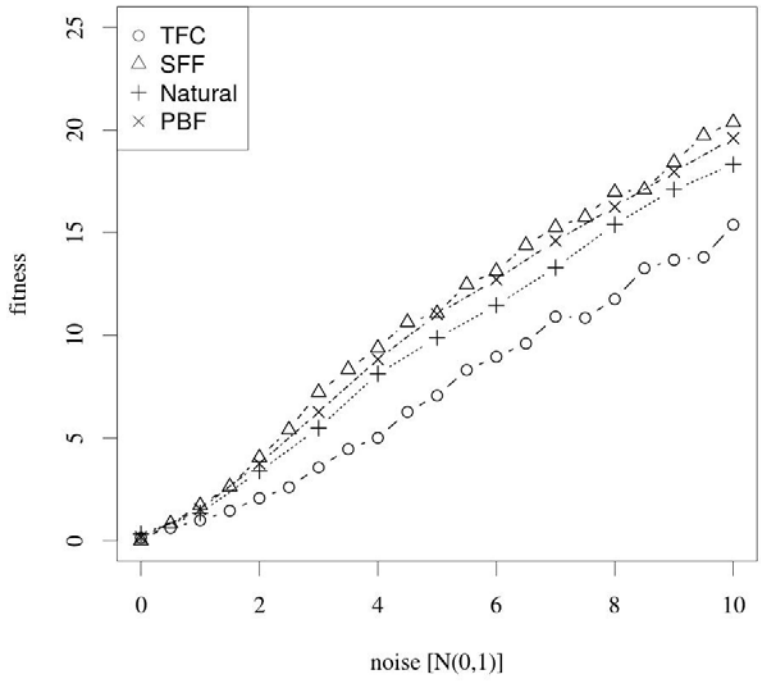



***Figure 5:*** *Comparison of approaches (smaller fitness values are better).*

## 5 Discussion

Our experiments provide some interesting points for discussion. First, we need to address the obvious critique regarding the use of a surrogate function instead of an actual simulation. There is a strong indication that our theoretical results can be applied to actual SBO, since the employed function is comparable in behaviour and complexity. Furthermore, the work of Frank et al. (2013) suggests the transferability of the results. Initial experiments we performed using actual simulation supports our assumption. An additional point for discussion is the simple error model we employed. Granted, we cannot expect normally distributed errors sharing one standard deviation for all candidate solutions. Therefore, further experiments using different error models are required.

The discussed optimisation approaches were mainly compared to the classic reference approach. Currently, we have not tested the OCBA approach (Optimal Computing Budget Allocation) approach by Chen (2003). Therefore, we cannot provide any insights on its effectiveness in comparison to the discussed approaches.

One important result of our experiments is the large influence of the parameter setup. For example, the good performance of the phase based approach observed in Frank (2013) can probably be attributed to a suboptimal setup of the reference approach. Additionally, the phase-based approach is very hard to set up properly, since it needs a huge amount of parameters. Switching phases based on the slope is very sensitive. It might be beneficial to consider other measurements to determine the appropriate time of switching to the next phase. For example, the observed variance of the fitness of a given population can be used instead. A study evaluating this approach should also consider experiments with regard to the amount and setup of phases. For further use of the phase-based approach, it would be beneficial to develop a deeper understanding on the principles of its parameterisation. Otherwise, the amount of parameters makes the phase-based approach rather impracticable.

Regarding the other approaches, we can also consider further improvements. For a discussion on the t-test based approach, we refer you to Uhlig and Rose (2015). The natural approach can modified to keep the results of simulation runs instead of discarding them completely. For example, it could simply update the fitness for each generation instead of completely resetting it. Potential changes should probably try to preserve the simple nature of this approach, since the simple parameterisation of the approach is its biggest drawing point.

## 6 Summary and Outlook

In this paper, we compared three advanced approaches to the traditional comparison strategy used for simulation-based optimisation. All of them outperformed the classical approach that relies on a fixed amount of simulation runs. Nevertheless, our experiments show that the traditional approach can operate quite efficiently, if it is parameterised correctly. Determining adequate parameters, however, is challenging. We either have to rely on intuition and previous experience, which can be misleading, or we have to determine the parameters by performing extensive experiments, which is costly. This predicament is also true for the phase-based approach. With this in mind, it is important to point out that the other two approaches are much more robust with respect to their parameterisation. The “Natural” approach returns

good results, while only requiring a minimum of setup. The statistical approach shows the best performance and significantly exceeded all other approaches.

Future research in this area needs to evaluate other advanced approaches in comparison to the proposed techniques. Furthermore, the proposed methods have to be evaluated in other scenarios. For a final verdict, we have to consider the use of actual simulations in contrast to the surrogate benchmark function and we have to evaluate the influence of different kinds of uncertainty.

## Acknowledgement

We want to acknowledge the work of our student Dominik Schwindling, who has performed initial experiments for our studies using actual simulation models. His results provide a strong indication for the transferability of the presented results.

## References

Arnold, D. V.; Beyer, H.-G.: A comparison of evolution strategies with other direct search methods in the presence of noise. Computational Optimization and Applications 24(2003) 1, pp. 135-159.

Chen, C.-H.: Efficient sampling for simulation-based optimization under uncertainty. Fourth International Symposium on Uncertainty Modeling and Analysis, College Park, Maryland (USA), 2003, pp. 386-391.

Frank, M.; Laroque C.; Uhlig, T.: Reducing computation time in simulation-based optimization of manufacturing systems. In: Pasupathy, R.; Kim, S.-H.; Tolk, A.; Hill, R.; Kuhl, M. E. (eds.): Proceedings of the 2013 Winter Simulation Conference, Washington (USA), December 8th-11th2013, pp. 2710-2721.

Hagendorf, O.: Simulation based parameter and structure optimization of discrete systems. ARGESIM Report AR 67. Wien: ARGESIM / ASIM 2010.

Jin, Y.; Branke, J.: Evolutionary optimization in uncertain environments – A survey. IEEE Transactions on Evolutionary Computation 9 (2005) 3, pp. 303-317.

Law, A. M.; Kelton, D. M.: Simulation Modeling and Analysis, 3rd ed. Boston: McGraw-Hill Higher Education 1999.

Mühlenbein, H; Schomisch M.; Born J.: The parallel genetic algorithm as function optimizer. Parallel Computing Volume 17 (1991) 6, pp. 619-632

Uhlig, T.: Serein – a framework to Model Metaheuristics, website: http://sourceforge.net/projects/serein/, 2013.

Uhlig, T.: Self-replicating individuals – An approach to model metaheuristics and hyper-heuristics. Phd. Thesis, UniBw München, 2015.

Uhlig, T.; Rose, O.: Efficient use of simulation runs in simulation-based optimization. In: Yilmaz, L.; Chan, W. K. V.; Moon, V.; Roeder, T. M. K.; Macal, C.; Rossetti, M. D. (eds.): Proceedings of the 2015 Winter Simulation Conference, Savannah (USA), December 7th-10th 2015 [submitted].

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# A Simheuristic for the Waste Collection Problem with Stochastic Demands in Smart Cities

## *Simheuristischer Ansatz zur Optimierung der Müllentsorgung mit stochastischer Nachfrage in Smart Cities*

Aljoscha Gruler, Angel A. Juan, Universitat Oberta de Catalunya, Barcelona (Spain), aljoscha.gruler@gmail.com, ajuan@uoc.edu

Christian Fikar, Patrick Hirsch, University of Natural Resources and Life Sciences, Vienna (Austria), christian.fikar@boku.ac.at, patrick.hirsch@boku.ac.at

Carlos Contreras Bolton, Universidad Andrés Bello, Santiago de Chile (Chile), ccontreras@gmail.com

**Abstract:** This paper describes a simheuristic to improve urban waste collection processes. Growing city-populations around the world lead to an increased importance of an effective organisation in basic municipal activities such as waste management. The Waste Collection Problem (WCP) can be formulated as a special instance of the well-known Vehicle Routing Problem (VRP), which has been subject to extensive research over the last decades. While traditional VRP and WCP solution approaches consider deterministic problem input variables, the complexity and uncertainty of real-life problem settings such as the collection of urban waste calls for new approaches. As such, the contribution of this paper is the development and description of a simheuristic methodology for the WCP through the combination of an efficient heuristic with Monte Carlo simulation, leading to realistic and reliable routing solutions.

## 1 Introduction

Cities are becoming ever-more complex systems. About half the world's population is currently living in urban areas, and it is projected that this percentage will increase to 70% in 2050 (United Nations, 2008). Many municipalities struggle to cope with newly arising challenges, as the existing urban infrastructure and organization does not meet the necessary standards concerning economic, social and environmental changes in urban areas (European Comission, 2011).

Nowadays, many cities across the world develop so called Smart City initiatives, which include a wide range of aspects to help municipalities make better use of their resources. A typical application field for such initiatives is waste management, as inefficient garbage collection can lead to several negative impacts such as traffic

congestions, unnecessary pollution and high noise levels, combining to a decreased urban standard of living (Mattoni et al. 2015; Tranos and Gertner 2012; Neirotti et al. 2014). Furthermore, waste management is a major cost factor with annual per capita costs of up to 200€ only for the collection of urban waste in European cities (The World Bank 2012).

Increasingly, the research society has turned its attention to the Waste Collection Problem (WCP), which is usually formulated as a special instance of the well-known Vehicle Routing Problem (VRP) (Beliën et al. 2011). The aim of the 'classical' VRP consists of optimising the costs of serving a range of customers with a capacitated vehicle fleet stationed at a central depot, whereby (i) every node is served by only one vehicle, (ii) all routes start and end at central depot, (iii) no customer is served twice, and (iv) the vehicle capacity cannot be exceeded (Daneshzand 2011). In this context, the WCP can be seen as reversed VRP, whereby a certain demand (i. e. waste) is collected from a range of customers (i.e. waste containers). A more detailed description of the WCP formulated as VRP is given in Sector 2.

The VRP has been subject to extensive research, especially in the development of different optimisation methods (e.g. metaheuristics). While the VRP was usually formulated as deterministic problem in the past, new powerful solution methods for stochastic VRP problem settings have been presented in the past years, which can be applied to a range of real-life instances. Caceres-Cruz et al. (2015) define such real-life VRP applications as rich VRP's, characterized by additional problem constraints (e.g. time windows) and a high level of uncertainty, i.e. stochastic inputs. A typical example of a rich VRP is the WCP, as urban waste collection is characterized by different uncertain input variables such as travel times within cities and garbage levels in waste containers. Furthermore, constraints such as opening hours of waste disposal sites and driver lunch breaks need to be considered for a realistic representation of the problem.

This paper presents a simheuristic approach as proposed by Juan et al. (2015) to solve the WCP. In line with the current trend in research to develop powerful hybrid approaches to solve optimisation problems, simheuristics combine simulation methods with (meta-) heuristics to consider the stochastic nature in real-life optimisation problems. After outlining the WCP as VRP and current research on the WCP in Sector 2, the fundamentals of our simheuristic methodology are presented in Sector 3. In the following, a simheuristic for the WCP is outlined in Sector 4, before the last sector gives a short conclusion and outlook on possible further work on the problem.

## 2 The WCP in the Context of VRP Research

The VRP is one of the most famous combinatorial problem settings and has been subject to intensive research over the last decades. As a special VRP instance, the WCP can be outlined on a graph *G = (N, E)*. Thereby, *N* defines a set of vertices $N = \{n_0, n_1, \dots, n_i, n_{i+1}, n_{i+m}\}$. A number of (homogeneous) collection vehicles with vehicle maximum capacity VMC are located at the central depot $n_0$. Furthermore, *i* waste containers with waste levels $w_i$ are presented through the node sub-set $\{n_1, \dots, n_i\}$. An additional constraint of the WCP in comparison to the traditional VRP is that waste collection vehicles have to empty themselves at a landfill (waste disposal site), once they have reached their vehicle capacity or before

returning to the depot. The landfills are presented by the nodes $\{n_{i+1}, n_{i+m}\}$. The travel costs (which might be actual costs, distances or environmental impact) $c_{i,j}$ of moving between two nodes are represented by a set *E* of edges.

In Figure 1 below, a WCP with two routes can be seen. Note that in route 1, the collection vehicle serves a range of waste containers, empties itself at a landfill and then returns to the central depot. In route 2 however, the collection vehicle collects waste until its vehicle capacity is reached, empties itself and then continues its route. Thus, a vehicle might empty itself and then continue its route (which is usually the case). A time window constraint at the central depot, defining the daily operating hours, prevent one vehicle from collecting the waste of all containers on a single route.

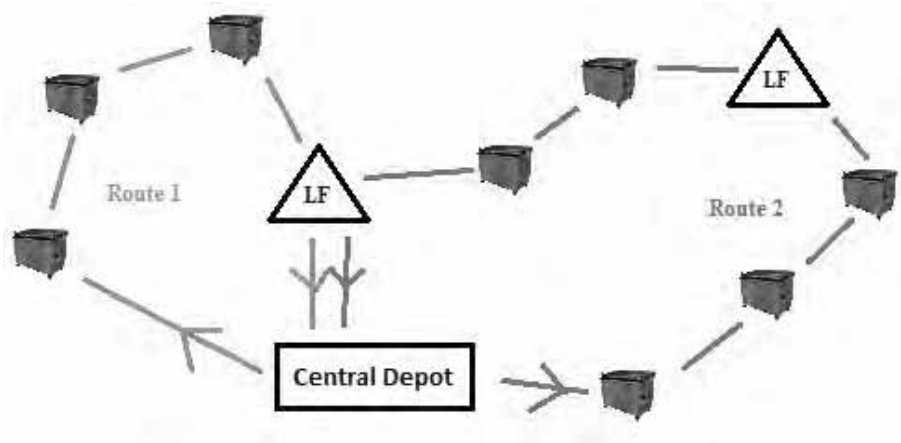



*Figure 1: Example of a WCP with two Routes*

The WCP as described above has been subject to a wide range of research. Most publications are based on different metaheuristic approaches such as Tabu Search, GRASP or Simulated Annealing. For example, Kim et al. (2006) present a metaheuristic approach based on Simulated Annealing which they implement in a commercial route planning software. In corporation with a North American waste collection service provider, they report route savings of up to 10% using their metaheuristic approach. To our best knowledge, their publication also provides the only publicly available benchmark for the WCP, which is based on real-life problem settings taken from their case studies and includes 10 different instances ranging from 102-2100 nodes per instance (these benchmarks are used to implement and test the proposed simheuristic as presented in the following sectors).

Using the same benchmarks to validate their findings, Benjamin & Beasley (2010) propose a hybrid approach combining variable neighbourhood- with tabu search approach, whereas Buhrkal et al. (2012) apply an adaptive large neighbourhood search. While these papers present competitive WCP solution methods, a major drawback is that they assume the WCP to be a deterministic problem setting. Considering the high level of uncertainty in urban waste collection, this is an unrealistic assumption.

In a case study presented by Johansson (2006), analytical modelling is connected with discrete-event simulation to evaluate different scheduling and routing policies considering real-time input data. The paper suggests savings of up to 26% in waste collection mileage through more informed routing decisions, leading to an operational waste management cost decrease of 10% in the Swedish city of Malmö. Similar to the approach presented in this paper, Johansson thereby considers stochastic waste levels to improve routing decisions.

## 3 Fundamentals of our Simheuristic Methodology

The simheuristic to solve the WCP is based on the metaheuristic framework combined with simulation as proposed by Juan et al. (2015) and outlined in Fig. 2 below. This framework can be applied to a wide range of stochastic combinatorial optimisation problems (COP's). Their approach makes use of the fact that there are many powerful solution methods (e.g. metaheuristics) to deterministic COP instances, and that the solution of a deterministic COP leads to acceptable and promising solutions to their stochastic counterparts.

Consider a stochastic COP in which certain input variables experience some level of uncertainty. In a simplification process, the deterministic counterpart to this problem is considered by replacing random/stochastic input variables with deterministic values. This might be done for example by estimating a stochastic value based on past experience. Note that these deterministic estimates have to be more or less reliable. In realistic problem settings like the WCP, we believe that reliable experience values exist to obtain good estimates.

In the following, the interaction between metaheuristic optimisation and simulation starts (highlighted by the light-blue background). Using any efficient (meta-) heuristic optimisation algorithm, new solutions to the deterministic COP are created. Each newly created solution is tested on its solution quality. If it is considered a promising solution, a fast simulation process is run with the problem setting, meaning that the input variable which was 'simplified' to create a deterministic COP instance is simulated a defined number of times. Using the results of the first simulation run, the quality and feasibility of promising deterministic COP solutions in a stochastic environment can be estimated. Note that it might happen that a very promising deterministic COP solution is discarded at this point due to the low quality or infeasibility of its stochastic counterpart.

Depending on the results, the most promising stochastic COP results (the 'elite solutions') are chosen to be run in a larger, more extensive simulation process. After the second simulation process, another quality and feasibility evaluation is made. In the following, the most promising elite solutions are subject to a reliability analysis, which is another advantage of the suggested simheuristic. A reliability analysis extends the information that can be retrieved through the optimisation process, as the simulation runs are used to evaluate the number of times that a suggested solution can be completed without any failures when considering simulated stochastic inputs. Using this additional information, decision makers can base their final decision not only on solution-costs, but also on the solution-reliability.

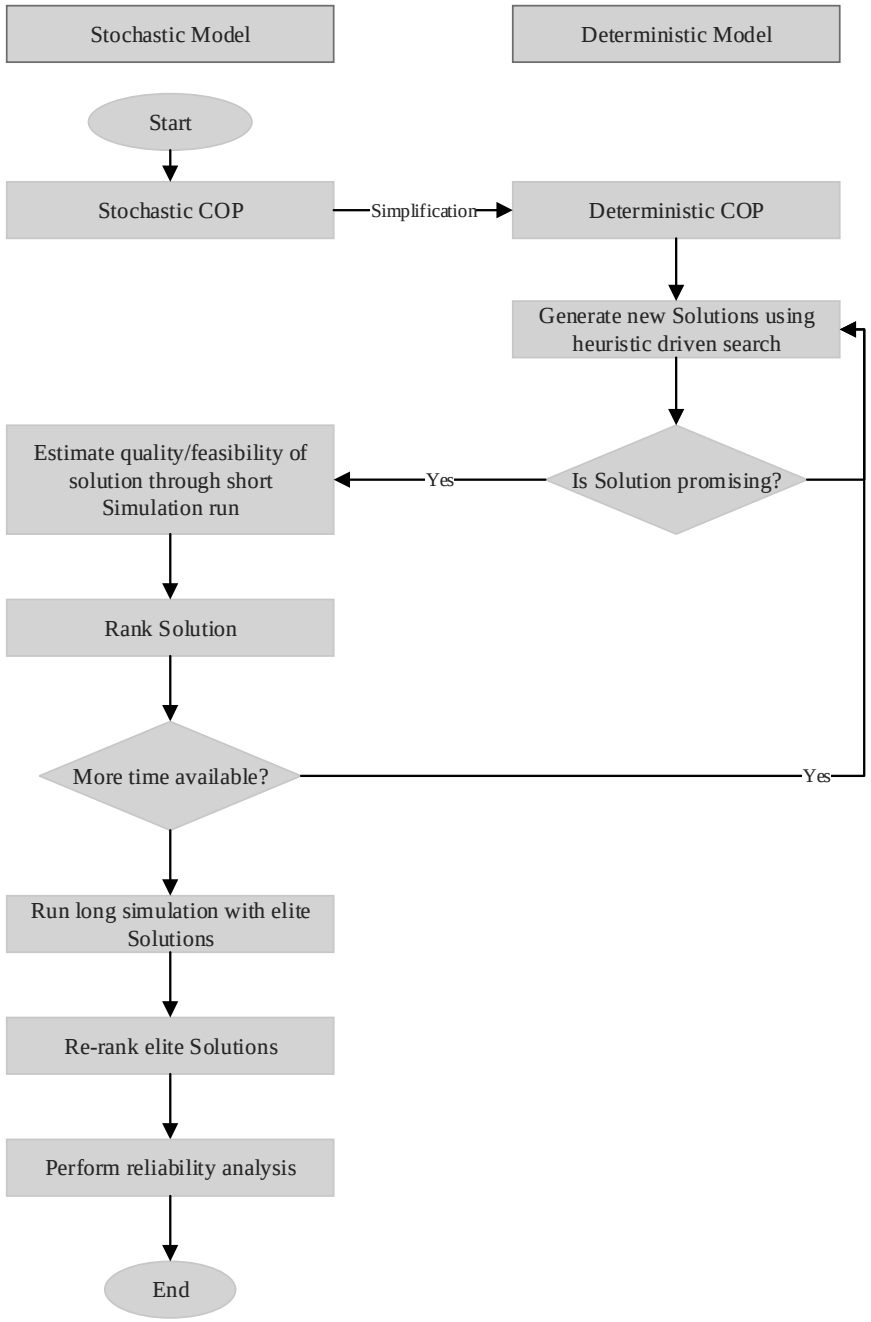



*Figure 2: Simheuristic Methodology as suggested by Juan et al. (2015)*

# 4 A Simheuristic for the WCP

In urban waste collection, route planners can never be sure about the actual travel times or waste levels at the containers. This paper will present a simheuristic methodology to cope with the stochastic nature of waste levels in the collection of garbage in the following.

The amount of waste in urban garbage containers are constantly frequented by a high number of people disposing very different amounts of waste. Furthermore, there are no fixed time schedules for waste disposal, leading to a high uncertainty concerning actual waste amounts to be collected by a waste collection vehicle when it arrives at a collection site. In the following, a heuristic to solve the deterministic WCP is presented in Sector 4.1. A simheuristic approach combining this heuristic with simulation is then proposed in Sector 4.2.

## 4.1 A Heuristic Approach to Solve the Deterministic WCP

The first step in our simheuristic approach is the simplification of the stochastic WCP into a deterministic problem instance. Instead of considering stochastic waste levels $w_i^{st}$ at each waste container *i*, the expected waste levels based on past experience is used. To solve the deterministic WCP with waste levels $w_i^{det}$, we propose the use of an extended version of the well-known Clarke-and-Wright Savings heuristic (CWS). This algorithm starts of by building an initial solution by supplying each customer on a single route with a single vehicle, as seen in (a) in Fig. 3 below. Note that this is the worst possible feasible solution to a VRP, as it will include the highest possible costs. Based on this initial solution, the CWS iteratively merges some of the routes in order to serve more than one customer on a single route, as seen below above (b) (Laporte 2007).

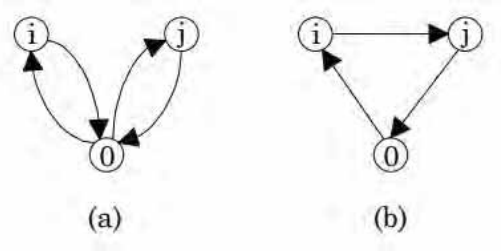



*Figure 3: Merging of Routes through the CWS*

The merging process is based on savings associated to a pair of customers. These savings define the reduction in the overall cost function of serving each customer with a single vehicle, as proposed in the initial solution, compared to combining these customers on an edge and serving them with the same vehicle. For example, savings $s_{ij}$ of connecting customers i and j are computed through $s_{ij} = c_{0i} + c_{0j} - c_{ij}$. All the computed possible savings are listed on a savings list. During the following merging process of a feasible VRP solution, the algorithm iteratively selects the edge with the highest possible savings potential from the savings list as long as (i) no capacity constraints are violated, and (ii) the nodes defining the edges are next

to the depot. As always the edge with the highest possible savings is chosen to construct VRP solutions, the CWS is a so called greedy algorithm.

The original CWS is improved through a randomization technique obtained by combining the CWS with Monte Carlo simulation (MCS), which can be seen as a set of techniques that use statistical distributions and random numbers to solve stochastic and deterministic problem settings (Law 2007). While the original CWS always chooses the edge with the highest potential savings, we use a biased randomization approach at each iteration during the solution construction. While the edge with the highest savings is chosen with the highest probability, all feasible edges are potentially available in the route construction. Using this approach, the algorithm calculates a number of different problem solutions while avoiding getting stuck in a local minima through the edge selection process improved by MCS.

Furthermore, the original CWS algorithm is improved by clustering nodes that are close to each other, making sure they are served by the same customer and a cache memory technique as proposed by Juan et al. (2011a). Finally, all routes of the obtained WCP solution are further optimised by applying the Lin-Kernighan heuristic (LKH), which is a famous adoption of the λ-opt algorithm (Helsgaun 2009).

The suggested heuristic is tested and verified using the benchmarks provided by Kim et al. (2006) and the results of Benjamin & Beasley (2010) and Buhrkal et al. (2012). Their results over all 10 problem instances in comparison to our competitive heuristic approach can be seen in Table 1. With an average percentage gap of 0.8%, our heuristic leads to the current best-known solution in 4 of the 10 benchmark instances.

*Table 1: Results of the proposed Heuristic*

| **Instance** | **Kim et al.** | **Benjamin & Beasley** | **Buhrkal et al.** | **Our Heuristic** | **%-Diff. to Best Benchmark** |
|---|---|---|---|---|---|
| **102** | 205.1 | 183.5 | **174.5** | 157.7 | -9.6 |
| **277** | 527.3 | 464.5 | **447.6** | 463.3 | 3.5 |
| **335** | 205.0 | 204.5 | **182.1** | 193 | 6.0 |
| **444** | 87.0 | 89.1 | **78.3** | 85.3 | 8.9 |
| **804** | 769.5 | 725.6 | **604.1** | 616.4 | 2.0 |
| **1051** | 2370.4 | **2250.2** | 2325.7 | 2244.7 | -0.2 |
| **1351** | 1039.7 | 915.1 | **871.9** | 968.9 | 11.1 |
| **1599** | 1459.2 | 1364.7 | **1337.5** | 1219.9 | -8.8 |
| **1932** | 1395.3 | 1262.8 | **1162.5** | 1170.3 | 0.7 |
| **2100** | 1833.8 | **1749.0** | 1818.9 | 1652.4 | -5.5 |
| *Average* | | | | | *-0.8* |

### 4.2 Using Simulation to Consider Stochastic Waste Levels

For the simulation part of the simheuristic, not the complete vehicle capacity VMC is considered for the solution calculation of WCP. It is assumed that a certain safety capacity (1-k) is considered by route planners in the construction of the route, as it is possible that the actual stochastic demands exceed the deterministic demand levels considered earlier. For this reason, the homogenous VMC for all vehicles is multiplied by a load factor k, leading to a new vehicle maximum capacity $VMC^* = VMC \times k$ with which the deterministic solutions are calculated. Note that a high value of k will lead to a low safety stock and a lower total number of routes and vice versa.

Using vehicle maximum capacities $VMC^*$ for the calculation, the heuristic proposed in Sector 4.1 is applied to get a high-quality deterministic WCP solution. For each solution that is considered as promising, waste levels for each customer are simulated using MCS various times. Thereby, different customer demands $w_i^{st}$ are modelled with a log-normal distribution based on a mean equal to each customer's deterministic waste level $w_i^{det}$ and a certain variance (e.g. 10%).

Based on the load factor k, each stochastic solution consist of the base costs provided by the heuristic for the deterministic WCP and the estimated corrective actions necessary for route failures. These expected route failures are calculated based on the number of simulation runs in which the stochastic waste levels exceed the original vehicle capacity VMC, which is penalized with the costs for an additional trip to a landfill. Note that higher values for k will lead to low base costs (through the use of a higher load factor/lower safety capacity) but higher additional costs through more route failures and vice versa. Furthermore, this approach leads to an estimate of each solution reliability by dividing the amount of route failures with the number of simulation runs.

Based on the results after the first simulation run, the most promising stochastic elite solutions are defined. Waste levels for these solutions are simulated in an extensive second simulation run, leading to a list of solution proposals ranked according to their overall expected costs and reliabilities. As the whole process can be run with low computational effort, different values for k and the influence on the solutions can be tested, as outlined in the detail by (Juan et al. 2011b).

## 5 Conclusion and Future Outlook

This paper proposes a simheuristic approach to solve the WCP through the combination of an efficient heuristic with Monte Carlo simulation. The methodology allows the consideration of stochastic waste levels in garbage containers, leading to more realistic and reliable routing solutions in waste management. Using the results of the competitive heuristic based on the famous CWS for the VRP to solve the simplified deterministic WCP as outlined in this paper, Monte Carlo simulation is used to simulate different waste levels and obtain results for the more realistic stochastic WCP. On the one hand, the simulation procedure leads to results based on the expected additional penalty costs for route failures (i. e. additional trips to waste disposal sites) and the basic route costs. On the other hand, a further result dimension is added by estimating the reliability of each route depending on a

collection vehicle load factor, giving waste management decision takers additional information in their routing choices.

In the development of Smart Cities, municipalities can make use of cheap and efficient tracking and tracing technologies such as volumetric sensors implemented into waste containers to obtain reliable initial values for the suggested simheuristic. Instead of basing the mean value for the simulation on experience values, such volumetric sensors (which are already being applied in various European cities) could be used to obtain more reliable initial waste levels, allowing decision takers to calculate their routes with lower variance and a higher load factor for the collection vehicles. For an extensive overview over tracking and tracing technologies used to improve waste management see Faccio et al. (2011). In the future, a case study in which volumetric sensors are implemented in waste containers in the Catalonian city Sabadell to improve their municipal waste management using the presented simheuristic is planned.

## References

Benjamin, A.; Beasley, J.: Metaheuristics for the waste collection vehicle routing problem with time windows, driver rest period and multiple disposal facilities. Computer & Operations Research 37 (2010) 12, pp. 2270-2280.

Buhrkal, K.; Larsen, A.; Ropke, S.: The waste collection vehicle routing problem with time wondows in a city logistics context. Procedia – Social Behavioral Sciences 39 (2012), pp. 241-254.

Caceres-Cruz, J.; Pol, A.; Guimarans, D.; Riera, D.; Juan, A.: Rich vehicle routing problem: Survey. ACM Computing Surveys 47 (2015) 2, pp. 32:1-32:28.

Daneshzand, F.: The Vehicle Routing Problem. In: Farahanni, R.; Rezapur, S.; Kardar, L. (eds.): Logistics Operations and Mangement. Amsterdam: Elsiver 2011, pp. 127-153.

European Comission: Cities of tomorrow – Challenges, visions, ways forward. Luxembourg: Publications Office of the European Union 2011. Retrieved from: http://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/citiesoftomorrow/citiesoftomorrow_final.pdf

Faccio, M.; Persona, A.; Zanin, G.: Waste collection multi objective model with real time traceability data. Waste Management 31(2011) 12, pp. 2391-2405.

Helsgaun, K.: General k-opt submoves for the Lin-Kernhigan TSP heuristic. Mathematical Programming Computation 1 (2009) 2-3, pp. 119-163.

Johansson, O.: The effect of dynamic scheduling and routing in a solid waste management system. Waste Management 26 (2006) 8, pp. 875-885.

Juan, A.; Fauli, J.; Jorba, J.; Riera, D.; Masip, D.; Barrios, B.: On the use of Monte Carlo simulation, cache and splitting techniques to improve the Clarke-and-Wright savings heuristic. Journal of the Operational Research Society 62 (2011a) 6, pp. 1085-1097.

Juan, A.; Faulin, J.; Grasman, S.; Rabe, M.; Figueira, G.: A review on simheuristics: Extending metaheuristics to deal with stochastic combinatorial problems. Operations Research Perspectives 2 (2015), pp. 62-72.

Kim, B.; Kim, S.; Sahoo, S.: Waste collection vehicle routing problem with time windows. Computers & Operations Research 33 (2006) 12, pp. 3624-3642.

Laporte, G.: What you should know about the vehicle routing problem. Naval Research Logistics 54 (2007) 8, pp. 811-819.

Law, A.M.: Simulation modeling and analysis, 4. ed. Boston: McGraw-Hill 2007.

The World Bank: What a waste – A global review of solid waste management. Washington: Urban Development Unit 2012. Retrieved from: http://siteresources.worldbank.org/INTURBANDEVELOPMENT/Resources/336387-1334852610766/What_a_Waste2012_Final.pdf

United Nations: World urbinization prospect: The 2007 revision. New York: UN Department for Economic and Social Affairs 2008. Retrieved from: http://www.un.org/esa/population/publications/wup2007/2007WUP_Highlights_web.pdf

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Development of a Simheuristic Approach for Solving Realistic Inventory Routing Problems

## *Entwicklung eines Simheuristikansatzes zur Lösung von realitätsnahen Inventory-Routing-Problemen*

Lena Pfeilsticker, Markus Rabe, TU Dortmund, Dortmund (Germany),
lena.pfeilsticker@tu-dortmund.de, markus.rabe@tu-dortmund.de

Angel A. Juan, IN3-Open University of Catalonia, Barcelona (Spain),
ajuanp@uoc.edu

**Abstract:** In order to stay or become even more cost-efficient, companies have to cooperate along the supply chain for improving the efficiency of transport and inventory systems. The majority of authors dealing with a problem in this research area either address the inventory problem or the transportation problem. Fewer published works deal with the integrated decision-making of inventory and transportation, called Inventory Routing Problem (IRP). In this paper, an approach is developed, that enables to find feasible low-cost solutions for complex IRPs in efficient time frames and under high levels of uncertainty. A suitable approach is to combine heuristic optimisation with simulation techniques, called 'simheuristics'. In this paper, a stochastic IRP with realistic work conditions is modelled and implemented. Furthermore, experiments are conducted, validating the implementation with an advantage in computation time and low costs in comparison to other conventional solving methods. The results are a significant contribution to create a link between companies and research in the field of IRPs.

## 1 Introduction

Within supply chains, independent decision-making processes are still a common approach. However, by cooperating along the supply chain, potentials to improve the transport and inventory systems can be exploited. If the supply chain were to be considered as a whole, an integrated decision-making could lead to significant decreases in total costs and operation durations. The general problem of integrating inventory and transportation (vehicle routing) decisions is called Inventory Routing Problem (IRP). Considering the reality of supply chain systems and real-life IRP applications, it is necessary to conduct research by considering the whole logistical system to find solutions with minimal cost and highest efficiency. A significant progress in this research field has already been reported by Juan et al. (2014a). The

authors propose an approach combining simulation and heuristics to a so-called 'simheuristic' (Juan et al. 2015), leading to solutions which are improving known solutions to IRPs.

In this paper, a stochastic IRP is introduced. It is explained how an approach combining simulation and heuristics is able to find a good and feasible solution. Also, the paper demonstrates the implementation of the solution process as a Java application. Figure 1 shows an exemplary Inventory Routing System.

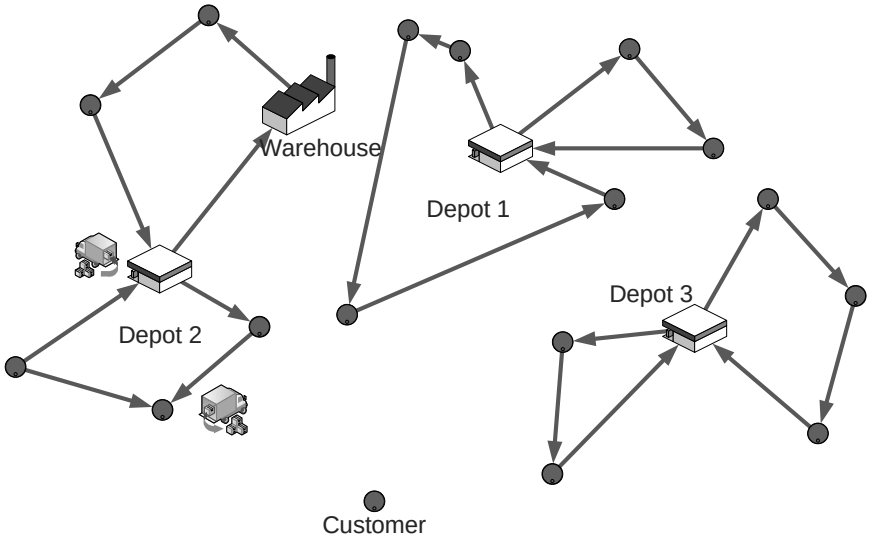



*Figure 1: Exemplary Inventory Routing System*

The paper is structured as follows: In Section 2, existing approaches to solve IRPs are discussed, with a special focus on the combination of simulation and heuristics. The methods additionally applied are introduced shortly. Section 3 introduces the given problem. Section 4 explains the proposed solution process and explains the benefit of the applied methods for solving stochastic IRPs. The implemented solution process is described in Section 5. For selected essential elements, the algorithm is explained. In Section 6, the paper evaluates several conducted experiments with the final implementation. Finally, the results are concluded in Section 7.

## 2 Simheuristics Approaches to Solve Inventory Routing Problems

Several studies have combined simulation and optimization approaches for solving complex optimization problems (van Dijk and van der Sluis 2008). The supply chain process has been a popular target for this type of techniques. In the past years, several approaches have been proposed for different variants of the IRP. The main factors to take into account when classifying the different works are: (a) whether they consider deterministic (Bertazzi et al. 2008) or stochastic demands (Federgruen and Zipkin 1984); (b) whether they consider single- (Powell and Godfrey 2002) or

multiple-periods (including an infinite horizon) (Trudeau and Dror 1992); (c) whether they allow for inventory shortages (Jaillet et al. 1992); (d) whether they consider single (Campbell et al. 1998) or multiple products (Mjirda et al. 2014); (e) whether they use the same replenishment policy for all nodes (Bertazzi et al. 2013) or personalized replenishment policies for each node (Juan et al. 2014); and (f) whether they use exact (Baldacci et al. 2011) or approximate methods (Subramanian et al. 2012) to solve the problem.

In the present paper the solution process for the given problem is embedded in a simheuristic framework. It is able to build an interface between real-life stochastic IRPs and deterministic methods commonly used for complex problems in this research area (van Dijk and van der Sluis 2008). Simheuristics are hybrid simulation-optimization methods used for NP-hard, combinatorial optimization problems (COPs) with stochastic components (Juan et al. 2015). They allow for obtaining good feasible solutions in reasonable computing times. When applying a simheuristic approach, a stochastic COP is converted into a deterministic problem by using the expected mean values of the single input parameters. This problem is then solved using a heuristic approach and the deterministic solution is tested for feasibility. In a fast simulation process with a low number of replications, the promising solutions are evaluated by using stochastic parameters, leading to a stochastic solution to the COP. The generated solutions are ranked according to their best outcome and the best results are again tested for their feasibility and quality in an intensive simulation process with a large number of replications to determine the best solution (Juan et al. 2015). The method indirectly assumes that high-quality solutions for the deterministic version of the problem are likely to be high-quality solutions for the stochastic version. If this correlation is not perfect, it is possible to miss some good solutions for the stochastic version. The overall process is displayed in Figure 2.

During the process of fast simulation, the best solutions are stored in a cache. When the iterative process of finding solutions is repeated, the newly computed solutions are compared to the already computed best solutions in the cache. If the solutions in the cache are worse, the newly computed solution replaces the worst solution in the cache. Using this approach, promising solutions are selected over inferior ones. The approach does not guarantee finding the overall best solution for the problem it is applied to. But, it helps finding reasonably good solutions in a given maximum time.

During several steps of the solution process, a random sampling from skewed probability distributions is used in a Monte Carlo simulation. This approach is used to generate randomness in the heuristics (Juan et al. 2013). When applying a heuristic that is employed in a simulation process, the next constructive movement is selected based on a previously selected criterion. In the here treated IRP, the elements that can be chosen to construct the next movement are often sorted in a list from which they are selected. The order of this list is based on the benefit for the solution. E.g., for an IRP, a list of customers with the closest location to a depot is computed and the depot gets assigned to the customers with the top positions on the priority list, i.e. the customers with the minimal distance to the depot.

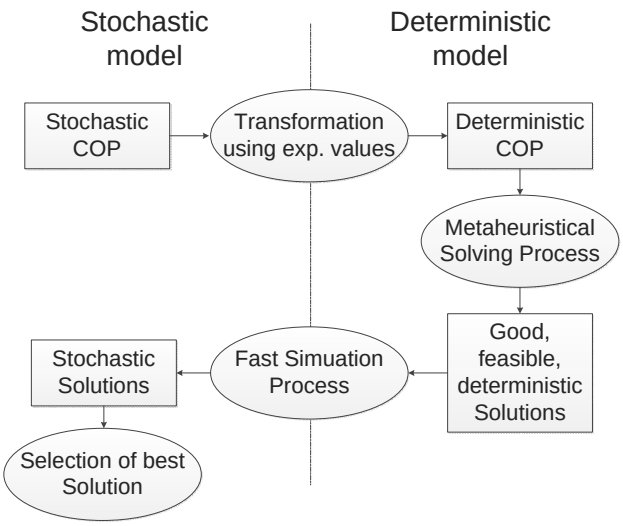



*Figure 2: Simheuristic approach (based on Juan et al. 2015)*

However, heuristics are deterministic approaches: using a specific input will always lead to the same results. To obtain several solutions from which the best solution can be selected, the elements of the priority list are combined with skewed distributions: A probability to be selected is matched with each element of the priority list, generating a stochastic approach. In the following simulation, the order in which the elements of the priority list are selected is different at every step of the iterative solution process and multiple solutions are obtained. A uniform distribution of the priority list of elements would destroy the logic of creating a priority list. Therefore, a skewed distribution has to be employed, assigning higher probabilities to be picked to elements higher on the priority list. This approach has been proven to outperform other common approaches used to solve this type of problems (Juan et al. 2014a). Their problem is described with the following conditions: single-period, single-product, single-depot, homogeneous fleet, centralized control unit, possible stockouts, stochastic demands and no time windows for the deliveries.

Further methods are the round-robin scheduling applied in the assignment of customers to depots by which they are served. A randomized variant of the Clarke and Wright savings (CWS) algorithm is used for determining the routing of vehicles from the depots to the assigned customers. An iterated local search (ILS) process is employed for refining solutions during a step of assigning customers to depots (Juan et al. 2014b).

## 3 The Given Inventory Routing Problem

The environmental work conditions for the problem in the present work are determined previously to the problem definition. For this purpose, an analysis of current literature and information provided by several companies has been conducted. Requirements for the treated problem should not be new to the field of IRPs but have been considered before in other research, to be sure that an impact on the problem exists. In order to achieve the goal to solve a realistic IRP, a significant

number of working conditions or requirements (see below for some examples) have to be chosen, reflecting the reality of complex inventory routing systems. This is opposed to the approach of many other authors focusing on a low number of working conditions (van Dijk and van der Sluis 2008). In order to present a solution to a new IRP, the requirements are chosen in a combination that is not similar to characteristics in other authors' works. When choosing requirements, it is beneficial when the information provided by companies and the results from the literature research match. Finally, characteristics that are already used in Juan et al. (2014a) are extended and applied to the here treated IRP.

After reviewing all proposed characteristics and after a check for the accordance of them with all considered sources, a combination of several requirements is defined to reflect the reality of IRPs sufficiently while at the same time being implementable in a code. Based on the reviewed literature (see section 2) and the analyzed information provided by real-life companies, the requirements for the treated IRP are set as follows: It is considered with multiple products, multiple depots, one warehouse, a heterogeneous fleet, a single period, stochastic customer demands, a centralized control unit, possible stockouts and without time windows for the deliveries. The warehouse in this scenario is considered as a storage unit with virtually unlimited stock levels and can be accessed by vehicles sent from the depots. The solution comprises the completion of the delivery of all demanded products in a network of customers, depots and a warehouse, while minimizing the overall costs. The cost function consists of the set-up costs for the sending of one vehicle, of the inventory costs resulting from a stock cost rate and the number of stored products and a variable distance-based cost for travelled ways between nodes in the system. Thus, comparing with Juan et al (2014a), this research extends the problem to multiple products and depots as well as a heterogeneous fleet.

## 4 Proposed Solution Process

The solution process is divided into three steps, which are computed in a simheuristic framework: assignment, inventory planning and routing. The goal is to find the overall minimum costs. In this section, an overview is given on the process. The details are paid closer attention to during the implementation description.

The inputs for the computation of the solution are as follows: location of depots, location of warehouse, location of customers, stochastic customer demands, current inventory levels of every depot for every product, routing-costs function, inventory-costs function, number of vehicles by type and their associated number of compartments.

During the first step of the solution process, customers are bound to a specific depot. The result of this step is a single-depot IRP for each depot. A number of sub-maps is created, equal to the number of depots. This approach is based on a 'splitting policy' and decreases the problem complexity exponentially (Juan et al. 2014b). During this assignment step, a priority list of nodes is generated for each depot. The list is based on marginal savings in distance costs and considers the maximum capacity of the depots with regard to the combined customer demand for every product. To obtain multiple solutions for the comparison of the best outcome, i.e. lowest overall costs, a random sampling with a skewed distribution is introduced to assign customers to depots (cp. Section 2). Random sampling from skewed distributions is applicable

whenever a new order of given elements needs to be constructed based on a sortable criterion during a Monte Carlo simulation, as described above. The order of the elements can be randomized and thus new solutions can be computed, based on a logical order. When solving an IRP, a lot of elements need to be arranged and skewed distributions favour the beneficial choices during the process. Then, a first assignment solution is generated using a round-robin process. At each round of the round-robin process the depot with the most (remaining) overall serving capacity selects a node from the generated list. After obtaining a first solution, a multi-start process generates thousands of maps with feasible solutions of assignment (customers to depot) in a short time period.

In the next step, the optimal inventory levels are estimated for each product of every depot, by considering the stochastic customer demands. For each of the depots, different replenishment policies are tested to determine the inventory levels with the lowest costs. For this purpose, the overall customers' demands of each depot are estimated. This is performed based on the results from the previous assignment step. Then, the expected inventory costs associated with each combination of depot and replenishment policy are calculated, including stockouts. With a basic CWS heuristic, the routing costs resulting from each replenishment policy are calculated. The policy with the lowest overall costs is chosen to serve for a computation of an initial base solution for the routing process.

The routing step generates routes from each depot to each of its assigned customers, employing a CWS heuristic with skewed distributed probabilities. Then, a first solution is computed. An unlimited number of the smallest trucks available is provided to form routes directly to and from the depots to every customer assigned to it. Then, the randomized CWS heuristic is performed to improve the first solution (Juan et al. 2011). While performing the CWS heuristic, the vehicle with the biggest surplus of overall capacity is chosen first, accounting for its number of different compartments and their individual capacity as opposed to the customers' demands. If a customer's number of demanded products is higher than the number of compartments in the biggest available truck, the customer is served by a minimum of two different trucks. The CWS heuristic chooses the next nodes based on their highest saving values. These values are the base of a priority list which is again randomized. Thus, a different outcome of routes generates thousands of feasible solutions in short computing times and the best solution found in a defined maximum time can be chosen. The result of the randomized CWS heuristic is improved with the help of a route cache. As described in section 2, this step is part of the intensive simulation process on elite solutions.

After the whole process, the solution with the overall lowest costs is returned as best solution. The outputs after the previously explained steps are the overall costs for routing and inventory, the set inventory levels with their associated replenishment policies and the routing plan for each depot. A flowchart of the proposed solution process is illustrated in Figure 3.

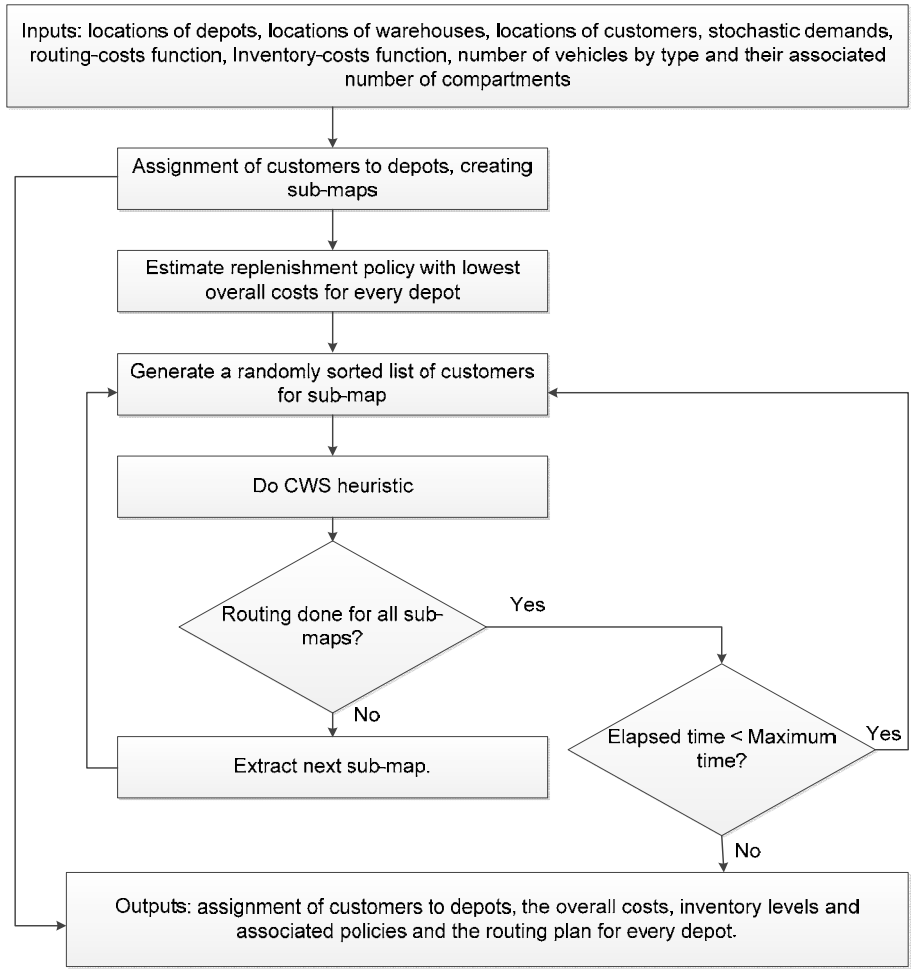



*Figure 3: Flowchart of the proposed solution process*

## 5 Implementation

The implementation is part of an iterative process. Since a lot of elements of the solution process are heuristic, there is no optimal solution and no reference values of comparable IRP solutions exist. Thus, to obtain a working code of the implemented solution process' utility, its development is conducted in several stages.

As in the proposed solution process, the customers are assigned to depots based on a capacitated round-robin scheduling during a multi-start approach. While assigning customers to the depots, the depot with the highest capacities left is chosen to be assigned to the next customer. The customers are chosen from a priority list based on marginal savings and with the help of oriented randomization. The distances between all customers, depots and warehouses are computed with coordinates and the Euclidean distance. Ensuing the assignment process, the comparison of inventtory policies leads to set inventory levels of each product for each depot. As described

in the proposed solution process, an element is embedded in the code that introduces randomness, called a random number generator (RNG). This is used for the randomization in the assignment and routing steps of the algorithm. When merging the routes during the randomized CWS heuristic, all the vehicles available at one depot are ideally used to generate the highest savings possible. Thus, when merging two nodes, the new route with the highest savings value chooses an available vehicle with the highest surplus of capacity. This leads to assigning the biggest vehicles to the routes with the highest savings. Since the merging is conducted with the help of oriented randomization, the matching of saving values and size of vehicles is expected to have a reasonably good result. By using oriented randomization of the decisions about choosing a vehicle for a route, alternative scenarios are considered and the solution with the best result is returned as an output when running the code.

## 6 Experiments

The code was implemented in several stages. A basic structure with deterministic inputs was developed and finally, after a number of extensions, stochastic customer demands and calculations to consider those demands, when setting the inventory levels, were added.

Several experiments were performed, validating and verifying the code at each developed stage. The final implementation is constructed as described in the proposed solution process. It is easily and quickly possible to adapt the input values and probability distributions used to compute results. The algorithm is able to consider different replenishment policies $p = \{0.25,0.5,0.75,1.0\}$ for every depot, indicating the percentage of refilled inventory (e.g. 0.5 = 50%) and to set the individual inventory levels according to the used probability distribution, based on the lowest-costs scenario. The code is also able to test different scenarios with varying fluctuation of customer demands while considering different stock cost rates. On the basis of the code output, a risk analysis can be conducted, allowing to evaluate the ensuing costs associated with different variances of customer demand.

In the following, an analysis of the final implementation outputs is conducted. 15 different scenarios were tested and compared regarding their outcome. The scenarios were based on different variances $k$ in customer demands and different stock cost rates $\lambda$. The experiments were performed with three depots (D1, D2, D3).

The overall costs and the ratio between inventory and routing costs are displayed in Figure 4. It shows the overall costs of every combination of variance $k$ and stock cost rate $\lambda$. The best computed policies $p$ for every depot are listed below in Figure 4. The numbers 1 to 5 indicate the assigned possibility (1 = 0.0, 2 = 0.25, 3 = 0.5, 4 = 0.75, 5 = 1.0).

The results reflect the costs adequately: The higher the stock cost rate $\lambda$, the higher the overall inventory cost. Furthermore, as can be abstracted from Figure 4, in the case of $p = 0.0$ in all depots, no inventory costs are incurred or if one or two depots out of three have an inventory set at $p = 0.0$, the overall inventory costs are lower, respectively.

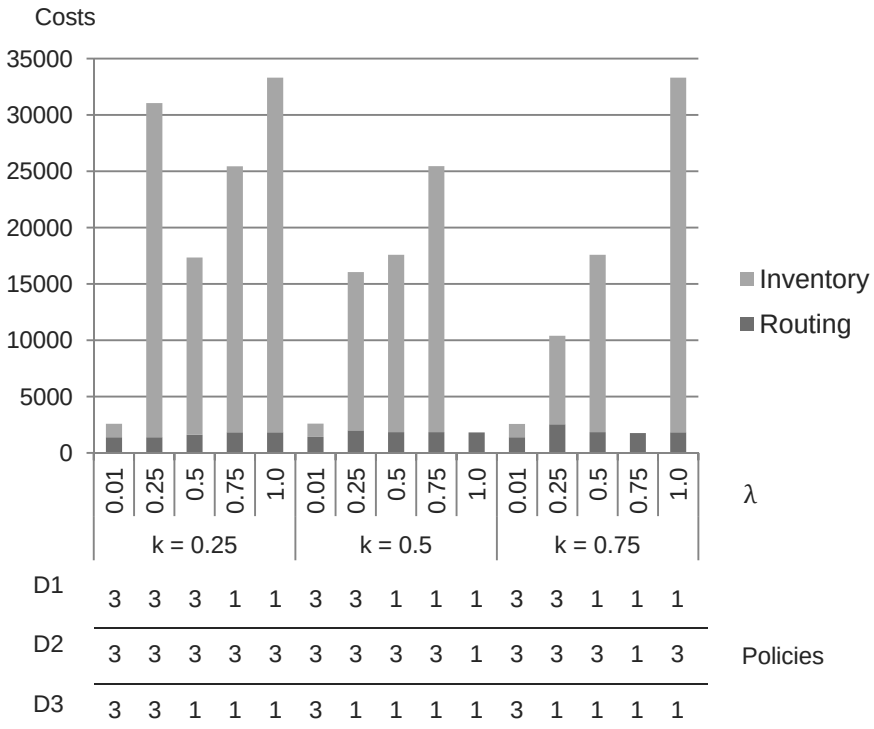



*Figure 4: Results of Stochastic Experiments*

## 7 Conclusion

In this research, a simheuristic algorithm was presented for solving a realistic rich single-period, multi-depot IRP with stochastic customer demands, stockouts and a heterogeneous fleet. The proposed methodology for the solving process combines several approaches to enable an integrated decision-making process for the simultaneous planning of inventory levels and routing. The research area of implementing rich IRPs is challenging because of highly complex requirements and their interdependencies, especially when introducing stochastic customer demands. Due to the combination of several methods, the developed program is able to deal with a high number of conditions in realistic IRPs under great levels of uncertainty. It can be used with any probability distribution to design realistic customer demands, is able to different replenishment policies for every depot and can test scenarios with different variances and ensuing costs thereby possessing the ability to analyze risks. The overall conclusion is that the program contributes significantly to the research field of IRPs due to its flexible applicability and possibility to be used on real-life cases.

## Acknowledgements

This work has been partially supported by the Martin-Schmeißer Foundation.

## References

Baldacci, R.; Bartolini, E.; Mingozzi, A.; Valletta, A.: An exact algorithm for the period routing problem. Operations Research 59 (2011) 1, pp. 228-241.

Bertazzi, L.; Savelsbergh, M.; Speranza, M.G.: Inventory routing. In: Golden, B.; Raghavan, S.; Wasil, E. (Eds.): The Vehicle Routing Problem. New York, London: Springer 2008, pp. 49-72.

Bertazzi, L.; Bosco, A.; Guerriero, F.; Lagana, D.: A stochastic inventory routing problem with stock-out. Transportation Research Part C 27 (2013), pp. 89–107.

Campbell, A.; Clarke, L.; Kleywegt, A.; Savelsbergh, M.: The Inventory Routing Problem. In: Crainic, T.G.; Laporte, G. (Eds.): Fleet Management and Logistics. Boston: Springer 1998, pp. 95-113.

Federgruen, A.; Zipkin, P.: A combined vehicle routing and inventory allocation problem, Operations Research 32 (1984) 5, pp. 1019–1036.

Jaillet, P.; Huang, L.; Bard, J.; Dror, M.: Delivery cost approximations for inventory routing problems in a rolling horizon framework. Transportation Science 36 (2002) 3, pp. 292–300.

Juan, A.A.; Faulin, J.; Jorba, J.; Riera, D.; Masip, D.; Barrios, B.: On the use of Monte Carlo simulation, cache and splitting techniques to improve the Clarke and Wright savings heuristics. Journal of the Operational Research Society 62 (2011) 6, pp. 1085-1097.

Juan, A.A.; Faulin, J.; Ferrer, A.; Lourenço, H.R.; Barrios, B.: MIRHA: multi-start biased randomization of heuristics with adaptive local search for solving non-smooth routing problems. TOP 21 (2013) 1, pp. 109-132.

Juan, A.A.; Grasman, S.E.; Carceres-Cruz, J.; Bektas, T.: A simheuristic algorithm for the Single-Period Stochastic Inventory-Routing Problem with stock-outs. Simulation Modelling and Practice Theory 46 (2014a), pp. 40-52.

Juan, A.A.; Pascual, I.; Guimarans, D. Barrios, B.: Combining biased randomization with iterated local search for solving the multidepot vehicle routing problem. International Transactions in operational research 35 (2014b), pp. 1-21.

Juan, A.; Faulin, J.; Grasman, S.; Rabe, M.; Figueira, G.: A review of simheuristics: extending metaheuristics to deal with stochastic optimization problems. Operations Research Perspectives 2 (2015), pp. 62-72.

Mjirda, A.; Jarboui, B.; Macedo, R.; Hanafi, S.; Mladenovic, N.: A two phase variable neighborhood search for the multi-product inventory routing problem. Computer & Operations Research 52 (2014) Part B, pp. 291-299.

Powell, W.B.; Godfrey, G.A.: An adaptive dynamic programming algorithm for dynamic fleet management I. Transportation Science 36 (2002) 1, pp. 21-39.

Subramanian, A.; Penna, P.H.V.; Uchoa, E.; Ochi, L.S.: A hybrid algorithm for the heterogeneous fleet vehicle routing problem. European Journal of Operational Research 221 (2012) 2, pp. 285-295.

Trudeau, P.; Dror, M.: Stochastic inventory routing: route design with stockouts and route failures, Transportation Science 26 (1992) 3, pp. 171–184.

van Dijk, N.M.; van der Sluis, E.: Practical optimization by OR and simulation, Simulation Modelling and Practice Theory 16 (2008) 8, pp. 1113–1122.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulation-based Optimisation in Industry 4.0

## *Simulationsbasierte Optimierung für die Industrie 4.0*

Fabian Dunke, Stefan Nickel, Karlsruhe Institute of Technology, Karlsruhe (Germany), fabian.dunke@kit.edu, stefan.nickel@kit.edu

**Abstract:** Driven by massive improvements in information and communication technologies throughout the past two decades, there is an ongoing trend towards endowing physical entities in logistics networks and production facilities with information transmitting devices. As a result of continuous data collection throughout the process chain, new opportunities with respect to optimisation in (near) real time arise. In this paper, a simulation-based approach is presented which allows for the assessment of the optimisation potential granted by information trans–mitting devices. We first relate real time problems to the online optimisation paradigm with lookahead and to discrete event simulation. Based on a holistic approach to performance measurement of algorithms, we then show in a simulation study for the truck access control at factory sites how simulation and optimisation can be linked in order to evaluate technological scenarios related to Industry 4.0.

## 1 Introduction

Technological developments and breakthroughs from the last two decades such as RFID technology, geographic information systems or sensor/actor systems allow us nowadays to collect and process data continuously over time (Stadtler and Kilger 2015, Ghiani et al. 2004). As a result of these opportunities, new statistical disci–plines (Big Data, Smart Data or Predictive Analytics amongst others) were designed to take advantage of the available data and provide better informational input to solvers of related optimisation problems. In production and logistics, the ultimate vision of this information- and technology-driven approach to optimisation is the creation of cyber-physical systems where physical and virtual representations of entities are amalgamated into self-optimising smart entities (Bauernhansel et al. 2014). What we obtain is called the Internet of Things, i.e., a net–work of embedded intelligent objects interacting with each other (Bullinger and Ten-Hompel 2007). In this context, the German government launched Industry 4.0 in 2012 as their high-tech innovation strategy for manufacturing industries with the goal of combining information and communication technologies with mechanical engineering into smart factories (Bundesministerium für Bildung und Forschung 2012).

A crucial issue to success or failure in combining the physical world with the virtual world lies in the quality of the interface between these two worlds (Bauernhansel et al. 2014, Stadtler and Kilger 2015): Processes in Industry 4.0 can only be executed effectively and efficiently when intelligent methods of control are provided which additionally yield a substantial surplus of virtually augmented physical devices as compared to today's conventional production and logistics systems. Hence, ade–quately interlocking physical systems with data analytics represents the methodo–logical key to establishing the Internet of Things and cyber-physical systems on a sustainable basis for the next decades.

This work addresses one of the most fundamental questions concerning the value of continuous data and information gathering throughout day-to-day operations: What is it worth to provide decision makers (or algorithms) with additional information that has been gathered by new information technologies? To answer, the arising problems are first classified as online optimisation problems with lookahead, and then it is shown how a simulation-based approach to online optimisation with look–ahead can be applied to decide upon the value of information in a specific practical setting. A case study will illustrate how the approach can be put into practice to obtain results about the benefit of information in the respective problem settings.

## 2 Online Optimisation with Lookahead and Discrete-event Systems

According to the theory of online optimisation, decisions have to be made repetitively while not all information of a problem is yet available (Albers 2003, Borodin and El-Yaniv 2005). This is in sharp contrast to offline optimisation where all information is known initially. Formally, an offline algorithm knows the entire input sequence $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ with $n \in \mathbb{N}$ in advance, but in online optimi–sation only $\sigma_{\leq i} = (\sigma_1, \sigma_2, \ldots, \sigma_i)$ is known when an online algorithm has to decide upon input element $\sigma_i$ with $i = 1, \ldots, n$. However, most types of problems occurring in industrial applications on the operational level are online since typically not all customer orders, processing requirements or service requests are known at the start of a day. On the other hand, by improved possibilities of electronic data interchange, large amounts of the information (such as processing durations, service locations, or object dimensions) can be made accessible earlier to the decision maker when this data is transmitted continuously. This can be achieved by tagging entities with transmission devices which is considered an essential prerequisite of the Internet of Things. As a result, the decision maker learns about the existence and content of upcoming requests much earlier than in the case without information transmission, and he can hopefully elaborate a better plan. This setting is called online optimi–sation with lookahead and it is characterised by having available $\sigma_{\leq j} = (\sigma_1, \sigma_2, \ldots, \sigma_i, \sigma_{i+1}, \ldots, \sigma_j)$ with $j > i$ when one has to decide upon input element $\sigma_i$. In the past, online optimisation with lookahead has been investigated problem-specifically, e.g., in routing (Jaillet and Wagner 2006; Allulli et al. 2008).

From a formal point of view, the set of all online optimisation problems is parti–tioned into two sets (Dunke 2014): First, the set of (pure) online optimisation prob–lems where one only sees a request when it becomes known without any additional information gathering device, and second, the set of online optimisation problems

with lookahead where information is made accessible earlier through the usage of some additional information gathering device. Hence, the problems in the first set are the baseline problems, and the problems in the second set are variants with improved information policies. Different information policies thus lead to different forms of online optimisation problems with lookahead, and preponed information as compared to the pure online case can be considered the basic requirement of online optimisation with lookahead. Additionally, upon earlier information retrieval, new possibilities for the processing of the requests or orders may occur which put us in a better position as compared to the pure online setting. Thus, potential new processing opportunities have to be incorporated in the classification as well. Subsuming, the following relation between online problems, online problems with lookahead and offline problems is obtained:

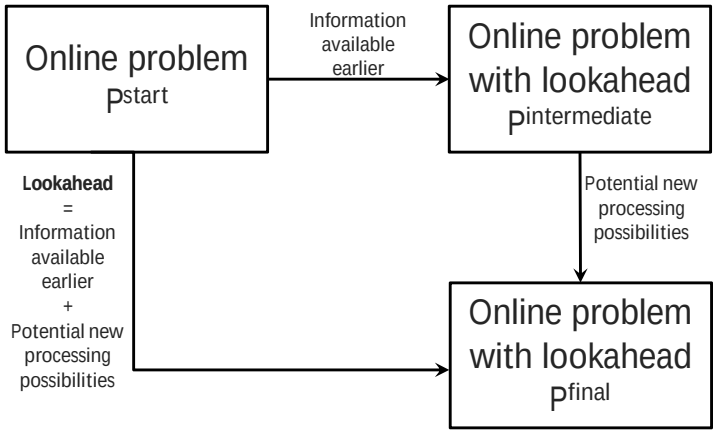



*Figure 1: Online problems with lookahead evolve from pure online problems*

Taken together, information made accessible earlier and resulting potential new pro–cessing possibilities comprise our notion of lookahead in online optimisation, i.e., lookahead and its effects can always be decomposed into an informational compo–nent and a processual component. Clearly, for the design, maintenance and restruc–turing of production and logistics systems it is essential to know which performance improvements could be made in operations through the introduction of lookahead.

Since in online optimisation it is assumed that the release of input elements occurs at discrete points in time, the solution process induced by the execution of an online algorithm can also be modelled as a discrete event system as well (Dunke 2014). Clearly, the release of an input element at a discrete time instant in online optimi–sation corresponds to an event occurrence in a discrete event system. In terms of how system control is achieved, an online algorithm is executed repetitively in on–line optimisation, whereas in a discrete event system the state transition is evaluated repeatedly. The main difference between online optimisation and discrete event systems is that in an online system an algorithm proactively controls the state transi–tion, whereas in a discrete event system this mechanism is implicitly modelled in the state transition function (Lunze 2012). Based on this observation, it is obvious to encapsulate the algorithm behaviour in the state transition function such that a direct

connection between online optimisation and discrete event systems is obtained. It is well-known that discrete event systems can be analysed with modelling techniques such as automata, Markov chains or Petri nets (Cassandras 2008). However, due to the complexity of real world systems these methods exhibit severe computational issues such that we choose the method of discrete event simulation in order to replicate the functional logics of a real world application (Cassandras 2008). In addition, we explicitly integrate online optimisation algorithms in order to cope with the dynamic character and decision need in industrial applications. The result com–prises a hierarchical relation between simulation and optimisation as has already been outlined in previous publications (März 2011). In this case, the simulation is the leading module and repetitively calls an online optimisation (sub-) module to elicit decisions that are needed to continue with the functional logic of the system.

## 3 Performance Measurement of Online Algorithms in Simulation Environments

The classical approach to measuring algorithm performance is the worst-case perspective that is taken on in Computer Science. Here, an algorithm is said to be a *c*-approximation algorithm, if its outcome is never more than *c*-times worse than that of an optimal (exact) algorithm (Cormen et al. 2009). Since this guarantee has to hold for any input, also the worst case guarantees this quality. Yet, there may be a vast majority of (typical) input sequences exhibiting a much more friendly behaviour. Transferred to online algorithms which have to operate dynamically, an online algorithm is said to be *c*-competitive, if its outcome is never more than *c*-times worse than that of an optimal (exact) offline algorithm which would have known the entire input in advance (Borodin and El-Yaniv 2005). In addition to the worst-case input sequence, comparison is sought against an omniscient offline algorithm which adds another dimension of "worst case". This is why there are large criticisms of the competitive approach (Hiller and Vredefeld 2012, Boyar et al. 2009). Despite of that, it is the only measure that found acceptance so far due to its clear interpretation. Nonetheless, in applications there is a need for more versatile performance measure displaying an algorithm's overall behaviour including the average case and the worst case. To this end, we have proposed the use of a distributional approach where each input sequence is given the same weight (Dunke 2014, Dunke 2013). Since online optimisation assumes that no probabilities for instances are given, the uniform distribution as the maximum entropy distribution is the best way to emulate the state of informational nescience, and counting results are obtained. By denoting the set of all (sampled) input sequences by $\Sigma$ and letting $\sigma \in \Sigma$ be an input sequence which is processed by some algorithm $ALG_1$ the following two definitions are obtained:

*Definition 1: Counting distribution function of objective value*
Let $v_{ALG_1}(\sigma)$ be the objective value of $ALG_1$ on $\sigma$, then the function

$$F_{ALG_1}(v) := \frac{\sum_{\sigma \in \Sigma} \mathbf{1}_{(-\infty, v]}(v_{ALG_1}(\sigma))}{|\Sigma|} \tag{1}$$

is called the counting distribution function of the objective value for $ALG_1$.

Note that the indicator function $\mathbf{1}_A(x)$ is 1 if $x \in A$ and 0 otherwise.

*Definition 2: Counting distribution function of performance ratio*
Let $v_{ALG_1}(\sigma)$, $v_{ALG_2}(\sigma)$ be the objective values of $ALG_1$ and $ALG_2$ on $\sigma$, respectively, then the function

$$F_{ALG_1,ALG_2}(r) := \frac{\sum_{\sigma\in\Sigma} \mathbf{1}_{(-\infty,r]}(v_{ALG_1}(\sigma)/v_{ALG_2}(\sigma))}{|\Sigma|} \tag{2}$$

is called the counting distribution function of the performance ratio of $ALG_1$ relative to $ALG_2$.

In contrast to single key performance indicators, these two counting distribution functions offer the advantage of conveying a comprehensive view on algorithm per–formance: Typical input sequences are covered as well as outliers and algorithms are compared instance-wise by the performance ratio idea ($F_{ALG_1,ALG_2}(r)$) as well as glo–bally by the objective value idea ($F_{ALG_1}(v)$, $F_{ALG_1}(v)$). In the context of Industry 4.0 and concerning the question of information value, it is easy to see that taking as $ALG_1$ an algorithm without lookahead and $ALG_2$ an algorithm with lookahead, an indication of whether lookahead pays off in a problem setting or not is obtained.

## 4 Truck Access Control at Factory Sites

A number of simulation studies have been conducted in order to assess the value yielded by additional information. A detailed discussion of the impact of additional lookahead data in warehouse order picking and pickup and delivery systems can be found in Dunke (2014). However, these applications are somewhat basic with respect to the opportunities of Industry 4.0.

For this reason, the following discussion concentrates on another application: In order to relate to a larger number of the opportunities given through information gathering devices (e.g., sensor/actor systems, continuous data transmission, global positioning systems) as fundamentally supposed in Industry 4.0, a number of different technology equipment scenarios is considered for the truck arrival process at the main gate of a process manufacturing company. Throughout the day, a large number of trucks carrying loads of different raw materials arrive at the main gate in order to bring the raw materials to the manufacturing sites at the company grounds. Upon arrival at the gate, a truck has to enter one of the available lanes (when it is free) in order to handle all administrative requirements (e.g., weighing of transported materials, filling out administrative forms, determine the routing and order of sites through the company grounds). The goal of the company is to ensure that each fac-tory site has on stock the required amounts of the respective raw materials when needed according to the production schedules. In case that there is missing material, a penalty cost for lost production is charged. Clearly, the company wants to minimise overall penalty costs.

### 4.1 Technology Scenarios

Concerning the technological configuration, several scenarios are considered for the purpose of examining whether the installation and maintenance costs for the techno–logical improvements can be justified by substantial penalty costs reductions.

*Scenario 1 (Base case):* No additional information is provided to the operators at the main gate. Thus, drivers will select freely whatever lane they would like to use. Algorithm DRIVERSCHOICE reflects the free choice of the drivers by assigning an arriving truck to a random lane.

*Scenario 2:* Additional information about the occupation of the lanes (in terms of the number of trucks) is given to the operators, e.g., through a camera system. Using this information, it is possible to prescribe the drivers to a specific lane. Algorithm LEASTLOAD assigns an arriving truck to a lane with the least number of trucks.

*Scenario 3:* In addition to the information from scenario 2 also the current status of the check-in counters of the lanes is provided, e.g., through electronic data interchange. Based on these data, the decision maker is capable of factoring in recent breakdowns or disturbances. Algorithm LEASTLOADUNBLOCKED assigns an arriving truck to a currently operating lane that exhibits the least number of trucks; if all lanes are blocked, a lane with the least number of trucks is chosen.

*Scenario 4:* In addition to the information from scenarios 2, 3 also the expected (remaining) workloads (for gate service) of all waiting and currently processed trucks is transmitted to the main gate operators, e.g., through the combined usage of status protocols and track & trace technology. Algorithm LEASTREMAINING UNBLOCKED assigns an arriving truck to a currently operating lane that exhibits the least remaining overall workload; if all lanes are blocked, an arriving truck is assigned to a lane with the least remaining overall workload.

*Scenario 5:* In addition to the information from scenarios 2, 3, 4 also the types and amounts of raw materials in the truck loads is reported to the operators. This can be achieved by forwarding truck load information to the operator at the moment of release at the logistics carrier. Algorithm SAVINGS computes for each truck a savings value which reflects the amount of penalty costs that could be saved if the truck was granted immediate access. Then, a truck with the largest savings value is assigned to a currently operating lane that exhibits the least remaining overall workload; if all lanes are blocked, a truck with the largest savings value is assigned to a lane with the least remaining overall workload. If there are no stock shortages at all at the moment, algorithm LEASTREMAININGUNBLOCKED is invoked.

*Scenario 6:* In addition to the information from scenarios 2, 3, 4, 5 also the geographical position of the trucks is sent continuously. This can be achieved by track & trace technology based on global positioning systems. Thus, the operators learn about the estimated arrival time of a truck a fair amount of time earlier than in the case without lookahead. In this application, it is assumed that the information about the expected truck arrival date is submitted when the truck first appears in a specific radius of interest (measured in minutes remaining until the truck will reach the gate). Algorithm SAVINGS LOOKAHEAD also computes a savings value for a truck in terms of saved penalty costs if that truck was granted immediate access. In contrast to the previous scenario, the savings value is computed not only for each truck currently waiting but also for each truck currently approaching the gate from within the lookahead radius (i.e., the trucks in the lookahead set). If there is a truck with a savings value larger than a prescribed threshold savings value, then one lane is reserved and additional trucks will be entered in this lane only if the station is free, i.e., there will be no waiting line in this lane so that the savings value can be realised as soon as possible. Otherwise, the algorithm behaves identical to SAVINGS.

## 4.2 Computational Results

The simulation-based optimisation for scenario/technology evaluation has been put into practice in a case study which has been implemented in the simulation software AnyLogic 7.1.2 (see Fig. 2). The replication results were analysed with the discussed performance measurement techniques and average case analysis. Each of the 100 simulation replications considers a work day of 14 hours. Throughout the day, 300 trucks with (randomly generated) loads for three raw material (goods) types arrive. These are consumed in seven factories continuously at time-varying rates. At the main gate, there are four lanes where trucks first await individual gate service before they enter company grounds. Regular gate service takes between 2 and 5 minutes. However, random breakdowns of lane counters with durations between 15 and 100 minutes are possible (e.g., when there are problems due to missing material bills). The lookahead duration in the last scenario is 60 minutes, i.e., by using the track & trace device a truck's expected arrival is known 60 minutes ahead of time.

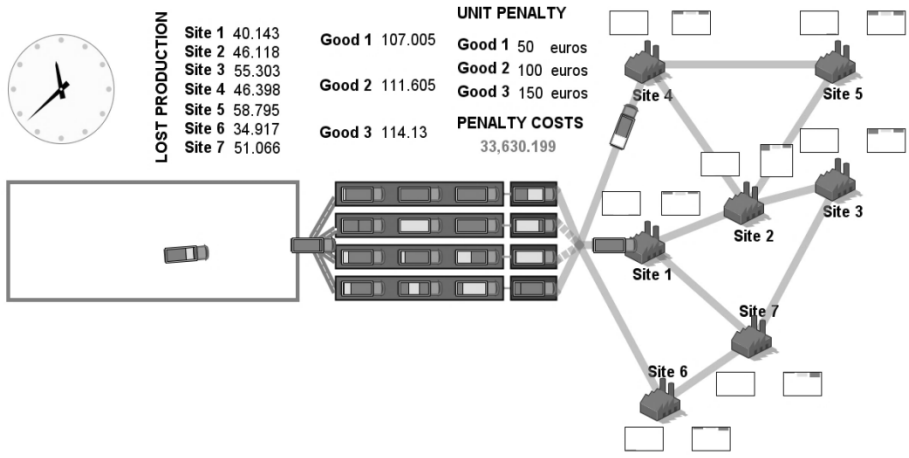



*Figure 2: Animation of the truck access control simulation model in AnyLogic*

### 4.2.1 Average Results

The average penalty costs for the six different scenarios with their respective algo–rithms are displayed in Table 1.

*Table 1: Average penalty costs in monetary units for the different algorithms / technological scenarios after 100 simulation replications*

| Algorithm | Penalty costs |
|---|---|
| DRIVERSCHOICE | 96055 |
| LEASTLOAD | 93414 |
| LEASTLOADUNBLOCKED | 93418 |
| LEASTREMAININGUNBLOCKED | 93396 |
| SAVINGS | 89116 |
| SAVINGSLOOKAHEAD | 89018 |

As the most important result, it is observed that there are two stages of fundamental improvement. The first one is achieved by restricting drivers to enqueue in a specific lane and not to let them choose freely. All of the three algorithms LEASTLOAD, LEASTLOADUNBLOCKED, LEASTREMAININGUNBLOCKED based on current know–ledge about each lane's workload yield improvements in the same order of magni–tude (around 2.75 %) relative to DRIVERSCHOICE. This result is especially important with respect to practical considerations since it suggests that the driver behaviour shall be controlled by the manufacturing company and not by the driver himself or by the logistics carrier. Since there is no substantial difference between technology scenarios 2, 3 and 4, we recommend installing the cheapest devices (with respect to procurement and maintenance) which facilitate a monitoring of the trucks in front of the gate. The second main result implies that there is another substantial improve–ment in terms of reduced penalty costs attainable when information about the load types and amounts is submitted to the gate operators. Algorithms SAVINGS and SAVINGSLOOKAHEAD both incur an additional reduction of the penalty costs of about 4.5 %. Although lookahead pays off at a slight margin, the additional impro–vement of SAVINGSLOOKAHEAD over SAVINGS is insignificant. In particular, the distributional analysis will reveal that there are several input instances where even SAVINGS outperformed SAVINGSLOOKAHEAD.

### *4.2.2 Distributional Results*

The analysis based on the counting distribution functions for the objective value (penalty costs) and the performance ratio (penalty costs of two algorithms relative to each other) provides us with additional insight compared to the average analysis. Figure 3 shows that the algorithms in the sets {DRIVERSCHOICE}, {LEASTLOAD, LEASTLOADUNBLOCKED, LEASTREMAININGUNBLOCKED} and {SAVINGS, SAVINGSLOOKAHEAD} can clearly be segregated from each other (similar to a sto–chastic ordering). At the same time, it is also observed that within the second and the third set, respectively, no algorithm consistently outperforms any other algorithm from the same set as seen from the plots frequently intersecting with each other.

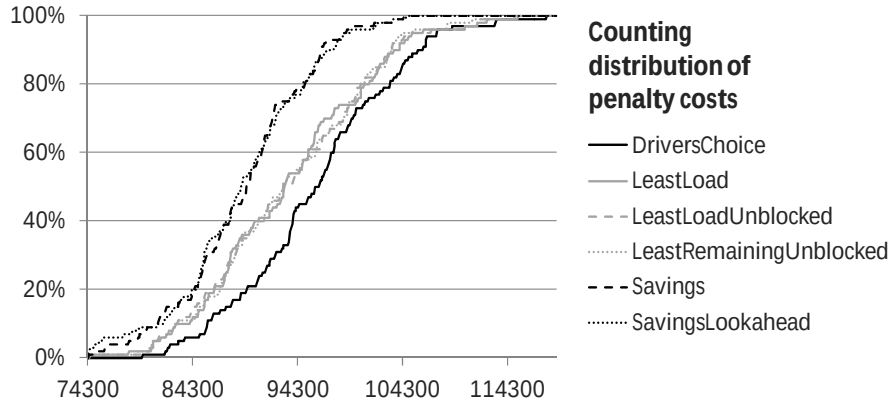



*Figure 3: Counting distribution functions of the penalty costs in monetary units for the different algorithms / technological scenarios after 100 simulation replications*

In Figure 4, the penalty costs attained by the respective algorithms are compared to those of DRIVERSCHOICE on the same input instances.

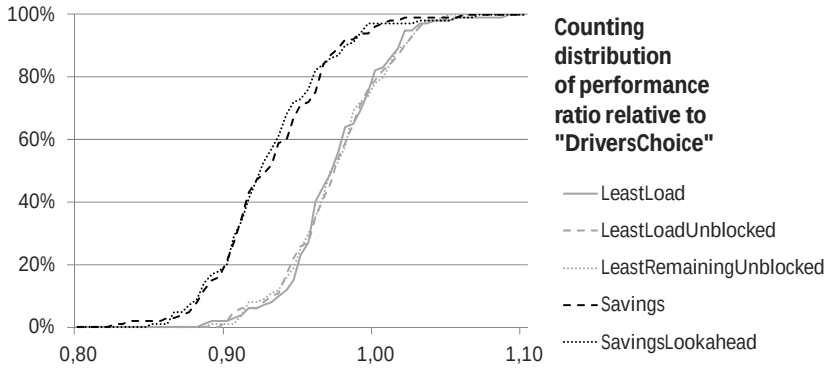



*Figure 4: Counting distribution functions of the performance ratios of the penalty costs for the different algorithms (technological scenarios) relative to algorithm DRIVERSCHOICE (scenario 1) after 100 simulation replications*

Although it is known that applying algorithm DRIVERSCHOICE represents the worst possible scenario, it is seen in the distributional analysis of the performance ratios that there is a small share of input instances where performance ratios are larger than 1 implying that DRIVERSCHOICE may outperform for example SAVINGSLOOKAHEAD occasionally. Again, the performance segregation of scenarios 2, 3, 4 on the one hand and scenarios 5, 6 on the other hand becomes apparent from the gap between the curves. Further, it is seen that each of the plots is relatively steep in a specific area where the main mass of all encountered inputs agglomerates. Thus, each algorithm (scenario) can be linked to a typical order of magnitude for improvement.

In conclusion, the overall recommendation consists of simultaneously installing information gathering devices according to scenario 2 concerning current lane statuses and according to scenario 5 concerning detailed load data of the trucks in front of the gate. As can be seen by algorithms LEASTLOAD and SAVINGS, these technological scenarios provide a considerable improvement in terms of penalty costs reductions. From the distributional analysis, it is also seen that there is no essential need to install the technological devices suggested in scenarios 3, 4 and 6.

## 5 Summary

In this article, we first outlined the need for a systematic approach to algorithm analysis in the context of Industry 4.0 as a consequence of the increasing number of possibilities arising by the introduction of information transmitting devices. Based on analogies between online optimisation and discrete event systems, simulation-based optimisation was established as a method to elicit the potential of different algorithm candidates under different information regimes. In a sample case study, these concepts were successfully applied as a first step to select the right techno–logies and algorithms for efficient operations in production and logistics.

## Acknowledgement

The research for this paper was financially supported by BASF (Ludwigshafen, Germany) within a research program.

## References

Albers, S.: Online algorithms: a survey. Math. Programming 97 (2003) 1-2, pp. 3-26.

Allulli, L.; Ausiello, G.; Laura, L.: On the power of lookahead in on-line server routing problems. Theoretical Computer Science 408 (2008) 2-3, pp. 116-128.

Bauernhansl, T.; Ten Hompel, M.; Vogel-Heuser, B. (Eds.): Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung, Technologien, Migration. Wiesbaden: Springer Vieweg 2014.

Borodin, A.; El-Yaniv, R.: Online computation and competitive analysis. Cambridge: Cambridge University Press 2005.

Boyar, J.; Irani, S.; Larsen, K. S.: A comparison of performance measures for online algorithms. In: Dehne, F; Gavrilova, M.; Sack, J.-R.; Tóth, C. D. (eds.): Algorithms and Data Structures. Berlin, Heidelberg: Springer 2009, pp. 119-130.

Bullinger, H.-J.; Ten Hompel, M.: Internet der Dinge. Berlin: Springer 2007.

Bundesministerium für Bildung und Forschung: Zukunftsbild „Industrie 4.0“, available at http://www.bmbf.de/pubRD/Zukunftsbild_Industrie_40.pdf, 2012.

Cassandras, C. G.; Lafortune, S.: Introduction to discrete event systems. 2. ed. New York: Springer Science+Business Media 2008.

Cormen, T.; Leiserson, C.; Rivest, R.; Stein, C.: Introduction to algorithms, 3rd edition. Cambridge: The MIT Press 2009.

Dunke, F.: Online optimisation with lookahead. Ph.D. thesis, Karlsruhe Institute of Technology, 2014.

Dunke, F.; Nickel, S.: Simulative algorithm analysis in online optimisation with lookahead. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (eds.): Simulation in Produktion und Logistik: Entscheidungsunterstützung von der Planung bis zur Steuerung. Paderborn: HNI-Verlagsschriftenreihe 2013, pp. 405-416.

Ghiani, G.; Laporte, G.; Musmanno, R.: Introduction to logistics systems planning and control. Chichester: John Wiley & Sons 2004.

Hiller, B.; Vredeveld, T.: Probabilistic alternatives for competitive analysis. Computer Science Research and Development 27 (2012) 3, pp. 189-196.

Jaillet, P.; Wagner, M. R.: Online routing: Value of advanced information as improved competitive ratios. Transportation Science 40 (2006) 2, pp. 200-210.

Lunze, J.: Ereignisdiskrete Systeme – Modellierung und Analyse dynamischer Systeme, 2nd edition. München: Oldenbourg 2012.

März, L.; Krug, W.: Kopplung von Simulation und Optimierung. In: März, L.; Krug, W.; Rose, O.; Weigert, G. (eds.): Simulation und Optimierung in Produktion und Logistik. Berlin, Heidelberg: Springer 2011, pp. 41-45.

Stadtler, H.; Kilger, C. (eds.): Supply Chain Management and Advanced Planning: Concepts, models, software, and case studies, 5th ed. Berlin, Heidelberg: Springer 2015.

*Simulation in Production and Logistics* 2015
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Beitrag zur Abfallreduktion mittels mathematischer Optimierung und simulationsbasierter Validierung

## *A Contribution to the Waste Reduction by Mathematical Optimisation and Simulation-based Validation*

Larissa Janssen, Thorsten Claus TU Dresden / IHI Zittau, Dresden (Germany), larissa.janssen@mailbox.tu-dresden.de, thorsten.claus@tu-dresden.de

Frank Herrmann, Ostbayerische Technische Hochschule Regensburg, Regensburg (Germany), frank.herrmann@oth-regensburg.de

Abstract: Perishable goods with a shelf life of 3-5 days require frequent inventory controls in the food retail trade. Maintaining a high service level for these products in stores is difficult due to demand uncertainty. A high service level often leads to a lot of stock in stores. In case it is not sold, it is waste. As an approach towards waste volume reduction, the economic order quantity (EOQ) problem is optimised in combination with the delivery frequency problem on the basis of the microperiods, i.e. morning, midday, afternoon, and evening. The presented model is designed for multiple stores with one regional warehouse. The simulation model measures the waste volume in stores. Experiments show that the combined optimisation of both problems leads to a reduction of total costs and waste volume.

## 1 Einführung und Abgrenzung der Arbeit

Das *operative Beschaffungsmanagement* ist eine wichtige Aufgabe in der betrieblichen Praxis. Es beschäftigt sich mit den Fragen: wann soll eine Bestellung aufgegeben und wieviel soll bestellt werden. Bei verderblichen Gütern kommt die Berücksichtigung der Haltbarkeit der Ware hinzu, die das Problem deutlich erschwert (Nahmias 2011). Eine aktuelle Literaturübersicht mit Publikationen zu verderblichen Gütern wurde von den Autoren dokumentiert (Janssen et al. 2015). Laut EHI Retail Institute spielt das Beschaffungsmanagement bei der Verringerung von Lebensmittelabfällen eine wichtige Rolle, insbesondere für Frischware mit kurzer Haltbarkeit (Kranert et al. 2012).

Es ist anzunehmen, dass bei der Nachfrage-Unsicherheit eine häufige Lieferung von Frischware in geringen Mengen zu weniger Abfällen führt als eine seltene Lieferung in großen Mengen. Es ist jedoch nicht klar, ob sich eine häufige Belieferung mit Frischware für Filialen finanziell lohnen würde und wie stark Abfälle reduziert werden. Das soll in dieser Arbeit unter Berücksichtigung der entstehenden Abfall-

entsorgungskosten in Filialen untersucht werden. Ein Beispiel ist das Angebot von vorgepackten frischen Fertigsalaten im Lebensmittelhandel. Die Idee ist, nicht auf Tagesbasis (*Makroperiode*), sondern in Tageszeiten wie Vormittag, Mittag, Nachmittag und Abend (*Mikroperioden*) zu optimieren, um die Liefermengen möglichst gering zu halten. Unter dem *Transportproblem* wird in dem vorliegenden Modell nur die Lieferhäufigkeit und Liefermengen durch ein Regionallager verstanden.

Arbeiten, die sich gleichzeitig auf das Beschaffungs- und Transportproblem für verderbliche Güter konzentrieren, unterscheiden sich in den Schwerpunkten und in den zugrunde liegenden Lagerhaltungsstrategien. Die meisten der Publikationen konzentrieren sich auf den Bereich Produktion oder das Bestellmengenproblem in Distributionszentren. Dagegen wird in dieser Arbeit das Bestellmengenproblem für mehrere Filialen betrachtet. Oft geht es um die Optimierung von Losgrößen und Transporten zwischen eigenen und/oder gemieteten Lagern sowie direkten Auslieferungen an Kunden (vgl. Arianezhad et al. 2013; Bhunia et al. 2014; Kopanos et al. 2012 u.a.). In Bhunia et al. (2015) wird u.a. auf die Inflationsbedingungen und in Gumasta et al. (2012) auf unterschiedliche Kundentypen eingegangen. Cai et al. (2013) berücksichtigen dieses Problem für firmenexterne Logistikdienstleister. Das Beschaffungsproblem und der Warenverderb während des Transports werden untersucht in: Ghare und Schrader (1963); Molana et al. (2012); Leśniewski und Bartoszewicz (2014). In Sazvar et al. (2014) wird eine Balance zwischen finanziellen und ökologischen Kriterien der beiden Probleme gesucht und je nach Menge von Treibhausgasen eine Entscheidung getroffen. In der Arbeit Amorim et al. (2013) wird, wie in dieser Arbeit, in Mikroperioden geplant, jedoch widmet sich diese Arbeit einem anderen Transportproblem (Vehicle-Routing-Problem mit Zeitfenstern).

Die Optimierung der beiden Probleme in kleineren Zeiteinheiten als ein Tag stellt ein wesentliches Unterscheidungsmerkmal zu anderen kombinierten Modellen dar, weil dort bei einer Periode in der Regel von einem Tag ausgegangen wird. Die Optimierung in Mikroperioden erfordert wegen der Umrechnung der Haltbarkeit die Entwicklung eines neuen Optimierungsmodells. In anderen Bereichen (wie zum Beispiel der Automobilindustrie) besteht das Problem des Warenverderbs nicht oder in der Produktion (Gebhard 2009) wird oft ein Teilproblem in Mikroperioden und das gesamte Problem in Makroperioden optimiert. In dieser Arbeit wird dagegen das kombinierte Problem ausschließlich in Mikroperioden betrachtet.

## 2 Optimierungsmodell

In diesem Abschnitt wird das Optimierungsmodell schrittweise entwickelt. Die Formulierung des bekannten Bestellmengenproblems von Wagner und Whitin (1958) ist als binäres lineares Optimierungsmodell in Häselbarth und Scholl (2003) angegeben.

### 2.1 Ausgangsmodell von Häselbarth und Scholl

Das Optimierungsmodell von Häselbarth und Scholl (2003) erweitert das Wagner-Whitin-Modell um dynamische Einstandspreise, Lagerkapazitäts- und Haltbarkeitsbeschränkungen von Frischwaren sowie Verlust aus Qualitätsgründen (Schwund, Mortalität, Schimmel und anderes). Alle Parameter sind in Tabelle 1 angegeben.

*Tabelle 1: Parameter und Entscheidungsvariablen*

| | |
|---|---|
| $c_t$: variable Einkaufskosten | $LK_t$: Lagerhaltungskosten in Periode $t$ |
| $BK_t$: fixe Bestellkosten | $q_t$: Bestellmenge in Periode $t$ |
| $d_t$: Nachfrage inklusive Abfallmenge | $\tau$: eine Folgeperiode nach $t$ |
| $G_t$: maximale Lagerkapazität | $X_{th}$: Lagerbestand in Periode $t$ von Alter $h$ |
| $\varepsilon_h$, $\varepsilon_j$: Faktor der verbliebenen Ware | $z_t$: binärer Bestellflag in Periode $t$ |

$$Min\ C = \sum_{t=1}^{T} BK_t \cdot z_t + LK_t \cdot X_t + c_t \cdot q_t \tag{1}$$

$$q_t \le \tilde{D}_t^{\bar{T}(t)} \cdot z_t \quad t = 1,...,T \tag{2}$$

$$X_{tH} = 0 \quad t = H+1,...,T \tag{3}$$

$$X_{t,h=0} = q_t - d_{t,h=0} \quad t = 1,...,T \tag{4}$$

$$X_{th} = \varepsilon_h \cdot X_{t-1,h-1} - d_{th}, \quad t = 2,...,T \quad h = 1,...,\bar{H}(t) \tag{5}$$

$$\sum_{h=0}^{\bar{H}(t)} d_{th} = d_t \quad t = 1,...,T \tag{6}$$

$$X_t = \sum_{h=0}^{\bar{H}(t)} X_{th} \quad t = 1,...,T \tag{7}$$

$$X_t \le G \quad t = 1,...,T-1 \tag{8}$$

$$X_{th} \ge 0, d_{th} \ge 0, X_t \ge 0,\ q_t \ge 0, X_T = 0 \quad t = 1,...,T \quad h = 0,...,\bar{H}(t)$$
$$z_t \in \{0,1\} \tag{9}$$

*Indizes:*

| | | |
|---|---|---|
| *h:* Alter | *H:* max. Alter | *h = 0..H* |
| *t:* Periode | *T:* Planungshorizont | *t = 1..T* |

Gleichung 2 ergibt sich aus der Beachtung der spätesten Verbrauchsperiode $\bar{T}(t) = min\{T, t+H\}$. $\tilde{D}_t^{\tau}$ ist die kumulierte modifizierte Bedarfsmenge, die wie folgt berechnet wird: $\tilde{D}_t^{\tau} = d_t + \Sigma_{k=t+1}^{\tau} \tilde{d}_t^k$. Die modifizierte Bedarfsmenge $\tilde{d}_t^{k\,\epsilon\,\tau}$ ist anstatt $d_t^{\tau}$

*Tabelle 2: Parameter und Entscheidungsvariablen*

| | |
|---|---|
| $BK_{ft}$: fixe Bestellkosten | $\omega^{fix}_{ft}$: fixe Transportkosten in Filiale $f$ |
| $c_{ft}$: variable Bestellkosten | $\omega^{var}_{ft}$: variable Transportkosten in Fil. |
| $d_{ft}$: Nachfragemenge in $t$ der Filiale $f$ | $\omega RL^{fix}_{t}$: fixe Transportkosten des RL |
| $d^{FIFO}_{fth}$: FIFO-Nachfrage vom Alter $h$ | $\omega RL^{var}_{t}$: variable Transportkosten des RL |
| $d^{LIFO}_{fth}$: LIFO-Nachfrage vom Alter $h$ | $q_{ft}$: Bestellmenge in Periode $t$ der Filiale $f$ |
| $\tilde{d}^{k \in \tau}_{ft}$: modifizierte Nachfrage in Periode $k$ der Filiale $f$ | $Q^{max}_{t}$: max. Kapazität eines LKW |
| $\tilde{D}^{\tau}_{t}$: kumulierte Nachfrage von $t$ bis $\tau$ | $\tau$: eine der Folgeperioden von $t$ |
| $\varepsilon_{fh}$, $\varepsilon_{fj \in h}$: Verbleibfaktor der Ware | *Vormittag:* erste Periode $t$ an einem Tag |
| $\varphi^{LIFO}$: Nachfrageanteil, LIFO-Prinzip | $x_{fh}$: Restlagerbestand nach Alter $h$ |
| $G_f$: maximale Lagerkapazität | $X_{fth}$: Lagerbestand in $t$ vom Alter $h$ |
| $LK_{ft}$: Lagerhaltungskosten in $t$ der $f$ | $X^{free}_{fth}$: abfallfreier Lagerbestand in $t$ |
| $LT$: Wiederbeschaffungszeit | $y_{ft}$: Lieferung in Periode $t$, $h$=0 |
| | $z_{ft}$: binärer Bestellflag in Periode $t$ |

mit zu bestellen. Die Bedarfsmenge $\tilde{d}^{k}_{t}$ enthält zusätzlich die Abfallmenge aus Qualitätsgründen: $\tilde{d}^{k}_{t}=d_{\tau} / \Pi^{\tau-t}_{j=1} \varepsilon_j$. Die Folgeperiode $\tau$ liegt im Intervall: $\tau \in [t+1,\ldots,\overline{T}(t)]$. Die maximale Lagerfähigkeit wird durch Gleichung 3 berücksichtigt, da die Frischware mit Lageralter $H$ nicht weiter gelagert werden darf. Die Lagerbilanzgleichung wird durch die Gleichungen 4-6 repräsentiert. Mit einer Bestellung trifft die Frischware ein, die gemäß Gleichung 4 mit dem Alter $h=0$ eingelagert wird, sofern sie nicht zur Nachfragebefriedigung genommen wird. Der bereits gelagerte Lagerbestand $X_{t-1,h-1}$ wird in Gleichung 5 von dem Abfall aus Qualitätsgründen bereinigt. Es verbleibt nur ein Anteil $\varepsilon_h$ der Frischware auf dem Lager (der abfallfreie Lagerbestand). Die maximal mögliche Haltbarkeit der Ware in Periode $t$ beträgt: $\overline{H}(t)=min\{H,t-1\}$. Durch Gleichung 6 wird garantiert, dass die Nachfrage in jeder Periode aus dem Lagerbestand vom Alter $h$ vollständig gedeckt wird. Der Bestand in $t$ ergibt sich gemäß Gleichung 7 als Gesamtmenge aller in $t$ lagernden Frischwaren verschiedenen Alters. Die beschränkte Lagerkapazität wird durch Gleichung 8 abgebildet. Die Variablen für Lager- und Verbrauchsmengen werden in Gleichung 9 auf nichtnegative Werte beschränkt. Gleichung $X_t=0$ bewirkt die Leerung des Lagers am Ende des Horizonts $T$. $z_t$ nimmt nur Werte 0 oder 1 an.

## 2.2 Optimierungsmodell ohne Transportkosten

Das entwickelte Modell erweitert das Bestellmengenproblem von Häselbarth und Scholl um Mikroperioden, Last-In-First-Out (LIFO)- und First-In-First-Out (FIFO)-Entnahmeart sowie Wiederbeschaffungszeit>0. Alle Parameter sind in Tabelle 2 angegeben. Vorausschauend wird für den Planungshorizont entschieden, in welchen Mikroperioden $t$ und wieviel bestellt wird. Falls es zu Fehlmengen in der rollierenden Planung kommt, wird der Fehlbestand nicht aufgebaut und er gilt als verloren. Die Lieferung $y_{ft}$ und der Restlagerbestand $x_{fh}$ werden jedoch in dem Optimierungsmodell berücksichtigt. Der Sicherheitsbestand (Herrmann 2011) wird in der rollierenden Planung ermittelt und zu der Nachfrage $d_{ft}$ aufaddiert.

Geltungsbereich für *Indizes* in allen Gleichungen, falls nicht anders angegeben ist:

| | | |
|---|---|---|
| *f*: Filiale | *F*: Filialenanzahl | *f = 1..F* |
| *h*: Alter | *H*: max. Alter | *h = 0..H* |
| *t*: Periode | *T*: Planungshorizont | *t = 1..T* |

Die Zielfunktion (Gl. 10) und alle Gleichungen des Ausgangsmodells von Häselbarth und Scholl (2003) werden unter anderem um einen Filialen-Index *f* erweitert. Das Optimierungsmodell sieht zum Schluss wie folgt aus:

$$Min\ C = \sum_{f=1}^{F}\sum_{t=1}^{T} BK_{ft} \cdot z_{ft} + LK_{ft} \cdot X_{ft} + c_{ft} \cdot q_{ft} \tag{10}$$

$$X_{ft,h=0}^{free} = 0 \quad t \in Vormittag \tag{11}$$

$$X_{ft,h>0}^{free} = \varepsilon_{fh} \cdot X_{f,t-1,h-1} \quad t \in Vormittag \tag{12}$$

$$X_{fth}^{free} = \varepsilon_{fh} \cdot X_{f,t-1,h} \quad t \notin Vormittag \tag{13}$$

$$X_{f,t=1,h>0}^{free} = x_{f,h>0} \tag{14}$$

$$X_{ft,h=0} = X_{ft,h=0}^{free} + y_{ft} - d_{ft0}^{LIFO} - d_{ft0}^{FIFO} \forall t = 1,...,LT \tag{15}$$

$$X_{ft,h=0} = X_{ft,h=0}^{free} + q_{f,t-LT} + y_{ft} - d_{ft0}^{LIFO} - d_{ft0}^{FIFO} \forall t = LT+1,...,T \tag{16}$$

$$X_{ft,h>0} = X_{ft,h>0}^{free} - d_{ft,h>0}^{LIFO} - d_{ft,h>0}^{FIFO} \tag{17}$$

$$d_{ft}^{LIFO} = d_{ft} \cdot \varphi^{LIFO} \tag{18}$$

$$d_{ft}^{FIFO} = d_{ft} - d_{ft}^{LIFO} \tag{19}$$

$$X_{fth} = 0 \vee d_{fth}^{LIFO} = d_{ft}^{LIFO} - \sum_{j=0}^{h-1} d_{ftj}^{LIFO} \tag{20}$$

$$X_{ft,H-h} = 0 \vee d_{ft,H-h}^{FIFO} = d_{ft}^{FIFO} - \sum_{j=H-h+1}^{H} d_{ftj}^{FIFO} \tag{21}$$

$$\sum_{h=0}^{H} d_{fth}^{LIFO} = d_{ft}^{LIFO}, \sum_{h=0}^{H} d_{fth}^{FIFO} = d_{ft}^{FIFO} \tag{22}$$

$$\sum_{h=0}^{H} X_{fth} \leq G_f \tag{23}$$

$$X_{fth}^{free} \geq 0, d_{fth}^{LIFO} \geq 0, d_{ft}^{FIFO} \geq 0 \tag{24}$$

Die Bestimmung des Makroperiodenwechsels ist notwendig, um das Alter der Frischware in Mikroperiode *t* zu ermitteln. Betrachtet man eine Mikroperiode relativ zu einer Makroperiode, so entspricht die relative Mikroperiode den Tageszeiten wie Vormittag, Mittag, Nachmittag und Abend. Der Wechsel einer Makroperiode findet am Vormittag statt und ist durch $t \in Vormittag$ festzustellen. Die relative Mikroperiode eines Tages wird aus dem Modulo bestimmt: *Ergebnis=t mod V*, wobei *V* die Anzahl von Mikroperioden in einer Makroperiode ist. In unserem Fall ist *V=4*. Ein Beispiel der Berechnung ist in Tabelle 3 gezeigt.

*Tabelle 3: Berechnung der relativen Mikroperiode aus t mod V*

| *Ergebnis* | *t* mod *V* | relative Mikroperiode |
|---|---|---|
| 1= | 5 mod 4: | Vormittag |
| 2= | 6 mod 4: | Mittag |
| 3= | 7 mod 4: | Nachmittag |
| 0= | 8 mod 4: | Abend |

Aufgrund der Mikroperiodenplanung ist die späteste Verbrauchsperiode aus Gleichung 2 wie folgt zu berechnen: $\bar{T}=min\{T,t+H\cdot V\}$. Das maximale Alter $\bar{H}$ der Frischware in Periode *t* aus Gleichung 5 gilt nicht mehr, da in dem entwickelten Modell der Restlagerbestand $x_{fh}$ aus der rollierenden Planung berücksichtigt wird und dieser kann verschiedene Alter *h* der Frischware enthalten.

Der abfallfreie Lagerbestand wird als $X_{fth}^{free}$ bezeichnet. Am Morgen gibt es bei der ganz frischen Ware ($h=0$) keine Abfälle (Gl. 11). Ansonsten ($h>0$) wird am Morgen nach einem Alterswechsel ($h-1$) der abfallfreie Lagerbestand ermittelt (Gl. 12). Am Mittag, Nachmittag und Abend wird $X_{fth}^{free}$ aus der vorigen Mikroperiode $t-1$ des gleichen Alters ermittelt (Gl. 13). Der Restlagerbestand $x_{fh}$ aus dem vergangenen Planungshorizont wird in $t=1$ angerechnet, wobei das Alter $h=0$ beim Restlagerbestand nicht vorhanden ist (Gl. 14).

Die Gleichung $X_T=0$ wird entfernt, da das Angebot der Frischware im Gegensatz zum Ausgangsmodell zeitlich unbegrenzt ist und zum Beispiel der Restlagerbestand aus dem letzten Planungshorizont die Nachfrage theoretisch übersteigen kann. Auf die Entscheidungsvariable $X_t$ wird verzichtet. Die Lagerbilanzgleichungen 4 und 5

werden durch Gleichungen 15-17 ersetzt. Ist die Wiederbeschaffungszeit $LT>0$ und $h=0$, trifft in $t \epsilon [1,\ldots, LT]$ keine Lieferung $q_{ft}$ ein (Gl. 15). In Periode $t \epsilon [LT+1,\ldots,T]$ kann eine Lieferung mit $h=0$ aus der Bestellperiode $t\text{-}LT$ eintreffen (Gl. 16). Ist das Alter $h>0$, so entfällt aus der Lagerbilanz die Bestellung $q_{ft}$ und die Lieferung $y_{ft}$ (Gl. 17).

Die Nachfrage $d_{ft}$ wird nach Entnahmearten in zwei Nachfragen $d_{ft}^{LIFO}$ und $d_{ft}^{FIFO}$ zerlegt (Gl. 18 und Gl. 19). Falls der Lagerbestand $X_{fth}$ mit der Frischware des Alters $h$ geringer oder gleich wie die benötigte Nachfragemenge aus $d_{ft}^{LIFO}$ ist, wird dieser Bestand ganz entnommen (Gl. 20). Falls der Lagerbestand $X_{fth}$ größer ist, wird aus dem Lager so viel entnommen, wie für die Nachfrage $d_{ft}^{LIFO}$ noch fehlt. Das Gleiche gilt auch für die Nachfrage $d_{ft}^{FIFO}$ (Gl. 21). $d_{fth}^{LIFO}$ und $d_{fth}^{FIFO}$ sind jeweils nach Alter $h$ aufgeteilte Nachfragemengen, die je nach verfügbarem Lagerbestand die Nachfrage $d_{ft}^{LIFO}$ und $d_{ft}^{FIFO}$ wiedergeben (Gl. 22). Die Lagerkapazitätsbegrenzung bleibt erhalten (Gl. 23). Weitere Nichtnegativitätsbedingungen sind in Gleichung 24 zusammengefasst.

## 2.3 Modellerweiterung um Transportkosten

Das Modell aus dem vorigen Abschnitt wird um Transportkosten des Regionallagers (*RL*) und Filialen (*Fil.*) erweitert:

$$Min\, C = \sum_{f=1}^{F}\sum_{t=1}^{T}(BK_{ft}\cdot z_{ft} + LK_{ft}\cdot X_{ft} + c_{ft}\cdot q_{ft} + \omega_{ft}^{fix}\cdot z_{ft} + \omega_{ft}^{\mathrm{var}}\cdot q_{ft}) + \sum_{f=1}^{F}\sum_{t=1}^{T}(q_{ft}/Q_t^{\max}\cdot \omega RL_t^{fix} + q_{ft}\cdot \omega RL_t^{\mathrm{var}}) \quad (25)$$

Die Transportkosten von Filialen setzen sich aus den fixen $\omega_{ft}^{fix}$ und variablen $\omega_{ft}^{var}$ Kosten zusammen. Die variablen Kosten sind von Bestellmengen und die fixen Kosten sind von der Lieferanzahl und LKW-Größe abhängig. Die Transportkosten des Regionallagers enthalten fixe $\omega RL_t^{fix}$ und variable $\omega RL_t^{var}$ Kosten. Die variablen Kosten beziehen sich auf die Menge der transportierten Frischware. Die Summe der Bestellmengen geteilt durch die Anzahl der maximalen LKW-Kapazität in Periode $t$ ergibt die Anzahl der hierfür benötigten LKWs. Die Anzahl der LKWs wird in realen Zahlen dargestellt. Nach der Optimierung muss die LKW-Anzahl auf die nächstgrößere Zahl aufgerundet werden.

# 3 Simulationsumgebung

Das entwickelte Optimierungsmodell wird unter einer rollierenden Planung in einer Simulationsumgebung für die unten beschriebene Fallstudie erprobt. Zum Aufbau der Simulationsumgebung wird das Simulationstool Tecnomatix Plant Simulation der Firma Siemens eingesetzt. Die Erzeugung von statischen und dynamischen Parametern (Nachfrage, Anfangslagerbestand, Kostensätze usw.) erfolgt in der Simulationsumgebung. Diese Informationen sind Input-Daten für das Optimie-

rungsmodell und werden über die Schnittstelle an den IBM ILOG Solver übergeben, der für das vorliegende Problem eine optimale Lösung ermittelt. Diese Lösung ist Optimierungs-Output und wird anschließend in der Simulationsumgebung weiterverarbeitet. Der Solver wird mit dem Aufruf der Datei oplide.exe gestartet. Die Pfade zu den Dateien mit dem Optimierungsmodell und den Input-Daten werden als Parameter beim Aufruf der Exe-Datei übergeben. Danach werden diese ausgelesen.

## 4 Numerisches Beispiel zur Fallstudie Fertigsalate

Es wird eine Fallstudie für vorgepackte Fertigsalate im Lebensmitteleinzelhandel betrachtet. Die Fertigsalate gibt es ganzjährig im Angebot. Es werden drei Optimierungsfälle betrachtet: (a) *getrennte* bzw. (b) *gemeinsame* Optimierung und (c) Optimierung bei einer *täglichen Belieferung*. In den Fällen a und c wird das Modell aus Abschnitt 2.2 und im Fall b aus dem Abschnitt 2.3 verwendet. D.h. in den Fällen a und c sind Transportkosten in dem Optimierungsmodell nicht enthalten. Dagegen werden im Fall b die Transportkosten in dem Modell berücksichtigt. In allen Fällen werden Transportkosten seitens der rollierenden Planung zu den Gesamtkosten hinzugerechnet, sobald eine Lieferung stattgefunden hat. In den Fällen a und b beträgt die Wiederbeschaffungszeit $LT$=1 Mikroperiode, im Fall c ist $LT=4$, d.h. die Frischware kann im Fall c max. einmal täglich geliefert werden.

Tabelle 4 liefert eine Übersicht von Parametern des RL und der Filialen. Alle Kosten sind in Geldeinheiten (*GE*) pro eine Mengeneinheit (*ME*) angegeben. Die Nachfrage in Filialen ist normalverteilt (Erwartungswert $\mu=E(d)=200$ mit Standardabweichung $\delta=50$). Die maximale Haltbarkeit beträgt mit dem Liefertag 4 Tage ($H=3$).

*Tabelle 4: Parameter*

| Kostenparameter in *GE*: | | Weitere Parameter: | |
|---|---|---|---|
| Fixe Transp.kosten Fil. $\omega_{ft}^{fix}$: | 1000 | Nachfrage in Filiale: | $\mu$=200, $\delta$=50 |
| Var. Transp.kosten Fil. $\omega_{ft}^{var}$: | 1 je *ME* | LIFO-Anteil $\varphi^{LIFO}$: | 0,4 |
| Fixe Transp.kosten RL $\omega RL_t^{fix}$: | 500/*LKW* | LKW-Kapazität $Q_t^{max}$: | 15000 *ME* |
| Var. Transp.kosten RL $\omega RL_t^{var}$: | 1 je *ME* | Planungshorizont $T$: | 12 |
| Bestellkosten $BK_{ft}$: | 1000 | max. Alter $H$: | 3 Tage |
| Lagerkosten $LK_{ft}$: | 1-3 je *ME* | Filialenanzahl $F$: | 1;2 |
| Abfallentsorgungskosten | | Wiederbesch.zeit $LT$: | 1;4 |
| in Filialen: | 5 je *ME* | Verbleibsfaktor $\varepsilon_h$ der Ware bei $h \in \{0,1,2,3\}$: | {1,1,1,1} |

Tabelle 5 zeigt die Simulationsergebnisse ohne und mit der gemeinsamen Optimierung des Beschaffungs- und Transportproblems sowie Ergebnisse bei einer täglichen Belieferung von Filialen. Die Gesamtkosten sind bei einer *gemeinsamen* Optimierung am niedrigsten. Der Anteil der Transportkosten (1.115 GE) an den Gesamtkosten (3.292 GE) beträgt in Filialen mit der *gemeinsamen* Optimierung 34%. Die durchschnittliche Lieferhäufigkeit pro Tag beträgt 1,42 und die durchschnittliche

Bestellmenge steigt gegenüber der getrennten Optimierung. Die Kosten bei der *täglichen Belieferung* sind am höchsten. Dies liegt hauptsächlich an den deutlich gestiegenen Lagerhaltungskosten (wegen der LT=4) und an den Abfallentsorgungskosten. Die hohe Zunahme der Abfallmengen (23%) ist in diesem Fall auf die reduzierte Lieferhäufigkeit zurückzuführen, die dazu beigetragen hat, dass die durchschnittliche Bestellmenge (1075) größer wurde. Die Abfälle entstehen aufgrund der Nachfrage-Unsicherheit, die durch den Sicherheitsbestand in der Simulationsumgebung kompensiert wird. Die *getrennte* Optimierung führt zu Mehrkosten von ca. 10% im Vergleich zum *gemeinsamen* Optimierungsfall.

*Tabelle 5: Durchschnittliche Optimierungsergebnisse*

| Mengeneinheit (*ME*)<br>Geldeinheiten (*GE*) | getrennte Optimierung | gemeinsame Optimierung | tägliche Belieferung |
|---|---|---|---|
| Bestellmenge *q* (*ME*): | 322 | 472 | 1075 |
| Abfallmenge (bei *h*>*H*): | 0% | 0,25% | 23% |
| Lieferanzahl pro Tag: | 2,15 | 1,42 | 1 |
| Fil. Beschaffung (*GE*): | 1.626 | 1.792 | 5.739 |
| Fil. Transport (*GE*): | 1.520 | 1.115 | 655 |
| RL Transport (*GE*): | 475 | 385 | 398 |
| Gesamtkosten (*GE*): | 3.621 | 3.292 | 6.792 |

## 5 Zusammenfassung

Die durchgeführten Experimente haben gezeigt, dass die Optimierung von Bestellmengen zusammen mit der Lieferhäufigkeit zur Minimierung von Gesamtkosten im Lebensmitteleinzelhandel führen kann. Bei einer häufigeren Belieferung als einmal pro Tag nehmen in Experimenten die Abfallmengen bei Frischwaren mit einer kurzen Haltbarkeit (3-5 Tage) ab. Außerdem haben die Experimente gezeigt, dass bei dieser Frischwarengruppe die gemeinsame Planung in Mikroperioden insgesamt günstiger ist als in Makroperioden (Tagen), weil dadurch die Lagerhaltungs- und die Abfallentsorgungskosten gesenkt werden können.

In weiteren Untersuchungen werden die Abfallentsorgungskosten und die unsichere Nachfrage (auf Basis von Szenarien) in dem Modell selbst berücksichtigt. Darüber hinaus wird eine Tourenplanung in Mikroperioden entwickelt, um die Gesamtkosten und die Abfallmengen weiter zu senken.

## Literatur

Amorim, P.; Belo-Filho, M.; Toledo, F.; Almeder, C.; Almada-Lobo, B.: Lot sizing versus batching in the production and distribution planning of perishable goods. International Journal of Production Economics 146 (2013) 1, S. 208–218.

Arianezhad, M.; Makuie, A.; Khayatmoghadam, S.: Developing und solving two-echelon inventory system for perishable items in a supply chain: Case study

(Mashhad Behrouz Company). Journal of Industrial Engineering International (2013) 9:39, S. 1-10.

Bhunia, A.; Jaggi, C.K.; Sharma, A.; Sharma, R.: A two-warehouse inventory model for deteriorating items under permissible delay in payment with partial backlogging. Applied Mathematics und Computation (2014) 232, S. 1125–1137.

Bhunia, A.K.; Shaikh, A.A.; Gupta, R.K.: A study on two-warehouse partially backlogged deteriorating inventory models under inflation via particle swarm optimisation. International Journal of Systems Science 46 (2015) 6, S. 1036-1050.

Cai, X.; Chen, J.; Xiao, Y.; Xu, X.; Yu, G.: Fresh-product supply chain management with logistics outsourcing. Omega 41 (2013) 4, S. 752–765.

Gebhard, M.: Produktion und Logistik: Hierarchische Produktionsplanung bei Unsicherheit. Wiesbaden: Gabler 2009.

Ghare, P.M.; Schrader, G.H.: A model for exponentially decaying inventory system. The Journal of Industrial Engineering 14 (1963) 5, S. 238–243.

Gumasta, K.; Chan, F.T.; Tiwari, M.: An incorporated inventory transport system with two types of customers for multiple perishable goods. International Journal of Production Economics 139 (2012) 2, S. 678–686.

Häselbarth, L.; Scholl, A.: Dynamische Bestellmengenplanung für verderbliche Luxusgüter. Arbeitspapier Friedrich-Schiller-Universität Jena, Wirtschaftswissenschaftliche Fakultät, 2003. ISSN 1611-1311.

Herrmann, F.: Operative Planung in IT-Systemen für die Produktionsplanung und -steuerung. Wiesbaden: Springer 2011.

Janssen, L.; Claus, T.; Herrmann, F.: Review of inventory systems with deterioration 2012-2015. Zittau: IHI 2015, http://bit.ly/1fez7UM.

Kopanos, G.M.; Puigjaner, L.; Georgiadis, M.C.: Simultaneous production and logistics operations planning in semicontinuous food industries. Omega 40 (2012) 5, S. 634–650.

Kranert, M.; Hafner, G.; Barabosz, J.; Schuller, H.; Leverenz, D.; Kölbig, A.; Schneider, F.; Lebersorger, S.; Scherhaufer, S.: Ermittlung der weggeworfenen Lebensmittelmengen und Vorschläge zur Verminderung der Wegwerfrate bei Lebensmitteln in Deutschland. Stuttgart: Institut für Siedlungswasserbau; Wassergüte- und Abfallwirtschaft 2012, http://bit.ly/1JHrAu7.

Leśniewski, P.; Bartoszewicz, A.: LQ optimal sliding mode control of periodic review perishable inventories with transportation losses. Advances in Intelligent Systems and Computing 240 (2014), S. 45–55.

Molana, S.M.H.; Davoudpour, H.; Minner, S.: An (r, nQ) inventory model for packaged deteriorating products with compound poisson demand. Journal of the Operational Research Society 63 (2012) 11, S. 1499–1507.

Nahmias, S.: Perishable inventory systems. New York, Dordrecht, Heidelberg, London: Springer 2011.

Sazvar, Z.; Mirzapour Al-e-hashem, S. M. J.; Baboli, A.; Akbari Jokar, M. R.: A bi-objective stochastic programming model for a centralized green supply chain with deteriorating products. International Journal of Production Economics 150 (2014), S. 140–154.

Wagner, H.M.; Whitin, T.M.: Dynamic version of the economic lot size model. Management Science 5 (1958) 1, S. 89-96.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Sensitivitätsanalyse eines Kostenmodells zur Ableitung von Kriterien für die simulationsbasierte Optimierung von Planungsprozessen

## *Sensitivity Analysis of a Cost Objective Function to Derive Criteria for the Simulation-based Optimisation of Planning Processes*

Mathias Kühn, Michael Völker, TU Dresden, Dresden (Germany),
mathias.kuehn@tu-dresden.de, michael.voelker@tu-dresden.de

Evangelos Angelidis, Daniel Bohn, Zhugen Zou, Universität der Bundeswehr München, Neubiberg (Germany), evangelos.angelidis@unibw.de, daniel.bohn@unibw.de, zhougen.zou@unibw.de

Abstract: In order to reduce additional costs due to increased planning effort of customized products, an effective production planning is necessary to realize optimised total production costs. The current work proposes a simulation based optimisation approach of such NP-hard multimode resource constrained problems for a novel comprehensive cost model. This paper demonstrates interdependence of various cost factors influencing total production cost with the conviction of dependence between objective and choice of priority rules. Methods for determining the interactions of cost-scenarios for robustness checks for the model-based derivation of priority rules in the simulation do not exist as a standard procedure. The following article presents the current state of research of the method for the analysis of influencing factors in a process cost model using the sensitivity analysis with the objective function optimised selection of priority rules.

## 1 Motivation und Zielstellung

Der Trend zur Nachfrage individueller Produkte führt zu einem Anstieg des Planungsaufwandes und der Produktionskosten in allen Branchen. Um diese Produkte ohne Mehrkosten für den Kunden herzustellen, ist eine Optimierung der Produktionskosten durch eine effektive Produktionsplanung notwendig. Dabei gilt es vor allem die zunehmende Anzahl der zu berücksichtigenden Einflussfaktoren auf die Zielerfüllung zu beherrschen und entsprechend bei der optimierten Ablaufplanung zu berücksichtigen. Dies gilt speziell für die Montage von Großerzeugnissen, die vorwiegend durch kostenintensive und hochspezialisierte Ressourcen realisiert wird (Carl und Völker 2011). Auf Grund des Produktionsumfangs werden die Kundenaufträge als Projekte initiiert. Die einzelnen Aufträge innerhalb der Projekte konkur-

rieren dabei um beschränkte Ressourcen, weshalb ein Ressourcenbelegungs- und Ressourcenzuteilungsproblem vorliegt. Dieses kann als Multi-Mode Ressource-Constrained Multi-Project Scheduling Problem (MRCMPSP) spezifiziert werden, welches NP-schwer ist.

Zur Lösung solcher Probleme liefert die Theorie der Simulationsbasierten Optimierung (SBO) vielversprechende Ergebnisse. Dazu wird die modellgetriebene SBO-Plattform verwendet, mit der unter minimalen Entwicklungs- und Zeitkosten entsprechende Lösungsverfahren für verschiedene Domänen entwickelt werden können (Angelidis et al. 2013). Basierend auf dieser Plattform wurde eine Heuristik für die komplexe Montage entwickelt. Das Optimierungskriterium ist die Minimierung von Kosten auf Basis eines Prozesskostenmodells, welches Zeit- und Ressourcenzielkategorien berücksichtigt (Carl und Angelidis 2015). Die Lösungsqualität jedes Simulationsmodells ist dabei abhängig von den Prioritätsregeln, die zur Auftragsauswahl in einer Warteschlange verwendet werden. Die Plattform bietet die Möglichkeit zur Auswahl und Kombination von mehr als 40 Regeln, die auf den Arbeiten von Haupt (1989) und Pinedo (2008) basieren. Zur Auswahl einer zielfunktionsoptimalen Prioritätsregel ist es notwendig die Wirkzusammenhänge und die Robustheit der Einflussgrößen im Modell zu kennen. Speziell der Einfluss von Kostenfaktoren in der Zielfunktion und wie diese in Projektszenarien korrelieren ist weitestgehend unbekannt, sodass die Auswahl der Prioritätsregeln in der Regel zufallsbasiert erfolgt (Angelidis et al. 2013). Zudem können Kostendaten meist nicht vollständig und genau bestimmt werden (Bsp.: Abschätzung Energieverbrauch) und sind durch unkontrollierbare Faktoren dynamisch (Bsp.: Auftragseingang bestimmt Maschinenabschreibungskosten), wodurch eine Szenarienbetrachtung im Rahmen möglicher Extremwerte notwendig wird.

Einzelne Untersuchungen zur Ermittlung der Wirkzusammenhänge in der Ablaufplanung mittels Sensitivitätsanalyse wurden durchgeführt (Hall 2004; Penz 2001). Darauf aufbauende Methoden zur modellbasierten Ableitung von Prioritätsregeln in der Simulation sind nicht erforscht. Der nachfolgende Beitrag stellt den Stand der Forschung zur Untersuchung von Einflussfaktoren in einem Prozesskostenmodell und zur zielfunktionswertoptimierenden Auswahl von Prioritätsregeln mittels Sensitivitätsanalyse vor. Durch die Beantwortung der Forschungsfragen ist zu bewerten, inwiefern sich die Sensitivitätsanalyse als Hilfsmittel in der Simulation eignet.

Forschungsfragen:

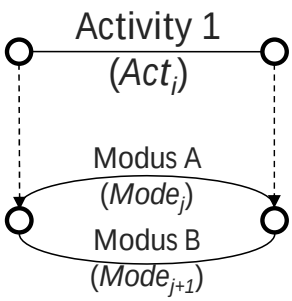
- Wie korrelieren die Zielfunktionsgrößen in verschiedenen Prozesskostenszenarien?
- Welche Optimierungskriterien des Prozesskostenmodells lassen sich aus der Sensitivität der Eingangsparameter ableiten und wie robust sind diese?

# 2 Modellierung und Lösungsverfahren

## 2.1 Modellierung

Die Modellierung zur Beschreibung der Montage komplexer Großerzeugnisse orientiert sich an einem realen Fallbeispiel mit der Zielstellung Ablaufpläne hinsichtlich minimaler Gesamtkosten aufzustellen. In einem Produktionssystem sind dabei mehrere Aufträge (Bsp. Druckmaschinen) mit einer Vielzahl von Arbeits-

gängen unter limitierten Ressourcen parallel abzuarbeiten. Jeder Auftrag wird als eigenständiges Projekt mit spezifischen Produkt-, Prozess- und Ressourcendaten und einem individuellen Netzplan initiiert. Die Vorgangspfeilnetzpläne (engl. activity-on-the-arc) enthalten die definierten Start- und Endtermine des Projekts. Zur Durchführung einzelner Aktivitäten im Netzplan wird die mögliche Auswahl zur Kombination von Anzahl und Art der den Qualifikationen zugeordneten Ressourcen festgelegt, woraus eine unterschiedliche Aktivitätendauer resultiert (Abb. 1). Diese alternativen Durchführungsmöglichkeiten einer Aktivität werden als Modi bezeichnet.




| Mode-Option for Activity 1 | | |
|---|---|---|
| | Demand Quali. A | Duration |
| Modus A | max | min |
| Modus B | min | max |

| Qualification-Matrix | | | |
|---|---|---|---|
| | Amount | Quali. A | Quali. B |
| Resource A | e.g. 6 | X | - |
| Resource B | e.g. 3 | X | X |

*Abbildung 1: Darstellung des Modi-Problems*

Ein vereinfachtes Fabrikmodell mit alternativen Prozessfolgen und bereits gewählten Modi ist in Abbildung 2 dargestellt. Durch die dargestellten Knotenarten wird bestimmt, welcher Vorgang als nächstes abgearbeitet wird. Handelt es sich um einen Knoten mit paralleler Abarbeitung, sind alle nachfolgenden Arbeitsvorgänge durchzuführen. Bei einem alternativen Knoten wird nur ein bestimmter Arbeitsvorgang in Abhängigkeit des Auswahlkriteriums (Bspw. Verfügbarkeit der Ressource) durchgeführt. Die dem Arbeitsvorgang zugeordneten Ressourcen sind durch Quadrate unterhalb des Vorgangs abgebildet.

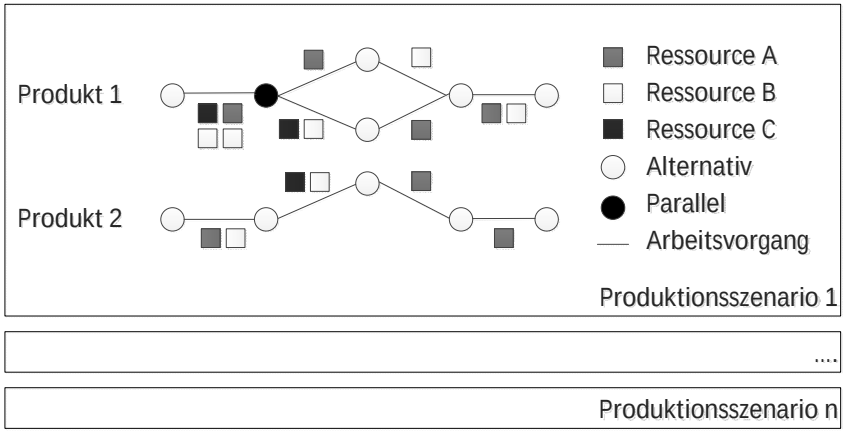



*Abbildung 2: Vereinfachtes Fabrikmodell*

Das in dieser Publikation untersuchte Modell einer komplexen Montagelinie wird nachfolgend durch die Spezifizierung charakteristischer Daten beschrieben. Grundlegend können die Daten in statische und dynamische Daten eingeteilt werden. Im betrachteten Modell sind statische Daten Ressourcendaten (Verfügbarkeit zu Anzahl und Art), Arbeitszeitmodelle (Schicht- und Pausenpläne), systemseitige Kostendaten (Bsp.: Produktionsgemeinkosten, Ressourcenkosten) und Projektstrukturdaten (Hierarchische Gliederung der Projekte). Dynamische Daten sind Auftragsdaten und beinhalten Angaben zur Projektlaufzeit (frühester Anfangs- und Endzeitpunkt), Projektcharakteristik (Anzahl und Vorrangbeziehungen der Aktivitäten), Modi (Aktivitätendauer in Abhängigkeit zur Ressourcenkombination) und kundenseitigen Kostendaten (Strafzahlungen bei verspäteter Lieferung, Bonuszahlungen bei vorzeitiger Lieferung). Die Zielfunktion zur Bewertung der Lösungen beinhaltet alle Faktoren, die durch die Montageablaufplanung beeinflussbar sind. Diese sind u.a. die Maximierung der Ressourcenauslastung, Minimierung der Durchlaufzeit und Maximierung der Termintreue. Mit der Normierung dieser Zielgrößen auf Basis von Kosten können diese in einer multikriteriellen Zielfunktion zusammengefasst werden. Das Prozesskostenmodell wird nachfolgend kompakt zusammengefasst. Für vollständige Erläuterungen zu Annahmen, Nebenbedingungen und mathematischen Formulierungen des Modells wird auf Majohr (2008) sowie Carl und Völker (2011) verwiesen. In der Zielfunktion sind Kosten für die Ressourcennutzung, Opportunitätskosten, Mietkosten, Strafkosten und Bonuskosten sowie zeitabhängige und zeitunabhängige Gemeinkosten berücksichtigt. Die Kostenarten sind in Tabelle 1 detailliert aufgeführt:

*Tabelle 1: Übersicht Prozesskostenparameter der Zielfunktion*

| Kostenparameter | Beschreibung |
|---|---|
| Strafkosten $v_i$<br>Bonuskosten $b_i$ | Kosten für die zeitliche Differenz zwischen Soll-Liefertermin und Ist-Liefertermin multipliziert mit dem spezifischen Kostenfaktor. |
| Zeitunabhängige Gemeinkosten $c_h$ | Kosten, die über den gesamten Planungshorizont anfallen und unabhängig von der Aktivitätendurchführung sind (Bsp. Verwaltungskosten). |
| Zeitabhängige Gemeinkosten $c_a$ | Kosten, die über die Gesamtheit aller Aktivitätendauern anfallen (Bsp. Heizkosten). |
| Ressourcenkosten regulär $c_k^r$<br>Ressourcenkosten Überstunden $c_k^o$<br>Ressourcenkosten Sonderschichten $c_k^s$ | Kosten, die bei der Durchführung von Aktivitäten anfallen. Abhängig vom jeweiligen Zeitabschnitt in dem diese anfallen (reguläre Arbeitszeit, Überstunde, Sonderschichten) erfolgt die Verrechnung mit dem spezifischen Kostenfaktor. |
| Ungenutzte Ressourcenkosten $c_o^k$ | Kosten, die anfallen wenn interne Ressourcen nicht ausgelastet sind. |
| Mietkosten $c_e^k$ | Kosten, die bei der Durchführung von Aktivitäten mit externem Personal anfallen. |

## 2.2 Fallbeispiel

Die Charakteristik des für die Forschung verwendeten Fallbeispiels eines Großanlagenbauers basiert auf dem im Kapitel 2.1. vorgestellten Datenmodell. Das Basisszenario enthält 10 Kundenaufträge mit insgesamt 13 herzustellenden Produkten, die in einem Zeithorizont von 20 Arbeitstagen abzuarbeiten sind. Es sind zwischen 126 und 276 Aktivitäten notwendig um ein einzelnes Produkt herzustellen. Insgesamt umfasst der Netzplan 2.737 Aktivitäten, die zwischen 1 und 2 Modi ausgeführt werden können. Der Ressourcenpool umfasst 13 Mitarbeiter mit 3 Personaltypen sowie einen Ausrüstungstyp Kran. Der Pool kann bei Bedarf um je eine Ressource pro Personaltyp und eine Ressource pro Ausrüstungstyp erweitert werden. Die Produktion erfolgt im 3-Schichtsystem. In den regulären Schichten von Montag bis Freitag steht der komplette Ressourcenpool pro Schicht zur Verfügung. Für Überstunden in der 2. Schicht stehen 2 Mitarbeiter pro Personaltyp und in der 3. Schicht 1 Mitarbeiter pro Personaltyp zur Verfügung. Für Sonderschichten am Wochenende stehen pro Schicht 2 Mitarbeiter pro Personaltyp zur Verfügung. Überstunden am Wochenende sind unzulässig.

## 2.3 Simulationsbasierte Optimierung

In der Optimierung von Montagelinien bedarf es neuer Lösungsstrategien, um in großen Suchräumen innerhalb kurzer Laufzeiten optimierte Ablaufpläne zu generieren. Eine Plattform, die sich mit dieser Problemstellung beschäftigt, ist in Angelidis et al. (2013) beschrieben. Die erste Herausforderung besteht darin, das reale System in ein sogenanntes Fabrikmodell zu transformieren. Dieses Modell beschreibt die Produktionsanlage und enthält alle notwendigen Informationen, um alle möglichen Produktionsszenarien abzubilden. Der Optimierer erstellt basierend auf der Lösungsstrategie und bereits simulierten Szenarien weitere Simulationsmodelle. Durch die Einbeziehung der Ergebnisse von vorangegangenen Szenarien wird versucht, eine Verbesserung zu erzielen. Ziel des Optimierungszyklus ist die Erstellung eines Ablaufplans, der die Gesamtkosten in dem betrachteten Zeitraum minimiert. Die Ergebnisse einer Simulation werden in einem Simulationslog gespeichert. Das Simulationslog enthält alle notwendigen Informationen, um das Szenario zu beschreiben und zu bewerten, wie bspw. zu Zuständen der Ressourcen und zum Zeitplan zur Durchführung der Arbeitsgänge. Aufbauend auf diesen Informationen werden die Key Performance Indicators (KPI) berechnet, damit der Optimierer den kompletten Simulationszyklus effizienter auswerten kann. Bspw. werden die KPIs zur Restzeit (Differenz zwischen Liefertermin und Fertigstellungszeitpunkt) von Aufträgen, Produkten oder Vorgängen ermittelt. Semi (2004) gibt einen Überblick aller verfügbaren KPIs. Der gesamte Ablauf zur Erstellung optimierter Produktionspläne mit der vorgestellten Plattform ist in Abbildung 3 dargestellt.

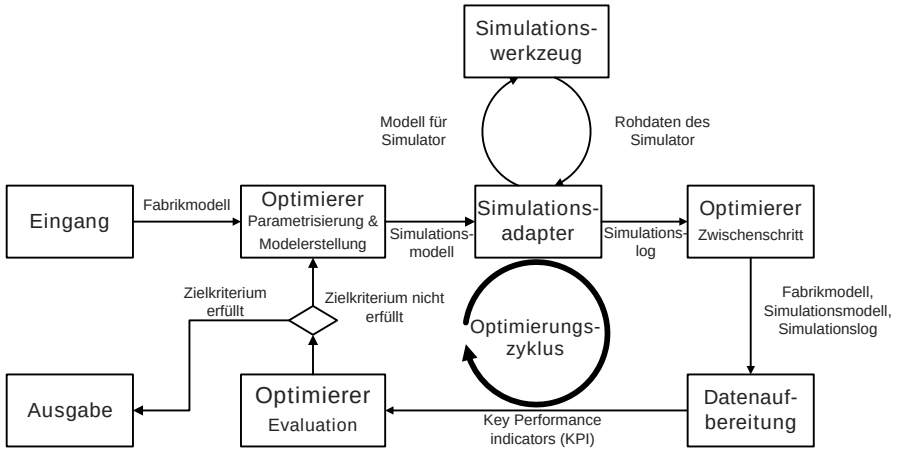



*Abbildung 3: Ablaufplanung mittels simulationsbasierter Optimierung*

# 3 Sensitivitätsanalyse zur Ableitung von Optimierungskriterien

Zur Beantwortung der Forschungsfragen werden computergestützte Laborexperimente durchgeführt. Dazu wird zunächst der Versuchsplan mit den zu variierenden Parametern aufgestellt. Der nächste Schritt ist die Erstellung des kostenminimalen Ablaufplans pro Experiment mit der vorgestellten Software SBOP. Nach Durchführung der Experimente erfolgt die Sensitivitätsanalyse zur Ermittlung der Interaktionseffekte. Auf Basis der ermittelten Einflussfaktoren mit der größten Sensitivität in der Zielfunktion ist ein optimales Prioritätsregel-Set zusammenzustellen.

## 3.1 Simulationsexperimente

Die Auswahl der zu untersuchenden Parameter erfolgt auf Basis des in Kapitel 2.1 vorgestellten Prozesskostenmodells sowie des in Kapitel 2.2 vorgestellten Szenarios. Aus dem Kostenmodell und dem Szenario ergeben sich 48 Parameter. Für den Parameterraum werden branchenübliche Werte für unterschiedliche Kostenszenarien angenommen. Das Inkrement wird mit 1 Geldeinheit (GE) festgelegt. Für die Untersuchung von Interaktionseffekten der Parameter und Auswertung dieser mittels einer Sensitivitätsanalyse eignet sich ein vollständig randomisierter Versuchsplan. Um das Auftreten und die Untersuchung dieser Effekte zu gewährleisten wird die Versuchsreihe mit 1000 zufällig erzeugten Produktionsmodellen festgelegt. Bei dem als Blackbox betrachteten Optimierer handelt es sich um einen genetischen Algorithmus. Die Einstellungen zur Mutationsrate (10 %), Anzahl der Generationen (20), Crossoverpunkt (5) und Populationsgröße (10) werden als gegeben betrachtet. Für dieses Experiment wird die Prioritätsregel FIFO angewendet.

Die Experimente wurden mit einem PC mit 4 Kernen und 2,4 Gigahertz Taktfrequenz und einem Arbeitsspeicher von 24 Gigabyte durchgeführt. Die Rechenzeit für die Optimierung und Ergebnisausgabe der Versuchsreihe betrug ca. 30 Stunden. Es wurden 1000 Produktionsmodelle erzeugt, dies führte zu einem Szenario-Aufkommen von 60000 Modellen.

## 3.2 Sensitivitätsanalyse

Für die Sensitivitätsanalyse wird die auf Python basierte Software PySen verwendet. PySen ist ein meta-modellbasiertes Tool zur Durchführung von globalen Sensitivitätsanalysen (Reuter et al. 2011). In der Tabelle 2 sind die Ergebnisse und spezifische Auswertungen zusammengefasst.

*Tabelle 2: Ergebnisse der Versuchsreihe*

| Kosten–parameter | Kosten–minimum $K_{min}$ (GE) | Kosten–maximum $K_{max}$ (GE) | Extremwert–verhältnis $K_{max} / K_{min}$ | Kosten–mittelwert $K_{mw}$ (GE) | Ø Anteil A Kostenart an Gesamtkosten (%) | Sensitivitäts–wert S (%) |
|---|---|---|---|---|---|---|
| $v_i$ | 626 | 14.776 | 23,6 | 10.702 | 3,01 | 18,15 |
| $b_i$ | -7.294 | -16.154 | 2,21 | -11.554 | -3,25 | 17,91 |
| $c_h$ | 34.425 | 53.622 | 1,55 | 43.311 | 12,24 | 1,33 |
| $c_a$ | 12.898 | 30.882 | 2,3 | 22.089 | 6,22 | 1,43 |
| $c_k^r$ | 99.132 | 211.732 | 2,13 | 159.782 | 45,05 | 29,04 |
| $c_k^o$ | 0 | 344,54* | 344 | 90 | 0,025 | 5,38 |
| $c_k^s$ | 18.019 | 34.409 | 1,90 | 25.990 | 7,32 | 7,44 |
| $c_o^k$ | 19.579 | 54.742 | 2,79 | 36.435 | 10,27 | 10,13 |
| $c_e^k$ | 54.249 | 84.078 | 1,54 | 67.829 | 19,12 | 9,19 |
| $K_{ges}$ | 281.331 | 420.827 | 1,49 | 354.673 | 100 | 100 |

*Auf Grund des Werts von $K_{min}$=0 GE wird für Berechnung von $K_{max}/K_{min}$ $K_{min}$ mit 1 GE angenommen.

Die Sensitivitätsanalyse ergab die höchsten Sensitivitätswerte S für die Eingangsfaktoren der regulären Ressourcenkosten (29,04 %), für die Strafkosten (18,15 %), für die Bonuskosten (17,91 %), für die Kosten ungenutzter Ressourcen (10,13 %) sowie für die Mietkosten (9,19 %). Der Vergleich zwischen Sensitivitätswert S und dem durchschnittlichen prozentualen Anteil der betrachteten Kostenart an den Gesamtkosten A zeigt, dass hier die größten Unterschiede auftreten. Für die nachfolgenden Untersuchungen wird festgelegt, dass nur Einflussgrößen betrachtet werden, bei denen der durchschnittliche prozentuale Anteil an den Gesamtkosten A oder der Sensitivitätswert S größer als 10 % ist. In Abbildung 4 sind die Kostenverläufe der Versuchsreihe zu den betrachteten Einflussgrößen dargestellt, aufsteigend sortiert nach den Gesamtkosten.

## 3.3 Auswahl Prioritätsregeln

Der zu betrachtende Stellhebel zur Kostenoptimierung ist die Auswahl der Prioritätsregeln. Dazu wird die Prioritätsregel FIFO mit der Wirkungsweise drei weitere Prioritätsregeln entsprechend den Optimierungskriterien verglichen. Zur Erhöhung der Termintreue (Reduzierung Strafkosten) wird die Regel Earliest Operation Due Date (ODD), zur Minimierung der Varianz der Verspätung die Crititcal Ratio-Regel

(CR) und zur Minimierung der Ressourcenkosten die Regel Shortest-Processing-Time (SPT) verwendet. Die Auswahl der Prioritätsregeln basiert auf den ermittelten Wirksweisen auf eine einkriterielle Zielstellung, die u.a. durch Corsten und Gössinger (2009), Fischer (2007) und Iringova et al. (2012), ermittelt wurden.

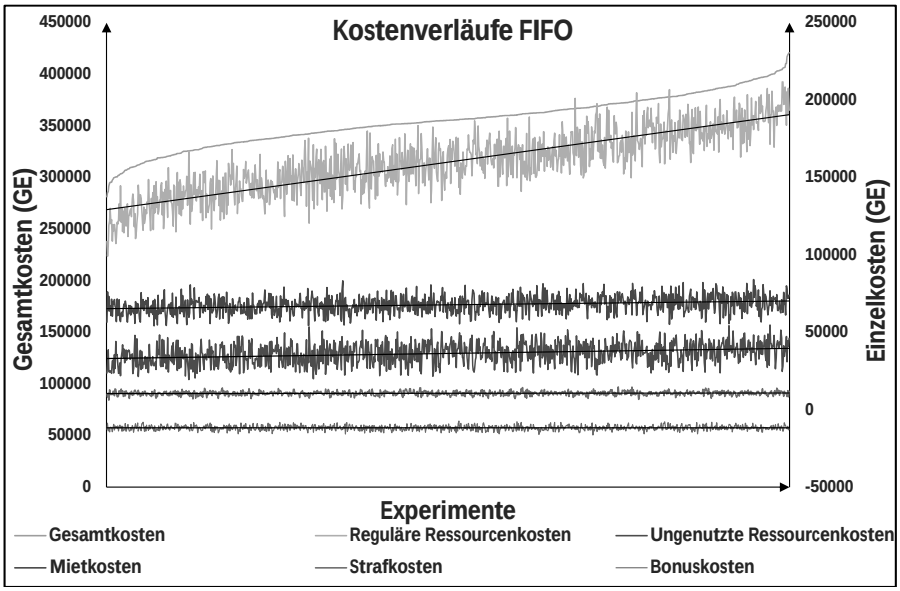



*Abbildung 4: Kostenverläufe der Versuchsreihe*

Für die Untersuchung zur Wirkungsweise der Einzel-Prioritätsregeln werden aus den im Vorfeld durchgeführten Versuchen 150 Prozesskostenszenarien zufällig ausgewählt. Die Generierung der Ablaufpläne erfolgt mit den identischen Einstellungen des genetischen Algorithmus zur Optimierung.

Die Auswertung erfolgt exemplarisch für die zwei Kostenarten mit dem höchsten Sensitivitätsanteil an den Gesamtkosten. Die ODD-Regel weist mit Abstand die beste Optimierung der Gesamtkosten und Strafkosten auf. Die CR-Regel bewirkt Einsparungen zu den regulären Ressourcenkosten und Strafkosten, führt jedoch zu einer geringen Verschlechterung der Gesamtkosten. Dies kann durch geringere Bonuskosten begründet werden, da diese Regel die Varianz der Verspätung reduziert. Mit Anwendung der SPT-Regel werden die regulären Ressourcenkosten optimiert, indem in Summe mehr kostengünstigere Ressourcen verwendet werden und somit die Auslastung steigt (Einsparung ungenutzte Ressourcenkosten gegenüber FIFO Ø 5.334 GE). Tabelle 3 fasst die Ergebnisse der Versuchsreihe zusammen.

Aus den Erkenntnissen zu den Wirkungsweisen der Prioritätsregeln wird ein optimiertes Prioritätsregel-Set mit der Anwendungsreihenfolge ODD-Regel, CR-Regel und SPT-Regel zusammengestellt. Zur Bewertung der Perfomance wird die Reihenfolge der Prioritätsregeln im Set variiert. Abbildung 5 stellt den Gesamtkostenverlauf der Prioritätsregel-Sets und den Kostenverlauf des optimalen Prioritätsregel-Sets im Vergleich zu den Single-Prioritätsregeln dar.

*Tabelle 3: Wirkungsweise ausgewählter Prioritätsregeln gegenüber FIFO*

| | SPT | ODD | CR |
|---|---|---|---|
| | | Gesamtkosten | |
| Ø Einsparung (GE) | -421 | 6.319 | -153 |
| Verbesserung | 69 | 134 | 73 |
| Verschlechterung | 81 | 16 | 77 |
| | | Reguläre Ressourcenkosten | |
| Ø Einsparung (GE) | 6.022 | -665 | 5.086 |
| Verbesserung | 147 | 66 | 135 |
| Verschlechterung | 3 | 84 | 15 |
| | | Strafkosten | |
| Ø Einsparung (GE) | -5.159 | 10.816 | 4.776 |
| Verbesserung | 1 | 150 | 150 |
| Verschlechterung | 149 | 0 | 0 |

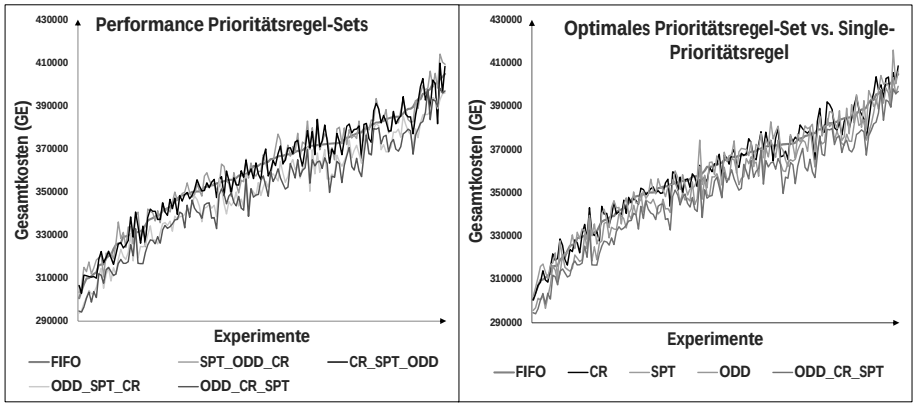



*Abbildung 5: Wirkungsweise der Prioritätsregel-Sets*

Die größte Einsparung gegenüber FIFO wird mit dem im Vorfeld als optimal angesehen Prioritätsregel-Set erzielt. Bemerkenswert ist, dass eine Einsparung in allen der betrachteten Kostenarten eintritt (Vgl. Tabelle 4). Mit 10.816 (GE) wird die größte Einsparung bei den Strafkosten erzielt. Die einzige Set-Reihenfolge mit negativen Auswirkungen auf die Gesamtkosten ist das Set SPT, ODD, CR. Bei dem betrachteten Set ist jedoch die Einsparung der regulären Ressourcenkosten am größten.

*Tabelle 4: Wirkungsweise der Prioritätsregel-Sets gegenüber FIFO*

| | ODDCRSPT | ODDSPTCR | SPTODDCR | CRSPTODD |
|---|---|---|---|---|
| | | Gesamtkosten | | |
| Ø Einsparung (GE) | 10.197 | 8.869 | -1.039 | 210 |
| Verbesserung | 148 | 146 | 63 | 75 |
| Verschlechterung | 2 | 4 | 87 | 75 |
| | | Reguläre Ressourcenkosten | | |
| Ø Einsparung (GE) | 4.885 | 4.784 | 5.874 | 5.539 |
| Verbesserung | 135 | 137 | 144 | 141 |
| Verschlechterung | 15 | 13 | 6 | 9 |
| | | Strafkosten | | |
| Ø Einsparung (GE) | 10.816 | 10.816 | -3.968 | 4.840 |
| Verbesserung | 150 | 150 | 6 | 150 |
| Verschlechterung | 0 | 0 | 144 | 0 |

## 4 Wesentliche Erkenntnisse

Zur Ableitung des Optimierungspotentials im Beispiel auf Basis der Sensitivitätsanalyse ist festzustellen, dass dieses nicht der Kostenart mit dem höchsten Sensitivitätswert (reguläre Ressourcenkosten 29,04 %) sondern den Strafkosten mit 18,15 % zugeordnet werden kann. Auch die üblicherweise verwendete Regel zur Optimierung der einzelnen Kostenart mit dem größten prozentualem Anteil an den Gesamtkosten ist hier unzutreffend (reguläre Ressourcenkosten 45,05 %, Strafkosten 3,01 %). In den betrachteten Versuchsreihen ist die Einsparung am größten, wenn das Verhältnis von Sensitivitätswert zu prozentualem Anteil der einzelnen Kostenart an den Gesamtkosten am größten ist. Je kleiner dieses Verhältnis ist, desto geringer ist das Optimierungspotential. Die Zusammensetzung des ermittelten optimalen Prioritätsregel-Sets entspricht dieser Erkenntnis. Auf Basis des Verhältnisses zwischen Sensitivitätswert und prozentualem Anteil der Kostenart an den Gesamtkosten können somit die zu optimierenden Parameter und die Auswahl der Prioritätsregeln eingegrenzt und gezielt ausgewählt werden. Grundlegend deutet jedoch ein hoher Sensitivitätswert auf ein hohes Optimierungspotential hin. Die Zusammenhänge sind in weiteren Experimenten zu untersuchen. Zur Wirkungsweise der Verwendung optimierter Prioritätsregel-Sets gegenüber einzelnen Prioritätsregeln konnte festgestellt werden, dass einzelne Ziele in einer multikriteriellen Zielfunktion berücksichtigt werden können. Zudem weisen die abgeleiteten zielfunktionsoptimalen Prioritätsregel-Sets eine größere Optimierungs-Robustheit über die Versuchsreihe gegenüber einzeln angewendeter Prioritätsregeln auf. Zusammenfassend können durch die Anwendung der Sensitivitätsanalyse simulationsunterstützende Auswertungen zur Optimierung durchgeführt werden. Unsere weitere Forschung fokussiert fortführende Untersuchungen zur zielfunktionswertoptimalen Auswahl von Prioritätsregeln mit der Übertragung in ein Standardverfahren zur Entscheidungsunterstützung.

## Förderhinweis

Die vorgestellten Ergebnisse sind Teil des Forschungsprojekts „Simulationsbasierte dynamische Heuristik zur verteilten Optimierung komplexer Mehrziel-Multiprojekt-Multiressourcen-Produktionsprozesse“ (Gefördert durch die Deutsche Forschungsgemeinschaft (DFG), Laufzeit 06/2014-06/2017) und wurden durch Voruntersuchungen im Rahmen von studentischen Abschlussarbeiten (u.a. Stumpf 2015) unterstützt.

## Literatur

Angelidis, E.; Bohn, D.; Rose, O.; Carl, S.: Simulation-based Optimization for Complex Assembly Lines with Workforce Constraints. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik 2013. Paderborn: Heinz-Nixdorf-Inst. Univ. Paderborn 2013.

Carl, S.; Angelidis, E.: Simulationsbasierte Prozesskostenrechnung zur Bestimmung kostenminimaler Ablaufplanvarianten in der Montageplanung bei KMU. Abschlussbericht: BMWi / IGF-Nr. ZBR 393 2015.

Carl, S.; Völker, M.: Prozesskostenorientiertes Montageplanungssystem (ProMoPs). In: Institut für Technische Logistik und Arbeitssysteme (Hrsg.): Planungsprozesse-Innovation & Nachhaltigkeit : Tagungsband Instituts–kolloquium 2011 des Instituts für Technische Logistik und Arbeitssysteme 2011.

Corsten, H.; Gössinger, R.: Produktionswirtschaft: Einführung in das industrielle Produktionsmanagement. München: Oldenbourg 2009.

Fischer, A.: Modellbasierte Wirkbeschreibung von Prioritätsregeln. Garbsen: PZH, Produktionstechn. Zentrum 2007.

Hall, N.G.; Posner, M.E.: Sensitivity Analysis for Scheduling Problems. Journal of Scheduling 7 (2004) 1, S. 49–83.

Haupt, R.: A survey of priority rule-based scheduling. OR Spektrum 11 (1989) 1, S. 3–16.

Iringova, M.; Vazan, P., Kotianova, J.; & Jurovata, D.: The Comparison of Selected Priority Rules in Flexible Manufacturing System. In: Proceedings of the World Congress on Engineering and Computer Science (Vol. 2) 2012.

Majohr, M.F.: Heuristik zur personalorientierten Steuerung von komplexen Montagesystemen. Dissertation. Dresden: Technische Universität Dresden 2008.

Penz, B.; Rapine, C.; Trystram, D.: Sensitivity analysis of scheduling algorithms. European Journal of Operational Research 134 (2001) 3, S. 606–615.

Pinedo, M.: Scheduling: Theory, algorithms, and systems. Dordrecht, New York: Springer Verlag 2008.

Reuter, U.; Mehmood, Z.; Gebhardt, C.; Liebscher, M.; Müllerschön, H.: Using LS-OPT for meta-model based global sensitivity analysis. 8th European LS-DYNA Conference 2011.

SEMI: Specification for Definition And Measurement Equipment Reliability, Availability And Maintainability: North American Regional Standards Committee, Semiconductor Equipment and Materials International. 2004.

Stumpf, M.: Konzipierung einer Entscheidungsgrundlage zur multikriteriellen Optimierung von Projektplanungsproblemen. Masterarbeit. Dresden: Technische Universität Dresden 2015.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Ein für die Wissenschaft freies Werkzeug zur Modellierung, Simulation und Optimierung

## *A Free Tool Suite for Modelling, Simulation and Optimisation for Academic Institutions*

Oliver Schönherr, Elias Winter, Christian Fabig, Michael Völker, TU Dresden, Dresden (Germany), Oliver.Schoenherr@tu-dresden.de, Elias.Winter@tu-dresden.de, Christian.Fabig@tu-dresden.de, Michael.Voelker@tu-dresden.de

Oliver Rose, Institut für Technische Informatik, Universität der Bundeswehr München, Neubiberg (Germany), Oliver.Rose@unibw.de

Abstract: This paper presents the new simulator OptimSim which is free for academic institutions and their industrial projects. The framework consists of multiple, extensible components like for simulation, optimisation, visualisation or interfaces that can be used standalone or in integrated ways. For industrial projects, OptimSim contains an integrated optimisation method which tries to create plans that are optimal, but robust with respect to uncertainties. We also use a new modelling approach as simulation model, which is partitioned into nine sub-models. This paper provides an overview of the entire concept, the simulator, the modelling approach and the optimisation environment. It summarises the current development state and sketches future plans.

## 1 Einleitung

Diese Arbeit beschreibt ein System zur Simulation und Optimierung im Bereich Produktion und Logistik: „Unter Produktion und Logistik werden Fertigungs-, Montage- und Produktionseinrichtungen einschließlich ihrer Prozesse sowie alle Aufgaben der Beschaffungs-, Produktions- und Distributionslogistik verstanden. Die Logistik bezieht sich dabei sowohl auf produzierende Unternehmen als auch auf nicht produzierende Betriebe wie Handelsunternehmen, Flughäfen und Krankenhäuser“ (März et al. 2011, S. 6). OptimSim versteht sich durch ein ganzheitlich umgesetztes Plug-in-Konzept als Baukasten, der alle notwendigen Komponenten für das Optimieren und Simulieren in der Produktion und Logistik bereitstellt und beliebig durch neue Komponenten erweitert werden kann. OptimSim ist ein System, das aus der Wissenschaft heraus entwickelt wurde und versucht, aktuelle Erkenntnisse der Simulation und Optimierung ganzheitlich umzusetzen. OptimSim wird für

Universitäten und andere wissenschaftliche Einrichtungen vollständig frei zur Verfügung gestellt, auch für deren industrielle Projekte.

Die Konzeption des Simulators wurde in einer Dissertation entwickelt, die das Ziel hatte, ein einheitliches, umfassendes, gut strukturiertes sowie einfach begreifliches Modellierungskonzept für die diskrete Prozessmodellierung bereitzustellen (Schönherr 2014). Als Modellierungssprache wurde die von der *Object Management Group* (OMG) veröffentlichte *Systems Modeling Language* (SysML) genutzt, die sich im Bereich des *Systems Engineering* als standardisierte Beschreibungssprache etablieren soll.

In der wissenschaftlichen Praxis zeigt sich immer wieder, dass die Forschung im Bereich der diskreten Simulation durch den Mangel an geeigneten Werkzeugen an ihre Grenzen geführt wird. So können geeignete Werkzeuge zu kostenintensiv sein, um wissenschaftliche Projekte durchzuführen, oder es gibt Anforderungen, die sich nicht mit den aktuellen Werkzeugen durchführen lassen, wie beispielsweise das zielgerichtete Anpassen der Simulationslogik, das Einbinden eines speziellen Optimierungsalgorithmus oder einer neu entwickelten Ergebnisausgabe. Ein weiteres Problem, das durch die Vielzahl genutzter proprietärer Tools entstanden ist, ist der Mangel einer einheitlichen Beschreibungssprache (Schönherr & Rose 2009). Mit OptimSim wollen wir dem Wissenschaftler in unserem Bereich eine Alternative geben und bieten ihm an, innovative Ideen mit dem pluginbasierten OptimSim zu entwickeln und das Werkzeug zudem frei zu nutzen.

## 2 Das Gesamtkonzept

OptimSim besteht aus sieben Komponenten, die pluginbasiert unabhängig voneinander einsetzbar sind und zudem vollständig dem wissenschaftlichen Umfeld frei zur Verfügung gestellt werden. Zentral steht dabei der Simulator im Vordergrund, der mit dem Ziel entwickelt wurde, die Anforderungen eines Simulations-/Optimierungswerkzeuges ganzheitlich abzudecken. Zu den Anforderungen gehören vor allem Flexibilität bei der Modellierung, schnelle Simulation insbesondere bei großen Modellen, frei definierbare und übersichtliche Darstellung der Ergebnisse, mögliche Datenbankanbindung zur Modell- und Ergebnisgenerierung und frei definierbare, integrierte Optimierungsmöglichkeiten. Um die genannten Anforderungen umsetzen zu können, ist vor allem die konsequente Verfolgung eines möglichst geeigneten Modellierungskonzeptes von Bedeutung. Wie in diesem Artikel ausführlich beleuchtet, unterscheidet sich das Modellierungskonzept deutlich von herkömmlichen Arbeiten. Das Eingabemodell für den Simulator kann unter anderem mit SysML modelliert werden, aber auch mittels Import aus einem Excel Spreadsheet (XLS), um industriellen Anforderungen zu genügen. Die SysML-Modelle können mit dem frei verfügbaren Modellierungswerkzeug TOPCASED Engineer erstellt werden, das speziell für die Domäne Produktion und Logistik angepasst wurde. Die XLS-Modelle können in industriellen Projekten beispielsweise von ERP-Systemen wie SAP mit Daten gefüllt werden. Weitere Modellformate sind durch den erweiterbaren Modellimport integrierbar. Aktuell wird an einem umfangreichen, individuell anpassbaren Optimierer gearbeitet, der voraussichtlich im Oktober 2015 in einer ersten Version zur Verfügung steht. Die Ergebnisse der Simulation können in Excel, aber auch mit Plantafeln eingesehen werden, zudem können die Ergebnisse automatisch als HTML-Datei generiert werden, um regelmäßige Ergebnisausgaben

online zur Verfügung zu stellen. Generell ist die Ergebnisaufbereitung, wie konzeptionell alle Teile des Systems durch jeweilige Plug-ins individuell gestaltbar. Abbildung 1 zeigt, aus welchen Plug-ins sich OptimSim aktuell zusammensetzt.

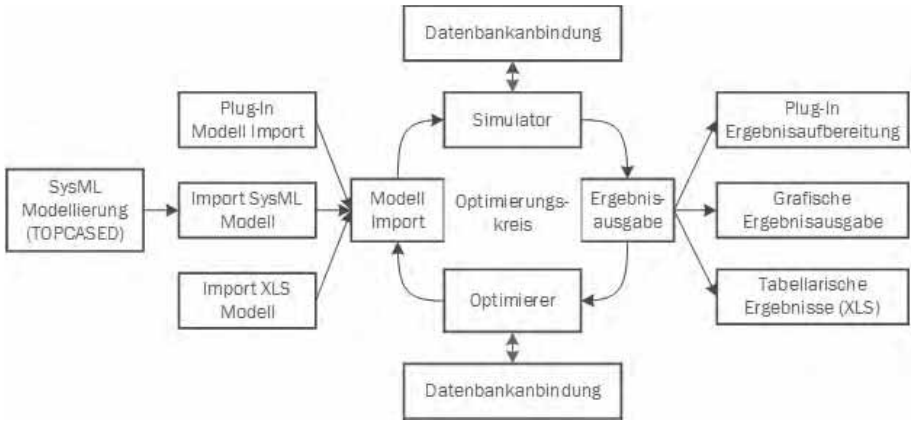



*Abbildung 1: Gesamtsystem OptimSim*

## 3 Das Modellierungskonzept

Das Modellierungskonzept basiert auf Arbeiten zur generellen Modellierung von Simulationsmodellen der Domäne Produktion und Logistik (Schönherr & Rose 2010). Es wurde besonders darauf geachtet, die Erstellung der Simulationsmodelle flexibel, strukturiert und nachvollziehbar (vollständig offengelegt und einsehbar) zu gestalten. Das Gesamtmodell des Untersuchungsgegenstandes wurde in der Literatur bisher weniger differenziert betrachtet. Oft wurde grundlegend nur zwischen Struktur und Verhalten unterschieden und teilweise kaum zwischen Simulator und Modell getrennt. „Meistens wird durch die Simulation ein Entscheidungsprozess unterstützt, bei dem mehrere Systemvarianten analysiert werden, die sich in Struktur oder Verhalten unterscheiden“ (VDI 1997; März et al. 2011, S. 13). Die Ergebnisse einer Dissertation zeigen auf, dass die semantische Strukturierung von Simulationsmodellen differenzierter zu betrachten ist (Schönherr 2014). Die Konzeption basiert auf einem neuen Ansatz, der den Simulator und sein Modell in neun Teilmodelle unterteilt.

Das Systemmodell besteht aus sieben Teilmodellen: Drei davon kann der Anwender bearbeiten, die übrigen vier dienen ausschließlich der Information. Das Systemmodell zeigt alle für den Anwender relevanten Informationen auf. Zudem legt es sämtliche Elemente, Verhaltensweisen, Steuerungsmechanismen, Zustände, Ereignisse und Kommunikationsabläufe fest. Ferner werden das experimentelle Modell und das Analysemodell außerhalb des Systems definiert. Beide legen fest, wie der Simulator auf dem Systemmodell arbeiten soll und welche Simulationsergebnisse (Statistiken) dem Anwender zurückgegeben werden (vgl. Abb. 2).

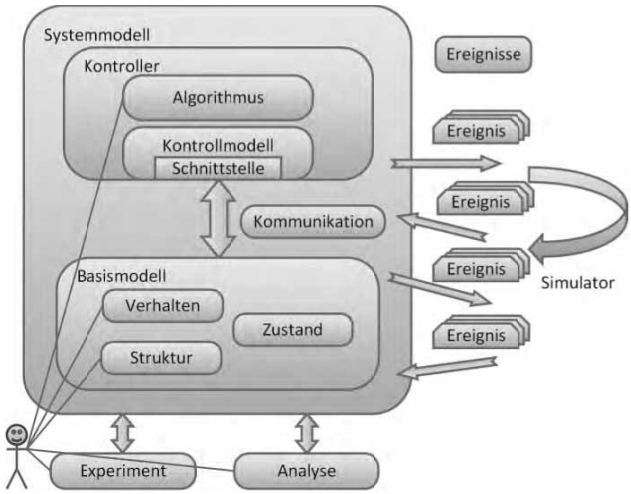



*Abbildung 2: Das Modellierungskonzept von OptimSim (Schönherr 2014, S.90)*

Das strukturelle Modell beschreibt die statische Struktur des Systems, wie die Elemente, deren Attribute sowie deren Beziehungen. Beispielsweise sind Ressourcen, Prozesse und deren Beziehungen Teile des statischen Modells. Das Verhaltensmodell beschreibt das dynamische Verhalten der Elemente bzw. zwischen ihnen, beispielsweise das Wandern des Werkstückes durch die Produktionsanlage einer Fertigung. Das Zustandsmodell beschreibt, welche Verhaltenszustände die Elemente der Simulation annehmen können. Das strukturelle Modell, das Verhaltensmodell und das Zustandsmodell bilden die Grundstruktur eines Systems ab und werden im Folgenden als Basismodell bezeichnet.

Die algorithmischen Entscheidungen können den vollständigen Systemzustand einbeziehen. Alle Elemente des Systems sind mit dem Steuermodell verbunden, das den kompletten Systemzustand kennt. Wenn eine Entscheidung innerhalb eines Systemelements getroffen werden soll, führt das Steuermodell einen Algorithmus aus und gibt das Ergebnis / die Entscheidung zum betreffenden Element zurück. Der Anwender kann zwischen verschiedenen vorhandenen Algorithmen im algorithmischen Modell wählen, aber auch neue Algorithmen definieren (im Simulator frei mit Java), die das Steuermodell dann ausführt. Für jeden schon vorhandenen Algorithmus kann der Anwender bestimmte Werte festlegen – beispielsweise, welche Anzahl an Elementen notwendig ist, um in einer *Constant Work in Process* (CONWIP) Regelstrecke neue Elemente einzuschleusen. Die gewählten Algorithmen werden dann über die zugeordneten Elemente vom Kontroller berechnet. Das Steuermodell definiert spezielle Elemente (beispielsweise Monitor, Kontroller) sowie Schnittstellen zwischen dem Steuermodell, dem algorithmischen Modell und dem Basismodell.

Das Kommunikationsmodell ist rein informativ für den Anwender. In ihm werden die Kommunikationsabläufe des Systems mit Sequenzdiagrammen festgelegt. Das Ereignismodell ist die Schnittstelle zwischen dem Systemmodell und dem Simulator. Genauso wie das Analysemodell ist auch das experimentelle Modell kein Teil des Systemmodells. In diesen Modellen wird festgelegt, wie der Simulator auf dem

Systemmodell arbeiten soll und welche Simulationsergebnisse der Anwender zurückbekommt. Im experimentellen Modell definiert der Nutzer unter welchen experimentellen Aspekten die Simulation ablaufen soll (beispielsweise der Simulationszeitraum). Im Analysemodell bestimmt der Anwender, welche Faktoren (Statistiken) betrachtet werden sollen.

Vor allem ermöglicht die differenziertere Betrachtung eines Simulationsmodells den Anwender Modelle semantisch strukturierter zu sehen und flexibler zu modellieren. Beispielsweise ermöglicht die Trennung zwischen Kontroll- und Basismodell dem Anwender, flexibel eigens definierte Algorithmen an beliebigen Stellen im Basismodell einzupflegen.

# 4 Simulator

Der Kern von OptimSim besteht aus dem im Plug-in-Konzept integrierten Simulator. Entsprechend der Klassifikationsmöglichkeiten für Simulationsmodelle nach (März et al. 2011, S. 16) ist ein diskreter, dynamischer, ereignisgesteuerter Simulationsansatz umgesetzt, der deterministische und nichtdeterministische Simulationsdurchläufe unterstützt. Die im Simulator integrierte Bibliothek für Zufallskomponenten kann über Plug-ins beliebig erweitert werden.

## 4.1 Umsetzung des Simulators

Das beschriebene Modellierungskonzept wurde im Simulator mit der objektorientierten Programmiersprache Java umgesetzt. Für alle Modellelemente der einzelnen Teilmodelle (Basismodell, Steuermodell, etc.) wurden Java-Schnittstellen definiert. Durch das konsequente Schnittstellenkonzept sind die Elemente voneinander entkoppelt und können mit geringem Aufwand angepasst werden. Auf diese Weise wird die geforderte flexible Erweiterbarkeit bezüglich des Modellierungskonzeptes erreicht. OptimSim enthält in seinem aktuellen Umfang bereits unterschiedliche Implementierungen für Algorithmen, die für verschiedene Elemente des Basismodells an sogenannten Kontrollpunkten integriert werden können. Kontrollpunkte sind somit die Schnittstelle vom Basismodell zur Steuerung und bei allem möglichen Zustandsänderungen zu finden. Neben den lokalen Steuereinheiten der einzelnen Modellelemente übernimmt der Kontroller globale Steuerungsaufgaben und ist durch den Monitor, der als Observer implementiert ist, stets über den vollständigen Systemzustand informiert.

## 4.2 Simulationsanalyse

Während der Simulation können verschiedene Werte aufgezeichnet werden, die jeweils unterschiedlich komplex in ihrer Berechnung sein können. Der Anwender legt anfangs fest, welche Werte er beobachten möchte. OptimSim realisiert eine modulare Berechnung. Nach dem Entwurfsmuster der Beobachter (Starke 2011, S. 169) ist das Werkzeug so entwickelt, dass es kleine lokale Wertberechnungseinheiten an gewünschten Ereignispunkten einzelner Elemente durchführt. Durch diese Designentscheidung kann die Simulationsanalyse vor der Simulation zielgerichtet und feingranular vorbereitet werden, um während der Laufzeit unnötige Berechnungen zu vermeiden. Es bleibt dem Anwender überlassen, welche Werte er bei welchen Elementen durch die Simulation ermitteln möchte. Die Simulations-

ergebnisse werden in einem weiteren Verarbeitungsschritt intervallbasiert aufbereitet. Der Anwender kann das Intervall selbst bestimmen. Beispielhaft sind so wochenfeine Kenngrößen für eine Jahressimulation oder minutenfeine Werte in einer Tagessimulation zielgerichtet einstellbar.

## 4.3 Ausgabemöglichkeiten

Die endgültigen Simulationsergebnisse liegen im Laufe der Simulation in einer Objektstruktur innerhalb des Arbeitsspeichers vor und können auf unterschiedliche Weise aufbereitet werden. Wie die Ergebnisse ausgegeben werden, kann in Plug-ins eigens definiert werden. In der derzeitigen Version existieren zwei Ausgabe-Plug-ins: zum einen als *Excel Spreadsheet* (XLS) und zum anderen in der *Java Script Object Notation* (JSON). Die tabellenbasierte Ausgabe aller Werte als XLS-Dateien ermöglicht es dem Anwender, die entstandenen Daten in verbreiteten Programmen zu visualisieren und aufzubereiten. Die Ausgabe der Daten im JSON-Format erfolgt derzeit vorwiegend für eine Zwischenablage der Werte zur Laufzeit, um den benötigten Arbeitsspeicher der Anwendung zu reduzieren. Aktuell wird an einer dritten, grafischen Ausgabevariante gearbeitet. Mit dieser wird es möglich, die Ausgabe als Gantt-Diagramme zu visualisieren. Die grafische Darstellung erfolgt plattformunabhängig mit HTML5 im Browser. Die Anwendung kann damit auf einem lokalen PC oder auf einem Server ausgeführt werden und über einen beliebigen Browser gesteuert werden. So kann das Ergebnis eines Simulationslaufes Mitarbeitern einer Fertigung über das Intranet bereitgestellt werden.

## 4.4 Performance

Der Simulator wurde mit dem Ziel entwickelt, bei größtmöglicher Flexibilität, auch umfangreiche Modelle schnell zu simulieren. Bereits durch die feingranular einstellbare Simulationsanalyse werden unnötige Berechnungen reduziert (vgl. Abschnitt 4.2). Die Simulationsgeschwindigkeit ist unter anderem von Einfüge-Operationen von Ereignissen in den sortierten Ereigniskalender abhängig. Um Performanceverlusten entgegenzuwirken, werden zustandsunabhängige Ereignisse, wie Dienstpläne der Mitarbeiter, vor der Simulation berechnet und bereits sortiert eingefügt. Je nach Anteil solcher Ereignistypen ermöglicht diese Initialisierung das Einsparen von Sortieroperationen. Während der Simulation muss der aktuelle Modellzustand wiederholt zur Berechnung von Algorithmen und neuen Ereignissen herangezogen und ausgewertet werden. Bedingt durch die Trennung der Teilmodelle, sind viele Kommunikationswege innerhalb des Simulators nötig. Es wurde bei der Umsetzung darauf geachtet, die Zahl der Kommunikationsaufrufe zu minimieren. Dies wurde durch Caching-Verfahren und der Verwendung von Datenstrukturen erreicht, die das Zusammenfassen mehrerer Werte unterstützen.

Bei der objektorientierten Simulation werden zahlreiche Objekte erzeugt, die nach kurzer Zeit wieder freigegeben werden können (z.B. Produkte, Ereignisse, Ressourcenanforderungen usw.). Beim Anlegen eines Objektes durch den Aufruf des Konstruktors, wird der für das Objekt benötigte Arbeitsspeicher zeitaufwendig reserviert (Shiraizi 2003, S.101 ff.). Wenn Objekte nicht mehr benötigt werden, sorgt die Java-Laufzeitumgebung dafür, dass zuvor reservierter und nicht mehr benötigter Arbeitsspeicher freigegeben wird. Für die Realisierung dieser sogenannten *Java-Garbage-Collection* stoppt die Umgebung den Programmablauf und setzt

ihn nach Abschluss dieses Prozesses fort. Die *Garbage-Collection* arbeitet zyklisch und kann nur begrenzt vom Programm kontrolliert werden (Flanagan 2003, S.103). Der in OptimSim gewählte Ansatz vermeidet kurzlebige Objekte, indem nicht mehr benötigte Objekte für die spätere Wiederverwendung durch explizites aktiv bleiben reserviert bleiben. Die *Garbage-Collection* wird dadurch wesentlich entlastet und die Anzahl von Konstruktoraufrufen wird reduziert, was einen wesentlichen Performancegewinn erzielt (Shirazi 2003, S.102 ff.). Tabelle 1 zeigt einen Auszug aus erfolgten Geschwindigkeitstests für eine Jahressimulation. Es wurden vier Modelle simuliert, deren Größe durch Produkte, Prozesse, Ressourcen und Ereignisse bestimmt wurde. Aufgezeigte Ergebnisse sind die Zeiten zum Einlesen, Initialisieren, Simulieren und Aufräumen der Modelle und erfüllten unsere Vorstellungen.

*Tabelle 1: Performance Tests*

| | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| # Products | 1 | 7 | 11 | 7 |
| # Tools | 83 | 97 | 73 | 35 |
| # Worker | 32 | 97 | 0 | 0 |
| # Process | 485 | 1875 | 4602 | 352 |
| Parse Time[ms] | 451 | 911 | 1484 | 385 |
| Init [ms] | 1971 | 2704 | 3666 | 1042 |
| Simulation[ms] | 781 | 2058 | 6855 | 361 |
| Clean Up [ms] | 123 | 140 | 544 | 50 |
| Total Time[ms] | 3337 | 5813 | 12549 | 1838 |
| # Events | 306804 | 1170975 | 6332916 | 74761 |
| # Moves / Min | 3974750 | 3837289 | 8604551 | 1485042 |

## 5 Das Eingabemodell

Entsprechend des ganzheitlichen Plug-in-Konzepts ist auch der Import von Modellen durch Plug-ins realisierbar. Es besteht daher keine Beschränkung auf eine bestimmte Modellbeschreibungssprache. OptimSim enthält bereits eine Unterstützung für das Einlesen von SysML- und UML-Modellen. Zusätzlich sind weitere Parser enthalten, die das Einlesen beispielsweise von XML, JSON und XLS unterstützen.

### 5.1 SysML-Modellierung

Mit OptimSim ist es möglich, SysML-Modelle, die mit dem TOPCASED Engineer entwickelt wurden, einzulesen, zu simulieren und zu optimieren. In den diesem Paper zugrundeliegenden Arbeiten wurde SysML als mögliche Sprache für einen Modellierungsstandard in der diskreten Simulation von Produktions- und Logistikprozessen vorgeschlagen (vgl. Schönherr & Rose 2009).

## 5.2 Alternative Möglichkeiten

Das Werkzeug ermöglicht mit der durch Plug-ins erweiterbaren, Modellimportkomponente, Modelle aus anderen Formaten wie zum Beispiel XLS oder XML einzulesen. Das Metamodell und die Import-Schnittstelle des Simulators sind offengelegt. Somit kann der Anwender auch selbst Werkzeuge zur Modellerstellung erstellen und an den Simulator koppeln. Durch den integrierten XLS-Parser können industrielle Systeme komfortabel an den Simulator angekoppelt werden. Durch die so verfügbare Schnittstellenoffenheit ist es zudem möglich, OptimSim an externe ERP-Systeme oder Datenbanken (bspw. ODBC) anzukoppeln. Hybride Ansätze sind umsetzbar: Bei industriellen Projekten ist es möglich, das Metamodell mit SysML zu definieren, und die Daten aus Excel einzulesen.

# 6 Der Optimierer

Wie alle Komponenten von OptimSim, ist die Optimierungskomponente durch Plug-ins erweiterbar. Der Rahmen des Optimierers besteht aus der Definition von Schnittstellen für das zu optimierende Problem, Ausgangskonfigurationen, Zielfunktionen, Bewertungsfunktionen, Heuristiken, Abbruchkriterien, Optimierungsverfahren und Optimierungsergebnisse. Der Aufbau macht es dem Nutzer einfach, eigene Optimierungsverfahren zu entwickeln. Die aktuell enthaltene Optimierungskomponente ist sehr flexibel gehalten und wird in den folgenden Ausführungen beschrieben.

Bereits umgesetzt sind iterative und problemunabhängige Optimierungsverfahren wie unter anderen *Hill-Climbing*, *Simulated-Annealing* und *A*-Suche*. Im Hinblick auf hochdimensionale Parameterräume, wie sie bei der simulationsbasierten Optimierung auftreten können, wurde zudem ein generativer Algorithmus in den Optimierer integriert. Zudem wurde ein Verteilungsverfahren in den Optimierer integriert, das es erlaubt, mehrere Durchläufe parallel auszuführen, so dass die Vorteile von Mehrkern-Systemen ausgenutzt werden können. Um die einzelnen Iterationsschritte des Optimierers nachvollziehen zu können, werden alle Eingabekonfigurationen und deren Simulationsergebnisse gespeichert.

## 6.1 Die Funktionsweise des Optimierers

Die Funktionsweise des in OptimSim integrierten Optimierer-Plug-in ist in Abbildung 3 dargestellt. Grundlegend stößt der Optimierer den Simulator iterativ an Simulationsdurchläufe durchzuführen. Nach jedem Durchlauf erhält der Optimierer die Simulationsergebnisse, wertet diese aus und bestimmt anhand eines Algorithmus Stellschrauben, mit denen ein neues Simulationsmodell erstellt wird, um die nächste Iteration zu starten. Anhand einer Abbruchbedingung wird der iterative Prozess gestoppt. Die Ergebnisse jedes Durchlaufs werden gespeichert.

Der Anwender hat die Möglichkeit, zwischen vorgegebenen Algorithmen wie Simulated-Annealing oder Hill-Climbing zu wählen sowie eigene zu definieren. Zudem kann der Anwender zwischen verschiedenen Erfolgsfaktoren wie Auslastung oder Zykluszeiten wählen. Außerdem hat der Anwender bei den Stellschrauben sehr weitreichende Möglichkeiten, beispielsweise könnte er verschiedene eigene Dispatching-Regeln für Warteschlangen oder Ressourcenzugriff definieren und

diese als Stellschrauben kombinieren. Es wurde möglichst viel Wert darauf gelegt, dem Anwender größtmögliche Freiheit bei der Optimierung zu geben.

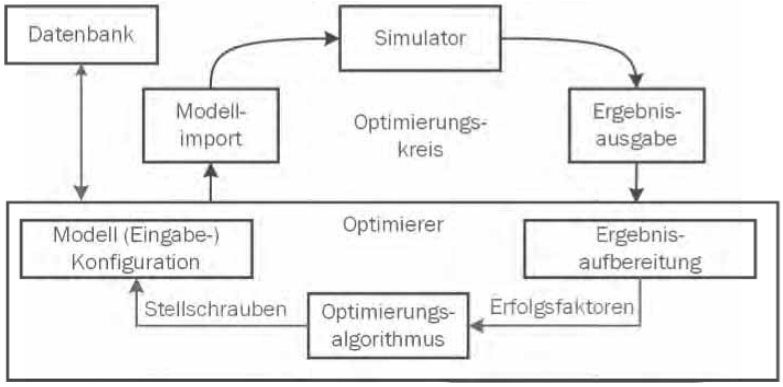



*Abbildung 3: Funktion des bei OptimSim integrierten Optimierers*

## 6.2 Ein integrierter Optimierungsalgorithmus für industrielle Projekte

Da die Optimierungskomponente von OptimSim es ermöglichen soll, industrielle Problemstellungen bei praktischen Projekten zu lösen, wurde ein besonders vielversprechender Algorithmus implementiert, über den im Folgenden ein kurzer Ausblick gegeben wird. Der Algorithmus basiert auf der Kombination von Ansätzen, welche im Rahmen der anwendungsnahen Forschungsprojekte ProMoPs (Völker et al. 2010) und PIP (Fabig et al. 2013) entwickelt wurden. Der Planungsansatz im Forschungsprojekt PIP zielt durch die Berücksichtigung von stochastischen statt deterministischen Prozessgrößen vor allem auf einen robusten Ablaufplan ab. Dadurch kann häufigen Neuplanungen durch Berücksichtigen von Störungen entgegengewirkt werden, zudem werden die Auswirkungen von ungenauen und schwankenden Eingangsdaten berücksichtigt, dafür aber keine Kapazitätsgrenzen. Der ProMoPS-Ansatz arbeitet mit festen Prozessgrößen und versucht unter der Berücksichtigung von Kapazitätsgrenzen, einen kostenminimalen Ablaufplan zu erstellen. Durch die Kombination der beiden Ansätze können die jeweiligen Nachteile (fehlende Berücksichtigung von Kapazitätsgrenzen bzw. störungsanfällige Planung durch feste Prozessgrößen) ausgeglichen werden, um einen kostenminimalen robusten Ablaufplan zu ermitteln.

# 7 Aktueller Stand und Ausblick

Mit OptimSim wird versucht, ein freies Werkzeug für die Wissenschaft bereitzustellen, das dem aktuellen Forschungsstand entspricht und zudem ganzheitlich erweiterbar ist. Aktuell scheitern wissenschaftliche Innovationen oft an den Restriktionen kommerzieller Werkzeuge; für die Umsetzung industrieller Projekte stellen Lizenzkosten oft Barrieren dar. Ein weiteres Problem sind die vielen Modellierungssprachen, die von den verschiedenen Simulatoren im Untersuchungsgegenstand genutzt werden. Durch OptimSim ist die SysML einen Schritt weiter,

sich als Standard für die diskrete Modellierung durchzusetzen. Innerhalb der nächsten Jahre wird OptimSim kontinuierlich weiterentwickelt. Es wird spannend, in wie weit sich OptimSim in universitären Bereich als Simulations- und Optimierungswerkzeug durchsetzen kann. Zudem ist zu hoffen, dass in Projekten die OptimSim nutzen, neue innovative Plug-ins entstehen. Besonders interessant wären Komponenten für die zwei- und dreidimensionale Visualisierung sowie verschiedenste Modellimportkomponenten zur Schnittstellenkopplung mit ERP-Systemen.

## Literaturverzeichnis

Fabig C.; Derno M.; Borkenhagen S.; Schmidt T.; Lasch R.: Planung und Risikobewertung projektspezifischer Produktionsnetzwerke. Logistics Journal, Vol. 2013.

Flanagan, D.; JAVA IN A NUTSHELL. Köln: O'Reilly Verlag 2013.

März, L.; Krug, W.; Rose, O.; Weigert G. Simulation und Optimierung in Produktion und Logistik. Praxisorientierter Leitfaden mit Fallbeispielen. Berlin: Springer 2011.

Schönherr, O.; Rose, O.: First Steps towards a general SysML model for discrete processes in production systems. In: Rossetti, M.D.; Hill, R.R.; Johansson, B.; Dunkin, A.; Ingalls, R.G. (Hrsg.): Proceedings of the 2009 Winter Simulation Conference (WSC). Austin (USA), 13.-16.December 2009, S. 2206-2218.

Schönherr, O.; Rose, O. (2010): Modelling of Production Systems with SysML. In: Zülch G.; Stock P. (Hrsg.): Proceedings of the 2010 ASIM-Fachtagung Simulation in Produktion und Logistik (S. 453-460). Karlsruhe: KIT Scientific Publishing 2010.

Schönherr, O.: Modellierung, Simulation und Transformation von diskreten Prozessen in der Produktion und Logistik auf der Basis von SysML: Dissertation 2014 Verfügbar unter: http://d-nb.info/1052026486/34 [aufgerufen am 26.01.2015]

Shirazi, J.: Java Performance Tuning. Sebastopol, CA. O'Reilly Media Inc 2003.

Starke, G.: Effektive Software-Architekturen, Ein Praktischer Leitfaden. München. Carl Hanser Verlag 2011.

VDI. VDI Richtlinie 3633. 1997. Verfügbar unter: http://www.beuth.de/de/technische-regel-entwurf/vdi-3633/1199512 [aufgerufen am 26.01.2015]

Völker, M.; Rose, O.; Carl, S.: Simulationsbasierte Optimierung der Planung und Steuerung von Montage- und Logistikprozessen in komplexen Montagesystemen. In: Strukturwandel in der Logistik: Wissenschaft und Praxis im Dialog: 5. BVL Wissenschaftssymposium Logistik 2010, S. 143-153.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Verknüpfung von Simulation und Optimierung: Kategorien und Beispiele – Ein Bericht über die VDI Richtlinie 3633 Blatt 12

## *Combination of Simulation and Optimisation: Categories and Examples - A Report of the VDI Standard 3633.12*

Thomas Hanschke, TU Clausthal, Clausthal (Germany),
thomas.hanschke@tu-clausthal.de

Horst Zisgen, SWZ Clausthal-Göttingen, Clausthal (Germany),
horst.zisgen@gmx.de

Abstract: The increasing complexity of manufacturing and logistics requires a combination of simulation and optimisation techniques to come to appropriate business decisions, since besides modelling itself an "optimal choice" of parameters is crucial. This paper summarizes the recent results of the VDI work group "simulation and optimisation" for the VDI standard 3633.12. The focus of this paper is the presentation of the different categories to combine simulation and optimisation which have been developed in the work group. Furthermore, examples of concrete business decisions are given to illustrate these categories and to show how a combination of simulation and optimisation could work in practice.

## 1 Einleitung

Die kombinierte Verwendung von Simulations- und Optimierungsverfahren wird bei der Beantwortung von Fragestellungen in der Produktion und Logistik zunehmend wichtiger, da zum einen die Modellierung eines dynamischen Prozesses und zum anderen Verfahren zur Suche von optimalen Lösungen benötigt werden. Eine solche Verknüpfung bietet die Möglichkeit, die u. a. von Davenport und Harris (2007) in ihrem Buch „Competing on Analytics" als getrennt betrachteten Verfahren „Predictive Modeling" und „Optimization" nutzbringend zu kombinieren (Abb. 1).

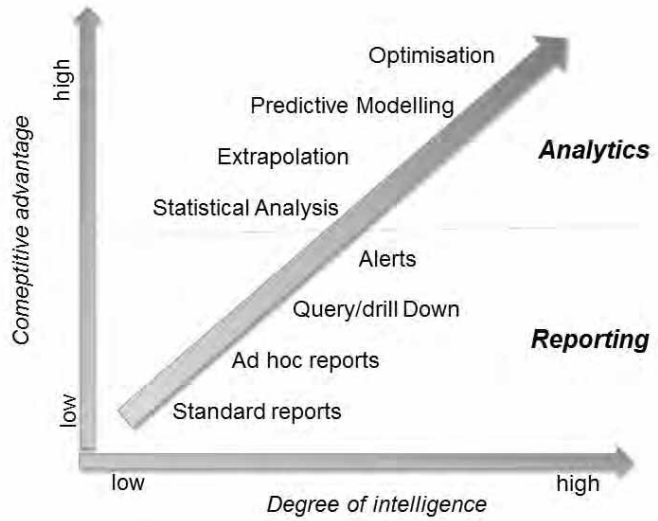



*Abbildung 1: Einteilung analytischer Verfahren nach Davenport und Harris (2007)*

Eine Form des von Davenport und Harris als „Predictive Modeling" bezeichneten Ansatzes ist die Simulation im Sinne, wie sie in der VDI Richtlinie 3633, Blatt1, definiert ist. Dort wird Simulation als das „Nachbilden eines dynamischen Prozesses in einem System mit Hilfe eines experimentierfähigen Modells, um zu Erkenntnissen zu gelangen, die auf die Wirklichkeit übertragbar sind" definiert. Somit steht die Entscheidungsunterstützung durch Simulationsexperimente an zentraler Stelle. Diese Entscheidungsunterstützung beruht aber bei einem Simulationsexperiment in erster Linie auf der Fähigkeit der Simulation, den Systemverlauf, und damit die Leistungskennzahlen des Systems, unter der Vorgabe von Modellparametern, vorherzusagen. Damit gilt es zuerst einmal im Rahmen der Simulation eine möglichst gute Modellbildung zu erreichen, damit die Aussagen über das mögliche Systemverhalten die Realität möglichst gut widerspiegeln. Eine detaillierte Einführung in die dazu benötigten Techniken finden sich z. B. in den Handbüchern von Banks (1998) und von Law und Kelton (2000).

Während die Ausführung eines Simulationslaufs eine Bewertung der gewählten Parametrierung liefert, ermöglicht die Optimierung die Auswahl des bestmöglichen Lösungsvorschlages aus einer Menge von Alternativen mit Hilfe algorithmischer Verfahren.

Durch eine Kopplung beider Ansätze kann man die Optimierung der Leistungskennzahlen von Systemen durch die systematische Variation von Modellparametern im Simulationsmodell erreichen.

Um Anwendern eine Hilfestellung an die Hand zu geben, in welcher Form die Verknüpfung von Simulation und Optimierung in der Praxis umgesetzt werden kann und wie der Anwender dadurch zu optimalen oder zumindest zu möglichst guten

Entscheidungen gelangt, hat sich der VDI Fachausschuss „Simulation und Optimierung“ mit dieser Fragestellung auseinander gesetzt und eine Richtlinie zum Thema „Simulation und Optimierung“ erarbeitet.

Dieser Beitrag stellt eine grobe Übersicht über die in dieser Richtlinie erarbeiteten Konzepte und Vorgehensmodelle dar. Dazu werden zum einen die im Rahmen der Richtlinienarbeit erstellte Kategorisierung von möglichen Arten der Kopplung von Simulation und Optimierung vorgestellt und in den Kontext der VDI Richtlinie 3633 zur Simulation gestellt. Zum anderen werden verschiedene Praxisbeispiele gezeigt, welche die unterschiedlichen Arten der Verknüpfung illustrieren. Die Beispiele kommen aus verschiedenen Bereichen der Produktion und Logistik und sind aus ganz unterschiedlichen Branchen hervorgegangen.

## 2 Kategorisierung

Die Kategorisierung der Kopplung von Simulation und Optimierung dient dazu, die unterschiedlichen Aufgaben der Simulation und der Optimierung im Zusammenspiel beider zu systematisieren. Diese unterschiedlichen Rollen ergeben sich aus den unterschiedlichen Fragestellungen, die mit Hilfe der kombinierten Simulation und Optimierung in der betrieblichen Praxis gelöst werden sollen. Letztlich dient die Kategorisierung auch dazu, die entsprechenden Vorgehensweisen und Implementierungen in der Praxis entsprechend zu planen und umzusetzen.

Grundsätzlich unterscheidet die Richtlinie in zwei Arten der Kopplung:

- Sequentielle Kopplung
- Hierarchische Kopplung

### 2.1 Sequentielle Kopplung

Bei der sequentiellen Kopplung werden das Simulationsexperiment und die Optimierung als zwei eigenständige und getrennt voneinander ablaufende Schritte betrachtet und durchgeführt. Dabei liefert das eine Verfahren jeweils die Eingangsparameter für das nachfolgende Verfahren.

#### *2.1.1 Simulation folgt der Optimierung*

Dieser Ansatz findet immer dann Anwendung, wenn es für die Optimierungsaufgabe zwar einen mathematischen Ansatz gibt, der aber auf Grund der Komplexität des zu optimierenden Systems nicht alle Facetten der Realität abbilden kann. Hier ist es oft sinnvoll, zuerst einmal mit einem stark abstrahierenden Optimierungsmodell zu arbeiten, um erste Verbesserungsvorschläge zu erzielen. Dabei kommt es in der betrieblichen Praxis nicht auf die mathematisch exakte Lösung an, sondern in der Praxis ist eine deutliche Verbesserung gegenüber dem Ist-Zustand durchaus ein Zugewinn. Die durch die Optimierung gewonnenen Lösungsvorschläge müssen noch auf die Auswirkung auf das Gesamtsystem hin untersucht werden, da die Optimierung nur ein stark vereinfachtes Modell verwendet. Dazu kann dann in einem zweiten Schritt ein Simulationsmodell, das eine genauere Abbildung der Realität ermöglicht, herangezogen werden. Mit Hilfe dieses Simulationsmodells können die von der Optimierung vorgeschlagenen Verbesserungen nachgespielt und überprüft werden.

*Beispiel: Los-Verarbeitung*

Arbeitsstationen mit Stapelverarbeitung spielen in der Halbleiterindustrie eine herausragende Rolle (Zisgen et al. 2008). Diese Systeme sind dadurch gekennzeichnet, dass der Bediener erst dann tätig wird, wenn eine feste Größe *b* von Aufträgen angekommen ist (Hanschke (2006) und Zisgen (2009)). Sind *b* Aufträge vorhanden, werden sie gleichzeitig und für die Dauer *E[S]* bedient. Ist die Auslastung des Systems gering, braucht es in der Regel auch eine gewisse Zeit, bis die vorgeschriebene Beladungsgröße *b* zustande gekommen ist. Umgekehrt wird man bei hoher Auslastung eine lange Wartezeit beobachten. Zwischen diesen beiden Extrembereichen gibt es offensichtlich einen Auslastungsgrad, der die Wartezeit in der Warteschlange minimiert. Eine wichtige Aufgabe der Produktionsplanung ist es, diesen Punkt zu bestimmen. Leider ist diese Fragestellung einer exakten mathematischen Lösung nicht zugänglich. Allerdings existiert eine Näherungsformel (Bolch et. al. 2006), die den Zusammenhang zwischen mittlerer Wartezeit, Produktionsrate, Losgröße und Bearbeitungszeit aufschlüsselt (Abb.2). Diese Formel dient im Optimierungsschritt dazu, die Produktionsrate zu ermitteln, bei der die Wartezeit minimal ist.

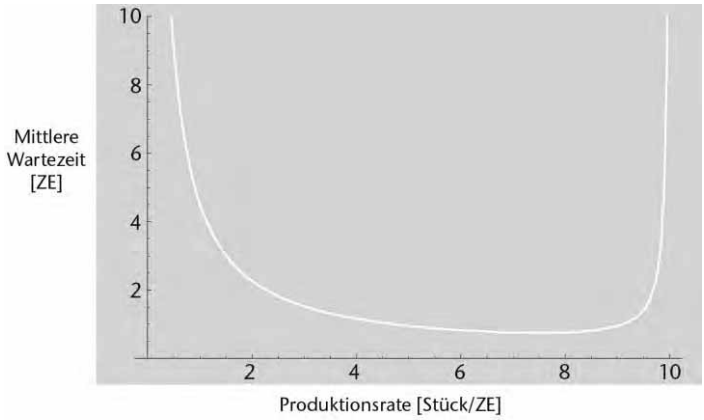



*Abbildung 2: Mittlere Wartezeit in Abhängigkeit von der Produktionsrate bei Batch-Verarbeitung*

Dieser Wert für die Produktionsrate wird als Startlösung in dem zugehörigen Simulationsmodell verwendet. Durch Betrachtung der benachbarten Werte kann auf diese Weise nun sukzessive die optimale Lösung des realen Problems angenähert werden.

### 2.1.2 Optimierung folgt der Simulation

Umgekehrt kann es in der Praxis schwierig sein, die Eingangsparameter für ein mathematisches Optimierungsmodell korrekt auszuwählen, da diese Werte oft nicht bekannt oder nur sehr schwierig zu ermitteln sind. In diesem Fall ist eine Kopplung in der umgekehrten Reihenfolge sinnvoll. So kann mit Hilfe der Simulation zuerst die für die Optimierung benötigten Eingangsparameter bestimmt und im Nachgang mittels Optimierungsverfahren die Verbesserung der Leistungskennzahlen erzielt werden, wie das folgende Beispiel aus der Luftfahrtindustrie zeigt:

*Beispiel: Revenue Management*

Fluggesellschaften bieten kein einheitliches Produkt an, sondern differenzieren ihre Produkte entsprechend den Bedürfnissen der verschiedenen Kunden. Grundsätzlich werden diese in Privat- und Geschäftsreisende unterteilt. Diese beiden Gruppen haben verschiedene Präferenzen, z. B. in Bezug auf den Zeitpunkt der Buchung, den Service und den Preis.

Die unterschiedlichen Anforderungen der Kunden erfordern spezifische Angebote. Um diesen gerecht zu werden, existieren Einteilungen nach Beförderungsklassen und Buchungsklassen. Erstere untergliedern das Flugzeug physisch in First, Business und Economy Class. Diese Einteilung legt den Komfort, den Service und die Ausstattung fest. Jede Beförderungsklasse unterteilt sich in Buchungsklassen. Die Buchungsklassen unterscheiden sich durch sogenannte Konditionen, welche überwiegend Einschränkungen in der zeitlichen Gültigkeit des Tickets bedeuten (Vorausbuchungsfrist und Umbuchungsoptionen, Minimal- bzw. Maximalaufenthaltsdauer, Wochenendbindung).

Die sich aus dieser Produktdifferenzierung ergebenden Leistungsbündel werden zu unterschiedlichen Preisen angeboten. Da die Sitzplatzkapazität eines Flugzeuges fix ist, stellt sich an diesem Punkt die Frage, welche Menge der jeweiligen Produkte angeboten werden sollen, um den maximalen Ertrag zu erzielen.

Zu dieser Problemstellung existiert eine Vielzahl von Optimierungsverfahren, welche aber als Eingangsparameter Annahmen über das Buchungsverhalten der Fluggäste benötigen. Um diese Eingangsparameter zu erhalten, werden zuerst anhand historischer Buchungsdaten von Fluggesellschaften mittels stochastischer Simulation realistische Nachfrageströme erzeugt, welche als Eingabe für die verschiedenen Optimierungsverfahren dienen.

## 2.2 Hierarchische Kopplung

Im Gegensatz zu einer sequentiellen Verknüpfung erfolgt bei einer hierarchischen Kopplung eine Integration des einen Verfahrens in das jeweils andere.

### *2.2.1 Optimierung in die Simulation integriert*

Im Fall, in dem die Optimierung in die Simulation integriert ist, wird das Simulationsexperiment an Stellen angehalten, an denen eine Entscheidung zu treffen ist, zugunsten welcher Alternative der Verlauf des Simulationsexperiments fortgesetzt werden soll. Dabei dienen dann die Systemzustände des Simulationsexperiments zum Entscheidungszeitpunkt als Eingangsparameter für das Optimierungsmodell. Das Ergebnis der Optimierung wiederum dient der Simulation als Entscheidungsparameter, um den weiteren Verlauf des Simulationsexperiments festzulegen. Bei der Integration der Optimierung in die Simulation ist darauf zu achten, dass das Optimierungsverfahren im Wesentlichen zwei grundsätzliche Eigenschaften besitzt. Zum einen muss das Optimierungsverfahren einen kontinuierlichen Zustrom an Daten erlauben und darf nur solche Daten für die Optimierung verwenden, die zum Simulationszeitpunkt (an dem die Entscheidung ansteht) bekannt sind. Und zum anderen müssen die Optimierungsverfahren hinreichend schnell Ergebnisse liefern können, um das Simulationsexperiment nicht merklich zu verzögern.

*Beispiel: Optimierung einer Walzstraße (Busch 2005)*

Das Beispiel behandelt die Reihenfolgeplanung in Walzstraßen. Die Warmzone eines Walzwerkes umfasst die Einrichtungen zur Wärmebehandlung (kontinuierlich betriebene Herdwagenöfen, in Losgröße eins betriebene Stoßöfen und Einlegeofen), zum Walzen (Vorgerüst und Fertiggerüst) und zum Abkühlen (Kühlstrecke und Kühlbett) von Walzerzeugnissen (Abb. 3). Zur Optimierung des Fertigungsprozesses innerhalb der Warmzone ergeben sich verschiedene Zielsetzungen für die einzelnen Abschnitte, die letztlich zur Minimierung der Prozesszeit einer Walzfolge führen sollen.

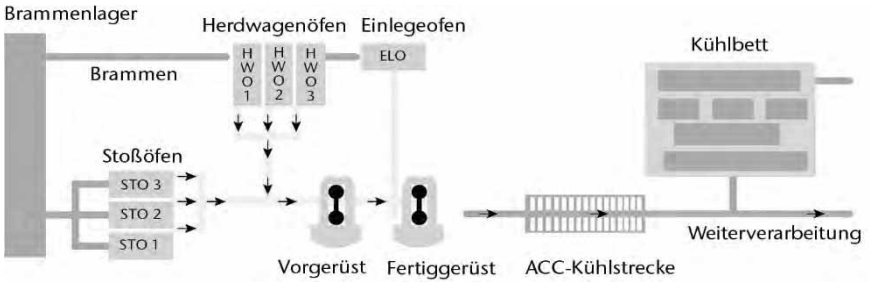



*Abbildung 3: Fertigungsfluss der Warmzone einer Walzstraße*

Die Zeitpunkte, zu denen die Brammen aus den Öfen gezogen werden, werden Ziehzeiten genannt. Zur Minimierung der Prozesszeit einer Walzfolge genügt es, die Abstände zwischen den Ziehzeiten zu minimieren. Im ersten Schritt wird anhand der technologischen Restriktionen eine bezüglich Prozesszeit und Qualität optimale Ausgangs-Walzreihenfolge mittels eines stochastischen Algorithmus bestimmt. Für diese Walzfolge werden dann die optimalen Ziehzeitpunkte und die durch die Stoßöfen und das Kühlbett entstehenden Ziehzeitverlängerungen mittels gemischt-ganzzahliger linearer Optimierung bestimmt. Anhand der Offline-Daten werden in einem dritten Schritt die optimalen Einfügepositionen der Brammen aus den Herdwagenöfen bzw. dem Einlegeofen mittels Vollständiger Enumeration über zuvor festgelegte Zeitintervalle berechnet. Nach diesen drei Optimierungsschritten steht die optimale Reihenfolge aller Brammen fest. Die exakten Ziehzeitpunkte und Haltezeiten der Brammen während der laufenden Walzung müssen jedoch mittels Simulation ständig an den aktuellen Prozessverlauf angepasst werden, um auf zufällige und unberücksichtigte Prozessschwankungen zu reagieren.

### 2.2.2 Simulation in die Optimierung integriert

Im umgekehrten Fall, die Simulation ist in die Optimierung integriert, dient die Simulation als Ersatz für eine funktionale Beschreibung der Zielfunktion innerhalb der Optimierung. Damit ist es auch in praktischen Problemstellungen, in denen sich keine funktionale Beschreibung der Zielfunktion formulieren lässt, möglich, Optimierungsverfahren einzusetzen. Dazu wird das zu optimierende System als Simulationsmodell abgebildet. Dabei sind die Freiheitsgrade der Optimierung veränderliche Parameter des Simulationsmodells. Gleichzeitig muss das Modell so aufgebaut sein, dass es die zu optimierenden Zielgrößen ermitteln kann. Nach einem initialen Simulationslauf werden die Leistungsgrößen ermittelt und anschließend

wird die Simulation mit neuen Startwerten durchgeführt. Die Bestimmung der neuen Startwerte erfolgt durch Suchalgorithmen und soll gegenüber den vorangegangenen Simulationsläufen eine Verbesserung der zu optimierenden Leistungsgrößen erwirken. Hierzu eignen sich insbesondere Meta-Heuristiken zu Bestimmung der neuen Startwerte. Dieses Verfahren wendet man so lange an, bis sich keine Verbesserung der Zielgrößen mehr einstellt.

*Beispiel: Allokation von Pufferplätzen*

Bei der optimalen Allokation von Puffern zu aufeinander folgenden Arbeitsstationen besteht allgemein das Problem darin, dass der Zusammenhang zwischen Durchsatz und Anzahl bzw. Zuordnung der Puffer nicht mehr mathematisch aufgelöst werden kann, d.h. auch nicht als Zielfunktion in einer Optimierung formuliert werden kann. Eine Bewertung, wie sich die Zuordnung insgesamt auf den Durchsatz auswirkt, kann dann z. B. durch eine Simulation erfolgen. Bevor es aber zu einer Bewertung kommen kann, werden mittels eines stochastischen Suchverfahrens sukzessive geeignete Lösungsvorschläge zur Pufferbelegung ausgesucht und anschließend die erzielte Verbesserung hinsichtlich des Durchsatzes mittels Simulation ermittelt.

## 3 Zusammenfassung

Die hier vorgestellte und in der VDI Richtlinie noch detaillierter ausgeführte Kategorisierung der Verknüpfung von Simulation und Optimierung bietet eine Hilfestellung für den praktischen Einsatz beider Verfahren zur Lösung betriebswirtschaftlicher Fragestellungen, insbesondere in der Produktion und Logistik. Weitere Beschreibungen und Fallbeispiele finden sich zu diesem Thema in März et. al. (2011).

Insbesondere die Kopplung über eine Integration der Simulation in die Optimierung findet sich in der englischsprachigen Literatur als *„simulation optimization“* wieder. So wurde dieser Begriff etwa schon von Carson und Maria (1997) verwendet. Weitere Literatur zu diesem Thema, insbesondere auch zur Integration dieser Verfahren in Simulationssoftware findet sich in Chau et. al. (2014) und einen aktuellen Überblick zur *„simulation optimization“* liefern Fu et. al. (2014).

## Literatur

Banks, J.: Handbook of simulation: Principles, methodology, advances, applications, and practice. New York: John Wiley & Sons 1998.

Bolch, G.; Greiner, S.; de Meer, H.; Trivedi, K.S.: Queueing Networks and Markov Chains, 2. Aufl. Hoboken (NJ): Wiley Interscience 2006.

Busch, H.: Stochastische Optimierungsverfahren zur Produktionssteuerung. Dissertation , TU Clausthal 2005.

Carlson, J.; Maria, A.: Simulation optimization: Methods and applications. In: Andradottir, S.; Healy, K.J.; Withers, D.H.; Nelson, B.L. (Hrsg.): Proceedings of the 1997 Winter Simulation Conference. Piscataway: IEEE 1997, S. 118-126.

Chau, M.; Fu, M.C.; Qu, H.; Ryzhov, I.O.: Simulation optimization: A tutorial overview and recent developments in gradient-based methods. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): Proceedings of the 2014 Winter Simulation Conference. Piscataway: IEEE 2014, S. 21-34.

Davenport, T.G.; Harris, J.G.: Competing on analytics. Harvard Business School Publishing Cooperation 2007.

Fu, M.C.; Bayraksan, G.; Henderson, S.G.; Nelson, B.L.; Powell, W.B.; Ryzhov, I.O.; Thengvall, B.:Simulation optimization: A panel of the state of the art in research and practice. In Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): Proceedings of the 2014 Winter Simulation Conference. Piscataway: IEEE 2014, S. 3696-3706.

Hanschke, T.: Approximations for the mean queue length for the $GI^X/G^{(b,b)}/c$ queue. Operations Research Letters 34 (2006) 2, S. 205-213.

Law, M.; Kelton, W.: Simulation Modeling and Analysis, 3. Aufl. New York: McGraw-Hill 2000.

März, L.; Krug, W.; Rose, O.; Weigert, G. (Hrsg.): Simulation und Optimierung in Produktion und Logistik – Praxisorientierter Leitfaden mit Fallbeispielen. Heidelberg u.a.: Springer: 2011.

VDI Richtlinie 3633, Blatt 1: Simulation von Logistik-, Materialfluss- und Produktionssystemen – Grundlagen. Berlin: Beuth 2014.

Zisgen, H.; Meents, I.; Wheeler, B.R.; Hanschke, T.: A queueing network based system to model capacity and cycle time for semiconductor fabrication. In: Manson, S.J.; Hill, R.R.; Mönch, L.; Rose, O.; Jefferson, T.; Fowler, J.W. (Hrsg.): Proceedings of the 2008 Winter Simulation Conference. Picataway: IEEE 2008, S. 2067-2074.

Zisgen, H.: A diffusion approximation for batch queues of type $GI^X/G(b,b)/1$. Performance Evaluation 66 (2009) 10, S. 557-563.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Inhouse-Logistikplanung unter Verwendung von Ansätzen aus der Verkehrsflusssimulation

## *Amending Inhouse Logistic Planning by Merging with Traffic Flow Simulation Capabilities*

Franz-Joseph König, ZIP Ingenieurbüro Industrieplanung und Organisation, München (Germany), Franz-Josef.Koenig@ZIP.de

Dr.-Ing. Stefan Roth, Volkswagen AG, Wolfsburg (Germany), Stefan.Roth@volkswagen.de

Marek Heinrich, Deutsches Zentrum für Luft- und Raumfahrt e.V., Berlin (Germany), Marek.Heinrich@dlr.de

Abstract: Production systems rely on efficient and robust intra-logistic transport systems. The ongoing change of premises and methodologies in the design urges planners to validate the planning results systematically. Industrial applications, such as MALAGA, do support the planners of intra-logistic systems by flanking their existing workflow with dynamic validation capabilities based on analyses performed with multiple simulation applications. This paper outlines how the dynamic validation of logistic transport systems during the design phase is realized within the Digital Factory of Volkswagen. Strong emphasis is on the need for simulation in general and in particular on the vehicle movement and traffic flow simulation. The deployed dynamic approach allows for investigating the impact of vehicle congestions on the reliability of a intra-logistic system. Finally, the paper outlines the benefits of the combined geometrical and dynamical validation in the context of the digital factory.

## 1 Logistik- und Materialflussplanung

Kalkulatorische Planungsansätze befriedigten in der Vergangenheit die wesentlichen Bedürfnisse der Logistikplanung. Das Hauptmerkmal solcher Ansätze ist eine systematische Erfassung planungsrelevanter Daten, auf deren Basis logistische Kenngrößen bestimmt wurden.

Kalkulatorischen Ansätzen liegt die Dokumentation von Lösungsansätzen und darauf basierend die Berechnung von Mengengerüsten zu Grunde. Problematisch dabei ist die Tatsache, dass eine derartige Methodik enorme Freiheitgrade liefert und damit auch Lösungsräume nahezu unendlich groß werden können. Restriktionen existieren kaum und die berechneten Ergebnisse hängen maßgeblich von der manu-

ellen Parametrisierung des Modells ab. Im Zuge der Parametrisierung können zufällige Fehler gemacht und widersprüchliche Annahmen getroffen werden. Eine Plausibilisierung von Kenngrößen und Mengengerüsten ist nur bedingt möglich.

Logistik ist in Deutschland häufig noch Vertiefungsfach betriebswirtschaftlicher Studiengänge. Excel-basierte Ansätze sind nach wie vor „state of the practice“ in der Logistikplanung. Eine konkrete Anlehnung an CAD-Layouts spielt heute eine noch untergeordnete Rolle. Materialfluss-Layouts werden heutzutage exemplarisch im Zuge der Überführung der Kalkulationsbasis in ein Simulationsmodell implementiert. Die Übergabe an den Simulationsexperten durch den Logistikexperten erfolgt zu einem späten Zeitpunkt mit dem Ziel einer simulativen Absicherung mit Hilfe der Methode der Ablaufsimulation (Mayer 2013). Abbildung 1 zeigt exemplarisch einen etablierten Planungsprozess, in dem der Simulationsexperte die Absicherung eines Planungsstandes verantwortet.

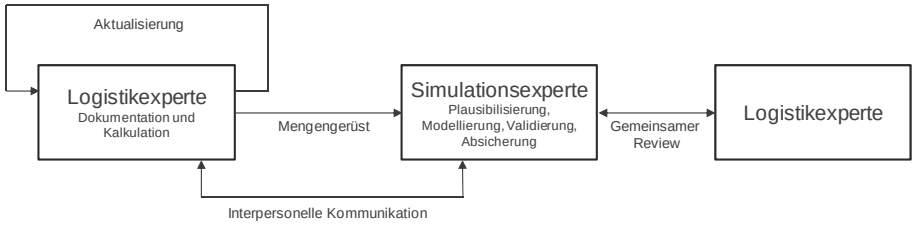



*Abbildung 1: Zusammenarbeitsmodell Logistik-/Simulationsexperte heute*

## 2 Digitale Fabrik von Volkswagen und neue Akzente

Im Gegensatz dazu kann die Digitale Fabrik von Volkswagen auf wesentliche neue Entwicklungen verweisen. In der VW-Markenplanung kombinieren die Logistikplaner bereits seit vielen Jahren die logistischen Prozessbeschreibungen mit den Materialfluss- und Fertigungslayouts der Fabrik. Die neuen Möglichkeiten – zumindest in der Inhouse-Logistik – sind enorm. Die Inhouse-Logistik fokussiert alle logistischen Prozesse, Operationen, Ressourcen und Kapazitäten, die auf dem Werksgelände zur Versorgung einer oder mehrerer Fertigungssysteme erforderlich sind (Abb. 2).

*Abbildung 2: Unterscheidung Inbound- und Inhouse-Logistik*

Die Digitale Fabrik verändert das Selbstverständnis der Logistikplanung. Die Verortung von Logistikoperation, Behältern und Baugruppen in der virtuellen Fabrik liefert die Basis für die Datendurchgängigkeit, die zur Validierung eines Planungsstandes erforderlich ist. Dem Logistikexperten ist es jetzt möglich, neben seiner

Produkt-, Prozess- und Ressourcenbeschreibung auch konsistente geometrische Ausprägungen wie Entfernungsbeziehungen zu etablieren. Auf dieser Basis ist unter anderem die Ableitung von Simulationsmodellen möglich.

## 3 Notwendigkeiten für neue, virtuelle Methoden

Die physischen Ausprägungen der logistischen Systeme in der Automobilindustrie sind seit jeher vielschichtig und komplex. Diese Situation hat sich mit der Einführung neuer Inhouse-Logistikkonzepte weiter verschärft. Die Konzepte einer Fließlogistik führen zu neuen Abhängigkeiten und zwingen den Logistikplaner zu einer ganzheitlichen systemischen Betrachtungsweise:

- Bauteile und Komponenten werden immer häufiger in Warenkörben oder Sequenzbehältern bereitgestellt. Dies ist eine Methode, um den stetig zunehmenden Produktvarianten Rechnung zu tragen und um platzsparende Anstellkonzepte realisieren zu können.
- Damit verbunden sind sinkende Materialreichweiten an den Verbauorten, was wiederum zu kürzeren Wiederbeschaffungszyklen und zu einer Zunahme der Transportaktivitäten führt.
- Lager- und Kommissionierflächen innerhalb der Werksgrenzen werden schrittweise zurückgebaut und in Produktionsflächen überführt. Es verringern sich Reichweiten und Wiederbeschaffungszyklen auch in den vorgelagerten Logistikbereichen. Die Transporthäufigkeit und damit das Verkehrsvolumen auf dem Wegenetz nehmen weiter zu.
- Die Allokation von Bauteilen und Komponenten auf Lager- und Kommissionierflächen muss unter Berücksichtigung der Transport- und Routenplanung erfolgen.
- Neue Transportkonzepte mit getakteten Routenverkehren und streng determinierten Fahrplänen kommen zum Einsatz (staplerfreie Fertigung).
- Transporte, Fahrkreise und Touren werden unter Berücksichtigung des gesamten Bauteilspektrums im Verbund ausgeplant.

Daraus folgt, dass in der Planungsphase kaum mehr lokale sondern globale Optima zu erwarten sind. Punktuelle Änderungen in der Produkt- oder Prozessinformation oder in der technologischen Ausprägung des Sollsystems führen meist zu einer notwendigen Anpassung des Gesamtsystems. Zu berücksichtigen ist, dass Änderungen häufig durch neue Planungsstände der Fertigungsplanung oder der Fabrikstruktur initiiert werden. Der Logistikexperte muss bei adäquater Detailierung seines Modells in der Lage sein, solche Änderungen zu bewerten.

## 4 Das Logistik-Planungssystem MALAGA

Die in Kapitel 3 angeführten veränderten Planungsprämissen erfordern den Einsatz angemessener Software-Tools zur Unterstützung des Logistikexperten. Die Marke VW und andere beteiligen sich seit 2003 aktiv an der Entwicklung des Inhouse-Logistikplanungssystems MALAGA (*Ma*terialfluss *La*yout und *g*raphische *A*uswertung) der Firma ZIP Industrieplanung, München. Seit 2012 investiert auch der Volkswagen-Konzern insgesamt in diese Entwicklung. Das Ergebnis ist eine Software-Lösung, die alle wesentlichen Anforderungen und Erfordernisse eines Automobilherstellers trifft. Maßgeblich für die Entwicklung war und ist der Grund-

satz, alle logistisch relevanten Entscheidungen am Fabriklayout zu spiegeln (Bracht 2008). Daher ist MALAGA voll in das CAD-Programm MicroStation (Bentley Systems) integriert. Seit 2007 wurden von der VW-Markenplanung zahlreiche Greenfield- und Brownfield-Projekte im In- und Ausland mit Hilfe des Systems MALAGA durchgeführt.

Als Eingangsdaten nutzt MALAGA um Logistikinformationen (Abb. 3 links) angereicherte Stücklisten. Das Einlesen manuell aufbereiteter Excel-Daten ist ebenso wie die Anbindung kundenspezifischer Hostsysteme möglich.

Der Schwerpunkt des Programms liegt in der logistischen Prozessmodellierung. In Form von generalisierten Prozessbeschreibungen werden sowohl der Vollgut- als auch der Leergutfluss MTM-konform modelliert. Die Prozesse sind als Graphen dargestellt und können in Bibliotheken verwaltet werden.

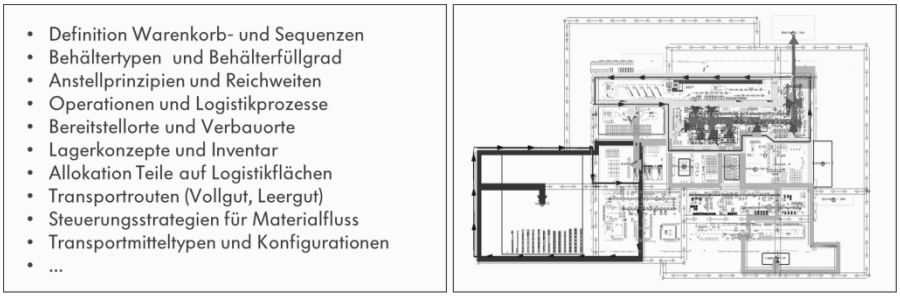



*Abbildung 3: Logistikinformationen und Materialflusslayout*

Die Aufprägung der Eingangsdaten auf die generalisierten Prozessbeschreibungen (Parametrisierung des Modells) liefert anschließend für jedes Bauteil in der Stückliste den konkreten Prozess. Die Darstellung des Materialflusses erfolgt im CAD-Layout in Form von Materialflusspfeilen (Abb. 3 rechts). Die Breite der Pfeile entspricht dabei der Häufigkeit der Transporte.

Im CAD-Layout hat der Logistikplaner Zugriff auf Strukturinformationen. Damit ist er in der Lage, das Wegenetz maßstabsgetreu nach diesen graphischen Vorgaben zu modellieren. Die Transportrelationen zwischen einzelnen Fertigungsstationen ermittelt MALAGA automatisch unter Berücksichtigung von Einschränkungen wie Einbahnstraßen oder Beschränkung auf bestimmte Transportmitteltypen (Stapler, Mitarbeiter, Routenzüge etc.).

Eines der wesentlichen Leistungsmerkmale von MALAGA ist eine Funktion zur Optimierung des getakteten Routenverkehrs. Ein von der Fa. Inpro GmbH, Berlin entwickelter Algorithmus ermittelt automatisch optimierte Routen auf Basis der Prozess- und Layout-Information (Dreher 2009). Der Verlauf der Routen wird direkt in das CAD übertragen. Dem Planer werden in MALAGA alle relevanten Ergebnisse wie Zykluszeiten, Fahrzeiten, durchschnittliche Auslastung etc. angezeigt. Anknüpfend an die Routenoptimierung kann unter Berücksichtigung von z. B. Pausen, Kapazitäten und Batterieladezyklen ein streng determinierter Fahrplan generiert werden.

Die Planungsprämissen einer Fließlogistik implizieren stets eine stabile Produktion. Unter der Maßgabe robuster Logistikprozesse und der Versorgungssicherheit versteht es sich von selbst, dass ein zeitdiskretes Verhalten des logistischen Systems sowie die Bewertung von Störungen an Wichtigkeit zunehmen (Juan 2013). Neben der Robustheit und Versorgungssicherheit gibt es zusätzliche Notwendigkeiten, die für eine zeitdiskrete Simulation des Systems sprechen:

- Wechselwirkungen und systemische Effekte sind in der statischen Modellierung nur eingeschränkt abbildbar.
- Aufgrund sinkender Bestände steigen die Transporthäufigkeit und damit die Verkehrsbelastung im Wegenetz. Die Wahrscheinlichkeit der gegenseitigen Beeinflussung von Transportmitteln (Stauung) nimmt zu.
- Wiederbeschaffungszeiten sind maßgeblich abhängig von der Auftragssteuerung. Nur die Simulation gestattet eine Berücksichtigung der Auftragssteuerung.
- Aufgrund sinkender Bestände in den Lagern muss eine Gegenüberstellung der Reichweiten und Wiederbeschaffungszeiten im gestörten System erfolgen.
- Den tatsächlichen Anteil an Leerfahrten liefert nur ein Simulationsexperiment.
- Unterschiedliche Frequenzen pro Sachnummer bei der Ein- und Auslagerung führen zu dynamischen Spitzen im Lagerverlauf. Diese Effekte können nicht immer vernachlässigt werden.
- Eine Kalkulation geht von unendlich großen Kapazitäten für den Materialumschlag aus. Eine zeitliche Belegung beschränkter Kapazitäten kann nur simulativ berücksichtigt werden.

Die hieraus resultierenden Simulationsanforderungen wurden im System MALAGA frühzeitig berücksichtigt. Ansatz und Fokus hierbei war und ist stets, dem Logistikexperten eine dynamische Betrachtung des zu planenden Logistiksystems ohne spezielle Softwarekenntnisse zu ermöglichen. Als Simulationskern wird die Software Plant Simulation von Siemens/Tecnomatix verwendet.

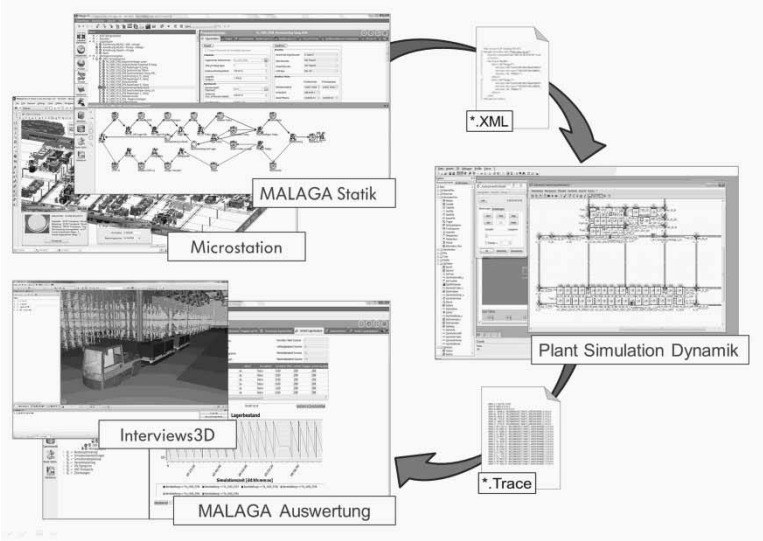



*Abbildung 4: Informationsfluss zwischen MALAGA und Plant Simulation*

Alle simulationsrelevanten Daten werden in MALAGA generiert und als XML-File exportiert. Sowohl der Aufruf als auch Aufbau und Initialisierung des Netzwerks erfolgen automatisch. Parallel zur Laufzeit schreibt die Dynamik Auswertungsdateien. MALAGA interpretiert diese Trace-Dateien und stellt dem Anwender die Ergebnisse in seiner gewohnten Planungsumgebung zur Verfügung (Abb. 4). MALAGA bietet die Möglichkeit zum Vergleich der Ergebnisse aus der statischen Kalkulation mit denen des Simulationsexperiments.

Bei wiederholter dynamischer Validierung und Korrektur des statischen Ausgangsmodelles können Konflikte frühzeitig erkannt und gelöst werden, was letztendlich zu einem robust ausgelegten statischen Modell des logistischen Gesamtsystems führt.

## 5 Motivation für eine integrierte Stausimulation

Aus den im Kapitel 4 formulierten Gründen konnte bereits aufgezeigt werden, dass die Transporthäufigkeit zunimmt und die Notwendigkeit zur dynamischen Validierung des statischen Modells gegeben ist. Die stärkere Belastung des Wegenetzes führt aber auch zu einer zunehmenden Wahrscheinlichkeit von gegenseitiger Beeinflussung und Behinderung der Transportmittel. Die daraus resultierenden dynamisch bedingten Verzögerungen ergeben folglich eine Verlängerung der durchschnittlichen Transportdauer. Dies führt zu einer Erhöhung der Wiederbeschaffungszeit und kann im Einzelfall einen Materialabriss verursachen. Insbesondere die Teilversorgung über Routenzugsysteme erfordert eine zuverlässige Konstanz in den Transportzeiten selbst unter veränderlichem Verkehrsaufkommen.

Die Simulationsstrategie muss gestaute Transportnetze daher zwingend mitberücksichtigen, Staueffekte abbilden und dem Logistikexperten plausible Anhaltspunkte für eine mögliche Problematik sowie Lösungsansätze aufzeigen.

Die minimale bzw. optimale und verzögerungsfreie Transportdauer ergibt sich aus der Fahrtdauer im unbelasteten Netz, also der Zeit, die ein Fahrzeug idealerweise benötigt um eine Tour, unter Einhaltung aller Regeln, abzuschließen, jedoch ohne durch andere Verkehrsteilnehmer behindert zu werden (Abb. 5).

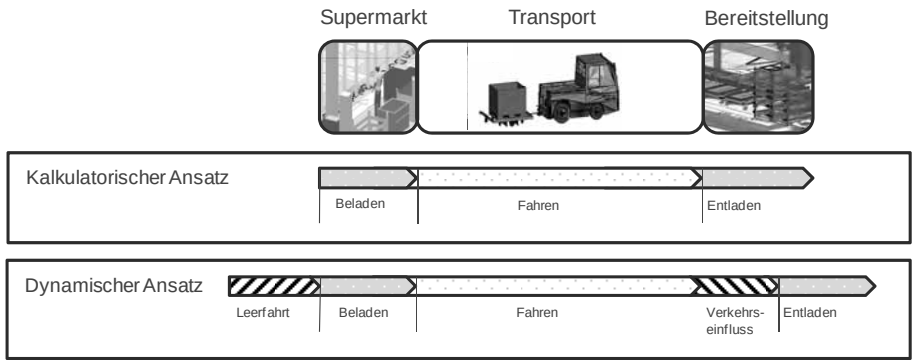



*Abbildung 5: Vergleich kalkulatorische und dynamische Transportzyklen*

Die dynamisch bedingten Verzögerungen gegenüber der Transportdauer im unbelasteten Netz können beispielsweise entstehen durch:

- die Blockade eines Wegabschnittes durch ein langsameres oder stehendes Fahrzeug (Ladevorgänge, Störung, Parken, …)
- das Warten auf Fahrzeuge auf bevorrechtigten Wegelementen (Rechts-vor-Links-Kreuzungen, Vorfahrtstraße, bevorrechtigte Straße, Übergänge, …)
- das Warten auf die Freigabe von Wegelementen (Tore, Schranken, …)

Die Transportwege eines Fertigungssystems werden ggf. auch von stochastischem, nicht logistikbezogenem Hintergrundverkehr belastet, beispielsweise durch Baustellenverkehr oder durch Personenverkehr zu Schichtwechseln.

## 6 Abbildung des werkinternen Verkehrs mittels der Simulationssoftware SUMO

Die infrastrukturellen Transportbedingungen in einem Fertigungsnetzwerk oder einer Fabrik, wie beispielsweise das Wegenetz und die geltenden verkehrlichen Regelungen, können mit den Ausprägungen des Straßenverkehrs übereinstimmen. Insbesondere auf räumlich ausgedehnten Fertigungsstandorten und u. U. auch in den Fertigungshallen, kann der Werksverkehr dem nationalen Verkehrsregelwerk unterstellt sein, bspw. durch den Hinweis: „Es gilt die StVO“ (Straßenverkehrsordnung) an den Werkspforten.

Zur Abbildung der Dynamik des durch die Inhouse-Logistik verursachten Verkehrsaufkommens können Methoden der Verkehrsflusssimulation verwendet werden.

Insbesondere haben die Topologie des Wegenetzes und die Ausprägung seiner Elemente, aber auch die Belastung des Netzes, das geltende Verkehrsregelwerk und die Fahrereigenschaften der zu modellierenden Fahrzeugführers Einfluss auf die Fahrtzeitdauer einzelner Transportaufträge.

Die Topologie des Wegenetzes eines Fertigungssystems ergibt sich unmittelbar aus den Planungsdaten (CAD-Layout). Der Transportbedarf ergibt sich aus dem Fertigungsumfang, der Fabrikstruktur und den Logistikprozessen. Bestehende Simulationsumgebungen halten jedoch nur beschränkte Möglichkeiten zur Abbildung von Standardverkehrsregeln und -fahrerverhalten vor (Staab 2013).

Eine leistungsfähige Simulationsumgebung muss das Fahrer- und Fahrzeugverhalten auch auf mehrspurigen Wegabschnitten wiederspiegeln können. Der simulierte Fahrer sollte über mehre Wegelemente hinweg planen und seinen Fahrverhalten strategisch anpassen können; beispielhaft sei hier die frühzeitige Spurwahl bei Abbiegevorgängen genannt. Darüber hinaus muss das Verhalten des Fahrers bei Annäherung und Überschreitung von Kreuzungsflächen regelkonform und unter Berücksichtigung vorhandener Bevorrechtigungen erfolgen. Gleichfalls beeinflussen Überholvorgänge entweder unter Verwendung mehrerer Spuren gleicher Richtung oder unter Mitbenutzung der Gegenspur die Kapazität des Netzes und die Kontinuität des Verkehrsflusses.

SUMO ist eine Verkehrsflusssimulationssuite, die auf diese Aufgabenstellungen zugeschnitten ist. SUMO wird am Deutschen Zentrum für Luft- und Raumfahrt seit 2001 kontinuierlich entwickelt (Bieker 2012). Wenngleich SUMOs Komponenten das gesamte Aufgabenspektrum eines Verkehrsplaners bzw. -ingenieurs abdeckt, so

liegt der funktionale Schwerpunkt der Software auf der Abbildung von Fahrzeugbewegungen durch ein Straßennetzwerk sowie auf der Interaktion von Verkehrsteilnehmern untereinander und mit Infrastrukturelementen bzw. Verkehrsmanagementsystemen.

Für die simulative Abbildung der Fahrzeugbewegungen in der Inhouse-Logistik wurde für die hier vorgestellte Lösung darauf verzichtet, alle wesentlichen Eigenschaften in das Plant Simulation Modell zu implementieren. Stattdessen wurde MALAGAs dynamisches Validierungsmodul um eine Kopplungsschnittstelle zu der Verkehrsflusssimulationssoftwaresuite SUMO erweitert.

Die Einbindung einer etablierten, ausgereiften Verkehrsflusssimulation erhöht die Nachvollziehbarkeit der Ergebnisse, erzwingt aber eine strikte Aufgabenverteilung der Simulationselemente. Mit der realisierten Implementierung der Schnittstelle ist die Übertragbarkeit der vorgestellten Entwicklung gewährleistet.

## 7 Kopplungsschnittstellen und Aufgabenverteilung zwischen SUMO und Plant Simulation

Neben den statischen Konfigurationen wie Netzwerkdatei und Routeninformationen stellt SUMO mit TraCI eine Socket-basierte Schnittstelle zur Kopplung externer Software-Module bereit. Mittels der TraCI-Schnittstelle lassen sich laufende Simulationsinstanzen weitreichend manipulieren, sodass Fahrzeuge de facto ferngesteuert werden. Mit TraCI können Routenführung, Haltepunkte und Fahrzeugzustände beeinflusst werden, jedoch ohne dabei die verkehrssimulationstypischen Modellierungen von Fahrzeugfolgemodellen und verkehrskonformen Fahrerverhalten im Wegenetz zu überschreiben.

Aufgrund des offenen und erweiterbaren Charakters von SUMO lassen sich die Schnittstellen der Verkehrssimulation an die durch den Hersteller festgelegten Schnittstellen von Plant Simulation anpassen.

Zwecks Performancemanagement verwendet die hier vorgestellte Kopplung verschiedene Kommunikationswege, je nach Zeck und Datenvolumen. So werden für dynamische Veränderungen bspw. am Netz oder den Fahrzeugzuständen der Datenaustausch über eine Socket-basierte oder ggf. In-Memory-Kommunikation realisiert. Umfangreiche statische Datenvolumen können hingegen über virtualisierte Laufwerke auch dateibasiert ausgetauscht werden.

Die Simulationshoheit für die dynamische Abbildung der Fertigungsprozesse und letztlich die Hoheit über die Transprotaufträge für die Inhouse-Logistik verbleibt bei Plant Simulation. Transportbedarfe werden durch die in Plant Simulation realisierte Transportsteuerung als Transportaufträge an die Flurförderfahrzeuge weitergereicht. Diese Fahrzeuge werden nun nicht mehr als Plant Simulation BEs (Bewegte Einheiten) dargestellt, sondern als autonome, bewegte Objekte im SUMO-Netz. Am Ende ihres Fahrauftrags übergeben die Objekte die Hoheit wieder an die Auftragssteuerung von Plant Simulation. Alle relevanten Ereignisse werden protokolliert und an MALAGA übergeben.

## 8 Kombinierte dynamische und geometrische Validierung im Fabrik-DMU

Das System MALAGA bietet die Möglichkeit, die aus der Dynamik erworbenen Bewegungsdaten als 3D-Animation auszuspielen. Bei Volkswagen wird das System Interviews3D von 3DI, Ilmenau eingesetzt um diese Animationssequenzen mit der 3D-Geometrie der virtuellen Fabrik zu überlagern (Bracht 2012). Dem Anwender gibt dies die Möglichkeit, Transportbewegungen und geometrische Randbedingungen gemeinsam zu visualisieren (Abb. 6). Die dynamische Validierung kann somit im Kontext aller planenden Gewerke stattfinden. Das Fabrik-DMU liefert die Möglichkeit der geometrischen Validierung um beispielsweise Anstell- und Anlieferkonzepte in Übereinstimmung zu bringen ohne Blockaden zu schaffen. Lagerflächen können visualisiert und mit Blick auf ihre zugedachte Funktion konzipiert und ausgestattet werden. Mit der hier vorgestellten Entwicklung können künftig auch Bereiche mit hoher Verkehrsdichte und Stausituationen einfach ausgewiesen werden. Dadurch kann der Logistikplaner organisatorische oder bauliche Maßnahmen ableiten.

*Abbildung 6: Fahrzeugbewegung im virtuellen Fabrik-DMU*

## 9 Chancen und Ausblick

Für den Logistikexperten bieten die mit dem System MALAGA bereitgestellten Funktionen neue Möglichkeiten. Das Fabriklayout stellt den gewerkeübergreifenden Kontext dar. Graphische Restriktionen wie Wegenetz und Flächenobjekte verringern den Freiheitsgrad seines Materialflussmodells systematisch. Innerhalb eines somit limitierten Lösungsraums können Optimierungsalgorithmen etabliert und plausible Lösungen automatisch vom System gefunden werden. Den entscheidenden Vorteil bieten die grafischen und tabellarischen Ergebnistypen, die im 3D-CAD-Layout und im System MALAGA jederzeit darstellbar und einsehbar sind. Die Auswirkungen von getroffenen Annahmen werden sichtbar und Sensitivitätsanalysen sind möglich.

Auf Basis des Fabriklayouts können konsistente Materialflussbeziehungen dargestellt werden. Mittels der Ausleitung von Simulationsmodellen, die nun auch realistisches Stauverhalten zeigen, ist die robuste Auslegung des Gesamtsystems möglich. Auch die Einbeziehung eines Simulationsexperten, der auf Basis seiner Standards für die methodisch ordentliche Absicherung eines finalen Planungsstandes zuständig wäre (Krug 2010), gelingt einfach. Ihm werden bereits lauffähige Simulationsmodelle bereitgestellt, die er verwenden und weiter untersetzen kann (Abb. 7).

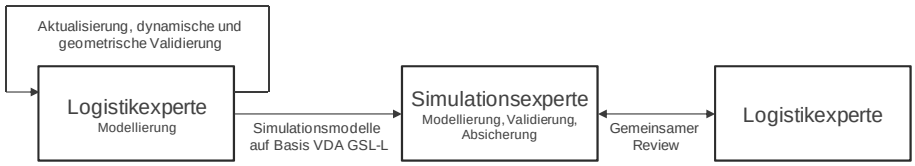



*Abbildung 7: Mögliches Zusammenarbeitsmodell mit Simulationsexperten*

## Literatur

Bieker, L., Krajzewicz, D; Erdmann, J., Behrisch; M.: Recent development and applications of SUMO - Simulation of Urban MObility. International Journal On Advances in Systems and Measurements 5 (2012) 3&4, S.128-138.

Bracht, U.; Rooks, T; Adrian, R: Virtuelle Logistikplanung für die Montage im Rahmen der Digitalen Fabrik. In: Markus Rabe (Hrsg.): Advances in Simulation for Production and Logistics Applications. Stuttgart: Fraunhofer IRB 2008, S. 439-447.

Bracht, U.; König, A.; Spies, J.: Integration des Fabrik-DMU in die Planungsprozesse der Fabrikstrukturplanung. WT-Online 03/2012, S. 103-108.

Dreher, S.; Nürnberger, A.; Kulus, D.: Routenoptimierung in der Produktionslogistik. ZWF 104 (2009) 3, S.119-123.

Juan, A.; Rabe, M.: Combining Simulation with Heuristics to solve Stochastic Routing and Scheduling Problems. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik - Entscheidungsunterstützung von der Planung bis zur Steuerung. Paderborn: HNI-Verlagsschriftenreihe 2013, S. 641-648.

Krug, W.; März, L.; Rose, O.; Weigert, G. (Hrsg): Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen. Berlin, Heidelberg: Springer 2010, S. 117-132.

Mayer, G.; Pöge, C.: Quo vadis Ablaufsimulation – Eine Zukunftsvision aus Sicht der Automobilindustrie, In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik - Entscheidungsunterstützung von der Planung bis zur Steuerung. Paderborn: HNI-Verlagsschriftenreihe 2013, S. 12-20.

Staab, T.; Galka, S.; Klenk, E.; Günthner, W.: Effizienzsteigerung für Routenzüge - Untersuchung des Einflusses der Routenführung auf die Auslastung und Prozessstabilität; In: Simulation in Produktion und Logistik Entscheidungsunterstützung von der Planung bis zur Steuerung. Paderborn: HNI-Verlagsschriftenreihe 2013, S. 168-176.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Ganzheitliche Supply-Chain-Optimierung mittels Kombination von mathematischer Optimierung und Simulation

## *Holistic Supply Chain Optimisation Using a Combination of Mathematical Optimisation and Simulation*

Alexander Klaas, Kamel Klibi, Miebach Consulting GmbH, Frankfurt (Germany),
klaas@miebach.com, klibi@miebach.com

Abstract: Supply Chain Optimisation in general views the different levels of a Supply Chain holistically, due to the high level of interdependencies, in order to reach a cost-optimal configuration. Commercially available tools, however, focus on certain parts of a supply chain such as the DC outbound level and furthermore, they do not consider dynamic effects such as delays in transport. Miebach Consulting has developed a new tool "MC Neo" that works coupled with Simulation software. It is able to derive a Supply Chain that minimises the total costs (including production and transport) and furthermore considers realistic and dynamic circumstances. First uses of the software resulted in a reduction of total costs of 10% compared to the use of conventional software.

## 1 Einleitung und Problemstellung

Die Supply Chain Optimierung besteht im Grundsatz aus der gleichzeitigen Betrachtung einzelner Fragestellungen auf jeder Stufe und speziell der komplexen Interdependenzen zwischen ihnen. Ziel ist es, die kostenoptimale Konfiguration der Standorte bzw. Teilnehmer entlang der Wertschöpfungs- und Lieferkette herzustellen.

Ausgehend von den Kundenbedarfen am Markt sind auszugsweise folgende Entscheidungen zu treffen:

- Welche Lieferanten beliefern welche Produktionswerke?
- Welche Produkte werden in welcher Menge in welchen Werken produziert?
- Welche Distributionszentren werden in welcher Menge von welchen Werken beliefert?
- Welche POS (Point of Sales) werden von welchen Distributionszentren versorgt?

Dabei sind die Gesamtkosten der Supply Chain als Summe von Beschaffungs-, Produktions-, Lager-, Abwicklungs- und Transportkosten auf jeder Ebene zu sehen. Es

ist offensichtlich, dass etwa eine Verlagerung der Produktion eines Produktes von einem Werk auf ein anderes eine Änderung der Kosten auf sämtlichen Stufen bewirkt. So können ggf. die Produktionskosten aufgrund eines geringeren Lohnniveaus am neuen Standort sinken, durch längere Transportwege aber die Transportkosten steigen. Aufgrund solcher Interdependenzen ist die Supply Chain Optimierung als mathematisches Problem hochkomplex und eine rein lokale Optimierung an einzelnen Stufen führt zu negativen Konsequenzen wie den Bullwhip-Effekt. Die hohe Rechenkomplexität der entstehenden Probleme ist ein Grund dafür, dass kommerziell erhältliche Tools nicht sämtliche Stufen der Supply Chain ganzheitlich betrachten (Sigmund 2014), sondern zumeist spezielle Themen wie die Belieferung der POS optimieren (Abb. 1). Es handelt sich dabei um eine reine Optimierung der Outbound-Transportkosten. Die Auswirkungen auf z. B. die Produktionskosten sowie auf die Inbound-Transportkosten werden dabei nicht betrachtet.

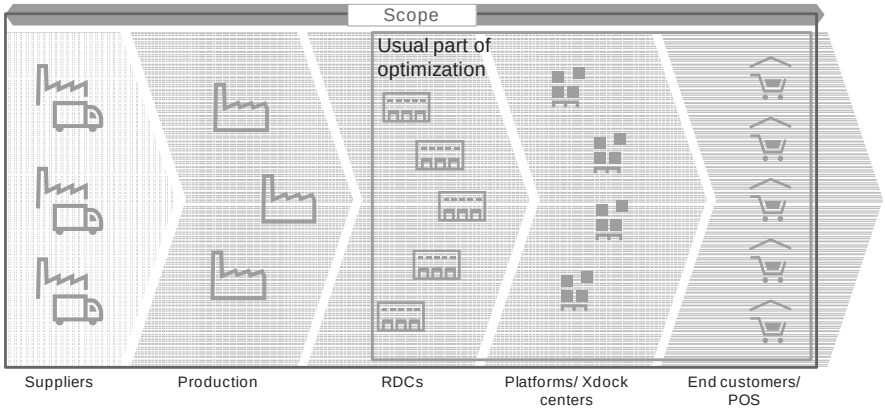



*Abbildung 1: Eine fünfstufige Supply Chain. Üblich ist eine Optimierung des umrandeten Bereiches*

Weiterhin ist es ein Problem, dass eine strategische Supply Chain Optimierung zumeist auf einer groben, gemittelten Abbildung des Materialflusses auf Jahresmengen basiert. Die gängigen Berechnungsmethoden wie die Lineare Programmierung sind in den kommerziellen Implementierungen nicht im Stande, dynamische Effekte wie Bestandsverläufe oder Bedarfsspitzen sowie Restriktionen wie Produktions- bzw. Verladekapazitäten zu berücksichtigen. Diese können allerdings die optimale Lösung stark beeinflussen, da z. B. die Einhaltung einer kurzfristigen Liefertreue erfordert, entweder genügend Kapazitäten auch für die Spitzen bereit zu halten, oder durch entsprechende Strategien hinreichende Puffer aufzubauen, für die allerdings ebenfalls Kapazitäten benötigt werden. Die ermittelte Konfiguration der Supply Chain könnte sich unter Berücksichtigung dieser Effekte als nicht optimal erweisen.

Die Supply Chain Simulation hingegen bedient sich ereignisorientierten Verfahren und ist durch die explizite Mitbetrachtung der Zeit in der Lage, genau diese Effekte abbilden. Im Gegensatz zu Optimierungsverfahren ist eine Simulation allerdings nicht in der Lage, eine optimale Konfiguration selbst zu bestimmen – sie kann im Sinne eines Experimentes nur eine Ausprägung untersuchen und ein Ergebnis in Form von Kennzahlen bestimmen. Die zu untersuchende Konfiguration der Supply

Chain kann allerdings im Sinne der Simulation und Optimierung übertragen werden (Campuzano 2011).

In diesem Beitrag wird ein neuartiger Ansatz vorgestellt, der die genannten Probleme löst. Einerseits handelt es sich um das eigens entwickeltes Optimierungstool MC Neo, das aufgrund einer hoch performanten Lösungsmethode in der Lage ist, Supply Chains von typischen Problemgrößen ganzheitlich zu betrachten. Andererseits wird eine Kopplung mit einem ereignisdiskreten Simulationstools vorgestellt, das eine Lösung der Optimierung mit dynamischen Effekten betrachtet und so iterativ zu einer optimalen Konfiguration führt.

## 2 Problemlösung

Der vorgestellte Ansatz kombiniert eine eigens entwickelte Implementierung zur mathematischen Supply Chain Optimierung (MC Neo) mit der Simulation. Das Verfahren ist in Abbildung 2 dargestellt und besteht aus drei Schritten.

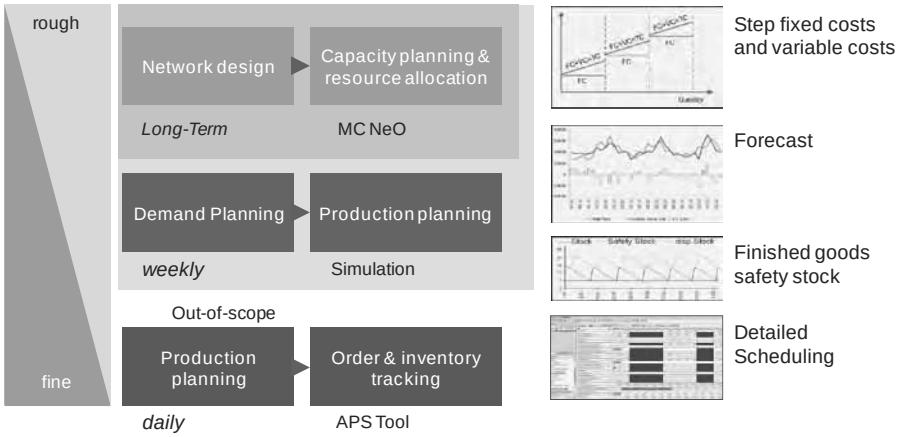



*Abbildung 2: Der Optimierungsansatz kombiniert strategische, taktische und operative Verfahren*

Das neu entwickelte Tool berechnet – je nach Zielstellung – die kostenoptimale Konfiguration der betrachteten Supply Chain. Im momentanen Entwicklungsstand liegt der Fokus auf der Problemstellung der Produktionsallokation – „In welchem Produktionswerk produziere ich wieviel von welchen Produkten, und über welche Distributionszentren beliefere ich meine POS?“

Die Software berücksichtigt dabei zahlreiche praktische Umstände und Restriktionen.

- *Pro Stock Keeping Unit (SKU) und Werk fixe und variable Produktionskosten*. Dabei sind die fixen Produktionskosten treppenförmig, um die Erweiterung von Produktionskapazitäten für diese SKU (Belegung einer weiteren Produktionslinie oder einer neuen Produktionsschicht) abzubilden. Die fixen Produktionskosten bilden Kosten wie Raum-, Kapital- und kalkulatorische Kosten wie kalkulatorische Abschreibung und kalkulatorische Zinsen ab. Die variablen Produktions-

kosten decken Personal-, Material-, und Energiekosten ab. Es können dabei lokale Unterschiede etwa im Arbeitslohn berücksichtigt werden, da die Kostenfunktion pro Werk spezifiziert werden kann.

- *Transportkosten auf jeder Relation von Produktionswerk zu Distributionszentrum.* Diese Kosten können dabei degressiv sein. Es muss eine Standardfunktion eingegeben werden und es kann zusätzlich für jede Relation eine eigene definiert werden. Dies erlaubt es etwa, die Transporte bei zollunionübergreifenden Relationen höher zu setzen, oder um länderspezifische Kosten zu berücksichtigen.
- *Maximale Entfernungen zwischen Produktionswerten,* um Restriktionen bezüglich der maximalen Transportdauer zu berücksichtigen. Die Güter müssen innerhalb einer gewissen Zeitspanne beim POS oder beim DC sein, um nicht zu verfallen.
- *Kapazitätsgrenzen in Produktionen und Distributionszentren*

Die Inputs der Berechnung werden aus einer Exceldatei eingelesen und in eine Datenbank unter einem eindeutigen Szenarionamen überführt (Abb. 3). Dies ermöglicht es, auch bei Änderungen an der Exceldatei im Nachhinein die genauen Eingangsfaktoren eines Ergebnisses nachvollziehen zu können. Die Daten können von MC Neo aus vor einer Berechnung auch direkt in der Datenbank manipuliert werden. MC Neo basiert auf C# und verwendet dabei die Microsoft Schnittstellen.

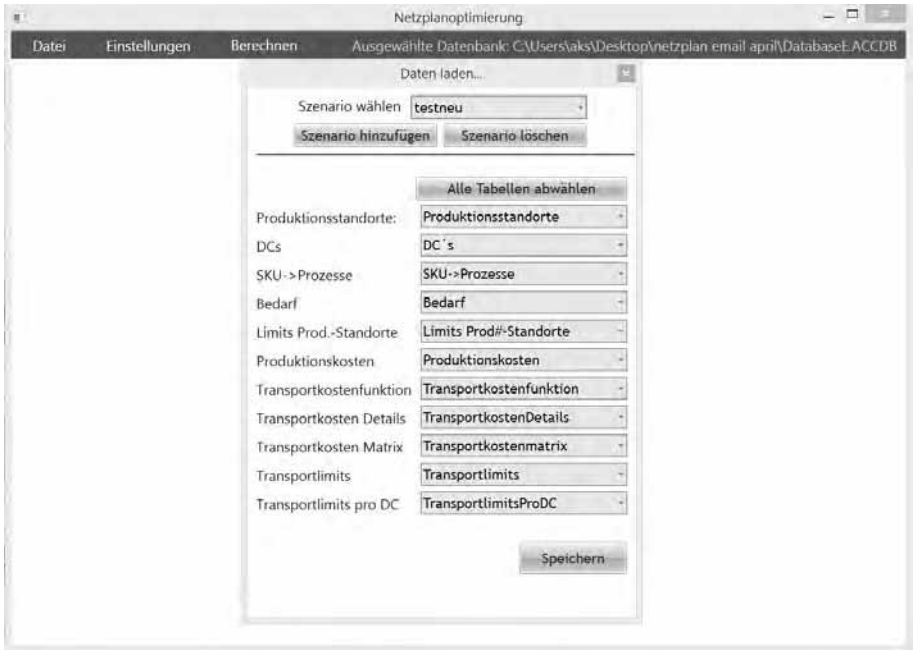



*Abbildung 3: Eingabemaske von MC Neo nach Excelimport*

Die Software ermittelt zunächst die Transportrouten auf allen Relationen Produktionswerk-Distributionszentrum kilometergenau (die postalischen Adressen sind vorgegeben) und unter Berücksichtigung der tatsächlichen Straßenverhältnisse,

indem eine Geokodierung sowie eine Straßennavigationssoftware eingebunden wird. Zusammen mit der Transportkostenfunktion können daher die Transportkosten für jede Menge auf der Relation berechnet werden. Aus den fixen und variablen Produktionskosten ergeben sich die Produktionskosten.

Anhand der Kostenfunktionen und unter Berücksichtigung der Restriktionen durchsucht die Software nun den Raum möglicher Konfigurationen, um die gesamtkostenoptimale Lösung zu finden. In diesem Fall ist dies die Zuordnung der Produktionen und Relationen zwischen Produktionswerk und Distributionszentrum. Dabei verwendet sie das iterative Frank–Wolfe Verfahren (Jaggi 2013), um schnell zu einer Lösung gelangen. Bei einer bereits hochkomplexen Supply Chain mit ca. 25 Produktionswerken und 50 Distributionszentren berechnet das Tool die Lösung in etwa einer Sekunde Rechenzeit.

Das Ergebnis der Berechnung wird in Form einer Tabelle in der Datenbank abgespeichert sowie grafisch dargestellt. Für jede Kombination von Produktionswerk, Produkt und Distributionszentrum wird die zu produzierende bzw. transportierende Menge errechnet, sodass die resultierenden Gesamtkosten über alle Relationen minimiert sind (Tabelle 1). Das Ergebnis wird in der Datenbank abgespeichert und lässt sich wiederum nach Excel exportieren. Die fixen Produktionskosten werden dabei für jede Kombination Produktionsstandort – SKU für nur eine Zeile aufgeschlagen. Die variablen Produktionskosten sowie die Transportkosten sind pro Mengeneinheit angegeben Grafisch stellen Linien zwischen Produktionswert und Distributionszentrum die Transporte dar, deren Dicke das Volumen anzeigen. Die Karte ist dabei verschieb- und zoombar.

| Fabrik | DC | SKU | Menge | fixe Prod. Kosten | var. Prod. Kosten | Transport-kosten | Summe Kosten |
|---|---|---|---|---|---|---|---|
| P4 | DC_1 | 4141 | 3101 | 1327734 | 31710 | 2061 | 106069028 |
| P4 | DC_2 | 4141 | 5870 | 0 | 31710 | 1844 | 196993659 |
| P4 | DC_3 | 4141 | 5283 | 0 | 31710 | 1869 | 177401245 |
| P4 | DC_5 | 4141 | 24248 | 0 | 31710 | 1630 | 808444805 |
| P5 | DC_4 | 4141 | 10247 | 43610579 | 41505 | 496 | 474015446 |
| P5 | DC_11 | 4141 | 35765 | 0 | 41505 | 1434 | 1535777064 |
| P5 | DC_18 | 4141 | 12619 | 0 | 41505 | 0 | 523798341 |
| P5 | DC_24 | 4141 | 8642 | 0 | 41505 | 1255 | 369576353 |

*Tabelle 1: Beispielhaftes Ergebnis der Optimierung von MC Neo*

Im zweiten Schritt wird die berechnete Konfiguration in ein Simulationsmodell der Supply Chain überführt. Schließlich ist die Berechnung mit über ein Jahr aufsummierten Mengen aufgrund von eventuellen Saisonalitäten unzureichend. Dabei wird

nun der tatsächliche Verlauf an Kundenbedarfen über ein Jahr tagesgenau zugrunde gelegt.

Die Konfiguration wird entsprechend des Outputs des Optimierungstools (siehe Tabelle 1) von einem Script der Simulationssoftware DOSIMIS-3 eingelesen. Das Script generiert automatisch das grundlegende Modell der Supply Chain, indem es für jeden Produktionsstandort und jedes DC entsprechende Objekte erstellt und die entsprechende Menge für den Produktionsplan und für die Simulation der Transportbewegungen verwendet.

Die eigentliche Supply Chain Simulation bildet nun den Verlauf jedes Artikels jeder Bestellung durch das Netzwerk detailliert ab. Je nach Artikelgruppe können verschiedene Lieferwege abgebildet sein, so kann etwa die Ware lagerhaltig an einem DC sein und von diesem wird der Kunde beliefert. Dieses und die anderen DCs werden entsprechend einer Bestellpolitik versorgt, die im Simulationsmodell hinterlegt wird. So kann beispielsweise eine s,S–Politik spezifiziert sein, nach der bei Unterschreiten eines Meldebestands s die Differenz zum Sollbestand S an der vorherigen Stufe bestellt wird. Weiterhin werden die Umschlagzeiten und Bereitstellzeiten an jeder Stufe der Supply Chain berücksichtigt. Im Modell kann schließlich der genaue Lieferzeitpunkt beim Kunden und damit der resultierende Servicegrad ermittelt werden, der sich durch die Konfiguration ergibt.

Die Produktionsplanung erfolgt rollierend wöchentlich mit einem Planungshorizont von x Wochen (x ist dabei als Parameter vom Anwender vorzugeben) unter Berücksichtigung von minimalen und maximalen Losgrößen und Rüstzeiten. Aufgrund der Kenntnis der definierten Bestellpolitiken an den hinteren Stufen und den Transportzyklen können diese ebenfalls in die Planung einfließen, wobei an dieser Stelle ein entsprechend der Aufgabenstellung geeignetes Basisverfahren aus dem Bereich PPS angewendet werden kann.

Das in Schritt 1 ermittelte Netzwerk wird daher auf die Praxistauglichkeit auch in Bedarfsspitzen getestet. Wird ermittelt, dass die Konfiguration nicht die Anforderungen erfüllt, kann der Anwender über das Setzen der Kapazitätsrestriktionen (Mindest- oder Maximalmengen) gezielt Schwachstellen ausschließen und den Optimierungslauf wiederholen. Beispielsweise könnte ein in der Jahresbetrachtung für ein Produkt kostenoptimales Produktionswerk unter Berücksichtigung der dynamischen Effekte doch nicht geeignet sein, die konfigurierte Menge der verschiedenen SKUs zu produzieren. Grund könnte ein Zusammenspiel von Bedarfsspitzen und Effekte durch die Losgrößen sein, wodurch in kritischen Momenten die Produktion nicht ausreicht und eine Versorgungslücke entsteht. Der Anwender könnte dann entweder die Produktion einer bestimmten an dem Produktionsstandort unterbinden, oder die Gesamtmaximalkapazität des Standortes reduzieren. Ein neuer Optimierungslauf mit MC Neo würde dann eine unter den neuen Restriktionen kostenoptimale Konfiguration ermitteln, die dann wiederum in der Simulation evaluiert würde. In dieser Iteration entsteht die gesuchte Lösung. Im Laufe der Iteration kann so sehr einfach ein neues Ergebnis aus MC Neo importiert werden, indem die Ergebnisdatei neu eingelesen wird.

Die in Schritt 2 ermittelte Konfiguration der Supply Chain auf wochenweiser Betrachtung der Produktion kann anschließend im Sinne des Advanced Planning and Scheduling weiter verfeinert werden. Dabei kann tagesgenau ein Produktions-

programm unter noch genauerer Berücksichtigung von Rüstzeiten und weiteren Einflussfaktoren ermittelt werden.

Der kombinierte Ansatz wurde in einem ersten Szenario verwendet. Es wurde dabei einerseits die Praxisfähigkeit und Usability getestet, andererseits der erzielte Nutzen ermittelt.

Aus Ist-Daten, die ebenfalls im Excelformat vorlagen, konnten die benötigten Daten extrahiert werden und durch eine Aufbereitung in die Datenbank von MC Neo importiert werden. Die grafische Ausgabe erlaubte eine schnelle Einschätzung und Plausibilitätsprüfung des Ergebnisses. Die erzielten Kosteneinsparungen lagen im Bereich von 10%, wobei die größte Ersparnis im Produktionsbereich lag. Die bisher verwendeten Supply Chain Tools wären nicht in der Lage gewesen, diese Optimierung unter Berücksichtigung vor allem der maximalen Transportdauer durchzuführen.

## 3 Fazit und Ausblick

Im Ergebnis zeigt sich von den Lieferanten bis zum Endkunden eine Supply Chain, bei deren Konfiguration stets die Gesamtkosten auf jeder Stufe ganzheitlich betrachtet wurden. Gleichzeitig wurden durch Kombination mit ereignisdiskreter Simulation dynamische Effekte berücksichtigt, sodass Schwankungen, die sich typischerweise durch die gesamte Supply Chain ergeben, bewältigt werden können.

Die eigens entwickelte Optimierungssoftware demonstriert dabei Tauglichkeit anhand eines ersten Beispielszenarios und wird um weitere Fragestellungen und Stufen der Supply Chain erweitert.

Zunächst wird das Tool um die Funktionalität erweitert, die Anzahl an DCs variabel zu gestalten. Dabei kann der Benutzer eine unveränderliche Kernmenge spezifizieren und darüber hinaus eine optionale Menge an potentiellen DC Standorten. MC Neo berechnet dann innerhalb eines Intervalls die optimale Anzahl an DCs sowie welche der potentiellen am effizientesten wären.

Weiterhin wird MC Neo um Kunden erweitert. Diese stellen zukünftig die Bedarfssenken dar und werden dann DCs zugeordnet. Dabei werden analog zur ersten Stufe Transportkostenfunktionen vorgegeben sowie Umschlagskosten (fix und variabel) an jedem DC.

## Literatur

Campuzano, F.; Mula, J.: Supply Chain Simulation: A system dynamics approach for improving performance. London: Springer 2011.

Jaggi, M.: Revisiting Frank–Wolfe: Projection-free sparse convex pptimization. 2013. Journal of Machine Learning Research: Workshop and Conference Proceedings 28 (2013) 1, S. 427-435.

Sigmund, J.: Netzwerkoptimierung in XCargo. 2014.
http://www.xcargo.de/fileadmin/user_upload/Hauptseite_Downloads/DE/XCargo_Praesentation_gesamt.pdf

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulative Planungsassistenz für die Disposition in mehrstufigen Distributionsnetzwerken

## *Simulation-based Assistance for Inventory Planning in Multi-echelon Distribution Networks*

Markus Witthaut, Josef Kamphues, Tobias Hegmanns, Fraunhofer-Institut für Materialfluss und Logistik IML, Dortmund (Germany),
markus.witthaut@iml.fraunhofer.de, josef.kamphues@iml.fraunhofer.de,
tobias.hegmanns@iml.fraunhofer.de

Abstract: The variety of articles and warehouses in complex distribution networks leads to a high effort for inventory planning. In this respect, the allocation of stock as well as the definition of stock levels are of particular importance for the efficiency of distribution networks. In order to support responsible staff, several analytical planning methods and tools are commonly used. However, these approaches do not provide the possibility to consider dynamic influences within the planning process (e. g. a varying forecast quality over course of time). For this reason, a tool has been developed which allows the user for analytically determining parameters for inventory planning (e.g. safety stocks or inventory targets) and, in addition, provides the functionality of a simulation-based validation of these parameters against the background of dynamic influences.

## 1 Einführung und Beschreibung des Anwendungsfalls

Die Produktion und Distribution von Gütern erfolgt zu einem großen Teil in weltweit verteilten Supply Chains, welche aus vielen Produktionswerken und Lagerstandorten bestehen, an denen unterschiedliche Artikel hergestellt und gelagert werden. Oft werden die Artikel über mehrere Stufen von den Werken zu den Kundenstandorten distribuiert. Ein Betreiber eines solchen Distributionsnetzes muss für jeden Artikel und jeden Lagerstandort festlegen, wie der Bestand zu disponieren ist bzw. ob der jeweilige Artikel gelagert werden soll oder nicht. Falls eine Lagerung erforderlich ist, muss ein Bestandsmanagementverfahren ausgewählt und parametriert werden. Bei vielen Verfahren der verbrauchsgesteuerten Disposition werden zum Beispiel Bestellrhythmen, Bestellmengen, Meldebestände und/oder Zielbestände festgelegt. In der Unternehmenspraxis werden diese Dispositionsverfahren und deren Parameter oft nur in großen Zeitabständen angepasst – beispielsweise bei Einführung einer neuen ERP-Software. Für die heutigen Supply Chains ist diese Vorgehensweise jedoch wenig geeignet: Änderungen im Nachfrageverhalten und

kurze Produktlebenszyklen erfordern eine häufige Anpassung der Dispositionsparameter.

In der hier vorgestellten Forschungsarbeit wurde die Auswirkung unterschiedlicher Dispositionsverfahren hinsichtlich der Bestände und des Lieferservices in einer mehrstufigen Supply Chain mit Hilfe der ereignisdiskreten Simulation (DES) untersucht. Insbesondere für die Bewertung des Lieferservices wurde die dynamische Betrachtung von Kunden- und Transportaufträgen durch ein Simulationsmodell als gut geeignet betrachtet, um belastbare Aussagen hinsichtlich Servicegrad und Kosten zu erreichen.

Im Anwendungsfall wurde die Bestandsplanung in einem Reifen-Distributionsnetzwerk fokussiert. In diesem Netzwerk werden Reifen an mehreren Stellen gelagert: An jedem Produktionswerk befindet sich ein Produktionslager, von welchem die Artikel über Zentrallager und/oder Regionallager an die Kunden geliefert werden. Diese Kunden – Autowerkstätten und Reifenhändler – sind 31 Absatzmärkten zugeordnet. In diesem Distributionsnetzwerk werden mehrere Tausend unterschiedliche Reifenmodelle gelagert. Hinsichtlich der Nachfrage müssen die Artikel in Sommer- und Winterreifen mit entsprechend unterschiedlichen Nachfrageverläufen eingeteilt werden.

Der Betreiber des Reifen-Distributionsnetzwerkes muss monatlich für jeden Artikel folgende Fragen beantworten:

1. Welche Bevorratungsebene ist zu wählen? Artikel können nur im Regionallager, nur im Zentrallager oder auch in beiden Lagertypen gelagert werden. Es ist auch möglich, einen Artikel in einer Teilmenge der Regionallager zu lagern.
2. Wie hoch sollen die Zielbestände für die jeweiligen Artikel-Lager-Kombinationen gewählt werden?
3. Welcher Servicegrad und welche Bestandskosten sind mit den Entscheidungen von Punkt 1 und 2 verbunden?

## 2 Bestandsplanungsstrategien für mehrstufige Distributionsnetzwerke

Die im Kontext der taktischen Bestandsstrategieoptimierung wichtigen Aufgaben der optimalen Bestandsallokation sowie die Bestimmung der optimalen Dispositionsparameter wurden bereits in vielen Arbeiten behandelt. Übersichten zu den existierenden Methoden finden sich z. B. in den Arbeiten von Axsäter (2007), Gudehus (2012), Tempelmeier (2012) oder Simchi-Levi et al. (2013). Jüngere Forschungsergebnisse zur optimalen Platzierung von Sicherheitsbeständen in mehrstufigen (multi-echelon) Distributionssystemen sind in den Arbeiten von Güller et al. (2011) oder Stumvoll et al. (2013) enthalten. Miranda und Garrido (2009) lösen Bestands- und Netzwerkdesign-Entscheidungen simultan. Kang und Kim (2010) fokussieren auf die Koordination von Bestands- und Transportmanagement.

Ansätze aus dem Bereich des Supply Chain Designs, welche sich mit der Optimierung der Netzwerkstruktur und Zuordnungsproblemen z. B. bei der Bestimmung optimaler Lagerstandorte und Verteilknoten befassen, sind generell nicht Teil der taktischen Planungsebene. Es bestehen jedoch einige Erweiterungen dieser Modelle, welche die Zuordnung von Transportrelationen zu Netzwerkknoten optimieren, so genannte Allocation-Location-Modelle (vgl. z. B. Heinrichmeyer (2009) oder Sucky

und Hönscheidt (2008)). Das zeitlich begrenzte Einführen bzw. Aussetzen bestimmter Transportrelationen im Netzwerk, z. B. von Direktrelationen, kann bei stark schwankenden Nachfrageverläufen oder im Saisongeschäft eingesetzt werden. Die Ergebnisse dieser Modelle beeinflussen demnach die Bestandsplanung der taktischen Planungsebene, weil sich Wiederbeschaffungszeiten ändern, wodurch Sicherheitsbestände und andere Dispositionsparameter angepasst werden müssen. Simulationsbasierte Ansätze zur taktischen Planung sind in jüngster Zeit vor allem für Aufgaben der Produktionsplanung – z. B. Englberger et al. (2013) oder Lange et al. (2013) – veröffentlicht worden.

# 3 Logistisches Assistenzsystem (LAS) und zugehörige Geschäftsprozesse

## 3.1 Prozessanalyse und Netzwerkstruktur

Für das untersuchte Distributionsnetzwerk wurden zunächst in enger Kooperation mit dem Anwendungspartner die Prozesse und Gestaltungsoptionen ermittelt. Die aufgenommenen Prozesse wurden im Anschluss mit etablierten und verbreiteten Prozessmodellen abgeglichen und verallgemeinert, um eine Grundlage zur Entwicklung der umzusetzenden Planungsservices zu schaffen. Mithilfe der generalisierten Prozessbeschreibungen konnten Planungsaufgaben abgeleitet und quantifizierbare Ziele definiert werden. Im Rahmen dieser Aufnahme und Analyse der Strukturen konnte die komplexe Netzwerkstruktur auf die wesentlichen Elemente reduziert werden. Das europäische Distributionsnetzwerk beinhaltet demnach acht Werke (mit jeweils einem direkt angebundenen Produktionslager), fünf Zentral- und 30 Regionalläger sowie 31 Absatzmärkte.

Der Nachschub für regionale und zentrale Läger erfolgt auf der Grundlage der eingegangenen Kundenbestellungen sowie einer monatlichen Prognose je Artikel und Lager. Die Erzeugung eines Nachschubauftrags im Produktionslager geschieht bei Unterschreitung eines Meldebestands, der anhand des für jede Artikelklasse festgelegten Ziel-Servicelevels berechnetet wird. Ist eine vollständige Erfüllung der Bestellmenge nicht möglich, werden Teilmengen ausgeliefert und es wird ein neuer Nachschubauftrag erzeugt.

Im betrachteten Distributionsnetzwerk werden hierbei übliche Logistikziele verfolgt:

- hoher Servicegrad
- niedrige Bestandskosten
- geringer Lieferverzug
- niedrige Transportfrequenz
- hohe Transportauslastung

## 3.2 Entwicklung einer verbesserten Planungsmethodik

Um eine Planungsmethodik entwickeln zu können, wurden zunächst die wesentlichen Planungsaufgaben in Anforderungsworkshops mit dem Netzwerkbetreiber definiert. Diese Aufgaben wurden dann gegen die heute zum Einsatz kommenden IT-Systeme abgeglichen und auf Grundlage der definierten technologischen Anforderungen zur Planungsunterstützung überführt. Im Anschluss erfolgte die Entwicklung eines methodischen Konzepts, in dem die operativen und taktischen

Planungsprozesse des Anwendungspartners abgebildet werden konnten. Dabei wurde definiert, wie zum einen das Zusammenspiel der definierten Services erfolgt und zum anderen eine skalierbare und erweiterbare Systemarchitektur sichergestellt werden kann. Das Ergebnis der Planungsmethodik stellt somit die Planungssysteme und -konzepte für die in der Prozessanalyse identifizierten Planungsaufgaben dar.

Die wesentlichen Fragestellungen, die mithilfe der entwickelten Planungsmethodik beantwortet werden sollten, beschäftigten sich zum einen mit der Wahl der optimalen Bevorratungsebene (Bestandsallokation) und zum anderen mit den Auswirkungen einer geänderten Absatz- oder Bestandsplanung auf die Transport- und Ressourcenplanung. Aus diesem Grund wurden für den Anwendungsfall Methodenbausteine entwickelt, welche wichtige Aufgabenbereiche der taktischen Netzwerkplanung abdecken und dabei insbesondere Unterstützung zur Beantwortung der o.g. Fragestellungen bieten.

Diese Methodenbausteine umfassen im Einzelnen folgende Aufgabenbereiche:

- Analyse des Nachfrageverhaltens verschiedener Artikel
- Dimensionierung von Ziel- und Sicherheitsbeständen
- Unterstützung der operativen Bestands- und Verfügbarkeitsplanung
- Bestimmung einer optimalen Bevorratungsebene
- Transport- und Ressourcenplanung im Distributionsnetzwerk

Für die Bereitstellung einer hinreichenden Planungsunterstützung in allen genannten Aufgabenbereichen wurde ein Simulationsansatz gewählt, mit dem die Auswirkungen der mithilfe der Methodenbausteine definierten Planungsszenarien in dem Distributionsnetzwerk untersucht werden konnten. Parallel zur Entwicklung der Methodenbausteine erfolgte daher die Entwicklung eines Simulationsmodells, das alle notwendigen Distributionsstrukturen und -prozesse abbilden kann. Die Planungsmethodik und das Logistische Assistenzsystem – zum Begriff Assistenzsystem siehe Blutner et al. (2009) – wurden mithilfe von vier Planungsservices realisiert (vgl. Abb. 1):

- *Stock Strategy Planner*: Unterstützung der taktischen Netzwerkplanung durch Methoden zur Dimensionierung von Ziel- und Sicherheitsbeständen sowie zur Bestimmung einer optimalen Bevorratungsebene
- *Availability Monitor*: Unterstützung der operativen Bestands- und Verfügbarkeitsplanung anhand der Quantifizierung von Auswirkungen des Produktionsprogramms auf die Artikelverfügbarkeit
- *Network Analyzer*: Analyse der methodisch erarbeiteten Planungsszenarien hinsichtlich der o.g. Planungsziele mithilfe einer Simulationskomponente
- *Reporting Service*: Aufbereitung der Simulationsergebnisse zur Generierung von Reports mit den relevanten Kennzahlen als Entscheidungsgrundlage

Der konzipierte Planungsprozess beginnt mit dem *Stock Strategy Planner*. Zum einen dient er der Definition von Artikelclustern. Der Planungsbaustein bietet dabei die Möglichkeit, die Artikel auf der Grundlage unterschiedlicher Analysen (bspw. ABC- oder XYZ-Analyse) zu segmentieren. Zum anderen werden mit diesem Service auf der Grundlage der definierten Artikelcluster individuelle Bevorratungsebenen definiert. Für jede Artikelkategorie kann so eine individuelle Bevorratungsstrategie definiert werden (bspw. regionale Bevorratung für A-Artikel und zentrale Bevorratung für C-Artikel). Darüber hinaus bietet der *Stock Strategy*

*Planner* die Möglichkeit, artikel- und lagerspezifische Zielreichweiten unter Berücksichtigung der Bestandsstrategie und definierter Dispositionsparameter (Ziel-Servicelevel, Bestelllosgröße, Bestellrhythmus etc.) zu berechnen.

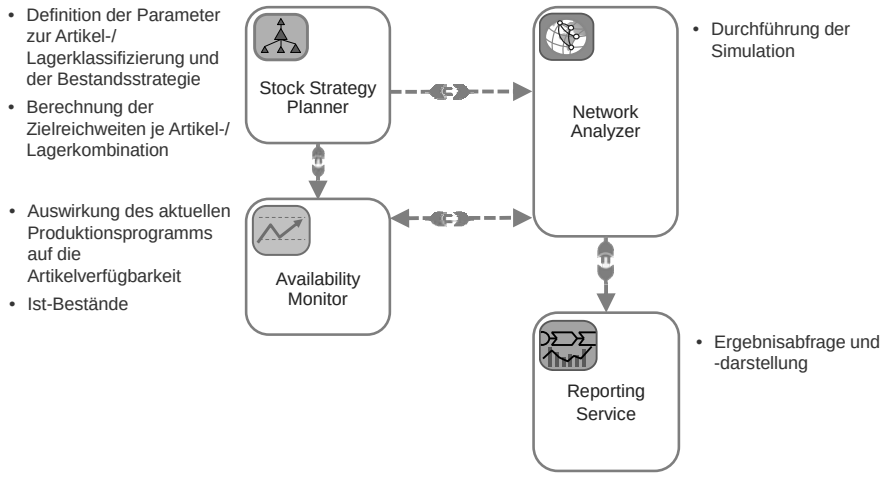



*Abbildung 1: Services des entwickelten Logistischen Assistenzsystems*

Für die operative Planung wird diese Planungskonfiguration dem *Availability Monitor* übergeben, der die erzeugte Konfiguration um Informationen zum Produktionsprogramm anreichert und durch die Kopplung mit dem *Network Analyzer* simulationsgestützt analysiert. Damit können die Auswirkungen des aktuellen Produktionsprogramms auf die Artikelverfügbarkeit des konfigurierten Planungsszenarios ermittelt werden.

Im Rahmen der taktischen Planung wird die vom *Stock Strategy Planner* erzeugte Planungskonfiguration direkt an den *Network Analyzer* übergeben, der die Definition von taktischen Planungsszenarien und die Überführung in Simulationsszenarien ermöglicht. Die Simulationskomponente dieses Bausteins ist die am Fraunhofer IML entwickelte Simulationstoolsuite OTD-NET (Wagenitz 2007).

Mit dem *Reporting Service* können die Ergebnisse des definierten Simulationsszenarios schließlich ausgewertet und aufbereitet werden. Dazu werden zum einen Bestände, Servicelevel und Lieferverzug artikel- und lagerspezifisch ausgewiesen und zum anderen die Simulationsergebnisse in Form von artikel- und relationsspezifischen Kennzahlen zur Transportplanung aufbereitet (Transportfrequenzen, Transportauslastung etc.). Aus diesen Ergebnissen lassen sich somit relevante Kosten- und Leistungskennzahlen ableiten, die dem Planer als Entscheidungsgrundlage dienen.

Die Funktionen der entwickelten Services werden von den derzeit in Unternehmen eingesetzten ERP- und SCM-Systemen nur teilweise abgedeckt (vgl. auch Witthaut 2013). Insbesondere die Möglichkeit, mithilfe der Simulation Planungsszenarien im Rahmen der operativen und taktischen Planung bereits vor der Einsteuerung in das reale System dynamisch bewertbar zu machen, ist mit den bestehenden Systemen nicht gegeben. Darüber hinaus wird den verantwortlichen Planern erstmals ein Methodenbaukasten zur Verfügung gestellt, der elementare fachliche Funktionen zur

Netzwerkplanung in Form von vier Services bündelt und es somit ermöglicht, Planungsprozesse fundiert, transparent und effizient zu gestalten.

Die Einbindung der vier entwickelten Services in die bestehenden Prozesse erforderte die Berücksichtigung von Schnittstellen zu anderen Systemen. So ist auf der einen Seite für die Erzeugung der Planungsszenarien eine Vielzahl an Stammdaten erforderlich (bspw. Absatzhistorie oder Soll-Servicelevel). Auf der anderen Seite ist die Möglichkeit zur Einbindung der Simulationsergebnisse in bestehende Planungssysteme notwendige Voraussetzung für eine erfolgversprechende Implementierung der entwickelten Services (vgl. Abb. 2).

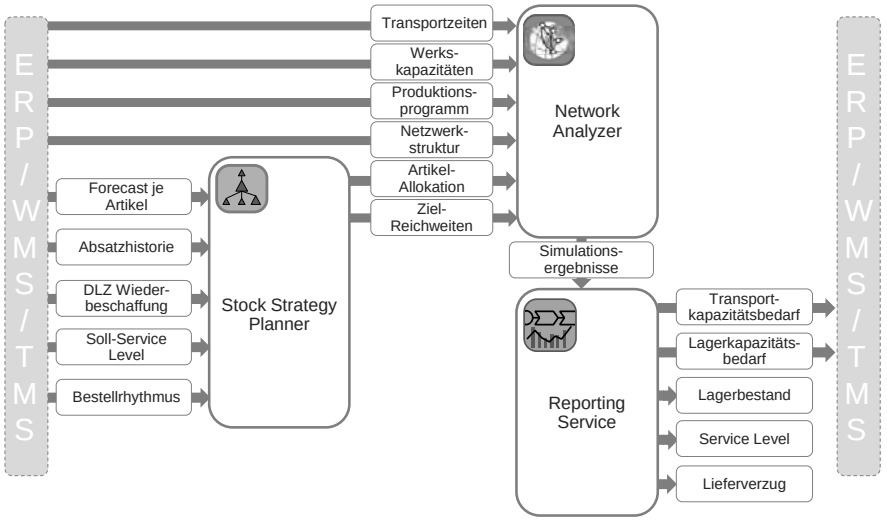



*Abbildung 2: Planungsservices und Schnittstellen zu bestehenden IT-Systemen*

# 4 Entwicklung und Einsatz des LAS

## 4.1 Systementwicklung

Im Rahmen der technologischen Entwicklung erfolgte zunächst die Spezifikation und Entwicklung der Systemstruktur für die einzelnen Komponenten des zu entwickelnden Assistenzsystems. Darauf aufbauend wurden die methodisch spezifizierten Planungsservices technologisch umgesetzt und in die entwickelte Systemstruktur implementiert. Das Ergebnis der Technologieentwicklung stellt ein Assistenzsystem zur taktischen Bestandsplanung dar. Das methodische Konzept wurde in eine Referenzarchitektur überführt, welche die Komponenten, Schnittstellen und Prinzipien festlegt, mit denen serviceorientierte logistische Assistenzsysteme realisiert werden können (Kamphues et al. 2013). Die entwickelte Referenzarchitektur bildet die technologische Basis, um die Simulation aufwandsarm in operative Planungsprozesse einzubinden.

Für den beschriebenen Anwendungsfall wurde darüber hinaus ein Frontend (GUI) entwickelt, das den Workflow für die taktische Entscheidungsunterstützung in der Bestandsplanung steuert (vgl. Abb. 3). Alle technisch notwendigen Kommunika-

tionsstrukturen bleiben für den Benutzer verborgen, sodass das Frontend als „Cockpit“ für die Assistenzsystemsteuerung fungiert und den Benutzer nacheinander durch die fachliche Parametrierung des Simulationsszenarios, die Simulationssteuerung und die Ergebnisauswertung führt. Dies gewährleistet eine intuitive Bedienbarkeit, die sich auf die Planungs- und Steuerungsfunktionalitäten beschränkt.

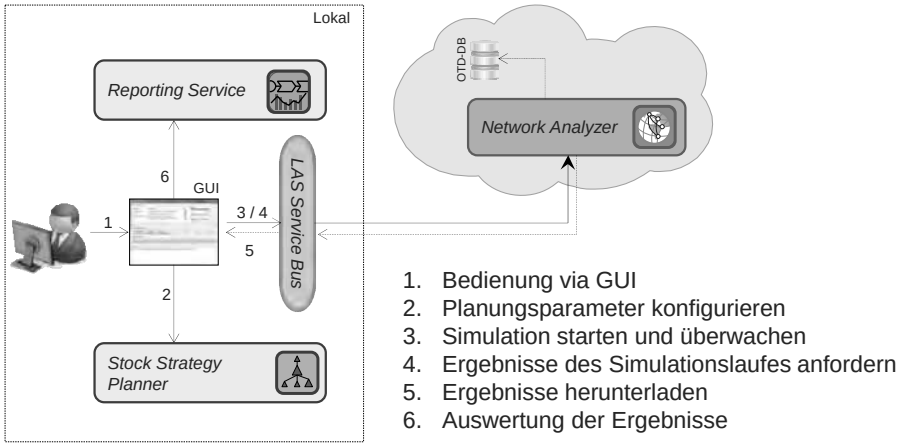



*Abbildung 3: Workflow für das Logistische Assistenzsystem*

## 4.2 Validierung und Testbetrieb

Die Validierung des entwickelten Logistischen Assistenzsystems erfolgte zusammen mit dem Anwendungspartner. Dazu wurden zunächst Fallstudien validiert, um die fachlichen Ergebnisse des Prototyps zu prüfen. Darüber hinaus erfolgten Funktionstests, um die technologische Einsatzfähigkeit zu ermitteln. Als Ergebnis der Validierung und des Testbetriebs wurden Verbesserungspotentiale abgeleitet, die für die Verwertungsabsichten der erzielten Ergebnisse von besonderer Relevanz sind.

Parallel zur methodischen und technologischen Entwicklung wurden im Rahmen von mehreren Treffen mit dem Anwendungspartner unterschiedliche Fallstudienvalidierungen vorgenommen, um zum einen die Planungsmethodiken selbst und zum anderen deren Ergebnisse mit den Prozessen und Gegebenheiten im Anwendungsfall abzugleichen. Ein Schwerpunkt war hierbei die Validierung des entwickelten Simulationsmodells, das die Grundlage für die Auswertungen der Planungsdienste darstellt. Aus diesem Grund erfolgte die Kalibrierung eines Referenzszenarios, das den realen Gegebenheiten im Distributionsnetzwerk des Anwendungspartners weitestgehend entsprach. Mithilfe von Kennzahlen zu Bestand, Servicegrad und Lieferverzug für verschiedene Testartikel konnten die realen Gegebenheiten der Vergangenheit mit den Ergebnissen der Simulation verglichen werden.

Von dem Referenzmodell ausgehend erfolgte dann die Validierung der Planungsmethodik, indem die mithilfe der Planungsbausteine definierten Szenarien anhand einer beim Anwendungspartner modellhaft umgesetzten Bevorratungsstrategie überprüft wurden. Dazu wurden die folgenden Schritte durchlaufen:

1. Abgleich der mithilfe des *Stock Strategy Planner* kalkulierten Ziel- und Sicherheitsbestände mit den beim Anwendungspartner implementierten Werten

2. Dynamische Validierung der analytisch ermittelten Dispositionsparameter hinsichtlich der Auswirkungen auf Servicelevel und Kosten mithilfe der Simulationskomponente (*Network Analyzer*)
3. Sukzessive Einsteuerung der kalkulierten und simulativ abgesicherten Ziel- und Sicherheitsbestände in die Operativsysteme.

Die Planungsmethodik konnte nach dem hier skizzierten Vorgehen erfolgreich validiert werden. Durch die Möglichkeit, die analytisch ermittelten Sicherheitsbestände mithilfe der Simulation dynamisch zu validieren, wurde dem Disponenten eine hilfreiche Planungsunterstützung geboten. Für dezidierte Standorte konnten Artikel gemäß des erforderlichen Handlungsbedarfes kategorisiert werden. So konnte im Rahmen einer modellhaften Bevorratungsstrategie aufgezeigt werden, dass lediglich 8 % der Dispositionseinheiten (entspricht 10 % des umgesetzten Volumens) einen Sicherheitsbestand aufwiesen, der maximal 10 % von dem kalkulierten und simulativ abgesicherten Wert abwich. Dem Gegenüber mussten die Sicherheitsbestände für 22 % der Dispositionseinheiten (entspricht 17 % des umgesetzten Volumens) erhöht und für 70 % der Dispositionseinheiten (entspricht 73 % des umgesetzten Volumens) verringert werden (vgl. Abb. 4).

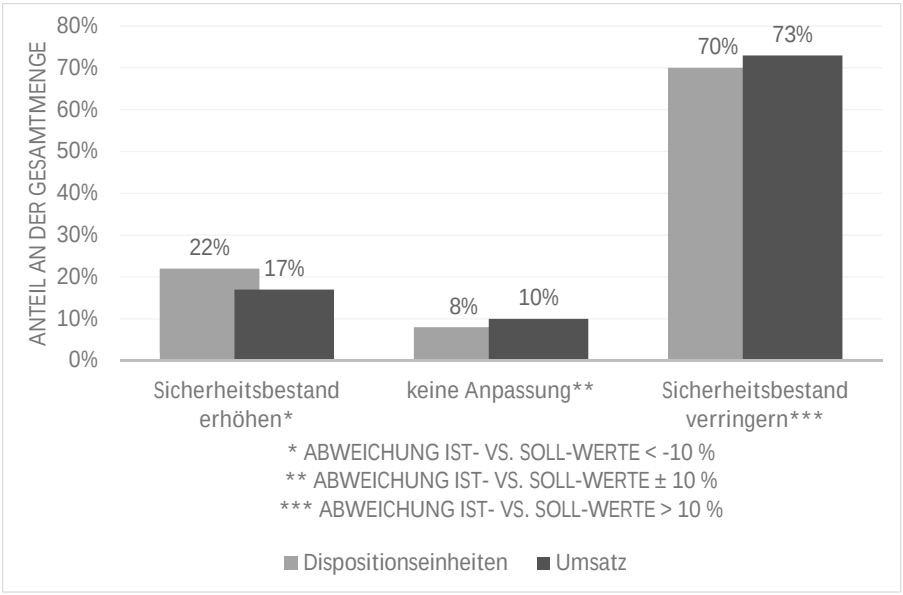



*Abbildung 4: Kategorien für den Handlungsbedarf im Rahmen der Disposition*

Nach Abschluss der Validierungsphase erfolgten die finalen Funktions- und Leistungstests des Prototyps. Dabei konnten die Stärken und Verbesserungspotentiale ermittelt und aufgezeigt werden. Der Prototyp erwies sich insbesondere aufgrund der im Folgenden aufgeführten Punkte als geeignet und zweckmäßig:

1. Die auf MS-Excel basierende Datenaufbereitung und -auswertung bietet den Vorteil eines individuellen und flexiblen Datenhandlings auf Basis einer weit verbreiteten Standard-Software. Dies führt auf der einen Seite zu einer vereinfachten Implementierung der methodischen Ansätze und bietet auf der

anderen Seite die Möglichkeit, die entwickelten Methoden zu adaptieren oder zu erweitern (Skalierbarkeit).
2. Die entwickelte Referenzarchitektur bietet eine systemunabhängige Integration des Simulators, sodass Planungsaufgaben abteilungs- und standortübergreifend durchgeführt werden können.
3. Das Frontend gewährleistet zum einen das reibungslose Zusammenwirken von lokalen und Web-Services und zum anderen eine intuitive Bedienbarkeit der entwickelten Systeme.
4. Die Möglichkeit, die Lastdaten für die Simulation einfach zu skalieren, führt zu einer erhöhten Akzeptanz für den operativen und taktischen Einsatz.

Die Validierung und die Funktionstests der entwickelten Methodiken und Werkzeuge konnten die Möglichkeit zur nutzbringenden Verwertung des Assistenzsystems demonstrieren, sodass das System für einen intensiven operativen Testbetrieb genutzt und ein Einsatz für andere Distributionsnetzwerke vorbereitet wurde.

## 5 Abschließende Bemerkungen

Die im Anwendungsfall gewonnen Erfahrungen haben gezeigt, dass mit einem simulationsbasierten Planungsansatz Bestandsreduktionen und Serviceverbesserungen in mehrstufigen Distributionsnetzwerken erzielt werden können. Eine Voraussetzung zur Bildung des Simulationsmodells war die Integration planungsrelevanter Daten aus unterschiedlichen IT-Systemen. Hierbei zeigten sich in der Praxis die größten Hürden hinsichtlich Datenverfügbarkeit und Datenkonsistenz. Für die weitere Forschung erscheint eine Erweiterung des LAS nutzbringend. Hierzu zählen GUI-Komponenten, mit denen Disponenten Änderungen an der Netzwerkstruktur oder den Materialflussprozessen vornehmen könnten.

## Danksagung

Diese Forschungsarbeit baut auf den Ergebnissen des vom Bundesministerium für Bildung und Forschung geförderten EffizienzCluster LogistikRuhr Projekt „Supply Chain Planning“ auf.

## Literatur

Axsäter, S.: Inventory control. Heidelberg et al.: Springer Science & Business Media 2007.

Blutner, D.; Cramer, S.; Krause, S.; Mönks, T.; Nagel, L.; Reinholz, A.; Witthaut, M.: Assistenzsysteme für die Entscheidungsunterstützung. In: Buchholz, P.; Clausen, U. (Hrsg.): Große Netze der Logistik. Berlin: Springer 2009, S. 241–270.

Englberger, J.; Herrmann, F.; Claus, T.: Simulationsbasierte Ermittlung von Kapazitätsbelastungsfaktoren zur Produktionsprogrammplanung. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik 2013. Paderborn: HNI 2013, S. 631–640.

Gudehus, T.: Dynamische Disposition: Strategien, Algorithmen und Werkzeuge zur optimalen Auftrags-, Bestands- und Fertigungsdisposition. Berlin: Springer 2012.

Güller, M.; Koc, E.; Noche, B.: A comparison study of multi-objective metaheuristic techniques for continuous review stochastic inventory system. In: 1st International Symposium on Computing in Informatics and Mathematics (ISCIM 2011), Tirana-Durres (Albania) 2011, S. 591–600.

Heinrichmeyer, H.: Leistungsbewertung verschiedener Optimierverfahren für das p-hub-Problem. In: Große Netze der Logistik: Springer 2009, S. 129–151.

Kamphues, J.; Groß, S.; Korth, B.; Zajac, M.; Hegmanns, T.: Serviceorientierte Referenzarchitektur für Logistische Assistenzsysteme zur simulationsbasierten Entscheidungsunterstützung. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik 2013. Paderborn: HNI 2013, S. 145–155.

Kang, J.-H.; Kim, Y.-D.: Coordination of inventory and transportation managements in a two-level supply chain. International Journal of Production Economics 123 (2010) 1, S. 137–145.

Lange, F.; Herrmann, F.; Claus, T.: Simulationsbasierte Kapazitätsabschätzung in einem Modell der Auftragsfreigabe mit Berücksichtigung von Servicegraden. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik 2013. Paderborn: HNI 2013, S. 591–600.

Miranda, P.A.; Garrido, R.A.: Inventory service-level optimization within distribution network design problem. International Journal of Production Economics 122 (2009) 1, S. 276–285.

Simchi-Levi, D.; Chen, X.; Bramel, J.: The logic of logistics: theory, algorithms, and applications for logistics management: New York: Springer Science & Business Media 2013.

Stumvoll, U.; Nehls, U.; Claus, T.: Ein simulationsbasiertes Assistenzsystem zur Pflege von Dispositionsparametern eines ERP-Systems im laufenden Betrieb. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik 2013. Paderborn: HNI 2013, S. 569–578.

Sucky, E.; Hönscheidt, I.: Dynamische Konfiguration von Logistiknetzwerken. In: Becker, J.; Knackstedt, R.; Pfeiffer, D. (Hrsg.): Wertschöpfungsnetzwerke. Heidelberg: Physica 2008, S. 133–152.

Tempelmeier, H.: Bestandsmanagement in Supply Chains. Norderstedt: Books on Demand 2012.

Wagenitz, A.: Modellierungsmethode zur Auftragsabwicklung in der Automobilindustrie. Dortmund: Praxiswissen 2007.

Witthaut, M.: Supply-Chain-Management-Systeme. In: ten Hompel, M. (Hrsg.): IT in der Logistik 2013/2014. Stuttgart: Fraunhofer Verlag 2013, S. 144–160.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Konzeptionierung eines integrierten modellbasierten Ansatzes zur Prognose von transportlogistischen und intralogistischen Ereignissen in Logistiknetzwerken

## *Design of an Integrated Model-based Approach to Predict Transport- and Intralogistical Events in Logistics Networks*

Michael Schröder, Neil Jami, Fraunhofer ITWM, Kaiserslautern (Germany),
michael.schroeder@itwm.fraunhofer.de, neil.jami@itwm.fraunhofer.de

Ulrike Beißert, Marco Motta, Fraunhofer IML, Dortmund (Germany),
ulrike.beissert@iml.fraunhofer.de, marco.moota@iml.fraunhofer.de

Abstract: The aim of this paper is to use an integrated approach of discrete event simulation and analytic prognosis models to allow a comprehensive investigation of complex supply chain processes in automotive industry and their interdependencies. The scenario considered is a supply chain of the car manufacturer Volkswagen with the task to manage bottlenecks of certain parts' supplies. Early anticipation of possible bottlenecks on first and second tier will be enabled by a detailed simulation model that is supported by analytical prognosis routines. The latter are calibrated from operational data collected in a longer time frame. The research is conducted in the collaborative research project “Visual Logistics Management (VILOMA)” funded by the Federal Ministry for Economic Affairs and Energy.

## 1 Einleitung

Im Rahmen des Managements von Wertschöpfungsnetzwerken sind unterschiedliche Fragestellungen zu beantworten, welche die Planenden aufgrund der Komplexität der Netzwerke (Wenzel und Bernhard 2008; Kuhn und Toth 2008) vor große Herausforderungen stellen. Strategische Fragen wie die Lokalisation, Größe und Konfiguration der Standorte lassen sich nur unter Berücksichtigung der Produktvielfalt und der Volatilität der Märkte beantworten. Taktische und operative Fragestellungen zielen auf die Konfiguration verschiedener Steuerungspolitiken z. B. für das Bestandsmanagement und die Distribution. Aus den genannten Fragestellungen leiten sich komplexe mathematische Probleme ab, zu deren Lösung vielfältige Daten, Einflussgrößen und Wechselwirkungen zu berücksichtigen sind.

In der Vergangenheit haben sich Prognosemodelle wie analytische Methoden und Simulationen zur Untersuchung des komplexen Systemverhaltens in Supply Chains etabliert. Analytische Methoden können genutzt werden, um z. B. zukünftige Systemzustände zu antizipieren und aufbauend hierauf geeignete Maßnahmen abrufbereit vorzuhalten. Demgegenüber eignet sich die Simulationstechnik zur Untersuchung der Wirkzusammenhänge in dynamischen Systemen (Reeker et al. 2011; Botta-Genoulaz et al. 2008). Beide Modellformen eignen sich für unterschiedliche Fragestellungen und finden in Logistiknetzwerken i.d.R. auf strategischer und taktischer Ebene Anwendung. Im operativ betriebsbegleitenden Einsatz konnte die Anwendbarkeit der Simulation prototypisch in der Automobilindustrie nachgewiesen werden (Liebler et al. 2013; Bockholt 2012); analytische Methoden sind im operativen Betrieb bislang jedoch wenig verbreitet, da die mathematischen Modelle für realitätsnahe Falluntersuchungen oftmals zu komplex werden.

Der Einsatz von Simulationen zur operativen Steuerung gewinnt im Kontext des Lieferkettenmanagements zunehmend an Bedeutung. Hierbei entscheiden u. a. die Datenqualität und gewünschte Ergebnisgranularität über die Verwendung der Simulation und bedingen den Abstraktionsgrad der Modelle. Gerade mit Hinblick auf die zunehmende Datenflut in den Netzwerken eignet sich hier der Einsatz analytischer Methoden, um die verfügbaren Datenmengen der Lieferkette zeitnah und detailliert analysieren und aufbereiten zu können. Die Verknüpfung der analytischen Methoden mit ereignisdiskreter Simulation zu einem vollständig modellbasierten Ansatz bietet im Kontext des Lieferkettenmanagements die Möglichkeit Ereignisse und Ereignisketten zu prognostizieren, aufbauende Systembewertungen durchzuführen und Konsequenzen abzuleiten, die in Entscheidungsprozesse der Disponenten einfließen. Im Rahmen des Forschungsprojektes Visual Logistics Management (VILOMA) werden die Möglichkeiten und Potenziale der Integration von analytischen Methoden und Simulation für Logistiknetzwerke anhand eines Anwendungsfalls aus der Automobilindustrie untersucht. Ein integrierter Forecast von transportlogistischen und intralogistischen Ereignissen soll möglich werden, als Grundlage zukünftiger Entscheidungsunterstützungssysteme für die Disponenten in der Lieferkette.

## 2 Konzept für die operative Steuerung von Lieferketten mittels Simulation und Prognose

Der Horizont operativer Entscheidungen in Lieferketten ist kurzfristig, auf der Skala von Minuten im intralogistischen Kontext bis zu einigen Tagen im Bereich der Transportlogistik. Disponenten orientieren sich an Erfahrungswissen und gleichen dieses mit den verfügbaren Daten ab. Dabei werden in der Regel mehrere Handlungsalternativen intuitiv überprüft und bewertet. Der Entscheidungsprozess wird in kurzer Zeit durchlaufen.

Im Gegensatz zu strategischen und taktischen Entscheidungen, bei denen mittels Simulationsstudien die Auswirkung der Entscheidung auf die zukünftigen Entwicklungen und die Systemperformance überprüft und abgesichert werden kann, muss die Einbettung von Simulationstechniken in operative Entscheidungskontexte anders gestaltet werden. Ein Disponent kann keine umfängliche Simulationsstudie vornehmen, bevor er eine steuernde Entscheidung in der Intra- oder Transportlogistik trifft. Die Herausforderung besteht daher in der Gestaltung der Simulations- und

Prognosemodelle in einer der kurzfristigen Entscheidungsfindung angemessenen Weise und ihre Integration in die Arbeitsumgebung des Disponenten.

Im Folgenden werden einige Aspekte dargestellt, wie mit einem integrativen Ansatz aus ereignisdiskreter Simulation und analytischen Prognosemodellen dieser Herausforderung begegnet werden kann.

## 2.1 Ereignisdiskrete Simulation zur Berechnung von Ablaufszenarien

Simulationsmodelle für logistische Abläufe werden in der Regel mit den Methoden der ereignisdiskreten Simulation gestaltet. Die sich *kontinuierlich* verändernden Variablen, wie die Bewegung eines LKWs auf seiner Route (Ortskoordinate) haben gegenüber den *diskreten logistischen Ereignissen* wie der Abfahrt, dem Erreichen von wichtigen Routenpunkten oder dem Eintreffen am Zielort geringere Bedeutung. Für die Simulation der Systementwicklung und ihre Bewertung reicht daher das Nachverfolgen diskreter logistischer Ereignisse aus.

Der Systementwicklungspfad in einer ereignisdiskreten Simulation wird durch eingebaute Zufallsgrößen und hinterlegte Steuerungsregeln beeinflusst. Die Zufallsgrößen modellieren die unbekannten Größen des stochastischen Umfeldes des simulierten Prozesses, z. B. die Transportzeit zwischen zwei logistischen Knoten. Die Steuerungsregeln simulieren dispositive Entscheidungen im Prozess. Hierbei findet eine Vereinfachung statt, da Disponenten nicht nur nach festen Regeln entscheiden, sondern Kontext- und Erfahrungswissen nutzen, das im Simulationsmodell nicht abgebildet ist. Gegebenenfalls werden die Steuerungsregeln durch Zufallsparameter randomisiert, um das nicht vollständig vorhersehbare Entscheidungsverhalten von Disponenten anzunähern.

Ein einzelner Lauf eines ereignisdiskreten logistischen Simulationsmodells kann nur *einen* möglichen Entwicklungspfad des Systems antizipieren. Dieser wird in der Realität fast sicher nicht exakt so eintreten. Verschiedene Entwicklungspfade entstehen durch Variation der Werte von randomisierten Variablen. Dazu muss ihre Zufallsverteilung zumindest ungefähr bekannt sein. Analytische Prognosemodelle können zur Gewinnung der Zufallsverteilungen aus gemessenen Prozessdaten genutzt werden (s. u.). Somit können mehrere Läufe durchgeführt werden und es stellen sich auch in den Ergebnisgrößen des simulierten Prozesses Verteilungen ein.

Anhand dieser Output-Verteilungen lässt sich die Vorhersagegenauigkeit zumindest in Grenzen beurteilen. Eine große Streuung eines Ergebniswertes (z. B. Reichweite des Lagerbestandes eines Einbauteiles) zeigt, dass dieser Wert in der Vorhersage sehr ungewiss ist. Der Disponent kann daher risiko-minimierende Maßnahmen ergreifen, im Beispiel für das Risiko eines Material-Engpasses. Eine kleine Streuung hingegen gibt ein gewisses Maß an Sicherheit, dass der Ergebniswert in der Realität nicht allzu weit von dem simulierten Ergebnis abweichen wird. Dabei ist vorausgesetzt, dass die Realität sich im Bereich der modellierten Einflussfaktoren bewegt, disruptive externe Faktoren wie Unfall des LKWs, Streik, katastrophale Wetterereignisse usw. führen zu Systementwicklungen, die sich nicht im Rahmen des Modells abspielen, und für die die Modellergebnisse somit auch keine Aussagekraft haben.

Zusammengefasst eignen sich ereignisdiskrete Simulationsmodelle gut, um zukünftige operative logistische Abläufe auf möglichen Systementwicklungspfaden zu

antizipieren. Je mehr Modellläufe unter Variation der Zufallsgrößen durchgeführt werden können, desto aussagekräftiger ist das Ergebnis, da sich dann auch die Prognosequalität von Outputgrößen über Streuungsmaße einschätzen lässt.

## 2.2 Analytische Prognosemodelle zur Gewinnung von Modellgrößen und die Generierung von Ereignissen

Nicht jede Variable eines ereignisdiskreten Simulationsmodells für operative logistische Abläufe muss als Zufallsgröße modelliert werden. Stabile Messwerte für Mengen oder Zeiten können als fixe (Durchschnitts-)Werte in das Modell eingebettet werden. Viele Einflussgrößen schwanken in der Realität jedoch, und müssen deshalb über Zufallsvariablen modelliert werden, z. B. die Transportdauer einer längeren LKW-Fahrt. Es muss in diesem Beispiel somit eigens eine Prognosefunktion realisiert werden, die aus dem Zeitpunkt des Abfahrtsereignisses den Zeitpunkt des Ankunftsereignisses prognostiziert:

$$T_{an} = T_{ab} + X \tag{1}$$

wobei $X$ die zufällige Fahrtzeit darstellt. Die stochastische Verteilung dieser Zufallsgröße muss dabei aus vorhandenen Daten abgeschätzt werden. Liegt eine repräsentative Menge an Fahrten mit den Zeitpunkten von Abfahrt und Ankunft vor, ist dies (bei Annahme stochastischer Unabhängigkeit) einfach über die empirische Verteilung zu machen. Andernfalls müssen aufwändigere statistische Verfahren eingesetzt werden, die $X$ in Korrelation mit anderen Zufallsgrößen des Modells abschätzen.

Neben Fahrtzeiten wie im Beispiel können auch die stochastischen Verteilungen für Umschlagzeiten, Bearbeitungszeiten, Zeiten intralogistischer Prozessschritte, usw. auf diese Weise ermittelt werden. Hinsichtlich Mengen-Größen ergeben sich analoge Ansätze, z. B. für die stochastische Modellierung der Nachfrage oder für schwankende Materialverbräuche.

Aber auch Zeitpunkte von externen Ereignissen lassen sich über analytische Prognosemodelle abbilden. Im Anwendungsfall, der in Abschnitt 3.2.3 dargestellt wird, sind dies sog. M100-Events, die den Synchronisationspunkt beim Übergang in die fahrzeugindividuelle Montagephase repräsentieren.

## 2.3 Integrierter Ansatz für operative Entscheidungsvorgänge

Der Einsatz von ereignisdiskreten Simulationsmodellen für operative Entscheidungsvorgänge unterscheidet sich von der Durchführung üblicher Simulationsstudien zur Absicherung von strategischen oder taktischen Entscheidungen. Die Simulationsläufe müssen im operativen Kontext weitgehend autark im Hintergrund ablaufen, Variationen von Modellgrößen über mehrere Läufe hinweg müssen automatisch generiert werden. Deshalb müssen in einem Großteil der Fälle analytische Prognosemodelle genutzt werden, um im Vorfeld der Nutzung des Simulationsmodells alle relevanten Zufallsgrößen abzuschätzen und darüber einen Experimentplan automatisch zu erzeugen.

Die notwendige Integration operativer Simulationsmodelle in Unterstützungssysteme für die Disposition bringt eine weitere Anforderung mit sich: auch die Parameter der Prognosemodelle gelten zumeist nicht konstant über längere Zeiträume

hinweg, und müssen deshalb laufend anhand der operativen Ist-Daten nachkalibriert werden.

Zusammenfassend stellt die Nutzung von ereignisdiskreten Simulationsmodellen in operativen Entscheidungsvorgängen vor die folgenden wesentlichen Herausforderungen:

- Integration der Simulationsmodelle in Unterstützungssysteme für die Disposition
- Automatische Ausführung im Hintergrund mit sehr kurzer Laufzeit
- Variation von stochastischen Modellgrößen über Verteilungen, die durch analytische Prognosemodelle aus den vorhandenen Daten laufend nachkalibriert werden
- Verdichtung der Simulationsergebnisse auf wenige, für den Disponenten aussagekräftige und entscheidungsrelevante Informationen.

Im Weiteren wird dargestellt, wie diese Ziele im Forschungsprojekt VILOMA umgesetzt werden.

# 3 Umsetzung für eine Lieferkette im Bereich der Automobilindustrie

## 3.1 Hintergrund: das Forschungsprojekt VILOMA

Das Projekt VILOMA (2015) wird seit 2013 vom Bundeswirtschaftsministerium mit dreijähriger Laufzeit gefördert. Es erforscht neue Ansätze für die Entscheidungsunterstützung von Disponenten durch Simulation, Prognose und Visualisierung.

Ein im Projekt untersuchtes Einsatzfeld ist das Engpassmanagement für die Lieferkette von Schaltermodulen für die Mittelkonsole des VW Golf 7. Je nach Ausstattung des Fahrzeugs weisen diese Schalter unterschiedliche Funktionen auf; insgesamt gibt es derzeit zwölf Varianten.

Das Projektkonsortium umfasst u. a. Volkswagen als OEM, die Tier1- und Tier2-Zulieferer für die Schaltermodule und die auf der Lieferkette operierende Spedition.

## 3.2 Drei Anwendungsfälle für die Integration im Bereich des Engpassmanagements

Für die Lieferkette der Schaltermodule wurde ein ereignisdiskretes Simulationsmodell mit der Software OTD-NET / OTD-InSite des Fraunhofer IML (Liebler et al. 2013) aufgebaut. Es umfasst transportlogistische und intralogistische Bausteine, letztere vor allem für Montageprozesse bei Tier1-Lieferant und OEM. Dieses Simulationsmodell der Lieferkette wird im weiteren Projektverlauf in eine prototypische operative Anwendung integriert, das VILOMA-System. Es richtet sich vorwiegend an Disponenten, mit dem Ziel, Transparenz in der Lieferkette über Unternehmen hinweg herzustellen und die Disponenten bei ihren steuernden Entscheidungen zu unterstützen.

Abbildung 1 zeigt den Komponentenaufbau des Simulationsmodells.

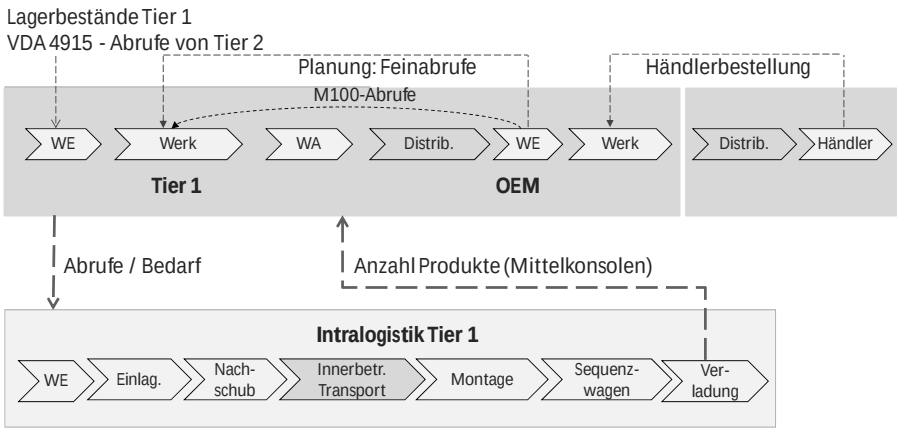



*Abbildung 1: Aufbau des ereignisdiskreten Simulationsmodells der Lieferkette.*

Das Simulationsmodell wird im VILOMA-System mit analytischen Prognosemodellen unterstützt, wie in Abschnitt 2.3 konzeptionell dargestellt. Dazu wird in mehreren Anwendungsfällen das Zusammenspiel von Simulation und Prognose konkretisiert, wobei unterschiedliche Prognosemodelle zum Einsatz kommen. Im Folgenden werden drei Fälle näher beschrieben.

### *3.2.1 Fall 1: Prognose dynamischer Einbauraten für Teilevarianten*

Die Schaltermodule werden in der Mittelkonsole links und rechts der Gangschaltung des VW Golf 7 eingebaut. Für das linke Schaltermodul gibt es derzeit neun Varianten, für das rechte drei. Sie gehen einher mit Ausstattungsmerkmalen des Fahrzeuges.

Das Simulationsmodell sieht beim Einbau von Teilevarianten den Parameter „Einbaurate“ vor. Dieser legt fest, mit welchem mengenmäßigen Anteil eine bestimmte Variante verbaut wird. Von Variante zu Variante gibt es erhebliche Unterschiede, manche werden täglich in mehrere hundert Fahrzeuge eingebaut, andere nur in geringen Mengen.

Ein einfacher Ansatz würde die Einbauraten aus Produktionsdaten der Vergangenheit ermitteln und für jede Variante des Schaltermoduls einen festen Wert der Einbaurate vorsehen. Das konzipierte Prognosemodell geht jedoch darüber hinaus, und prognostiziert für *jeden* Produktionstag die Einbauraten neu. Dies bezeichnen wir als *dynamische* Einbauraten.

Die Datengrundlage für die Prognose dynamischer Einbauraten sind die sogenannten *Feinabrufe*. Dies sind erwartete Teilemengen auf Tagesbasis. Sie liegen für drei Wochen im Voraus vor, und werden dem Tier1-Lieferanten vom OEM übermittelt. Allerdings schwanken sie bezogen auf einen festen Produktionstag nicht unerheblich, d.h. es gibt nicht *die* Feinabrufmenge für ein Produktionstag, sondern eine mit einem Vorlauf von drei Wochen startende Zeitreihe prognostizierter Bedarfe für jede Variante des Teils.

Das Prognosemodell für die dynamischen Einbauraten baut auf einer Analyse dieser Zeitreihen auf. Sie werden durch ein Modell abschnittsweise konstanter Ersatzwerte mit zufälligen Schwankungen um diese Werte charakterisiert (Abb. 2). Die Ange-

messenheit dieses Modells wurde durch eine umfangreiche datenanalytische Studie sichergestellt. Hinter Sprüngen des Ersatzwertes verbergen sich nach derzeitigem Kenntnisstand Umstellungen im Produktionsprogramm des OEM.

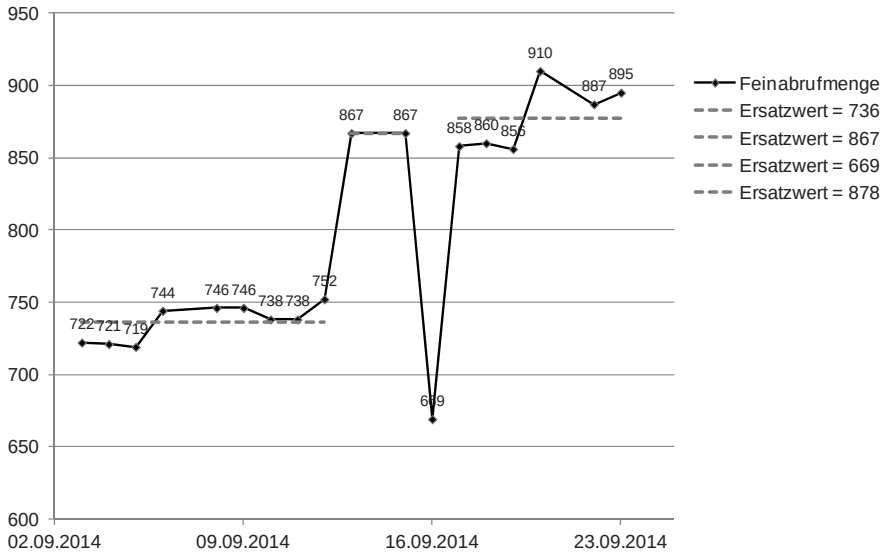



*Abbildung 2: Schwankende Feinabrufmengen werden vom Prognosemodell durch konstante Werte ersetzt, wobei Sprünge automatisch erkannt werden (Feinabrufe für Teil "5G1 927 137 T WZU" für den 23.9.14 und Ersatzwerte für das Prognosemodell)*

Vom Prognosemodell wird die Zeitreihe auf die konstanten Ersatzwerte reduziert, d.h. die als zufällig charakterisierten Schwankungen werden mittels eines fehlerminimierenden Anpassungsalgorithmus entfernt. Drei mögliche Werte bieten sich an, um die Bedarfsmenge / Einbaurate einer Teilevariante vorherzusagen:

- Letzter Ersatzwert (Aktualität)
- Maximaler Ersatzwert (Verringerung Engpassrisiko)
- Durchschnittswert der Ersatzwerte (Nutzung der gesamten Information)

Welcher dieser Werte für die Vorhersage genutzt wird, kann im Prognosemodell gemäß den Präferenzen des Disponenten eingestellt werden. Im Beispiel aus Abbildung 2 ergibt sich:

- Letzter Ersatzwert = 878
- Maximaler Ersatzwert = 878
- Durchschnittswert der Ersatzwerte = (736 + 876 + 669 + 878) / 4 = 789.75

### *3.2.2 Fall 2: Parameter für die Bestandsdisposition in der Simulation*

Wie in Abschnitt 2.1 dargestellt, werden im Simulationsmodell Entscheidungen, die in der Realität von Disponenten getroffen werden, durch (stochastische) Regeln abgebildet. Im hier betrachteten Fall ist es die Bestandsdisposition, d.h. die Nachbestellung von Teilen durch den Tier1-Lieferanten beim Tier2-Lieferanten. Letzterer

liefert die Schaltermodule, die auf Ebene des Tier1 in die fahrzeugindividuelle Mittelkonsole eingebaut werden.

Für die Bestandsdisposition werden in der Praxis Bestellmengen vom Disponenten festgelegt, wobei er in der Regel durch Algorithmen unterstützt wird, die Reichweiten und Sicherheitsbestände berücksichtigen. Die Festlegung einer Bestellmenge im Simulationsmodell kann diese zumindest in Teilen subjektive Entscheidungsfindung des Disponenten nicht vollständig nachbilden. Andererseits soll die Simulation dem Umstand Rechnung tragen, dass sich die Teilevarianten des Schaltermoduls hinsichtlich ihrer *Prognostizierbarkeit* unterscheiden.

Dazu wird die Simulation um ein Prognosemodell ergänzt, das eine XYZ-Analyse der täglichen Teilebedarfe vornimmt (Alicke 2005). Sie charakterisiert jedes Teil anhand des Variationskoeffizienten (Standardabweichung bezogen auf den Mittelwert) des täglichen Bedarfs in der Vergangenheit. Ein kleiner Wert des Variationskoeffizienten (kleiner als 0,2; X-Teil) erlaubt eine Bestandsdisposition ohne größere Sicherheitszuschläge, ein großer Wert des Variationskoeffizienten (größer als 0,5; Z-Teil) erfordert das Vorhalten eines größeren Sicherheitsbestandes. Mit dem Prognosemodell wird erreicht, dass in der Simulation für die Bestandsdisposition beim Tier1-Lieferanten nicht alle Varianten des Schaltermoduls gleich behandelt werden, sondern anhand ihres im zurückliegenden Zeitraum gemessenen Variationskoeffizienten des Bedarfs (Prognostizierbarkeit). Die Bestandsdisposition bildet somit nicht einfach die empirischen Bestellmengen mittels einer stochastischen Ersatzfunktion nach, sondern berücksichtigt die Prognostizierbarkeit des Bedarfs und entsprechende individuelle Sicherheitsbestände.

### 3.2.3 Fall 3: Abschätzung der zeitlichen Lage von M100-Events

Zwischen Tier1-Lieferant und OEM findet eine Just-in-Sequence-Belieferung statt. Mit einem Vorlauf von wenigen Stunden wird Tier1 über die zu liefernden Mittelkonsolen informiert. Dazu werden sog. M100-Events verwendet, dies sind Nachrichten welche die jeweils zu liefernde Mittelkonsole spezifizieren. Der M100-Event wird vom OEM ausgesendet, wenn der Produktionsprozess vom Karosserie-Rohbau in die fahrzeugindividuelle Montage übergeht. Der Tier1 montiert die Mittelkonsolen und liefert sie sequenzgerecht an das Band des OEM.

Im Simulationsmodell werden derzeit nur die Schaltermodule betrachtet, diese sind jedoch auch wie oben dargestellt von der fahrzeugindividuellen Ausstattung abhängig. Deshalb muss im Simulationsmodell berücksichtigt werden, dass die simulierten M100-Events unterschiedlichen Inhalt haben. Es wäre zum Beispiel fern der Realität, zunächst alle M100-Events für die erste Teilevariante zu generieren, dann für die zweite Teilevariante, usw. Vielmehr müssen die M100-Events der verschiedenen Varianten in einer realitätsgerechten Weise durchmischt werden.

Hierzu dient wiederum ein Prognosemodell, das die zeitliche Lage von M100-Events, bezogen auf eine jede Variante vorhersagt. Es basiert auf einer Analyse von typischen Mustern aus Vergangenheitsdaten. Die Analyse hat ergeben, dass eine Modellierung als *Poisson-Prozess* die Realität hinreichend gut abbildet (Abb. 3). Der Parameter des Prozesses kann aus der prognostizierten Gesamtmenge (prognostiziert wie in Abschnitt 3.2.1 dargestellt), für einen Produktionstag gewonnen werden, als mittlere Ankunftszwischenzeit. Die folgende Abbildung zeigt die gute Übereinstimmung von Beobachtung und der Prognose mittels Poisson-Prozess.

Es ist zu beachten, dass keine feste stochastische Ersatzfunktion zur Generierung der M100-Events herangezogen wird, sondern für jeden simulierten Produktionstag und jede Teile-Variante wird der Parameter des Poisson-Prozesses in der Simulation anhand der prognostizierten Produktionsmenge neu festgelegt.

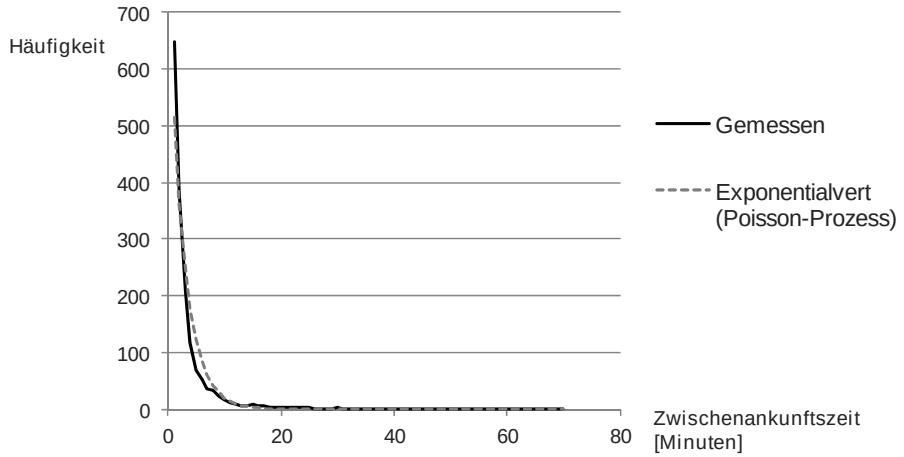



*Abbildung 3: Modellierung der M100-Abrufe für Teil "5G1 927 137 T WZ" als Poisson-Prozess. Die Verteilung der Zwischenankunftszeiten (Mittelwert = 2:48 Minuten) passt sehr gut zur Exponentialverteilung*

## 3.3 Umsetzung

Das VILOMA-System ist derzeit im Aufbau. Es wird das ereignisdiskrete Simulationsmodell für die Lieferkette der Schaltermodule auf Basis von OTD-NET und OTD-InSite (Liebler et al. 2013) enthalten, und dieses wird mit analytischen Prognosemodellen integriert. Dabei werden zunächst die in Abschnitt 3.2 beschriebenen Integrationsfälle beispielhaft umgesetzt, im weiteren Verlauf sollen dann weitere Prognosemodelle hinzukommen.

Das VILOMA-System wird laufend mit Daten aus den operativen Prozessen der Lieferkette versorgt werden. Diese werden über Integrationsroutinen in eine zentrale VILOMA-Datenbank eingespielt, die auf der In-Memory-Datenbank SAP HANA aufgesetzt wird. Benutzerseitig wird eine Webanwendung mittels SAP UI 5 realisiert, welche den verschiedenen Benutzungsfällen Rechnung trägt. Wie oben erwähnt, stehen Transparenz und Entscheidungsunterstützung für die Disponenten dabei im Vordergrund.

# 4 Zusammenfassung und Ausblick

Im vorliegenden Beitrag wurde das Konzept der Integration von ereignisdiskreter Simulation und analytischer Prognosemodelle für operative Abläufe in einer Lieferkette der Automobilindustrie vorgestellt. Hintergrund ist das derzeit laufende Verbundforschungsprojekt Visual Logistics Management (VILOMA) unter der Federführung von VW. Mit dem Ansatz, der anhand von drei Fallbeispielen konkretisiert

wurde, soll im Rahmen des sich in der Entwicklung befindlichen VILOMA-Systems Entscheidungsunterstützung im Engpassmanagement für Schaltermodule des VW Golf 7 realisiert werden. Disponenten sollen in die Lage versetzt werden, unternehmensübergreifend die nächsten Abläufe in der Lieferkette abzuschätzen und die Auswirkungen steuernder Maßnahmen einschätzen zu können.

## Danksagung

Das Projekt VILOMA wird vom Bundesministerium für Wirtschaft und Energie gefördert aufgrund eines Beschlusses des Deutschen Bundestages, Förderkennzeichen 19 G 13003D.

## Literatur

Alicke, K.: Planung und Betrieb von Logistiknetzwerken: Unternehmensübergreifendes Supply Chain Managment. Berlin, Heidelberg: Springer 2005.

Bockholt, F.: Operatives Störungsmanagement für globale Logistiknetzwerke - Ökonomie- und ökologieorientiertes Referenzmodell für den Einsatz in der Automobilindustrie. Dortmund: Praxiswissen 2012.

Botta-Genoulaz, V.; Lamothe, J.; Picard, F.; Riane, F.; Valla, A.: Discrete-event simulation for Supply Chain Management. In: Thierry, C.; Tomas, A.; Bel, G.: Simulation for Supply Chain Management. London: ISTE Ltd 2008, S. 69-102.

Kuhn, A.; Toth, M.: Assistenzsysteme für eine effektive Planung logistischer Netzwerke. In: Scholz-Reiter, B.: Technologiegetriebene Veränderungen der Arbeitswelt. Berlin: GITO 2008, S. 257-278.

Liebler, K.; Beißert, U.; Motta, M.; Wagenitz, A.: Introduction of OTD-Net and las: order-to-delivery network simulation and decision support systems in complex production and logistics networks. In: Pasupathy, R.; Kim, S.-H.; Tolk, A.; Hill, R.; Kuhl, M.E. (Hrsg.): Proceedings of the 2013 Winter Simulation Conference, Washington D.C. 2013, S. 439-451.

Reeker, C.; Hellingrath, B.; Wagenitz, A.: Bewertungsansatz zur ökologischen Gestaltung automobiler Logstiksysteme im Werksumfeld von Produktionsstandorten. In: Bogaschewsky, R.; Eßig, M.; Lasch, R.; Stölzle, W.: Supply Management Research. Wiesbaden: Gabler 2011, S. 139-168.

VILOMA: Visual Logistics Management. www.visuallogisticsmanagement.de, aufgerufen 16.08.2015.

Wenzel, S.; Bernhard, J.: Definition und Modellierung von Systemlasten für die Simulation logistischer Systeme. In: Nyhuis, P.: Beiträge zu einer Theorie der Logistik. Berlin, Heidelberg: Springer 2008, S. 487-509.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulation-based Decision Support for Vertical Supply Chain Segmentation Scenarios

## *Simulationsbasierte Entscheidungsunterstützung für Szenarien der vertikalen Supply-Chain-Segmentierung*

Sebastian Terlunen, Dennis Horstkemper, Martin Wölck, Bernd Hellingrath,
Westfälische Wilhelms-Universität Münster, Münster (Germany),
sebastian.terlunen@ercis.de, dennis.horstkemper@ercis.de, martin.woelck@uni-muenster.de, bernd.hellingrath@ercis.de

Abstract: In order to compete in high wage countries, many companies aim to deliver high quality products with specialised service offers to their customers instead of simply competing by employing lower prices. Therefore, these companies often have to segment their supply chains in multiple ways to satisfy the different expectations of a heterogeneous customer base – one size just doesn't fit all anymore. However, market dynamics and volatile customer behaviour make it hard to decide how exactly such a supply chain segmentation should be performed. These complex interactions make it impossible to solve this problem by applying exact optimisation algorithms. Thus, this work identifies an applicable simulation paradigm to provide decision support to viably define such segments. Moreover, needed adaptations of the chosen simulation paradigm are discussed conceptually and an implementation approach is presented. Hereby, our main research contribution lies in revealing the necessary adaptations and in integrating the appropriate planning approaches.

## 1 The Necessity of Evaluation Approaches for Vertical Supply Chain Segmentation

Supply chains nowadays operate in an increasingly complex, dynamic and uncertain business environment. Increased customer expectations, e.g. the demand for a high number of product variants and features, more responsive support services, and higher product availability, drive market dynamics. The resulting higher product variety causes unstable or unknown demands due to a higher multitude of products, high rates of new product introductions and shorter life cycles (Bozarth C. et al. 2009). To cope with these dual objectives of customer orientation and cost minimisation, especially in high-wage countries, the topic of supply chain segmentation has gathered substantial interest in research and practice (PricewaterhouseCoopersAG 2013; Hofman 2011). In this context, we focus on vertical supply chain segmentation (vSCS). A vertical segmentation separates the supply chain in two planning

areas, divided by means of the customer order decoupling point. The position of the customer order decoupling point massively influences multiple possibly contrary but interrelated target values, such as operational target values, e.g. cost minimisation, and market target values, e.g. the minimisation of delivery times (Aitken et al. 2005).

However, neither practical nor scientific literature offers appropriate best practices to determine the ideal vSCS due to the high number of factors influencing this decision. Furthermore, dynamic and turbulent market environments and the endeavour to align the supply chain processes more strongly towards the customer needs as well as the omnipresent target of cost minimisation imply that a specific vSCS will never be appropriate for long time frames. Consequently, a vSCS can only be applicable for a certain timeframe considering the specific environment of the producing company. Thus, a practical need for an efficient decision support for the determination of an optimal vSCS within applicable timeframes can be deduced.

For decision support throughout complex supply chains, several mathematical models exist (Hennies et al. 2014). These mathematical models can be used to analyse and benchmark the performance of different supply chain alternatives. They can be classified into analytical and simulation models (Law 2015). Analytical models, like queuing systems and linear programming models, describe the problem in a closed system of equations and can be solved exactly. They are, however, often difficult to formulate and usually very time consuming to solve (Page and Kreutzer 2005; Law 2015). Simulation is a method for analysing complex real or imaginary systems and facilitates experimenting with their variables to explore how the system will likely react to different changes (Page and Kreutzer 2005). As supply chains tend to become more complex and the demand situation uncertain, simulation is the tool of choice for many practitioners to test their supply chains in a fast manner (Terlunen et al. 2014).

For evaluating a specific vSCS, both types of mathematical models have been applied in research already. The first type of approaches uses linear equations to calculate the resulting material flow and costs (Hajfathaliha et al. 2010). As such, they are unable to take dynamic and volatile customer behaviour into account. To overcome this limitation, the second type of approaches employ discrete event-based simulation (Nienhaus 2004; Winkler 2010). Drawbacks of the simulation models applied in previous works are a negligence of customer needs and the non-existence of planning methods controlling the material and information flow within the two planning areas, leading to a loss of precision of these models. Besides, the named authors struggled to analyse complex networks within practicable timeframes.

Subsequently, to fulfil the practical requirements and to close the theoretical research gap, we devise the design of a simulation model being able to evaluate the level of customer orientation and cost effectiveness while overcoming the aforementioned limitations of previous work. To accomplish this goal, we structure our work along the following research questions: (1) How do the different supply chain segments differ in regards to material and information flows? (2) What planning methods and simulation paradigms are appropriate to be applied for the simulation of a segmented supply chain? (3) How can the chosen planning methods and the chosen simulation paradigm be integrated to enable an efficient simulation of a segmented supply chain?

To answer these questions, the remaining paper is structured as follows: First, section two explains the fundamentals of vSCS, describes the differences between order-driven and forecast-driven areas of a supply chain, and provides insights into the commonly applied planning methods for these areas. Furthermore, we discuss which simulation paradigms could be applied in theory for the given problem and argue for the most appropriate one. Within section three, we describe the required conceptual adaptations of the chosen simulation paradigm beyond already existing approaches that must be performed to foster the envisioned decision support. Afterwards, we describe our concrete implementation in section four and close with a summary and outlook towards further research in section five.

## 2 Planning and Simulation Approaches for Specific Supply Chain Segments

The vSCS is addressed by supply chain design task. It defines the separation of the supply chain into multiple sequentially connected value-adding steps, which are together either forming a forecast-driven or order-driven area. These two areas are dissociated by a buffer storage, the so-called customer order decoupling point. In the forecast-driven areas the value creation is being performed based on sales plans, which in turn were generated by using forecasts. Instead, in the order-driven areas, a direct relation to associated customer orders induces the value creation. The decision for the positioning of the customer order decoupling point should optimally be based on a previously performed identification of multiple logistical and economical customer requirements (Winkler 2010).

Generally, the evaluation of a specific vSCS scenario requires an analysis of the associated operational and market goals. While vSCS is considered as being a supply chain design task performed on the strategic level, the realisation of these goals should be evaluated within the tactical level to gain more detailed insights into the corresponding material flows and capacity utilisations. An analysis of the operational level is usually not possible yet, as exact customer orders cannot be forecasted with an appropriate accuracy and planning is thus performed on aggregated product groups.

In the following subsections, we describe (1) exemplary planning techniques which are appropriate to be used within these two different areas of value creation on the tactical level and (2) which simulation paradigm is appropriate to be employed in regards to the associated problem of evaluating a specific segmentation scenario.

### 2.1 Planning Methods within Supply Chain Segments

Scientific theory and practice both distinguish between forecast-driven and order-driven demand planning, in accordance to different supply chain segments. Different planning and controlling approaches are used within these distinguished areas and should be employed when simulating a segmented supply chain, to ensure that the overall simulation model acts accurately with regard to the real logistical system.

Forecast-driven demand planning and demand distribution throughout the supply chain on a tactical level is in most cases either performed through the interaction of local inventory policies or through centralised optimisation models. Systems based on inventory policies reissue supply processes whenever certain inventory amounts

fall below a predetermined level and can be considered as a heuristic (Dangelmaier 2009). While commonly used in practice, the limited exchange of customer demand information often leads to huge demand variations within a supply chain, which can culminate into the well-known Bullwhip Effect. In contrast, central planning approaches are based on detailed forecasts and try to find the optimal balance between inventory holding, production as well as transportation costs.

Order-driven demand planning and demand distribution throughout the supply chain is only performed based on concrete customer orders. In the easiest case, each participant of the supply chain simply propagates required secondary demands to fulfil their customer demands to the next upstream member of the supply chain. This approach however does not take any capacity bottlenecks into account (Albrecht 2010). In contrast, heuristics like capable-to-match can be used to find a feasible solution that considers capacity constraints. Employing a rule-based approach, capable-to-match propagates secondary demands to suppliers that can meet these demands within the required timeframes (Pibernik 2005).

All aforementioned approaches (as well as further specialised approaches which might be employed in practice) can be used to control the distribution of demands within a simulation of a supply chain. The choice of the used planning approaches hereby considerably influences the resulting operational and market target values. However, using these planning approaches alone does neither illuminate their interactions within a segmented supply chain, nor is it possible to attribute for demand volatility. Thus, it is useful to employ these planning approaches only as an element within an overall evaluation method for vSCS scenarios.

## 2.2 Evaluation of Paradigms for the Simulation of Supply chain Segmentation Scenarios

Conventionally, two fundamental classes of simulation concepts are considered in literature: discrete-event simulation (DES) and continuous simulation (CS) (Tolujew et al. 2010). While both approaches have received considerable attention, they have some fundamental drawbacks in the field of Supply Chain Management. Recently, a third approach, the discrete rate-based simulation (DRB) has also grown in popularity (Hennies et al. 2014).

According to a literature review conducted by Tako and Robinson, DES is the most often used simulation concept in Supply Chain Management (2012). It simulates the flow of individual objects, like different products, within a network of components (Page and Kreutzer 2005). State changes are caused by events, e.g. the change of the location of an object (Sokolowski 2009), and occur at discrete points of time. The state of the simulation model is defined by the state of all of its objects. It is assumed that no state changes occur between two discrete events, so that the simulation clock can be advanced from event to event (Law 2015). As DES can represent systems in various levels of detail, its areas of application are extensive. Still, it is particularly useful if the flow of individual goods, like the packet flow in a distribution centre, is considered (Scholz-Reiter et al. 2008). Nevertheless, DES can become very complex and time consuming if the number of entities is high.

In contrast, continuous simulation examines systems in which the state variables change continuously, i.e. uncountable infinite times during a fixed duration (Law 2015). In Supply Chain Management, CS is used to simulate the aggregated flow of

materials, information, and people (Reggelin 2011). Due to that, it is generally assumed that CS is used more often on a strategic level, which focuses on aggregated data and a long planning horizon, than on a tactical or operational level. This assumption, however, could not be proven in the literature review conducted by Tako and Robinson (2012). Still, compared to DES, CS is less often used in the domain of Supply Chain Management, unless the research is focusing on the Bullwhip Effect. CS is appropriate for systems, in which the state variables change continuously over time. Unlike DES, it does not model individual entities, but aggregated flow-rates. Practically, system states of CS models are recalculated in fixed time intervals $\Delta t$. The length of the $\Delta t$ is predetermined and influences runtime and accuracy.

While DES is particularly suitable if the flow of a limited number of discrete elements is considered, CS excels in scenarios with highly aggregated goods or liquids. Both approaches have, however, fundamental disadvantages: DES does not perform well if the number of discrete goods is particularly large, while CS might entail significant integration and rounding errors the longer the simulation runs.

In contrast, DRB simulation can be used to model systems, which are characterised by a huge amount of entities. This approach tries to mitigate the disadvantages of DES and CS by combining the event driven approach of DES with the continuous flows of CS (Krahl 2009). Possible examples are packaging processes, manufacturing processes, and the processing of liquids (Damiron and Nastasi 2008). The DRB paradigm was first described by Siprelle and Phelps (1997), who extended the commercial simulation software ExtendSim with the possibility to model high-speed bulk flow systems in a DRB manner. These extensions were additionally described by Krahl (2009) and by Damiron and Nastasi (2008).

Like CS, DRB does not model individual entities, but aggregated flows of homogeneous goods. While the flow-rates are assumed to change continuously over time in CS, DRB assumes that the flow-rates are piecewise constant and thus only change at discrete points of time. This assumption has decisive advantages, as it becomes possible to create linear-programming formulations for DRB systems (cf. Damiron and Nastasi (2008) for a detailed description). With this, it becomes possible to calculate the flow-rates of all model elements as well as to predict the exact points of time at which some events occur and advance the simulation clock in a next-event time approach. Furthermore, the DRB approach is based on model elements closely aligned to the logistical funnel model by Nyhuis (2003). Due to this inherent characteristics, DRB is particularly well-suited for supply chain simulations on tactical levels and will be used within this work (Terlunen et al. 2014).

## 3 Adaptation of the Discrete Rate-based Simulation Paradigm to Evaluate Segmented Supply Chains

In order to simulate a segmented supply chain the results of the aforementioned planning approaches should be employed to control the material flow in the overall simulation model. As explained in previous sections, the DRB simulation paradigm is considered as suitable for analysing segmented supply chains.

The most promising research results enabling the application of the DRB simulation paradigm was proposed by Reggelin, Hennies, Schenk and Tolujew, who named their approach “MesoScopic simulation” (Schenk et al. 2010; Reggelin and Tolujew

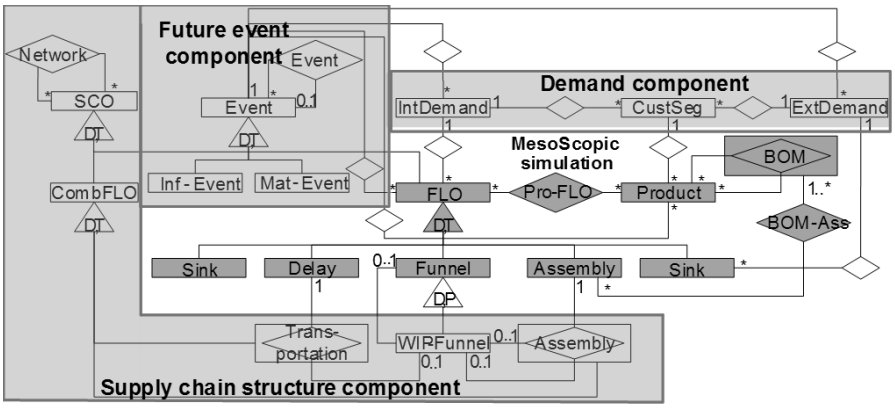



*Figure 1: Structural model of the discrete rate-based simulation paradigm*

2011). Summarising, the research group suggests adaptations in earlier proposed findings by other research groups, e.g. Krahl as well as Damiron and Nastasi. Their major contributions address the extension of DRB model elements in order to steer the material flows. Therefore, they introduced steering variables, which describe how the flow of goods is distributed, either by assigning quotas to model elements or by controlling the initial material flows at the first supply chain elements. The material flow cannot be routed according to customer demands, but only according to the predefined steering variables. Consequently, their approach cannot be used in conjunction with the aforementioned planning methods executed during a simulation run. Thus, it is impossible to evaluate a segmented supply chain accurately.

A demand-centric DRB approach for supply chains was first introduced by Terlunen et al. (2014), who extended the findings of Hennies et al. (2014) with the ability to model customer demands (Fig. 1). They proposed an orchestration of the originally proposed model elements to model complete supply chains by means of the standard processes production, transportation, storage, and supply (supply chain structure component). Further, they used the notion of external and internal demands to model the flow of demand information along the supply chain (demand component). While demand information is propagated upstream the network, products flow downstream the network in order to fulfill the propagated demands. For information about how the demand-centric DRB approach can be implemented to tactical supply chain simulation, we refer to Terlunen et al. (2014).

Using this approach, the results of different planning methods can easily be used to control the material flows within the supply chain. Lot-sizes generated by the planning approaches can be used to determine the internal demands at each supply chain element depicted by the simulation model. However, mechanisms to decide when and how often a planning method should be performed within a simulation are still needed.

A possible approach is to use the existing future events. As usually used in the DES simulation paradigm, e.g. by Kramer and Neculau (1998), we propose an implementation of a future event component in the DRB simulation paradigm, enabling a controlling of the information and material flow during a simulation run. The timespan between two events can be used to define variable time intervals. Furthermore,

control mechanisms and other operations can be performed at the occurrence of an event. We differentiate between three types of events: predefined events, control-based events and model-immanent events. Predefined events are used to control the information and material flow at simulation time and are determined before the simulation run. Control-based events are used to react when certain system states during a simulation run are reached, e.g. by issuing a resupply order when the information about low inventory amounts becomes known. Lastly, model-immanent events are employed to steer the material and information flow during a simulation run. E.g., when the material flow stops at one model element, model-immanent events are propagated downstream to further model elements in order to communicate the information that the material flow must be stopped.

## 4 Prototypical Implementation

To explain how the aforementioned event types can be facilitated to execute the implemented planning methods at the right point in time and to transform the results to internal demands into the simulation model, we exemplary describe such an implementation for the four planning methods explained in section two. After the initialisation of the simulation model at the beginning of a simulation run, the external customer demands enter the system at predetermined times. Afterwards, the external demands are transformed into internal demands, which can either be order-driven or forecast-driven. In accordance to the current supply chain segment and the aligned planning method, the internal demand is distributed in different manners upstream.

For forecast-driven segments of a supply chain we make use of inventory policies and network-planning models as possible planning methods. Concerning inventory policies, one can distinguish between continuous and discrete inventory level control mechanisms (Dangelmaier 2009). In order to implement a discrete inventory control mechanism, our proposed implementation makes use of predefined events. Hereby, the implementation is enabled to check existing inventory levels within the system at predefined points in time. If an inventory level of a product falls below a calculated report inventory level, two control-based event types are generated. The first control-based event causes an internal demand distribution until the next storage element upstream. Afterwards, the second control-based event initialises the start of the material flow at the corresponding storage, located upstream. In contrast, if a continuous inventory control mechanism is implemented, it has to be guaranteed that the inventory remains above a predefined level taking the incoming and outgoing material flow at a storage element into account at all times. Therefore, at every point in time at which the value of the incoming or outgoing material flow changes, a calculation is conducted determining the point in time when the predetermined level will be reached. Furthermore, the corresponding control-based events for the necessary demand distribution and the material flow calculation are generated in this case as well.

Centralised master planning in practice is often employed by using a rolling planning horizon (Fleischmann et al. 2005). Thus, the planning method is executed at constant intervals. Similar to the discrete inventory level control policies, predetermined events can be used to halt the simulation run, execute the master planning method, and to transform the results of the planning method into demands for the

different simulation elements. Furthermore, a recalculation of the material flows is also initiated afterwards by means of control-based events.

For order-driven segments of a supply chain we make use of an upstream-planning approach as well as the capable-to-match heuristic. Moreover, we assume that all customer demands are known as simulation parameters, which in term means that customer demands only enter the simulation model at predefined points in time. Thus, both of these planning approaches can be executed analogous to the centralised master planning in forecast-driven segments of the supply chain. For every point in time, when customer demands exist, the aforementioned two types of events are applicable. Predefined events are used to distribute demands to the corresponding supply chain elements and afterwards control-based events control the material flow.

Figure 2 gives an overview over the complete simulation run for a segmented supply chain in which the forecast-driven segments are controlled by a centralised master planning approach and the order-driven segments by a capable-to-match heuristic. After the decision maker starts the simulation run the supply chain structure and the corresponding necessary parameters are read from an input database. The centralised master planning and the capable-to-match heuristic can be implemented using external algorithms, for which an appropriate interface needs to be devised. For example, we have formulated a mathematical model for the centralised master planning, using the commercial solver CPLEX based on the findings of Pibernik and Sucky (2005). The simulation state can automatically be transformed into the necessary parameters to solve the optimisation problem, while the results are fed back into the simulation model as new demands. The capable-to-match heuristic acts analogous. Finally, results are written into an output database for further analysis.

## 5 Summary and Further Research

After explaining different characteristics of forecast-driven and order-driven segments of a supply chain as well as exemplary planning methods, we investigated multiple simulation paradigms in regard to their applicability to the task of simulating segmented supply chains. Due to its appropriate aggregation level, but especially because of the flexible event-management approach, we argued that the DRB simulation approach is applicable for this task. We described, how event management can be used to initialise and execute different planning methods, enabling us to create an accurate simulation model of an overall vSCS. When applying typical key performance values, it is possible to evaluate multiple segmentation scenarios efficiently. This wasn't possible in other approaches from literature that were designed for this task. Summarising, a viable decision support system for the determination and evaluation of an optimal vSCS scenario was devised.

However, the task of designing such a scenario has not been automated yet. Further research could be done to design decision support tools that automatically devise the position of the customer order decoupling point based on the simulation results. Furthermore, the implementation of alternative planning algorithms, e.g. to be able to simulate decentralised planning methods, should be considered to be able to accurately model more real-life supply chains. Lastly, computational experiments should be performed, to proof that the DRB simulation approach generates computational time savings in comparison to a DES approach.

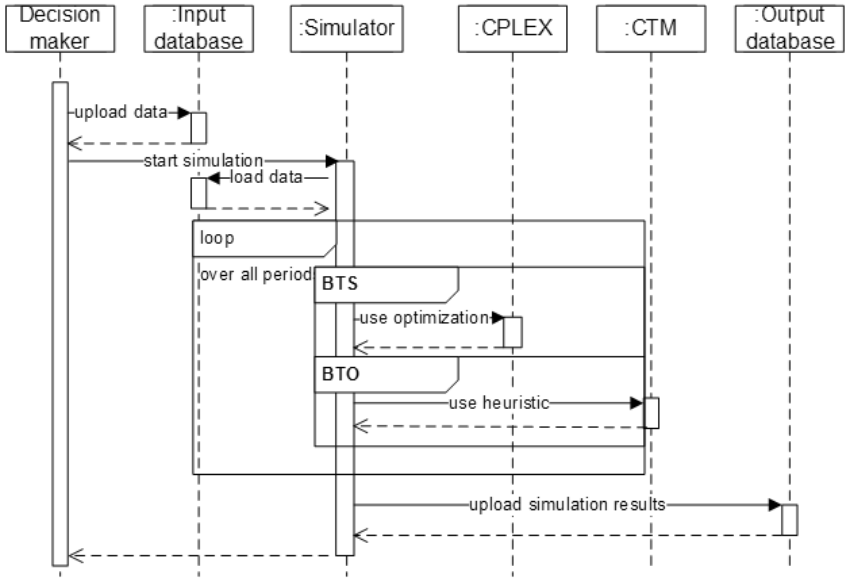



*Figure 2: Sequence diagram of the discrete rate-based simulation*

## References

Aitken, J.; Childerhouse, P.; Christopher, M.; Towill, D.: Designing and managing multiple pipelines. Journal of Business Logistics 26 (2005) 2, pp. 73-95.

Albrecht, M.: Supply chain planning and coordination. In: Supply Chain Coordination Mechanisms: Springer 2010, pp. 5-34.

Bozarth C.; Donald P. Warsing; Barbara B. Flynn; E. James Flynn: The impact of supply chain complexity on manufacturing plant performance. Journal of Operations Management 27 (2009) 1, pp. 78-93.

Damiron, C.; Nastasi, A.: Discrete rate simulation using linear programming. In: Proceedings of the 40th Conference on Winter Simulation, 2008, pp. 740-749.

Dangelmaier, W.: Theorie der Produktionsplanung und -steuerung: Im Sommer keine Kirschpralinen? Dordrecht, New York: Springer 2009.

Fleischmann, B.; Meyr, H.; Wagner, M.: Advanced planning. In: Supply Chain Management and Advanced 2005, pp. 81-106.

Hajfathaliha, A.; Teimoury, E.; Khondabi, I.G.; Fathi, M.: Using queuing approach for locating the order penetration point in a two-echelon supply chain with customer loss. International Journal of Business and Management 6 (2010) 1, pp. 258-268.

Hennies, T.; Reggelin, T.; Tolujew, J.; Piccut, P.-A.: Mesoscopic supply chain simulation. Journal of Computational Science 5 (2014) 3, pp. 463-470.

Hofman, D.: The supply chain top 25: leadership in action. Supply Chain Management Review (2011) 5.

Krahl, D.: ExtendSim advanced techology: discrete rate simulation. In: Winter Simulation Conference, Austin (TX) 2009, pp. 333-338.

Kramer, U.; Neculau, M.: Simulationstechnik. München: Fachbuchverlag Leipzig 1998.

Law, A.M.: Simulation modeling and analysis. New York, NY: McGraw-Hill Education 2015.
Nienhaus, J.: Modeling, analysis and improvement of supply chains – a structured approach. Zürich 2004.
Nyhuis, P.: Logistische Kennlinien: Grundlagen, Werkzeuge und Anwendungen. Berlin: Springer 2003.
Page, B.; Kreutzer, W.: The Java simulation handbook: Simulating discrete event systems with UML and Java. Aachen: Shaker 2005.
Pibernik, R.: Advanced available-to-promise: Classification, selected methods and requirements for operations and inventory management. International Journal of Production Economics 93 (2005) 1, pp. 239-252.
Pibernik, R.; Sucky, E.: Master planning in supply chains. In: Günther, H.-O.; Mattfeld, D.C.; Suhl, L.: (eds.) Supply Chain Management und Logistik. Heidelberg: Physica, pp. 69-93.
PricewaterhouseCoopersAG (Hrsg.): Next-generation supply chains: Efficient, fast and tailored. 2013. www.pwc.com/gx/en/consulting-services/supply-chain/global-supply-chain-survey/assets/pwc-next-generation-supply-chains-pdf.pdf
Reggelin, T.: Mesoskopische Modellierung und Simulation logistischer Flusssysteme. Magdeburg: docupoint 2011.
Reggelin, T.; Tolujew, J.: A mesoscopic approach to modeling and simulation of logistics processes. In: Winter Simulation Conference, Phoenix (AZ) 2011, pp. 1508-1518.
Schenk, M.; Tolujew, J.; Reggelin, T.: A mesoscopic approach to the simulation of logistics systems. In: Dangelmaier, W.; Blecken, A.; Delius, R.; Klöpfer, S. (eds.): Advanced Manufacturing and Sustainable Logistics. Berlin, Heidelberg: Springer 2010, pp. 15-25.
Scholz-Reiter, B.; Beer, C. de; Freitag, M.; Hamann, T.; Rekersbrink, H.; Tervo, J.: Dynamik logistischer Systeme. In: Nyhuis, P. (ed.): Beiträge zu einer Theorie der Logistik. Berlin: Springer 2008, pp. 109-138.
Siprelle, A.J.; Phelps, R.A.: Simulation of bulk flow and high speed operations. In: Winter Simulation Conference, Atlanta (GA) 1997, pp. 706-710.
Sokolowski, J.A.: Simulation: Models that vary over time. In: Sokolowski, J.A.; Banks, C.M. (eds.): Principles of Modeling and Simulation. Hoboken, NJ: John Wiley & Sons 2009, pp. 47-69.
Tako, A.A.; Robinson, S.: The application of discrete event simulation and system dynamics in the logistics and supply chain context. Decision support systems 52 (2012) 4, pp. 802-815.
Terlunen, S.; Horstkemper, D.; Hellingrath, B.: Adaption of the discrete rate-based simulation paradigm for tactical supply chain decisions. In: Winter Simulation Conference, Savannah 2014, pp. 2060-2071.
Tolujew, J.; Reggelin, T.; Kaiser, A.: Discrete Rate Simulation als grundlegendes Paradigma bei der Entwicklung von mesoskopischen Flussmodellen. In: Integrationsaspekte der Simulation: Technik, Organisation und Personal. - Karlsruhe : KIT Scientific Publ. 2010, S. 437-444.
Winkler, H.: Beitrag zur Positionierung von Kundenentkopplungspunkten in Produktionsnetzwerken. Dortmund: Praxiswissen 2010.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# A Procedure Model for the Credible Measurability of Data Warehouse Metrics on Discrete-event Simulation Models of Logistics Systems

## *Eine Vorgehensweise zur glaubwürdigen Messbarkeit von Data-Warehouse-Kennzahlen an ereignisdiskreten Simulationsmodellen von Logistiksystemen*

Markus Rabe, Felix Dross, Andrea Vennemann, TU Dortmund, Dortmund (Germany), markus.rabe@tu-dortmund.de, felix.dross@tu-dortmund.de, andrea.vennemann@tu-dortmund.de

Abstract: This paper presents a new procedure model for the credible measurability of data warehouse key performance indicators on simulation models of discrete-event logistics systems. The basis for the new procedure model was the procedure model of information acquisition by Bernhard and Wenzel (2005). The new model consists of two parts, an input data management and an output data management. The complete procedure model has been integrated into the procedure model for simulation including V&V by Rabe et al. (2008b). Furthermore, a concept for a control software for the automated processing of the input and output data has been developed and is presented in this paper.

## 1 Introduction

Discrete-event logistics systems (DELS) have been described as large-scale complex socio-technical systems which operate in an environment of uncertainty. Examples for such systems can be found in warehousing, manufacturing, and supply chains. Discrete-event simulation (DES) is the most commonly used analysis tool for evaluating DELS performance (McGinnis 2005). Data warehouses (DWHs) provide data structures in order to support tools for analytical decision making. In contrast to common databases, DWHs store data in redundant and aggregated ways, speeding up interactive analysis and providing data at sufficient aggregation levels (Ehmke et al. 2011). Online analytical processing (OLAP) software provides a fast, flexible and interactive access to the data in a DWH and enables the organisation, aggregation and visualisation of information. Data are presented in terms of hypercubes depicting multidimensional structures. The cubes visualise system performance measures (the cube's cells) in the context of their dimensions (the cube's borders) and therefore enable the flexible and multidimensional analysis of the data (Jarke et al. 2003;

Ehmke et al. 2011; Bauer and Günzel 2013). OLAP technology is often used to realise performance measurement systems (PMS). A PMS unites different performance measures that relate to each other in a hierarchical form and typically culminate in one key performance indicator (KPI) (Bauer and Günzel 2013; Dross and Rabe 2014). KPIs are used by the management to measure, control and steer DELS. While the combination of DES models with DWH technology has been described before, previous papers mainly focus on harnessing the technological power of OLAP on the input or output data of DES models. They do not describe the steps which are necessary to obtain a credible measurability of existing DWH-KPIs for the management on the output data of a DES model of a DELS. In this paper, the authors present a procedure model as well as a software architecture, which can be used to achieve a credible measurability of data warehouse metrics on discrete-event simulation models of logistics systems.

The paper is structured as follows: Section 2 provides an overview of related work. The authors list relevant literature which focuses on the combination of DES models with DWH technology as well as literature with related procedure models. Some of the presented procedure models have been used as a basis or have inspired the newly developed procedure model. Section 3 presents the developed procedure model. Section 4 introduces a software architecture which can be used to achieve a credible measurability of DWH-KPIs on a DES model. The procedure model presented in Section 3 can be used to support the implementation of the software architecture presented in Section 4. Section 5 closes the paper with a conclusion and an outlook on future research.

## 2 Related Work

Koutsoukis et al. (1999) and El-Darzi et al. (2001) describe the benefits of using DWH and OLAP software in the preparation of data for the estimation of input parameters for various decision models. The streaming of simulation output data to a data warehouse was mentioned by Banks (1997). The use of data warehouse technology as an approach for the interactive and flexible analysis of detailed state-transition data collected from DES models of logistics systems has been further described by Vasilakis et al. (2004) and Ehmke et al. (2011). However, none of them emphasises the importance of a credible measurability of management KPIs on a DES model. A situation where a credible measurability of DWH-KPIs on a DES model is particularly important has been extensively described by Dross and Rabe (2014). The authors describe a situation where a simulation model is used as the basis for a decision support system (DSS) which regularly receives the output of complex PMS. The DSS uses the outputs of the PMS and is aimed to evaluate possible actions regarding the DWH-KPIs in the DES model.

In order to ensure quality and increase the rapidity in DES projects in general, there already exist well-structured methodologies to follow (Law and Kelton 2000, Rabe et al. 2008a, 2008b; Wenzel et al. 2008; Verein deutscher Ingenieure e.V. 2014). Although it is a very important part of a simulation study, the process of collecting all relevant information for a DES model has been a comparably small topic in many classical discrete-event simulation books (e.g., Vincent 1998; Law and Kelton 2000; Robinson 2004; Pidd 2004; Banks et al. 2014). Consequently, dedicated procedure models for the data collection for discrete-event simulation models have been

developed. Lehtonen and Seppala (1997) present a methodology for data gathering and analysis in logistics simulation projects. Baron et al. (2001) developed checklists for the identification of necessary data for a simulation model. Based on Baron et al. (2001), Csanady et al. (2008) developed a methodology for the systematic information retrieval for simulation studies. Skoogh and Johansson (2008) present an extensive methodology for the input data management in DES projects. Their methodology can be embedded into a procedure model for simulation projects, such as the one described by Law and Kelton (2000). Bengtsson et al. (2009) present an approach to use a methodology to identify and collect data, then use an input data management software to extract and process the data. Another methodology for simulation input data management is presented by Hill and Onggo (2012; 2014). The methodology was derived from observations made in an action research conducted at a management consultancy company. The procedure model of information acquisition (PMIA) by Bernhard and Wenzel (2005) is another procedure model for the input data management phase of a simulation study. In contrast to the other procedure models described above, the PMIA is based on the idea of separating data and information. Furthermore, it provides a consistent V&V of the phase results. It has already been integrated into the procedure model for simulation including V&V by Rabe et al. (Bernhard et al. 2007; Jodin et al. 2009). Since the authors of this paper aimed for a credible measurability of DWH-KPIs, a consistent V&V of the phase results is a very important aspect. Therefore, the PMIA has been chosen as the basis for the newly developed procedure model presented in this paper. Unlike for the processing of the input data, there exists no concrete procedure model for the processing of the simulation output data. The VDI guideline 3633 sheet 3 provides some recommendations for the processing of the output data (Verein deutscher Ingenieure e.V. 1997). Actually, in most of the literature the emphasis lies on the statistical processing of the simulation output data (Banks et al. 2014; Law and Kelton 2000). Therefore, the output data management of the newly developed procedure model has been newly constructed according to the PMIA. Finally, most of these aspects have been integrated into the procedure model for simulation including V&V by Rabe et al. (2008a; 2008b).

## 3 Procedure Model

The developed procedure model can be divided into two logical parts, the input data management and the output data management. Figure 1 shows the complete procedure model as an extension to the procedure model by Rabe et al. (2008b). In the following sections, the authors will explain the new elements of the procedure model. For a detailed description of the older elements, the reader is referred to the respective literature. The authors assume that the desired DWH-KPIs have been documented in the task description of the simulation study.

### 3.1 Input Data Management

In this part of the procedure model, the required output data of the simulation model for the measurement of the desired DWH-KPIs are identified. Furthermore, the deviation to the respectively required raw data for the measurement of the desired DWH-KPIs is documented. This information is then used to identify and collect the necessary input data using the PMIA.

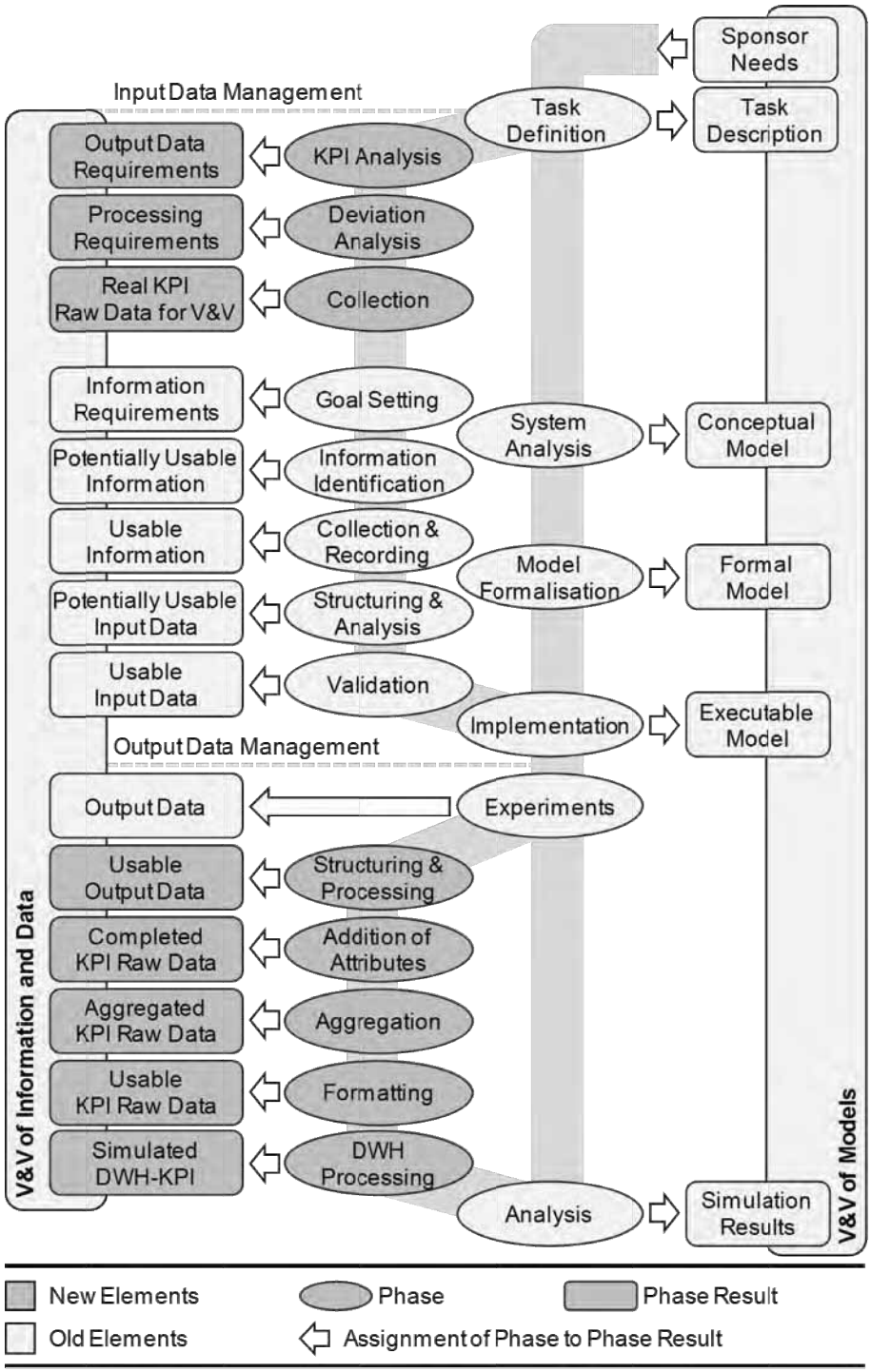



*Figure 1: Procedure model for a credible measurability of data warehouse metrics (Based on the procedure model for simulation including V&V by Rabe et al. 2008b)*

### 3.1.1 Identification of Necessary Simulation Output Data

The first new element in the procedure model is the KPI analysis. Within this phase, the first step is a general analysis of the DWH-KPIs, in which the focus lies on the logic of the KPIs. The technical aspects of the implementation in the DWH are initially ignored. The aim of the general analysis is to obtain an overview of the KPIs. The composition of the KPIs is determined by gradually dividing the calculation rules of the PMS in smaller parts in a top-down manner, which helps to resolve the complexity of the PMS. The analysis can be supported with KPI profiles, in which the structure of the PMS has been documented. If these are incomplete or non-existent, an interview with the PMS experts has to be conducted. In case of commonly known KPIs, the corresponding calculation rule can be taken from adequate literature, but it has to be ensured that the calculation rule has not to be adjusted in a company-specific way. If the analysis includes more than one KPI, the KPIs have to be analysed together, so that potential overlaps can be easily identified. To document the findings, the PMS can be represented graphically, e.g. in form of a tree structure. After the general analysis, the technical implementation of the DWH-KPIs in the DWH is examined. If an existing simulation tool is used for the simulation study, the simulation experts also have to identify which simulation output data can already be produced by the tool. The experts have to examine how the provided simulation output data fit to the raw data requirements of the DWH-KPIs. One important aspect to consider is that the lower layers of a PMS implementation in a DWH are mostly used to harmonise and aggregate the raw data which usually come from the operational systems. When using the output data of a simulation model as the data input source of a DWH, it might be possible to insert the data into much higher layers of the DWH. Hence, the simulation experts have to carefully identify which simulation output data will be used and in which specific input cube within the PMS implementations the simulation output data have to be written. Generally, choosing higher input cubes in the PMS implementation in the DWH leads to performance improvements in the calculation of the desired DWH-KPIs. With a suitable choice of the required simulation output data and the corresponding input cubes, the processing requirements can be minimised.

### 3.1.2 Identification of Processing Requirements

Ideally, the required output data of the simulation model and the corresponding input cubes in the DWH match in every aspect. In most cases though, the output data have to be processed to match the data requirements of the respective input cube. The goal of the deviation analysis is to identify those processing requirements. The processing requirements should be documented regarding attributes, aggregation and formatting. Regarding the attributes, the output data of the simulation model might not include certain attributes which are essential for the processing of the data in the DWH. In this case, the missing attributes have to be added to the output data. For example, the output data might include article numbers, but not corresponding article group numbers, which are required for further processing steps in the DWH. Regarding the aggregation, the output data may not meet the aggregation requirements of the DWH. For example, the output data may be recorded on a daily basis, but the input cube expects data on a monthly basis. Regarding the formatting, the output data might differ in terms of data types or in terms of the syntactical formatting.

### 3.1.3 Collecting Real KPI Raw Data for V&V

In order to achieve a sufficient match of simulated to real DWH-KPIs, it is mandatory to achieve a sufficient match of the simulation output data to the real raw data for the measurement of the DWH-KPIs. Hence, the real KPI raw data are useful to continuously validating the output data of the simulation model. In the collection phase, the real raw data for the DWH-KPIs are therefore collected and validated by calculating the desired DWH-KPIs manually, e.g. with a spreadsheet software.

### 3.1.4 PMIA

Based on the identified simulation output data requirements, the PMIA is now applied to determine the information needs for the simulation study and to obtain the relevant input data for the simulation model. These steps will not be discussed further here, because the steps have already been sufficiently described in the respective literature.

## 3.2 Output Data Management

Once the simulation model can produce the required output data, the output data have to be processed and loaded into the respective input cubes in the DWH. The output data management part of the procedure model focuses on the processing of the simulation output data in order to measure the desired DWH-KPIs. The authors assume that the simulation output data have already been validated e.g., using the real KPI raw data collected in the collection phase of the input data management.

### 3.2.1 Structuring and Processing of Simulation Output Data

First, the data needed to calculate the desired DWH-KPIs have to be selected from the overall simulation output data. If more than one input cube in the DWH is used, the selected output data have to be sorted respectively. Subsequently, a plausibility check with respect to the accuracy of the selection and sorting has to be performed. If more than one simulation run was performed, the output data of the different simulation runs now have to be combined into one data set using adequate statistical methods. Finally, a verification of the statistical processing steps and a validation of the processed output data have to be conducted.

### 3.2.2 Addition of Missing Attributes

If it is impossible to generate mandatory attributes of the simulation output data directly by an appropriate programming of the simulation model, these attributes have to be added to the output data. The missing attributes can usually be added using a relation to the simulation input data. For example, if the output data would only contain customer names, but customer numbers are required, the appropriate mapping of customer names to customer numbers could be obtained from the simulation input data. The completed data have to be checked for inconsistencies, e.g., using a desk test.

### 3.2.3 Aggregation

The aggregation of the data can be done along different dimensions. If the classification hierarchy of a dimension is generally known, such as the hierarchy of time, no further information for the aggregation process is required. If a company-

specific classification hierarchy is used, such as the classification into product groups, the classification levels have to be identified. As a prerequisite for the aggregation of the data, the completeness, disjointness and type compatibility of the data have to be ensured.

### 3.2.4 Formatting

As a final processing step, the data need to be formatted into a DWH-compatible form. For data types that are not strings, the data format is already defined by the respective data type. Therefore, formatting of the data is implicitly done with the conversion into the respective data type. For data types that contain a string, a syntactic formatting of the data might be necessary. This includes, for example, the adaptation of uppercase and lowercase letters within a string.

### 3.2.5 Measuring the DWH-KPI on the Simulation Output Data

Finally, the data are ready to be loaded into the input cubes in the DWH and the desired DHW-KPIs can be measured. The KPIs have to be validated with the real DWH-KPIs. If the simulated DWH-KPIs are sufficiently precise, the analysis of the DELS can be conducted now. Experimenting with the simulation model can lead to the desired insights regarding the connection of changes in the DELS to changes in the DWH-KPIs.

## 4 Software Architecture

Since the transformation of raw data into simulation input data and the processing of the output data into the required DWH data format is very time consuming if done manually, the authors have developed a concept for a control software which can perform the necessary steps automatically. The respective architecture is shown in Figure 2. The DWH system on the left is modelled after the reference architecture for DWH systems described by Bauer and Günzel (2013). Next to it is the simulation environment. The simulation of the DELS occurs in a data-driven simulation model. It consists of a generic simulation model with a connected database, from which the model is initialised. In the concrete implementation of the experimentation environment the software SimChain is used (SimPlan AG 2015). The architecture of SimChain has been described by Gutenschwager and Alicke (2004). The measurement of the DWH-KPIs from the output data of the DES model is realised with a logic copy of the original DWH, the shadowed DWH (Dross and Rabe 2014). An input data ETL module (Extract, Transform and Load) is used to automatically obtain and process the required input data. The raw input data might be obtained from the DWH or from additional data sources, which can be files or other databases. After each simulation experiment, the output data ETL module queries the simulation output data from the respective output data tables of the simulation tool, processes the data according to the steps described in the procedure model and loads the data into the shadowed DWH. The metadata repository is a database in which all additional information for the processing of the input or output data is stored, e.g. the mapping information for missing attributes and the attributes themselves. Some verification and validation checks (V&V) can be partly automated, so the simulation experts are supported by the control software. Possible V&V checks are the validation of the collected raw input data, the verification of the

input data processing steps, the validation of the data in the metadata repository and the verification of the processing steps in the output data ETL module. The software automatically performs the implemented V&V checks and shows an alert, if some irregularities have been found. In this case, the simulation experts have to examine the data and find the reason for the failing V&V checks.

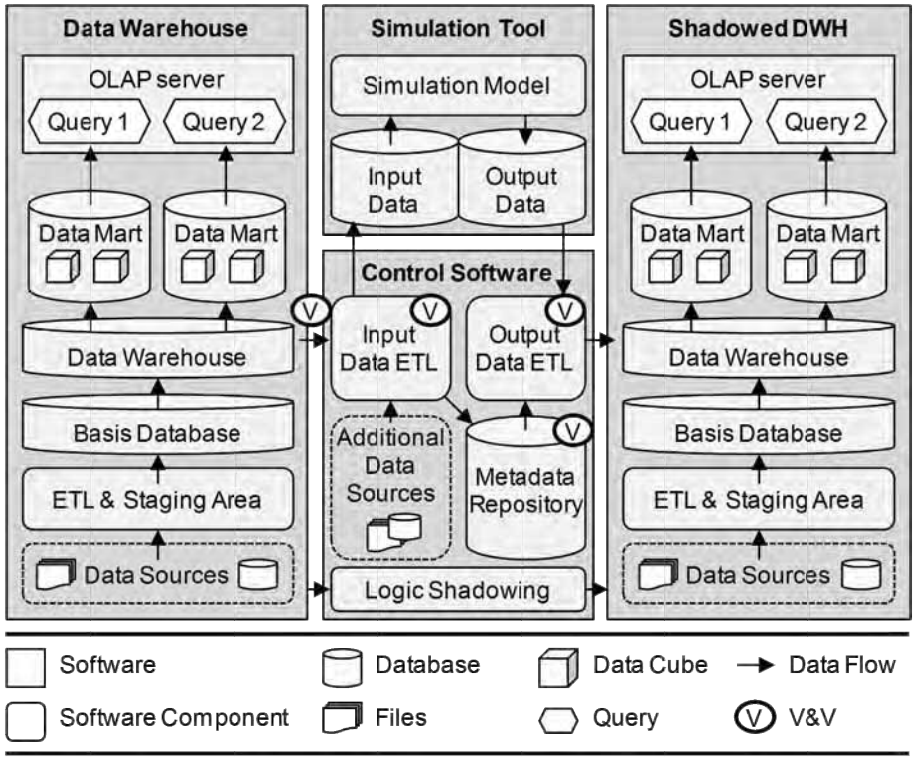



*Figure 2: Software architecture of the control software with V&V elements for the semi-automated updating of the simulation model*

## 5 Conclusion and Outlook

The procedure model described in this paper has already been successfully used by the authors in a simulation study. In this project it has already shown its usefulness and accurateness. The developed control software helped to automate some processing steps, but a manual V&V could not be completely substituted by the automated processes. In future projects, the authors will focus on advanced mechanisms for the automated V&V of the input and output data. This will be critical to achieve a sufficiently accurate measurement of the desired DWH-KPIs on the output data of the simulation model on a consistent basis. Only a credible measurability of DWH-KPIs will lead to a high acceptance of the simulation environment as a decision support instrument among the decision makers.

## References

Banks, J.: The future of simulation software: A panel discussion. In: Andradottir, S.; Healy, K.J.; Withers, D.H.; Nelson, B.L. (eds.): Proceedings of the Winter Simulation Conference. Piscataway: IEEE 1997, pp. 1155–1162.

Banks, J.; Carson, J.S.; Nelson, B.L.; Nicol, D.M.: Discrete-event system simulation. Harlow (UK): Pearson 2014.

Baron, C.P.; Dietel, U.; Kreppenhofer, D.; Rabe, M.: Handlungsanleitung Simulation. In: Rabe, M.; Hellingrath, B. (eds.): Handlungsanleitung Simulation in Produktion und Logistik. San Diego (CA): SCS Int. 2001, pp. 117–190.

Bauer, A.; Günzel, H.: Data-Warehouse-Systeme: Architektur, Entwicklung, Anwendung. Heidelberg: dpunkt 2013.

Bengtsson, N.; Shao, G.; Johansson, B.; Lee, Y.T.; Leong, S.; Skoogh, A.; Mclean, C.: Input data management methodology for discrete event simulation. In: Rossetti, M.D.; Hill, R.R.; Johansson, B.; Dunkin, A.; Ingalls, R.G. (eds.): Proceedings of the Winter Simulation Conference. Piscataway: IEEE 2009, pp. 1335–1344.

Bernhard, J.; Jodin, D.; Hömberg, K.; Kuhnt, S.; Schürmann, C.; Wenzel, S.: Vorgehensmodell zur Informationsgewinnung: Prozessschritte und Methodennutzung. Technical Report 06008, TU Dortmund (Germany), 2007.

Bernhard, J.; Wenzel, S.: Information acquisition for model based analysis of large logistics networks. In: Merkuryev, Y.; Zobel, R.; Kerckhoffs, E. (eds.): Proceedings of the 19th European Conference on Modelling and Simulation. Riga (Latvia) 2005, pp. 37–42.

Csanady, K.; Bockel, B.; Wenzel, S.: Methodik zur systematischen Informationsgewinnung für Simulationsstudien. In: Rabe, M. (ed.): Advances in Simulation for Production and Logistics Applications. Stuttgart: Fraunhofer-IRB-Verlag 2008.

Dross, F.; Rabe, M.: A SimHeuristic framework as a decision support system for large logistics networks with complex KPIs. In: Wittmann, J.; Deatcu, C. (eds.): Proceedings of the 22nd Symposium Simulationstechnik, Berlin (Germany) 2014.

Ehmke, J.F.; Großhans, D.; Mattfeld, D.C.; Smith, L.D.: Interactive analysis of discrete-event logistics systems with support of a data warehouse. Computers in Industry 62 (2011) 6, pp. 578–586.

El-Darzi, E.; Vasilakis, C.; Chaussalet, T.; Millard, P.H.: A simulation model to evaluate the interaction between acute, rehabilitation, long-stay care and the community. In: Zanakis, S.H.; Doukidis, G.; Zopounidis, C. (eds.): Decision Making: Recent Developments and Worldwide Applications. Berlin, Heidelberg: Springer 2001, pp. 475–485.

Gutenschwager, K.; Alicke, K.: Supply Chain Simulation mit ICON-SimChain. In: Spengler, T.; Voß, S.; Kopfer, H. (eds.): Logistik Management. Heidelberg: Physica 2004, pp. 161–178.

Hill, J.; Onggo, S.: Data identification and collection methodology in a simulation project: An action research. In: Tjahjono, B.; Heavey, C.; Onggo, S.; van der Zee, D.-J. (eds.): Proceedings of the 6th Operational Research Society Simulation Workshop 2012, Birmingham (UK): The Operational Research Society 2012.

Hill, J.; Onggo, S.: Data identification and data collection methods in simulation: A case study at ORH Ltd. Journal of Simulation 8 (2014) 3, pp. 195–205.

Jarke, M.; Lenzerini, M.; Vassiliou, Y.; Vassiliadis, P.: Fundamentals of Data Warehousing. Berlin, Heidelberg: Springer 2003.

Jodin, D.; Kuhnt, S.; Wenzel, S.: Methodennutzungsmodell zur Informationsgewinnung in großen Netzen der Logistik. In: Buchholz, P.; Clausen, U. (eds.): Grosse Netze der Logistik. Berlin: Springer 2009.

Koutsoukis, N.-S.; Mitra, G.; Lucas, C.: Adapting on-line analytical processing for decision modelling: the interaction of information and decision technologies. Decision Support Systems 26 (1999) 1, pp. 1–30.

Law, A.M.; Kelton, W.D.: Simulation modeling and analysis. Boston (MA): McGraw-Hill 2000.

Lehtonen, J.; Seppala, U.: A methodology for data gathering and analysis in a logistics simulation project. Integrated Manufacturing Systems 8 (1997) 6, pp. 351–358.

McGinnis, L.F.: Technical and conceptual challenges in organizational simulation. In: Rouse, W.B.; Boff, K.R. (eds.): Organizational Simulation. Hoboken (NJ): Wiley 2005, pp. 273–298.

Pidd, M.: Computer simulation in management science. Hoboken (NJ): Wiley 2004.

Rabe, M.; Spieckermann, S.; Wenzel, S. (2008a) A new procedure model for verification and validation in production and logistics simulation. In: Mason, S.; Hill, R.; Mönch, L.; Rose, O.; Jefferson, T.; Fowler, J.W.; Ingalls, R.G. (eds.): Proceedings of the Winter Simulation Conference. Piscataway: IEEE 2008, pp. 1717–1726.

Rabe, M.; Spieckermann, S.; Wenzel, S. (2008b) Verifikation und Validierung für die Simulation in Produktion und Logistik: Vorgehensmodelle und Techniken. Berlin, Heidelberg: Springer 2008.

Robinson, S.: Simulation: The practice of model development and use. Hoboken (NJ): Wiley 2004.

SimPlan AG, 2015: SimChain. http://www.simchain.net/ accessed on 17.03.2015.

Skoogh, A.; Johansson, B.: A methodology for input data management in discrete event simulation projects. In: Mason, S.; Hill, R.; Mönch, L.; Rose, O.; Jefferson, T.; Fowler, J.W.; Ingalls, R.G. (eds.): Proceedings of the Winter Simulation Conference. Piscataway: IEEE 2008, pp. 1727–1735.

Vasilakis, C.; El-Darzi, E.; Chountas, P.: A data warehouse environment for storing and analyzing simulation output data. In: Ingalls, R.G.; Rossetti, M.D.; Smith, J.S.; Peters, B.A. (eds.): Proceedings of the Winter Simulation Conference. Piscataway: IEEE 2004, pp. 703–710.

Verein deutscher Ingenieure e.V.: Richtlinie 3633 Blatt 3 "Simulation von Logistik-, Materialfluß- und Produktionssystemen - Experimentplanung und -auswertung". Berlin: Beuth 1997.

Verein deutscher Ingenieure e.V.: Richtlinie 3633 Blatt 1 "Simulation von Logistik-, Materialfluss- und Produktionssystemen - Grundlagen". Berlin: Beuth 2014.

Vincent, S.: Input data analysis. In: Banks, J. (ed.): Handbook of Simulation. New York (NY): Wiley 1998, pp. 54–91.

Wenzel, S.; Collisi-Böhmer, S.; Pitsch, H.; Rose, O.; Weiß, M.: Qualitätskriterien für die Simulation in Produktion und Logistik: Planung und Durchführung von Simulationsstudien. Berlin, Heidelberg: Springer 2008.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# A Simulation-based Supply Chain Risk Analysis Framework

## *Ein simulationsbasiertes Framework für die Analyse von Supply-Chain-Risiken*

Iris Heckmann, FZI Forschungszentrum Informatik, Karlsruhe (Germany),
heckmann@fzi.de

Stefan Nickel, Karlsruher Institut für Technologie (KIT), Karlsruhe (Germany),
stefan.nickel@kit.edu

Abstract: Supply chain networks are globalised and complex systems. Due to unexpected deviations and disruptions, which are subsumed under the notion of supply chain risk, the planning and optimising of such systems is harder than ever. So far no unanimous framework has been developed to explicitly quantify supply chain risks. Existing approaches focus on locally assessed risk probabilities and impacts or even reduce risk to the product of disruption probability and its consequences. In this work we present a simulation framework that ensures that supply chain risk management can be realised as a continuous process on top of short- to mid-term supply chain planning. Therefore, the framework models supply chain risk in a way that is both transferable to models of proprietary supply chain planning engines and suitable for quantitative supply chain risk analysis.

## 1 Introduction

The literature on supply chain risk analysis is mostly of anecdotal or case-based nature (Chopra and Sodhi 2004; Norrman and Jansson 2004; Zsidisin et al. 2005) and only few authors present empirical research (Wagner and Bode 2006). However, conclusions of both analyses are not transferable to other supply chains, other circumstances or other incidents. Insights based on empirical data have to be handled with care as they are often specific for the business, geographic and industrial environment of the supply chain affected as well as on internal dynamics. Similarly conclusions drawn by experts are subjective and often not re-producible in face of newly emerging supply chain risks. Quantitative, systematic and reliable analyses are scarce. Mathematical optimisation approaches demand for specific solution methods tailored to the underlying problem. Simulation is a suitable methodology to model numerous interacting characteristics using a flexible and modular design.

Therefore, simulation as a method to model and analyse complex systems is appropriate for the analysis of supply chain risks.

In this work we present a new simulation-based approach for the analysis of supply chain risk. We, therefore, briefly discuss the main requirements for the development of dependable simulation-based risk analysis. Next, we present main procedures of the simulation model. We illustrate how the supply chain risk analysis framework can be used in practice.

## 2 Requirements for Supply Chain Risk Analysis

Based on an extensive literature review we identified in a previous work that supply chain risk depends on several distinct conceptual entities, namely: risk exposition, risk objectives and risk attitude (Heckmann et al. 2015) (fig. 1).

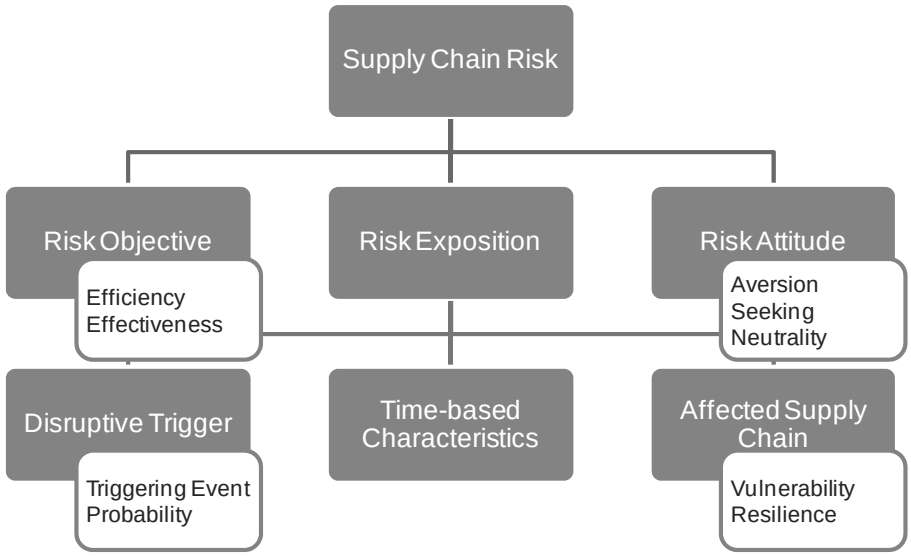



*Figure 1: Defining entities of supply chain risk (cp. Heckmann et al. 2015)*

The levels of all three entities determine the degree of present supply chain risks, whose evaluation is the purpose of risk analysis. When decision makers want to fully understand the genesis and the exacerbations of supply chain risks, then approaches are needed that support the identification and assessment of the underlying dynamics. More precisely, models are needed that support:

- the identification of influence of potential triggers on supply chain processes,
- the identification of dynamics between supply chain processes,
- the assessment of supply chain risk,
- the identification of influence of potential countermeasures on supply chain processes, and
- the assessment of the degree of supply chain risk reduction becoming effective due to the implementation of selective countermeasures.

We will refer to this bundle of characteristics as risk-aware characteristics.

Several studies identify further requirements for risk-aware decision tools. To ensure that supply chain risk management can be realised as a continuous process, supply chain managers, seek solutions that integrate supply chain risk analysis in their routines and planning processes (IBM 2009). The Risk Response Network of the World Economic Forum highlights the need for scenario-based and business continuity planning as core risk management priorities (Risk Response Network 2011). Most often scenarios are used to capture the shape of threating situations and evaluate their consequences on the supply chain. Those 'what-if'-analyses are largely used in planning, but they are often conducted by defining individual scenarios, rather than by elaborated scenario generation.

As risk largely depends on what the decision maker values or which impacts are considered as harmful, supply chain risk depends on the specific supply chain strategy and goals. Therefore, the importance of performance indicators in the supply chain risk assessment may differ from supply chain to supply chain and from decision level to decision level. Consider an example from daily operational supply chain planning. The operator uses a system for short-to-mid-term planning to manage changes and deviations. For him, supply chain vulnerability relates to order fulfilment or capacity utilisation. For a risk officer, who typically focuses on a mid-to-long-term horizon, supply chain vulnerability relates to performance indicators like costs of goods sold. The integration of both perspectives, roles and time scales provides reliable risk results and connect relevant planning levels.

Based on this discussion we derive the following characteristics for the development of supply chain risk analysis:

- risk-aware,
- scenario-based, and
- comprehensive for decision makers on different decision levels.

## 3 Scenario-based Procedure

The main challenge of the SimSCRF approach is to appropriately model the entities that define supply chain risk. We defined supply chain risk as a function that depends on the interaction between potential triggers, the constitution of the underlying supply chain, and the degree of target achievement, which is assumed to be influenced by the risk attitude of the decision maker. Figure 2 visualises the relationship between the conceptual approach and the methodological foundation, which is explained in the following paragraphs.

The theoretical basis of the SimSCRF is the assumption that potential triggers have an indirect and/or direct influence on supply chain processes. Supply chain processes are described by supply chain factors: a production process at production plants, for instance, can be defined through supply chain factors like production capacity, setup times, campaign duration, lot size, or production costs. Due to potential triggers these factors may change, which is depicted in different scenarios. A scenario is defined as a description of a prospective situation in terms of supply chain factors. Each scenario consists of a (predefined) set of supply chain factors and related factor values, which may or may not differ from their initial values.

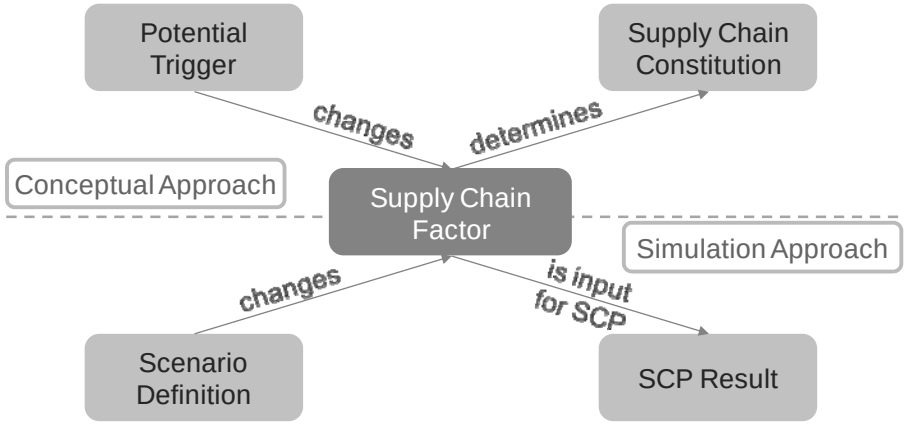



*Figure 2: Conceptual & simulation approach of a new supply chain risk analysis approach*

While potential triggers are modelled by the means of modified factor values, the resulting constitution of a supply chain is depicted by the result of a supply chain planning run. The purpose of supply chain planning systems is to balance demand and/or forecasts with available resources of production, transportation, and inventory. With the objective to optimise user-defined terms, e.g. cost minimisation, maximisation of profit or customer satisfaction etc., planning runs determine the optimal resource allocation. As these planning results are normally implemented and executed, they legitimately represent a good snapshot of the current workload of the underlying supply chain. A supply chain snapshot displays the structure and the current resource availability of the underlying supply network. It subsumes the description and quantification of supply chain factors.

Supply chain factors describe the degree of availability and the amount of operating costs related to supply chain processes, but they can additionally define the actual state of resource's usage. While the former set of supply chain factors defines the set of input parameters forwarded to the supply chain planning engine, the latter describes the set of result variables to be determined (optimised) by the planning engine. Modified input factor values, hence, lead to different planning results. The impact on performance is quantified by simulating a scenario in the supply chain planning engine so that the potential deterioration can be assessed. If the scenario performance is better than or equal to the accepted level, the deviations of supply chain factors applied within the scenario do not impose a supply chain risk.

## 4 Technical Entities of the Supply Chain Risk Analysis Framework

The SimSCRF approach builds modularly on existing proprietary supply chain planning engines and applies methods from factorial design of experiments to determine weaknesses of the underlying supply chain. To align the data model for supply chain planning with the simulation-based representation of a given supply chain, an objected-oriented information framework was developed. Introducing an middleware between planning engine and analysis algorithms offers the possibility

to easily switch between different planning engines and as a result conduct risk analysis for different supply chain planning problems. The approach presented is characterised by its technical entities and their interplay, namely, supply chain planning engine, information framework, analysis algorithms, and visualisation (fig. 3).

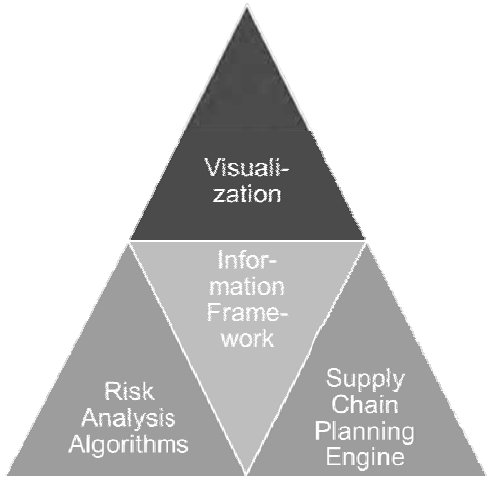



*Figure 3: Technical Entities of the simulation-based Risk Analysis Framework*

## 5 A Scenario-based Approach for Supply Chain Risk Identification and Assessment

Following the prevalent event-based approaches used in contemporary literature, often supply chain managers try to identify the most important potential disruptive events, such as labour strikes or earthquakes, and concentrate on defining adequate scenarios representing the characteristics of each potential trigger. This procedure consumes considerable amount of time and effort, and the results may not even be helpful. Instead of defining scenarios and evaluating their impact individually, we thus propose to start with the definition of accepted performance deterioration and to continue with the subsequent identification of those scenarios that make the specified deterioration effective.

If supply chain decision makers knew which supply chain processes have an influence on performance indicators and if they knew how much supply chain factors, that describe these processes, need to change such that performance is deteriorated, they would be able to concentrate solely on those processes. One weakness of the contemporary risk analysis is the dependence on decision maker's imagination or creativity. If he was not able to imagine causes for specific events, he would not have been able to define correspondent scenarios and most probably the supply chain would not have been prepared. The advantage of the SimSCRF is the possibility to overcome this dependence by a reverse simulation work flow. This process finds those input variable values that best fit a given objective without evaluating each assignment. This approach has just recently been coined by the misleading notion of simulation optimisation.

The work flow of the risk analysis framework is shown in Figure 4. A preliminary version of this work flow was discussed in Heckmann and Comes 2013. Three main pillars determine the work flow of the simulation: the decision maker, the simulation model and the planning engine. The simulation model is further divided into two sections. The front-end has an interface for communication with the decision maker and the back-end works as an interpreter between front-end, the simulation kernel generating the set of scenarios and the planning engine.

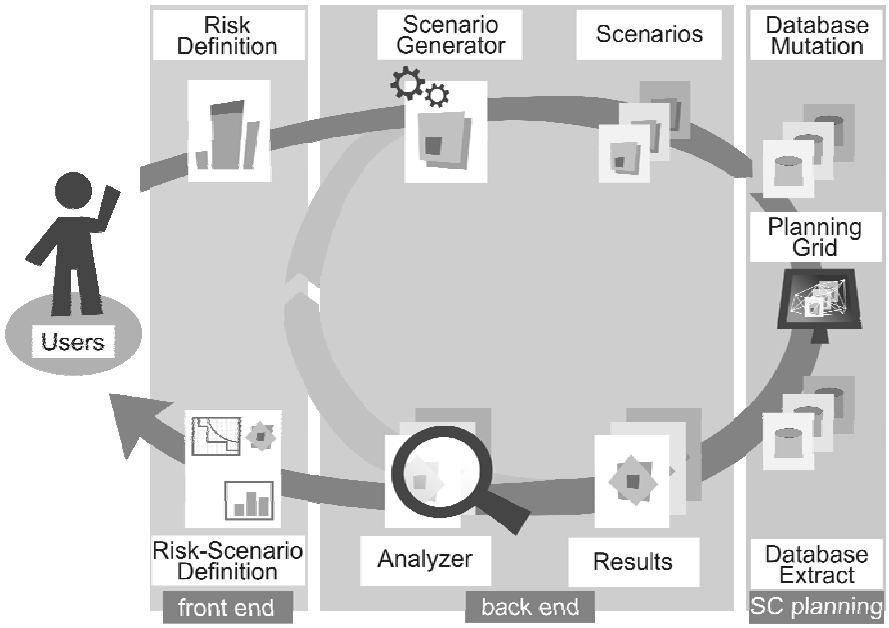
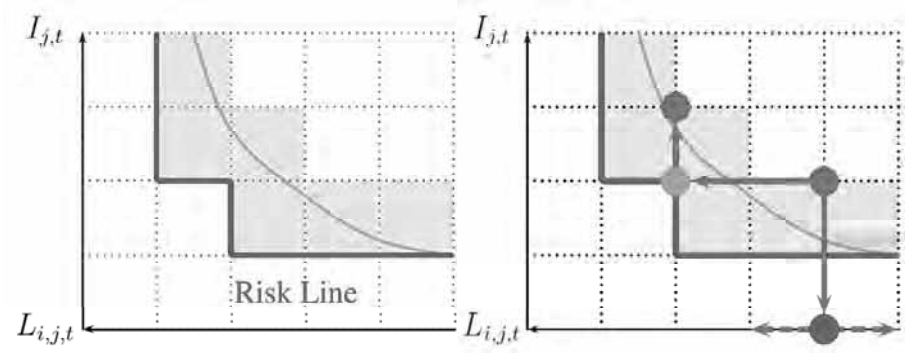



*Figure 4: Workflow of the Risk Analysis Framework*

The analysis approach has two objectives: First, it seeks to identify the set of supply chain factors that have a significant influence on relevant performance indicators (screening). The process flow of this type of analysis follows a sequential procedure; see the dark cycle of Figure 4. Second, for those supply chain factors that have been identified as influential, value levels need to be quantified that would result in unacceptable supply chain performance deterioration. Therefore, a set of scenarios is generated at the beginning of the analysis run and all scenario-based results are evaluated by the means of statistical methods referred to the Design and Analysis of Experiments (Montgomery 2009).

Figure 5 a) shows a result of a typical supply chain risk analysis run. For the sake of simplicity this figure shows only two factors: the x-axis shows factor values for the lead time between node i and j at time period t, $L_{i,j,t}$ from right to left, and the y-axis shows factor values for the inventory level at node j at time period t, $I_{j,t}$ from the bottom up. The decision maker is willing to accept a deterioration of 3% of the performance indicator service level.

Assume that the relationship between the factor values and the service level of 97% could be represented by the lighter line. Every factor value combination above this isoline would result in a service level greater than 97% which is good. Every factor value combination below this isoline would result in a service level less than 97% which is critical. The decision maker wants to identify the values defining the cube wherein the service level isoline is placed. These value combinations that are just critical determine the risk line.




*a) Result of Risk Analysis* *b) Mitigation Options*

*Figure 5: Risk Analysis*

The knowledge of the risk line provides the decision maker the opportunity to better understand which modifications of supply chain processes need his intervention. Assume, for instance, the lead time of the transportation process increases by two units compared to the actual lead time. In Figure 5 b) this modification of factor values is highlighted by the change from the dark to the lighter circle. This point is below the accepted level of target deterioration. If the decision maker is able to increase the inventory level of one unit of measurement by specific risk-reducing arrangements the service level could be traced back to acceptable levels, highlighted by the dark circle. The risk line gives also insight about factor modifications that cannot be handled by adjusting other factors. Assume the inventory level to drop from its actual level to a lower one, visualised by the red circle in Figure 5 b). Any decrease of the lead time would not be able to absorb the inventory fall. The risk line, thus, indicates which factor changes can be compensated on an operational basis and which must be considered at higher planning levels.

Figure 6 shows the user-friendly visualisation that supports the decision maker in defining his risk attitude based on the underlying supply chain structure and supply chain factors of interest.

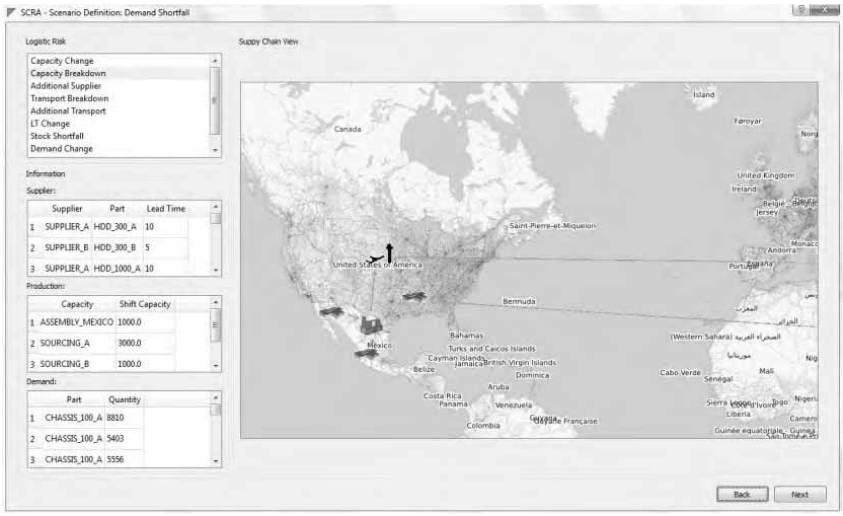



*Figure 6: User-friendly visualisation*

## 6 Conclusions

The main goal of SimSCRF is to provide the user with valid and credible implications on the dynamics that drive the underlying supply chain and that potentially make supply chain risk effective when disturbances occur. For the sake of conceptual and methodological consistency the risk analysis approach models and respects the defining entities of supply chain risk. As the simulation model is built around an operational planning system, the decision maker is able to establish a continuous improvement process: Lessons learned from the risk analysis can be adopted for risk-reducing measures in the operational system. Note that, this is only possible, because both systems access the same database. In addition to the property of data consistency, the scenario generation processes provide support for the identification and assessment of underlying dynamics. The screening procedure separates relevant from non-relevant supply chain factors. Based on the set of relevant or user-defined supply chain factors, the procedure for risk quantification identifies the combination of factor values that result in the non-achievement of targeted performance levels. The important performance indicators as well as their accepted level of deterioration are specified by the decision maker. He is thus able to express his risk attitude by specifying the level of deterioration, he is willing to accept.

Briefly, the main features of the simulation model are its capabilities of:

- respecting and modelling risk defining entities,
- providing data consistency between risk analysis and operational planning,
- offering risk quantification procedures.

The presented simulation model is a tool for the quantitative assessment of supply chain risks. It supports the user in his risk assessment through the definition of plausible scenarios, planning and scenario-evaluation according to his risk attitude.

The main goal of this work was to introduce the functionality of a new conceptual approach, whose objective is the analysis of supply chain risks. In order to demonstrate the functionality of the framework, we elaborate a case study in a forthcoming work, whose discussion is beyond the scope of this paper.

## Acknowledgement

Some of the authors have received financial support from BASF SE (Germany) within a research program.

## References

Chopra, S.; Sodhi M. M.: Managing risk to avoid supply-chain breakdown. MIT Sloan Management Review 46 (2004), pp. 53–61.

Heckmann, I.; Comes, T.: Modeling supply chain risk for operational supply chain planning. In: Proceedings of AMCIS - 19th Americas Conference on Information Systems, Chicago, Illinois, August 15-17 (2013).

Heckmann, I.; Comes, T.; Nickel, S.: A critical review on supply chain risk: defination, measure, and modeling. Omega 52 (2015), pp. 119-132.

IBM Global Services: Smarter Supply Chains: The Chief Supply Chain Officer Study. Technical report, IBM 2009.

Montgomery, D.: Design and Analysis of Experiments. Wiley 2009.

Norrman, A.; Jansson J.: Ericsson's proactive supply chain risk management approach after a serious sub-supplier accident. International Journal of Physical Distribution & Logistics Management 34 (2004), pp. 434–456.

Risk Response Network: New models for addressing supply chain and transport risk. Technical report, The World Economic Forum 2011.

Wagner, S.; Bode C.: An empirical investigation into supply chain vulnerability. Journal of Purchasing and Supply Management 12 (2006), pp. 301–312.

Zsidisin, G.; Ragatz G. L.; Melnyk S. A.: The dark side of supply chain management. Supply Chain Management Review 9 (2005), pp. 46–52.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Meta-Modellierung und Werkzeug-Integration in der Simulation von Seehafen Container-Terminals

## *Meta-modelling and Tool Integration in Simulation of Seaport Container Terminals*

Ann-Kathrin Lange, Carlos Jahn, Technische Universität Hamburg-Harburg, Hamburg (Germany), ann-kathrin.lange@tuhh.de, carlos.jahn@tuhh.de

Giovanni Pirovano, Politecnico di Milano, Mailand (Italy), gio.pirovano@gmail.com

Tommaso Rossi, Università Carlo Cattaneo - LIUC, Castellanza (Italy), trossi@liuc.it

Abstract: This document gives detailed information about the integration of different kinds of simulation in the planning process of maritime logistics systems. It demonstrates recent developments in maritime logistics and the consequential need for simulation in the optimization and planning of logistics systems. It introduces the different phases in planning processes and identifies adequate times for using simulation tools. Furthermore, different kinds of simulation and their uses are introduced. Afterwards, the document gives specific information about the method of meta-modelling and its possible integration with other kinds of simulation. The meta-model is used to automatically create a simulation after giving its specifications. Due to this, neither the personnel skills nor the time available to build the simulation model represent significant hurdles in the planning process.

## 1 Einleitung

Steigende Umsatzmengen, zunehmende Schiffsgrößen und begrenzte Erweiterungsflächen stellen den Betrieb von Container Terminals (CT) in Seehäfen vor große Herausforderungen. So erhöhte sich beispielsweise der weltweite containerisierte Handel von knapp 100 Mio. Twenty-foot Equivalent Units (TEU, deutsch Standardcontainer) im Jahr 2004 auf rund 160 Mio. TEU im Jahr 2013 (UNCTAD 2014). Die Kapazität des weltweit größten Containerschiffs verdoppelte sich von ca. 8.500 TEU im Jahr 2004 auf über 19.000 TEU im Jahr 2014. Ohne weitreichende Optimierungen der Abfertigungsprozesse würde dies zu erhöhten Liegezeiten der Schiffe im Hafen führen, die nicht nur zu steigenden Hafenliegegebühren für die Reeder führen, sondern die Effizienz des Schiffseinsatzes – gerade der sehr großen Containerschiffe

– signifikant senken können. Daher sind die Terminalbetreiber gezwungen, die Produktivität und Effizienz bei gleichbleibend niedrigen Preisen deutlich zu erhöhen (Wang und Cullinane 2006). Um den hohen Anforderungen bezüglich einer schnelleren Abfertigung in einem kürzeren Zeitfenster und einer höheren Qualität gerecht zu werden, ist es notwendig, dass die Terminals fortlaufend ihre operativen und administrativen Vorgänge begutachten und bei Bedarf entsprechend anpassen (Stahlbock und Voß 2008). Durch die begrenzten Erweiterungsflächen in den Häfen wird die Bedeutung von Optimierungen weiter verstärkt. Deshalb kommt es bei der Planung neuer und Umplanung vorhandener Terminals darauf an, Flächen und technische Systeme für Umschlag, Transport und Lagerung möglichst effizient einzusetzen. Die Methode der Simulation gewinnt vor diesem Hintergrund zur Absicherung und Optimierung von Lösungen bei Planungsprozessen in der Logistik im Allgemeinen (März und Weigert 2011) und insbesondere auch bei CT (Gibson et al. 1992; Koh et al. 1994; Merkuryev et al. 1998) immer mehr an Bedeutung.

Dabei kommt es mehr und mehr darauf an, die Simulation in frühen Planungsphasen und aufwandsarm einzubinden. So kann unter anderem ein Überblick über die zu erwartende Gesamtleistung des Terminals gegeben werden und mögliche Problemstellen vor der Umsetzung des Systems identifiziert werden. Ein großes Potenzial, dies zu meistern, besteht in der Meta-Simulation. Bisherige Anwendungen zur Simulation von CT werden aus einzelnen Elementen (Krane, Containerblocks, Fahrzeugen) zusammengesetzt. Für aussagekräftige Modelle sind dafür hohe Detailkenntnisse und ein großer Aufwand erforderlich, die ebenfalls ein hohes Fehlerpotenzial beinhalten (Novák 2011). Im Unterschied dazu können in sogenannten Meta-Modellen mit weniger Aufwand Simulationsmodelle insbesondere für frühe Planungsphasen generiert werden.

Bei der Betrachtung von CT aus Materialflusssicht stellen diese sich als offene Systeme mit zwei Schnittstellen zur Umgebung dar (Steenken et al. 2004). Eine dieser Schnittstellen ist die Wasserseite, an welcher das Laden und Löschen der Containerschiffe erfolgt. Die andere Schnittstelle ist die Landseite mit der Verladung der Container von und auf Bahn und LKW. Bei der Optimierung der Prozesse sollten im Zuge eines ganzheitlichen Ansatzes neben den terminalinternen Prozessen auch die wasserseitige und die landseitige Zuführung der Container betrachtet werden. Auch hier hat sich die Simulation als geeignetes Werkzeug erwiesen (John et al. 2014; Ben-Akiva und de Jong 2013).

In diesem Beitrag werden zunächst verschiedene Simulationsarten in der maritimen Logistik unterschieden und ihre jeweiligen Funktionen kurz dargestellt. Dabei wird auch auf die Rolle der Simulation im Planungsprozess von Logistiksystemen eingegangen. Anschließend erfolgt die Beschreibung der Umsetzung der Meta-Simulation im Allgemeinen und im Speziellen bei der Abbildung von CT. Dabei wird die Anwendung des Meta-Modells an einem einfachen Beispiel vorgestellt. Mögliche Schnittstellen der verschiedenen Simulationswerkzeuge und die Vorteile entsprechender Verknüpfungen werden aufgezeigt. Abschließend wird ein Ausblick auf mögliche Weiterentwicklungen des Ansatzes gegeben.

## 2 Der Einsatz von Simulation im Planungsprozess

Die Simulation im Bereich der maritimen Logistik beschäftigt sich prinzipiell mit drei Teilbereichen: Schiffsführung, Terminalprozesse und Verkehrsflüsse. Schiffsführungssimulatoren können sowohl zur Ausbildung des nautischen Nachwuchses als auch für Forschungsprojekte sowie im Rahmen von Hafenplanungsprojekten verwendet werden. So ermöglicht die Verwendung von Schiffsführungssimulatoren beispielsweise eine realistische Einschätzung der Einflüsse von Prozess- und Technologieänderungen auf die beteiligten Personen und minimiert so die Unfallgefahr bei der Umsetzung der Änderungen im realen Betrieb. Dabei birgt insbesondere die Kopplung mehrerer Simulatoren großes Potenzial (John et al. 2014). Ein anderes Beispiel ist der Einsatz von Schiffsführungssimulatoren für die abschließende Bewertung bei der Anpassung bzw. Neugestaltung von Wasserstraßen (PIANC 2014).

Zukünftige Erhöhungen der Gütermenge auf Hinterlandverkehrsträgern fordern einen entsprechenden Anstieg der Kapazitäten der jeweiligen Verkehrsinfrastruktur (Ben-Akiva et al. 2013). Güterverkehrsmodelle werden verwendet, um die möglichen Auswirkungen von politischen Maßnahmen und infrastrukturellen Änderungen auf den Verkehrsfluss zu untersuchen. Sie werden als wichtiges Instrument zur Begründung von infrastrukturellen Maßnahmen genutzt (Ben-Akiva und de Jong 2013). Weiterhin existieren spezielle Güterverkehrsmodelle zur Untersuchung des Hafenhinterlandverkehrs. Dabei dienen sie z. B. der Untersuchung der landseitigen Anbindung von Häfen in Bezug auf ihre Kapazitäten und Gestaltungsmöglichkeiten oder der Prognose der Entwicklung des Hafenhinterlandverkehrs oder des Modal Splits. Die Verknüpfung dieser Simulationsarten mit der Simulation von Terminalprozessen liegt im Fokus dieses Beitrages. Die Simulation von Terminalprozessen befasst sich unter anderem mit der Layoutplanung, der Bewertung verschiedener Betriebssysteme und der Analyse operativer Prozesse.

Der Entwicklungsprozess von Logistiksystemen, also auch von CT, kann generell in drei Phasen unterteilt werden. Dies sind die Planung, die Entwicklung und die Inbetriebnahme (Böse und Jahn 2011). Besondere Bedeutung hat die Simulation in der Planungsphase, genauer gesagt in der Vorplanung. Dazu gehören die Sammlung aller planungsrelevanten Daten, die Layoutplanung, die Kapazitätsanalyse von wichtigen Systembestandteilen und die Abschätzung von Geräte- und Personalanforderungen (Böse und Jahn 2011). Die Einbindung einer Visualisierungsumgebung in den Planungsprozess ermöglicht die einfache Betrachtung und Bewertung verschiedener Layoutvarianten. Da diese ohne großen Aufwand generiert werden können, ist eine umfangreiche Vorauswahl der Lösungsalternativen möglich. Dies reduziert den ggf. folgenden Simulationsaufwand deutlich (Jahn und Rauer 2011). Die Validierung der daraus gebildeten grundsätzlichen Systemvarianten kann unter Berücksichtigung von dynamischen und stochastischen Einflüssen des realen Betriebs und typischen Arbeitsanforderungen mittels ereignisorientierter Simulation durchgeführt werden (Böse und Jahn 2011).

Die frühzeitige Erstellung von Simulationsmodellen zur Bewertung der ersten Planungsentscheidungen verhindert somit mögliche Fehlplanungen und ermöglicht die Optimierung der Planung von Beginn an. Allerdings sind mit dieser Methode auch Schwierigkeiten verbunden, die die Nützlichkeit des Werkzeugs einschränken. So erfordert das Erstellen von Simulationsmodellen relativ viel Zeit und umfangreiche Softwarekenntnisse des Simulationsprogramms.

CT können unter anderem aufgrund ihrer hohen Anzahl an Organisationsformen, vieler Entscheidungsvariablen, statischer und dynamischer Nebenbedingungen und vieler Unsicherheitsquellen, wie z. B. Wetterbedingungen oder fehlerbedingten Ausfällen des Equipments, als komplexe Systeme betrachtet werden. Deshalb stellt die Simulation des Gesamtsystems CT hohe Anforderungen an den Ersteller und das Modell. Aus diesem Grund beschränken sich viele Simulationsmodelle auf einzelne Teilbereiche des Betriebs von CT, z. B. automatische Lagerblöcke (Kemme 2012; Canonaco et al. 2007), Betrieb der Containerbrücken (Liu et al. 2002; Dai et al. 2004) oder die Analyse des Horizontaltransportes (Duinkerken et al. 2007).

Die Verwendung eines Meta-Simulationsmodells umgeht diese Problematik, indem es den Nutzer unterstützt, Simulationsmodelle schnell mit relativ wenig Aufwand zu erstellen. In der Literatur werden Meta-Modelle auf zwei Wegen beschrieben. Zum einen werden sie als Hilfsmodelle für Simulationsmodelle definiert, die verwendet werden, um deren Komplexität zu verstehen (Friedmann 1996; Kleijnen 1987). Zum anderem werden Meta-Modelle als Satz an Simulationsmodellen beschrieben, die alle zu einer Kategorie an Systemen gehören. Das einzelne Simulationsmodell eines Systems der Kategorie, auf welche sich das Meta-Modell bezieht, wird durch eine Instanz des Modells abgebildet (Cigolini und Rossi 2010; Cigolini et al. 2011). Diese Instanz wird anhand von Eingabedaten des Nutzers automatisch generiert.

# 3 Meta-Modellierung

Das hier betrachtete Meta-Simulationsmodell wurde auf Basis der Simulations-Software „Enterprise Dynamics“ (ED) von Incontrol Simulation Solutions erstellt. Zur Ermittlung der relevanten Prozesse und Entscheidungsvariablen wurden zunächst Experteninterviews mit Vertretern verschiedener Terminalbetreiber und Terminalbesichtigungen durchgeführt. Durch die Betrachtung der detaillierten Abläufe konnten allgemein gültige Prozesse abstrahiert und in Form von Petri-Netzen dargestellt werden. Zur Nutzung dieser Prozesse für die Meta-Modellierung wurden die Schnittstellen zwischen den einzelnen Elementen untersucht und eine Clusterung dieser Elemente zu den verschiedenen Tätigkeitsbereichen von CT vorgenommen. Dadurch ist es möglich, die Meta-Modellierung durchzuführen, die sich nicht nur auf die Generierung einzelner Simulationsmodelle eines Betriebssystems aus einer Datenbank beschränkt, sondern die Modellierung aller standardmäßig vorkommenden Systeme von CT in Form von Simulationsmodellen ermöglicht.

Im Folgenden werden die Architektur des Meta-Simulationsmodells, die relevanten Objekte und die Software-Anwendung in Bezug auf die umfassende Darstellung in Lange et al. (2014) kurz vorgestellt, um anschließend die Verknüpfungsmöglichkeiten mit weiteren Simulationswerkzeugen aufzuzeigen.

## 3.1 Architektur des Meta-Modells

Das Meta-Modell erlaubt es dem Nutzer, die räumliche Struktur, die verwendeten Ressourcen und die besonderen Charakteristika des zu untersuchenden, realen CT zu definieren und daraus automatisch das entsprechende Simulationsmodell zu generieren. Durch Versuche mit diesem Model kann der Nutzer im Vorwege die Leistungsfähigkeit des Terminals untersuchen und Handlungen beschließen, die diese Leistungsfähigkeit steigern.

Das Meta-Modell besteht aus einer Benutzerschnittstelle mit zugehöriger Datenbank, einer Ad-hoc-Objektbibliothek und der Software-Anwendung. Durch die Benutzerschnittstelle kann der Nutzer sowohl die Elemente des CT als auch ihre Charakteristika, also die Werte der Parameter, durch die das Element charakterisiert wird, spezifizieren. Dabei werden sowohl die manuell eingegebenen Werte als auch die von der Schnittstelle daraus errechneten Werte in der Datenbank gespeichert.

Die Objektbibliothek enthält sowohl Standardatome aus ED als auch speziell erstellte Atome, die die operativen Bereiche des CT darstellen, wie z. B. die Abfertigung der Containerschiffe an der Wasserseite. Jedes Atom besteht dabei aus Daten, dargestellt durch die jeweiligen Attribute, und dem Verhalten, welches in 4D-Skript, der Programmiersprache von ED, modelliert wurde.

Die Software-Anwendung ermöglicht die automatische Generierung des Simulationsmodells. Dabei werden die Elemente des CT aus der Datenbank gelesen. Für jedes dieser Elemente wird das entsprechende Atom aus der Objektbibliothek entnommen, seine Attribute entsprechend der Werte in der Datenbank festgelegt, die richtige Zuweisung getroffen und in das Simulationsmodell eingefügt. Wenn dies für alle Elemente des CT durchgeführt wurde, wird das lauffähige Simulationsmodell erstellt, welches direkt für Experimente genutzt werden kann. Im weiteren Verlauf wird detaillierter auf die Objektbibliothek und die Software-Anwendung als komplexe Bestandteile des Meta-Modells eingegangen.

## 3.2 Objektbibliothek

In der Objektbibliothek liegen zwei verschiedene Arten an Objekten vor. Dies sind zum einen die Straßen, welche das Netzwerk erzeugen, in dem sich die Transportfahrzeuge des CT bewegen. Zum anderen gibt es die Elemente, die die Aktivitäten des Terminals erfüllen: Gate, Bahnabfertigung, wasserseitige Abfertigung, Container-Yard und aktives und passives Transportequipment. Zum aktiven Transportequipment gehören bspw. Straddle Carrier und Reach-Stacker, während dem passiven Equipment fahrerlose Transportfahrzeuge (AGV, Automated Guided Vehicles) und Traktor-und-Trailer-Kombinationen zuzuordnen sind. LKW, Züge, Schiffe und Container werden in dem Modell nicht als Objekte betrachtet, da es sich bei ihnen um die Entitäten handelt, die in Simulationsmodellen bewegt werden und dementsprechend automatisch vom Meta-Modell generiert werden.

Wie bereits in Abschnitt 3.1 dargestellt, werden Objekte durch Daten und Verhalten beschrieben. In Bezug auf die Daten handelt es sich bei den Attributen des Netzwerkes zum einen um die Paare an Knotenpunkten, die durch die Straße verbunden werden und zum anderen um die Parameter, welche das Setzen von Beschränkungen ermöglichen, wie z. B. Fahrtrichtung, Anzahl an Spuren oder Höchstgeschwindigkeiten. Die Attribute der Ressourcen-Atome quantifizieren prinzipiell die Auslastung der Terminal-Ressourcen, wie z. B. Lagerkapazität, Zwischenankunftszeiten der Schiffe oder Umschlagsraten der Krane. Beim Verhalten handelt es sich um Simulations-Submodelle, welche darstellen, wie die Elemente des CT miteinander und mit den beteiligten Entitäten, also LKW, Zügen, Schiffen und Containern, interagieren sollen. Aufgrund der Funktionsweise von ED wurde das Verhalten mittels State-Machine-Diagrams und Attributed-Petri-Nets modelliert.

## 3.3 Software-Anwendung

Die Software-Anwendung teilt sich in zwei Unterprozesse auf: Der erste fügt die entsprechenden Atome zu den Objekten der jeweiligen Ressourcen-Klasse in die Simulationsumgebung ein; der zweite generiert das Netzwerk durch Einfügen von Straßen-Atomen zwischen allen Paaren an Ressourcen-Atomen. Der erste Unterprozess beginnt mit dem Einlesen der Parameter des ersten Objektes der Ressourcen-Klasse, die erzeugt werden soll, aus der Datenbank. Anschließend wird das entsprechende Atom aus der Objektbibliothek entnommen und in die Simulationsumgebung eingefügt. Daraufhin versieht die Anwendung das jeweilige Atom mit den vorher eingelesenen Werten. Dieser Unterprozess wird für alle vom Nutzer definierten Objekte aller Ressourcen-Klassen wiederholt. Ist dies abgeschlossen, endet der erste Unterprozess und der zweite beginnt mit der Generierung des Netzwerkes. Dabei liest dieser Unterprozess die entsprechenden Parameter der Verbindung des ersten betrachteten Paares an Ressourcen-Objekten. Anschließend fügt die Anwendung das Straßen-Atom an entsprechender Stelle in die Simulationsumgebung ein und parametrisiert es entsprechend der eingelesenen Werte. Wurde dies für alle Paare an Ressourcen-Objekten durchgeführt, endet der zweite Unterprozess, da das Netzwerk komplett modelliert ist und das Simulationsmodell somit zur Nutzung zur Verfügung steht.

## 3.4 Anwendungsbeispiel des Meta-Modells

Die Anwendung des Meta-Modells zur Planung und Optimierung von CT soll anhand eines kurzen Beispiels dargestellt werden. Dabei wird ein fiktives Terminal mit einer wasserseitigen Abfertigung, vier Lagerblöcken und einer Bahnabfertigung betrachtet. Die LKW-Abfertigung wird zur Vereinfachung nicht berücksichtigt. Es werden folgende Entscheidungsvariablen zur Erstellung von acht Szenarien betrachtet: Die Anzahl der Containerbrücken (4 oder 8), die Anzahl der Straddle Carrier auf der Landseite (10 oder 20) und die Orientierung der Lagerblöcke in Bezug auf die Kaikante (parallel oder orthogonal). Zur Auswertung der Ergebnisse werden die durchschnittliche Abfertigungszeit der Züge und Schiffe und die zurückgelegte Strecke der Straddle Carrier herangezogen. Weitere Variablen können als Kontrollvariablen genutzt werden.

Im Folgenden werden die einzelnen Schritte zur Erstellung eines experimentierfähigen Simulationsmodells dargestellt:

1. Zunächst erfolgt die Sammlung aller benötigten Eingabedaten für das angestrebte Simulationsmodell.
2. Anschließend kann eine Darstellung des Layouts des CT im Softwaretool hochgeladen werden, um in der weiteren Erstellung des Simulationsmodells als eine Orientierungshilfe verwendet zu werden.
3. Im nächsten Schritt werden die benötigten Bereiche des CT per „drag and drop“ aus der Objektbibliothek auf die Oberfläche des Softwaretools übertragen. Die Verbindung dieser Bereiche erfolgt durch Netzwerke, die anhand ihrer Kreuzungspunkte definiert werden. Bei diesem Beispiel handelt es sich um die wasserseitige Abfertigung, den Container-Yard und die Bahnabfertigung.
4. Die Parametrisierung erfolgt durch die Eingabe der benötigten Daten in der Benutzerschnittstelle. Dabei existiert für jeden Bereich des Modells eine spezielle Eingabemaske. Weiterhin sind einige Felder bereits mit Standardwerten

ausgefüllt, die der Benutzer entweder übernehmen kann oder mit eigenen, genaueren Daten ersetzt.
5. Im nächsten Schritt wird durch die Software-Anwendung, wie in Kapitel 3.3 beschrieben, für jedes Szenario das entsprechende Simulationsmodell erstellt.
6. Abschließend können die Parameter des Simulationsexperiments, z. B. die Wiederholungsanzahl, festgelegt und das Experiment durchgeführt werden.

*Tabelle 1: Ergebnisse Simulationsexperiment Anwendungsbeispiel*

| Szenario | Abfertigungszeit Containerschiff | Abfertigungszeit Bahn | Strecke Straddle Carrier |
|---|---|---|---|
| A | 50,20 | 13,56 | 2.093 |
| B | 49,75 | 13,46 | 1.792 |
| C | 43,46 | 13,58 | 2.187 |
| D | 42,90 | 12,79 | 1.894 |
| E | 48,13 | 13,89 | 1.956 |
| F | 48,90 | 13,71 | 1.654 |
| G | 30,56 | 14,01 | 2.079 |
| H | 39,08 | 12,86 | 1.798 |

Anhand der Ergebnisse (Tab. 1) kann beispielsweise abgelesen werden, dass eine höhere Anzahl an Containerbrücken die Abfertigungszeit der Containerschiffe deutlich verringert oder das eine Verringerung der Straddle-Carrier-Anzahl zu einer erhöhten Abfertigungszeit der Züge und zu einer höheren zurückgelegten Strecke bei den Straddle Carriern führt. Es wurden also bei einem einfachen Modell innerhalb von weniger als 30 Minuten Arbeitszeit Ergebnisse ermittelt, die die Optimierung eines CT ermöglichen.

## 3.5 Umgang mit Unschärfe bei der Meta-Modellierung

Um ungewollte Unschärfen bei der Meta-Modellierung zu vermeiden, kann die Genauigkeit des lauffähigen Simulationsmodells durch den Nutzer selbst beeinflusst werden. Werden die voreingestellten Standardwerte und die vorhandenen Steuerungen verwendet, so kann mit geringem Aufwand eine grobe Abschätzung der Planungsentscheidungen getroffen werden. Dies ist insbesondere in einem sehr frühen Planungsstadium sinnvoll. Soll eine detaillierte Betrachtung vorgenommen werden, so können die Standardwerte durch exakte Werte ersetzt und die Steuerungen entsprechend der jeweiligen Abläufe angepasst werden. Dafür ist eine umfangreiche Informationsbeschaffung vor der Erstellung des Simulationsmodells notwendig. Dabei können insbesondere Rücksprachen mit Vertretern verschiedener CT oder Geräteentwicklern helfen, die exakten Werte für den Anwendungsfall zu ermitteln. Weiterhin ist eine Überprüfung der Einflüsse der verschiedenen Entscheidungsvariablen mit Hilfe einer Sensitivitätsanalyse und in Rücksprache mit Experten aus der Praxis empfehlenswert. Zur Überprüfung der Exaktheit des Simulationsmodells können Abgleiche mit Beispielen aus der Praxis durchgeführt werden.

## 4 Werkzeug-Integration

Böse und Jahn (2011) stellen das Konzept einer Tool-based Rapid Planning Environment vor. Dabei sollen verschiedene, in der Praxis erprobte Werkzeuge zur Planung von Logistiksystemen so gekoppelt werden, dass die Kosten- und Zeiteffizienz gegenüber der Planung mit alleinstehenden Systemen deutlich erhöht wird. Dies wird unter anderem damit begründet, dass sich Software-Werkzeuge zumeist auf einzelne, fest definierte Teilprobleme beziehen und somit als Ergebnis einzelne Teilmodelle erzeugen. Werden diese Teilmodelle durch angemessene Mechanismen gekoppelt, so führen Änderungen in einem der Modelle zu einer dynamischen Anpassung in den anderen Teilmodellen. Dies reduziert insbesondere sowohl den Zeitaufwand als auch die Fehleranfälligkeit und den manuellen Aufwand deutlich. Diese Planungsumgebung soll durch die Einhaltung verschiedener Prinzipien ermöglicht werden. Dazu gehören die Konsistenz der verwendeten Daten, die intuitive Bedienung der Werkzeuge, das Koppeln von verschiedenen Werkzeugen, die Verwendung bestehender Modelle und die Visualisierung der Planungsergebnisse. Als Werkzeuge wurden dabei ein haptischer Planungstisch, eine Tabellenkalkulation, ein Simulator für logistische Prozesse und ein Schiffsführungssimulator berücksichtigt.

Durch die Einbindung eines Werkzeuges zur Güterverkehrsmodellierung in eine solche Planungsumgebung könnten die wichtigsten Bereiche in der Planung von maritimen Logistiksystemen abgebildet werden. Dabei könnten, wie auch beim Schiffsführungssimulator, die Abhängigkeiten der Ergebnisse der unterschiedlichen Systeme ausgenutzt werden, um gekoppelt möglichst optimale Ergebnisse zu erzielen. Weiterhin würde sich ein Meta-Modell zur Simulation der logistischen Prozesse an einem Terminal sehr gut in das bestehende System einfügen, da es im Vergleich zur Verwendung bereits bestehender Modelle deutlich flexibler an neue Situationen anpassbar ist und dem Nutzer einen größeren Gestaltungsspielraum bietet. Weiterhin kann so ebenfalls die Zeiteffizienz gesteigert werden und die Fehlerwahrscheinlichkeit reduziert werden, da die Eingabemöglichkeiten des Nutzers klar vorgegeben sind und dieser von dem eigentlichen Modellierungsprozess entlastet wird. Zusätzlich wird die Benutzerfreundlichkeit durch übersichtliche Eingabemasken deutlich erhöht.

## 5 Fazit und Ausblick

Dieser Beitrag beschreibt verschiedene neue Entwicklungen bei der Simulation von Seehäfen. Insbesondere wird dabei auf die Meta-Modellierung und die Integration unterschiedlicher Werkzeuge eingegangen. Der Grund, warum diese möglichen Lösungen betrachtet wurden, sind die vielfältigen Herausforderungen bei der Planung von Logistiksystemen. Aufgrund des steigenden Drucks auf CT Betreiber ist es notwendig, Simulationsmodelle frühzeitig und aufwandsarm in den Planungsprozess einzubinden, um zu einem frühen Zeitpunkt mögliche Problemstellen zu identifizieren und einen guten Überblick über die Gesamtleistung des zu planenden Systems zu erhalten. Allerdings beinhaltet dieses Vorgehen ebenfalls Schwierigkeiten, da die Einbindung von Simulationsmodellen zum einen zumeist einen hohen Zeitaufwand und gute Detailkenntnisse der Simulationssoftware erfordert. Zum anderen müssen zur ganzheitlichen Betrachtung aller relevanten Akteure und Prozesse unterschiedliche Werkzeuge verwendet werden. Dies stellt bei einer einzelnen Betrachtung der

Werkzeuge eine deutliche Erhöhung des Aufwandes und der Fehlermöglichkeiten für die Beteiligten dar.

Durch die Verwendung von sogenannten Meta-Modellen, bei denen ein generisches Modell durch Spezifikation der Parameter in einer Benutzerschnittstelle ein Simulationsmodell erstellt und durch die Integration der verschiedenen Werkzeuge in eine definierte Planungsumgebung können diese Schwierigkeiten überwunden werden.

Zukünftig soll sowohl die Weiterentwicklung der einzelnen Elemente der Planungsumgebung als auch die Gestaltung der Planungsumgebung selbst weiter vorangetrieben werden. Die Anwendbarkeit soll anhand weiterer praktischer Beispiele verbessert werden.

## Literatur

Ben-Akiva, M; de Jong, G.: The Aggregate-Disaggregate-Aggregate (ADA) freight model system. In: Ben-Akiva, M.; Meersman, H.; Van De Voorde, E. (Hrsg.): Freight Transport Modelling. Bingley (UK): Emerald Group Pub 2013, S.69-90.

Ben-Akiva, M.; Meersman, H.; Van De Voorde, E.: Recent developments in freight transport modelling. In: Ben-Akiva, M.; Meersman, H.; Van De Voorde, E. (Hrsg.): Freight Transport Modelling. Bingley (UK): Emerald Group Pub 2013, S. 3-13.

Böse, J.W.; Jahn, C.: Innovative logistics planning. Development and use of an integrated planning environment for increase of effectiveness and efficiency of complex planning processes. In: Blecker, T.; Jahn, C.; Kersten, W. (Hrsg.): Maritime Logistics in the Global Economy. Current Trends and Approaches. Lohmar: Eul (Supply Chain, Logistics and Operations Management, 5) 2011, S. 73-107.

Canonaco, P.; Legato, P.; Mazza, R. M.; Musmanno, R.: A queuing network model for the management of berth crane operations. Computers & Operations Research 35 (2007) 8, S. 2432-2446.

Cigolini, R.; Rossi, T.: Sizing off-shore transshipment systems in dry-bulk transportation. Production Planning & Control 21 (2010) 5, S. 508–522.

Cigolini, R.; Pero, M.; Rossi, T.: An object-oriented simulation meta-model to analyse supply chain performance. International Journal of Production Research 49 (2011) 19, S. 5917-5941.

Dai, J.; Lin, W.; Moorthy, R.; Teo, C.-P.: Berth allocation planning optimization in container terminals. Working Paper: Georgia Institute of Technology (Atlanta, USA), National University of Singapore, 2004.

Duinkerken, M.B.; Dekker, R.; Kurstjens, S.T.G.L.; Ottjes, J.A.; Dellaert, N.P.: Comparing transportation systems for inter-terminal transport at the Maasvlakte container terminals. In: Kim, K.H.; Günther, H.-O. (Hrsg.): Container Terminals and Cargo Systems. Design, Operations Management and Logistics Control Issues: Berlin et al.: Springer 2007, S. 38-61.

Friedman, L.: Simulation metamodel. Norwell (USA): Kluwer Academic Publishers 1996.

Gibson, R.; Carpenter, B.; Seeburger, S.: A flexible port traffic planning model. In: Swain, J.J.; Goldsman, D.; Crain, R.C.; Wilson, J.R. (Hrsg.): Proceedings of the 1992 Winter Simulation Conference, Arlington (VA), 13.-16. Dezember 1992, S. 1296-1306.

Jahn, C.; Rauer, R.: Visualisierung in der Hafenlogistikplanung. Neue Potenziale durch integrierte Planungswerkzeuge. In: von Lukas, U.; Haase, K.; Mahnke, E.-M.; Malo, S. (Hrsg.): Tagungsband zur Konferenz Go-3D 2011: Computergraphik für die Praxis. Rostock, 30.08.2011. Fraunhofer-Institut für Graphische Datenverarbeitung IGD. Stuttgart: Fraunhofer Verlag, S. 47-55.

John, O.; Burmeister, H.-C.; Brödje, A.; Bornhorst, C.; Grube, C.: Assessing the MONALISA 2.0 Concept: Establishment of the European Maritime Simulation Network. In: Müller, R.; Deutsche Gesellschaft für Ortung und Navigation (DGON) (Hrsg.): Proceedings of the International Symposium "Information on Ships", Integrated Ship's Information Systems (ISIS) 2014, Hamburg, 04.-05. September 2014, S. 179-189.

Kemme, N.: Effects of storage block layout and automated yard crane systems on the performance of seaport container terminals. OR Spectrum 34 (2012) 3, S. 563–591.

Kleijnen, J.P.C: Statistical tools for simulation practitioners. New York (USA): M. Dekker 1987.

Koh,P.-H.; Goh, J. L. K.; Ng, H.-S.; Ng, H.-C.: Using simulation to preview plans of a container port operations. In: Sadowski, D.A.; Seila, A.F.; Manivannan, M.S.; Tew, J.D.: Proceedings of the 1994 Winter Simulation Conference, Lake Buena Vista (FL), 11.-14. Dezember 1994, S. 1109-1115.

Lange, A.-K.; Pirovano, G.; Pozzi, R.; Rossi, T.: Development of a container terminal simulation meta-model. In: Wittmann, J. (Hrsg.): Tagungsband 22. Symposium Simulationstechnik, ASIM 2014, 3. bis 5. September 2014, HTW Berlin. Wien (Österreich): ARGESIM 2014, S. 101-108.

Liu, C.-I; Jula, H.; Ioannou, P. A.: Design, simulation, and evaluation of automated container terminals. IEEE Transactions on intelligent transportation systems 3 (2002) 1, S 12-26.

März, L.; Weigert, G.: Simulationsgestützte Optimierung. In: März, L.; Krug, W.; Rose, O.; Weigert, G. (Hrsg): Simulation und Optimierung in Produktion und Logistik. Berlin, Heidelberg: Springer 2011, S. 3-12.

Merkuryev, Y.; Tolujew, J.; Blümel, E.; Novitsky, L.; Ginters, E.; Viktorova, E.; Merkuryeva, G.; Pronins, J.: A modelling and simulation methodology for managing the Riga harbour container terminal. Simulation 71 (1998) 2, S. 84-95.

Novák, P.; Šindelář, R.: Applications of ontologies for assembling simulation models of industrial systems. In: Meersman, R.; Dillon, T.; Herrero, P. (Hrsg.): On the Move to Meaningful Internet Systems: OTM 2011 Workshops, Hersonissos (Greece), 17.-21. Oktober 2011, S. 148-157.

PIANC MarCom Working Group 121: Harbour Approach Channels – Design Guidelines. MarCom report 121, January 2014.

Stahlbock R.; Voß, S.: Operations research at container terminals: A literature update. OR Spectrum 30 (2008) 1, S. 1-52.

Steenken, D.; Voß, S.; Stahlbock, R.: Container terminal operations and operations research - a classification and literature review. OR Spectrum 26 (2004) 1, S. 3-49.

UNCTAD Secretariat: Review of Maritime Transport. United Nations Publication, Schweiz, 2014.

Wang, T.-F.; Cullinane, K.: The efficiency of european container terminals and implications for supply chain management. Maritime Economics & Logistics 8 (2006) 1, S. 82-99.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Measuring the Performance Availability of a Logistics Application Controlled by a Wireless Sensor Network

## *Messung der Leistungsverfügbarkeit in einem über drahtlose Netzwerke gesteuerten logistischen System*

Moritz Roidl, Steffen Schieweck, Michael ten Hompel, TU Dortmund, Dortmund (Germany), moritz.roidl@tu-dortmund.de, steffen.schieweck@tu-dortmund.de, michael.ten.hompel@iml.fraunhofer.de

Abstract: Performance availability as a Key Performance Indicator (KPI) for facility logistics systems has become a topic of interest in recent years. It deals with the computation of an entire system's performance which is influenced by the availability of its subsystems. The importance of high-quality planning of facility logistics systems also increased dramatically regarding cost reduction and heightened demands. Simultaneously, decentralised logistics systems controlled by wireless sensor networks have taken a large step towards industrial application. An approach to measure performance availability in systems influenced by the effects of wireless sensor networks is presented here. A brief case study is provided to illustrate the effects of wireless sensor networks on performance availability.

## 1 Introduction

The two core elements that provide understanding of the desired performance of facility logistics systems are maximum performance and share of time at which the measured performance can be provided. Traditionally these two indicators have been tested separately at the launch of a facility logistics system (Maier 2012). In recent years, the concept of performance availability has been established which aims to combine both indicators. For this purpose, the VDI 4486 (2012) guideline provides handover testing procedures after the launch of the system as well as required definitions and agreements. The guideline is meant to be part of the contractual agreements between the vendor and the operator of facility logistics systems.

Typically, for the forecasting of facility logistics systems KPIs, a system gets subdivided into several manageable and isolated subsystems for which proven methods are available (ten Hompel et al. 2007). Consequently, interdependencies which occur between those subsystems are neglected. However, this may lead to insufficient results for strongly connected material flow systems, especially when

considering a system-focused KPI such as performance availability. Event-discrete simulation is highly suitable for the investigation of the timely behaviour of complex material flow systems (März et al. 2011). Though, the creation of simulation models is a complex, labour-intense and thus an expensive task.

For distributed systems, with resource constrained devices, the communication behaviour significantly influences the overall system performance. From the perspective of a classical logistics planner, the evaluation of a system would be merely a question of technical availability of the individual components. Nevertheless, there is currently no approach towards specific performance requirements for planning logistics systems that include wireless sensor networks.

A tool is presented here capable of measuring performance availability of material flow systems. To begin with, the basics of performance availability are described in chapter 2. Chapter 3 presents the implemented procedures for performance availability measurement. Special requirements for systems with wireless sensor networks are explained in chapter 4 subsequently the architecture of the tool is described in chapter 5. The results of some experiments comparing conventional systems to systems with wireless sensor networks are depicted in section 6. The article concludes with a brief summary and an outlook.

## 2 Performance Availability

Performance availability as a KPI has been established in the recent years. Practitioners agree that its principles strongly meet existing demands. In this section, an overview of the ongoing research is presented. A survey of relevant literature describes the definition of performance availability which is applied in this article.

### 2.1 Literature Overview

Before the term performance availability was established, research mainly focused on the technical availability of an entity, “the probability of finding the observed units in a correctly functioning state at any given time during the operating period” (VDI 3581 2004, p. 2). The performance of a facility logistics system, in the context of this paper, is defined as the throughput of material flow objects of either the whole facility logistics system or its subsystems or entities. Performance availability has first been brought to context by Wittenstein (2007) who establishes the term in the domain of customer specific product development in systems and plant engineering as a state in which a process is conducted according to its specifications. Hence, required resources need to be available at any time independent of disturbances or fluctuations in demand. The focus of the article lies on the design of a structured procedure for the fulfilment of performance as defined by a customer in development assignments. Maier (2012) proposes practice-oriented acceptance procedures regarding reliability and availability standards, considering the analysis of existing methods and thus deducing recommendations. She concludes that coexisting standards are not or only partially suitable for facility logistics systems. The work lays the foundations for the VDI 4486 (2012) whose definition of performance availability will be used for the remainders of this article and is described with further detail in section 2.2. It provides standard procedures for the measurement of performance availability of existing facility logistics systems.

The VDI 4486 (2012) guideline has served as a landmark for succeeding publications, being the first official standard to deal with performance availability. Based on a preliminary version of the VDI 4486 (2012), Maier (2010) proposes an analytical method for the forecasting of performance availability. A more detailed approach to that matter is given by Jung and ten Hompel (2013) who aim to determine the number of delayed objects when passing a branch in a conveyor system. Hegmanns et al. (2014) presents a concept for the determination of performance availability with varying focuses. In this concept, a number of analytical methods and simulation models are used to cover different aspects of performance availability for different material flow technologies. Some principles of simulation modelling for the measurement of performance availability are proposed by Schieweck et al. (2014). Aiming to push the research towards practical applicability Schieweck and Jung (2015) discuss the demands of practitioners and the current degree of fulfilment provided by the published research.

## 2.2 Definition

As stated in the previous section, the VDI 4486 (2012) guideline is considered the prevailing standard definition of performance availability. Hence, its definitions and measurement procedures are described briefly in this section. “The performance availability indicates the degree of fulfilment of processes agreed between contract parties (manufacturer and user) in accordance with the requirements and deadlines and in compliance with the agreed basic conditions” (VDI 4486 2012, p. 3). Thus, performance availability is defined as the degree of fulfilment of specified business goals. Moreover, measurement processes are defined which add a quantitative dimension to the term which is required for a verifiable contractual agreement.

The measurement as defined by VDI 4486 (2012) requires the definition of business goals. Those need to be either measurable in waiting times or the timeliness of the fulfilment of orders (VDI 4486 uses the term ‘running times’). Examples for business goals are the utilization of pickers or the service level of a picking process. When considering waiting times the degree of fulfilment is given by the ratio of the overall time $T_B$ minus the waiting time $T_W$ and the overall time $T_B$ (VDI 4486 2012, p.6).

$$\eta_W = \frac{T_B - T_W}{T_B} \tag{1}$$

It must be noted that the measurement of performance availability needs to be conducted in several observation periods for each of which the fulfilment of performance specifications is checked. In any case the desired performance was provided within the observation period, the performance availability value is set to 100%. For any other case the aforementioned formula is used. For the determination of the overall performance availability the average of the considered observation periods is calculated. The consideration of timeliness requires the counting of the overall number of orders N and the delayed orders n at the time of completion (VDI 4486 2012, p.7).

$$\eta_L = \frac{N - n}{N} \tag{2}$$

As before, first the achievement of specified performance needs to be checked for the observation period before applying the formula. Also, the overall performance availability is calculated as the average of the relevant values for the observation periods. The real-world acceptance procedure not only requires the specification of performance values but also the time in which a disturbance within the system needs to be resolved. Further conditions and agreements for the acceptance procedures can be found in VDI 4486 (2012). The guideline can also be referred to regarding useful examples and more formal definitions.

# 3 Performance Availability Measurement

Basic principles of the measurement of performance availability have been discussed by Schieweck et al. (2014). In this section, those principles will be concretized by describing the procedures implemented to the designed simulation tool. To begin with, the necessary input and output data are defined. Following, a description of the implemented structure for the proposed tool is provided.

## 3.1 Data

The defined data are subdivided into the categories of input and output data. The input data are necessary for the simulation and need to be created before the simulation runs. The output data are gained during the simulation runs, exported, analyzed and interpreted. Necessary input data sets in terms of the evaluation of performance availability are:

- picking list
- initial inventory
- inflow
- due dates

All those data sets are part of a real-world facility logistics system. The picking list defines the load of the picking system and comprises a number of picking lines. Usually, one picking line consists of the time of order creation, order number, number of the picking line, article number, and pieces per article. The initial inventory describes the amount and type of articles which are in stock at the time of simulation start. The inflow describes the flow of goods which arrive at the warehouse and are to be stored in the system. The due dates are defined for every order and specify their latest time of completion. Their creation and definition have considerable effect on the performance availability value when considering timeliness. Commonly available for simulation studies in industrial projects are the former three. An approach for the creation of practically applicable due dates for facility logistics systems is provided by Schieweck et al. (2014). Necessary output data for the assessment of performance availability are:

- waiting time
- throughput
- creation time

- completion time

The collection of the aggregated waiting time of a resource is necessary for the first of the aforementioned definition of performance availability (see formula 1). The throughput for every observation period is required as well, following the definition in section 2.2. For the assessment of timeliness, creation and completion time of every order has to be recorded.

## 3.2 Tool Structure

The developed simulation tool aims to use the advantages of two available programs, namely Microsoft Excel and Demo3D by Emulate3D Ltd. The structure of the warehousing system can easily and quickly be defined in Microsoft Excel by inserting characters and tokens into the cells. The entire warehousing system from receiving to shipping area may be modelled. The structure is then imported to the environment of Demo3D in which the model is instantly able to operate and create results. The results are automatically fetched and exported, again to the common environment of Microsoft Excel. Using a predefined folder structure, a user is able to define several scenarios with variable layouts and/or data which can be tested one after the other without any further interaction. A considerable reduction of model-building and testing time for the measurement of performance availability can be achieved.

The creation and/or preparation of the necessary input data depends highly on the individual situation and is therefore not treated within this article. The existing data needs to be brought to a standardized format. After the data was inserted to the tool, it automatically creates CSV-files which keep all of the required data for further usage. Analogically, the defined layout and its specifications are exported from Microsoft Excel into CSV-files. Within Demo3D scripting objects import the CSV-files, create the defined layout within the 3D-environment and store the data in Demo3D-internal tables. Now, the model allows instantly for running and testing. A more detailed description of the simulation model's internal functionality is given in section 4.

The recording of the data is coded within the individually created modules (e.g. conveyors, pickers) as well as in global scripting objects which handle the exporting of the data. The waiting time is recorded by each module individually. Every time a load on the module is willing to proceed but has to remain on the module the timeframe is recorded and added to the overall waiting time of the module. At the end of the simulation run a scripting object called "data fetcher" initiates a broadcast to all of the relevant modules requesting a waiting time report. Each module sends its own report. The fetched data is stored into a Demo3D-internal table and exported as an XLS-file. The measurement of timeliness can either be conducted internally in the model or externally in Microsoft Excel. When being conducted internally the method proposed by Schieweck et al. (2014) is implemented to the module which initiates the creation of order totes. Instantly, every order tote gets a due date assigned. At the shipping department, the due date is compared with the actual time and the order tote receives the tag "on time" or "delayed". The data fetcher also requests timeliness reports from the shipping department. It must be noted that the waiting times and timeliness require the assessment of the achievement of the

throughput specifications before finally being able to be used for the calculation of performance availability.

The various scenarios are defined within scenario folders. The scenarios are identified with a distinct number. Each scenario has its own data files (input and output) as well as one layout file each. Within the tool, the user defines the run duration of each scenario and selects the scenarios which are to be tested by entering the scenario number. At the beginning of each scenario run, its individual layout is imported. During the scenario run the output data are collected and after the scenario has finished they are exported into Microsoft Excel tables within the associated scenario folder.

## 4 Wireless Sensor Networks in Facility Logistics Systems

The application of cyber-physical and embedded systems has been developed in the field of logistics during recent years as a part of the Internet of Things (IoT) revolution. Current research in the field of logistics investigates the massive use of distributed embedded systems as a new basis for modern production, transportation and distribution strategies. These efforts can be summarized under the heading Industry 4.0. It describes a concept that builds on a cyber-physical system comprised of individual smart devices taking over control from the production planning system. These devices organize the transport in a self-sufficient manner while using services of external systems along the supply chain. While the research in these decentralized controlled production and logistics systems has produced promising results in the recent years, the idea of a completely self-sufficient individual smart device has not yet found a convincing and practical implementation. Recent developments in low power technology for wireless sensor networks make it apparent that a breakthrough might be imminent.

For systems of distributed and resource-constrained devices the communication behaviour significantly influences the possible overall system performance. In accordance with the requirements for developing smart industrial devices adjacent to the IoT compliance, the inBin (intelligent bin) has been developed. Not only does it stores data about its contents and knows its position, it works in an energy-neutral manner. In addition to interaction possibilities with the operator through its display and buttons it includes a wireless transceiver that enables communication with its materials handling central management system.

In an industrial real-world scenario, a large number of inBins will work together simultaneously in the manner of a wireless sensor network within one facility. inBins should be able to communicate within the system to assure its functionality. The communication infrastructure is based on a central server which handles all devices and there is no direct inter-device communication. Except for a centralized management of transport orders, the comprehensive system functions are provided by inBins. This includes regular status updates from the nodes to the management system, query capability for specific articles, as well as distributed decision strategies for delegating a transport request to individual bins.

In contrast to other applications for sensor networks this architecture is feasible due to the highly controlled environments in facility logistics.

## 5 Simulation of Wireless Sensor Networks

The simulation proposed in this article evaluates the performance availability of a large number of inBins within a material flow simulation tool. As such, the main focus lies on the correct modelling of the application layer protocol which controls the material flow system in the manner of a multi-agent-based system. There are two main agent types in the system: the mobile inBins and the stationary material flow equipment. The inBins share a limited radio resource channel with a simple infrastructure by using a rudimentary communication protocol while the rest of the system entities (i.e. conveyors) are considered to be connected by cable. A single transceiver on the server side can reach all inBins. The communication protocol is based on time slots of uniform length that cover a complete request-acknowledge cycle with status data exchanged between the devices and the server.

The main characteristic of wireless communication in relation to performance availability is the unreliable timing of messages. The limited radio resource channel means that any time a specific inBin tries to send a message it can be delayed by messaging activity of other inBins. In addition, the mobility of the inBins restricts the energy that is available for active radio communication, so there can also be communication delays due to energy shortage. These delays may increase further when communication protocols are considered that represent a many-to-many relationship between inBins during multi-agent-based decision-making. In this case, one inBin might have to wait for the answer from an unknown number of inBins to be able to decide, for example, the most time-efficient route through a material-flow system. This can be handled by using timeouts that restrict the waiting time after which the decision will be made. The timeout has to be chosen sufficiently long to get enough answers for a decision and, thus, has to be based on the maximum expected latency. The more inBins try to answer the request the higher latency will be since the communication medium is shared.

The Demo3D-based tool was extended to include a network object through which all messages between agents (i.e. inBins and conveyors, etc.) are transmitted. The messages are modelled as property maps which include the sender and receiver and message type information, as well as the payload data. The network has a lookup table with all registered agents and offers two services as script functions: the sending of a message (one-to-one) and a message broadcast (one-to-many). These functions can be called by an agent in a non-blocking way, so that the message will be processed independently by the network while the agent program continues to run. The network then applies a latency to the message and holds the message property table until the delivery time has been reached. It then looks up the recipient of the message and delivers the message property table through a script function call on the receiving agent.

Based on this architecture, all logistics processes are modelled decentralized and communication-based in the simulation tool. One advantage of using the central network object is the possibility to visualize the messages within the simulation tool, and thus combining material and information flow into a unified view of the facility logistics system (see figure 1). This can help while debugging the logistics processes which depend on the correct interaction between the physical objects and virtual agents.

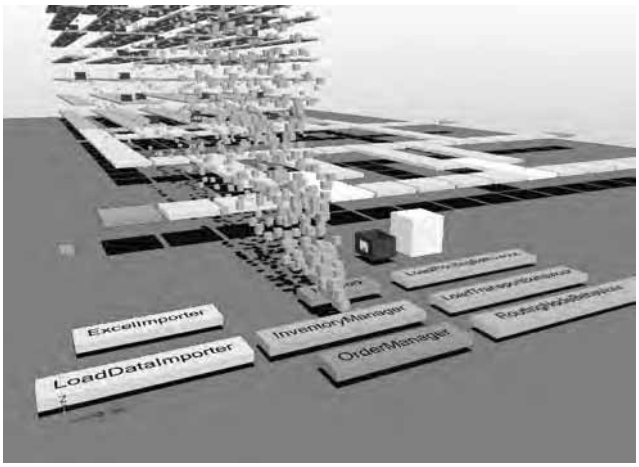



*Figure 1: Visualisation of the communication within a facility logistics system. Shown is the initialisation phase during which the intelligent containers (inBins) report to the inventory manager through wireless messaging.*

# 6 Case Study

The presented case study is aimed to demonstrate the functionality of the concept and presents some results about the impact of the network latency occurring in large logistics systems controlled by wireless sensor networks. First, the case study system and the data are defined. Following, some results will be presented and discussed.

## 6.1 System and Data

The case study system consists of a miniload storage area, a picking station, the connecting conveyor loop and a conveyor system for the order totes. The miniload system comprises 3 aisles with a storage capacity of 100 storage locations each. The overall storage capacity is 300 storage locations out of which 270 are occupied at the start of the simulation runs. The resulting storage utilization is 90%. At the picking station, each order line requires a time of 5 seconds for picking. The picking process may only be conducted when the required storage and order totes are at the desired positions at the picking station. The orders are defined as ‘completed’ as soon as the associated order tote arrives at the ‘shipping department’ which is shortly after the picking station.

Among the defined testing scenarios, the layout is not changed. For every hour of simulation time a mean of 50 orders are put into the system. The different scenarios are characterized by various sizes of the orders from 4 to 8 picking lines. All of the scenarios are tested incorporating a latency of 50 to 110 milliseconds for every message sent via the wireless sensor network. To assess the influence of the latency another set of testing runs are conducted with no latency at all. The due dates which are required for the investigation of the timeliness are created by the method proposed by Schieweck et al. (2014) with 4 routes, a departure pulse of 5 minutes and a critical value h which has a basic allowance of 40 and adds 5 seconds for every order line number greater than 4.

## 6.2 Results

For the assessment of the latencies influence two performance availability values are measured. First, every time a conveyor module requests a transport to the next conveyor module but has to wait due to missing capacity, its state is characterised as 'blocked'. Every occurring timeframe is recorded as a waiting time as defined in formula (1). For the presented case the performance availability of the conveyor system is tested. To achieve this, the average value of the waiting time of the conveying modules is used then the timeliness of an order is checked. For the investigation of the performance availabilities behaviour, it is assumed that the specified performance is not reached at any point in time. Hence, all of the waiting times and tardy jobs are included to the calculation. The results are depicted in figure 2.

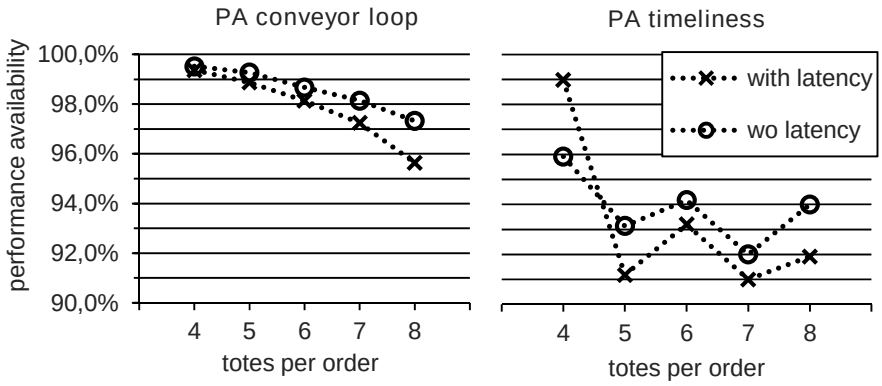



*Figure 2: Performance availability results*

Generally, the performance availability of systems with considerable latency is lower than in systems with insignificant latency. This is easily explainable as the latency causes slower processes which result in larger waiting and throughput times. Looking at the performance availability of the conveyor loop this thesis can be supported. Even more crucial is the observation that with increasing load to the system (aka higher number of totes per order) the performance availability for scenarios with latency decreases disproportionately. This has special significance in facility logistics systems when the maximum performance is required while maximum system load appears. The performance availability based on the timeliness of the orders also supports the thesis of lower values for systems with latency.

# 7 Summary and Outlook

In this article an approach to measure performance availability in systems influenced by the effects of wireless sensor networks is presented. A brief case study based on a simulation tool developed in Demo3D provides an illustration of the effects and a proof-of-concept for the developed tool. Future research has to be conducted to investigate influencing factors on performance availability for several system types

as well as the allowance on network latency for real-world systems. The latter has two key points of interest. Either the latency of a realistically sized decentralized system is measured or the maximum allowance for latency can be investigated for the achievement of satisfactory performance availability values. The simulation of real-world network latency might be conducted by simulation coupling (i.e. HLA).

## Acknowledgement

The presented research is funded by the German Research Foundation DFG within the subprojects C1 and C5 which are part of the project PAK 672 (HO 2403/8-1).

## References

Hegmanns, T.; Kuhn, A.; Roidl, M.; Schieweck, S.; ten Hompel, M.; Güller, M.; Austerjost, M.; Roßmann, J,; Eilers, K.: Planung und Berechnung der systemischen Leistungsverfügbarkeit komplexer Logistiksysteme. Logistics Journal: Proceedings. 10. Fachkolloquium der WGTL, Dortmund. Vol. 10(1) 2014.

Jung, E.-N.; ten Hompel, M.: Analytische Stauprognose in Stetigfördersystemen im Rahmen der Systemplanung. Logistics Journal: Proceedings. 9. Fachkolloquium der WGTL, Dortmund. Vol. 9(1) 2013.

Maier, M. M.: Prognose der Leistungsverfügbarkeit während der Planung. Logistics Journal: Proceedings. 6. Fachkolloquium der WGTL, München. Vol. 6 (1) 2010.

Maier, M. M.: Praxisgerechte Abnahmeprozeduren für intralogistische Systeme unter Berücksichtigung der Zuverlässigkeits- und Verfügbarkeitstheorie. Dissertation, Tübingen University. Düsseldorf: VDI Verlag 2012.

März, L.; Krug, W.; Rose, O.; Weigert, G.: Simulation und Optimierung in Produktion und Logistik. Berlin, Heidelberg: Springer Verlag 2011.

Schieweck, S.; Dregger, J.; Kaczmarek, S., ten Hompel, M.: Reliability of facility logistics systems - simulating performance availability. World Academy of Science, Engineering and Technology, International Science Index, Architectural and Environmental Engineering 2(1). pp. 135-142, 2014.

Schieweck, S.; Jung, E.-N.; ten Hompel, M.: Leistungsverfügbarkeit zwischen Theorie und Praxis. Proceedings Symposium "Leistungsverfügbarkeit in der Logistik". Dortmund, 2015 (accepted).

Ten Hompel, M.; Schmidt, T.; Nagel, L.: Materialflusssysteme: Förder- und Lagertechnik. Berlin: Springer 2007.

Verein Deutscher Ingenieure (VDI): Verfügbarkeit von Transport- und Lageranlagen sowie deren Teilsysteme und Elemente. VDI 3581. Berlin: Beuth 2004.

Verein Deutscher Ingenieure (VDI): Zuverlässigkeit in der Intralogistik - Leistungsverfügbarkeit. VDI 4486. Berlin: Beuth 2012.

Wittenstein, A.-K.: Bedarfssynchrone Leistungsverfügbarkeit in der kundenspezifischen Produktentwicklung. Dissertation, Stuttgart University. Heimsheim: Jost-Jetter 2007.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Performance Availability Analysis of Autonomous Intralogistic Systems: An Agent-based Simulation Approach

***Leistungsverfügbarkeit autonomer intralogistischer Systeme: Ein agentenbasierter Simulations-Ansatz***

Mustafa Güller, Elif Karakaya, Tobias Hegmanns, TU Dortmund, Dortmund (Germany), guller@lfo.tu-dortmund.de, karakaya@ lfo.tu dortmund.de, hegmanns@ lfo.tu dortmund.de

Abstract: In recent years, the ability to respond to real time changes in operations, agility in the turbulent market environment, and reconfigurability in equipment are likely to become essential characteristics for next generation intralogistics systems to deal with the dynamic environment. The Cellular Transport System (CTS) that aims to cope with these new requirements provides an efficient way to increase the flexibility and changeability of intralogistics systems. The main objective of this paper is to analyse the performance availability of the Cellular Transport System under various pre-defined scenarios in order to support the design process. Since analytical models have several simplifications and constraints, agent-based simulation (ABS) is used to evaluate the performance. Our results show that delays due to collisions between agents significantly impact throughput capacity, cycle times and performance availability.

## 1 Introduction

In recent years, the ability to respond to real time changes in operations, agility in the turbulent market environment, and reconfigurability in the equipment are likely to become essential characteristics for next generation intralogistics systems to deal with the dynamic environment. Hence, firms need to develop strategies and technologies that provide flexibility to succeed in reaction to changing conditions. However, it seems to be difficult to achieve within the rigid, specific and specialised world of automated material handling systems which offers static and inflexible solutions (Marin and Carrasco-Gallego 2013). In this regard, autonomous vehicles have been widely adopted as a key component to intralogistic systems in order to improve adaptability by physical flexibility in disposition, routing and space consumption (Güller and Hegmanns 2014). Furmans et al. (2010) offer an alternative approach called the FlexConveyor to conventional conveying systems, which is

a decentralised modular, unit-sized conveyor system. In this system, each module is able to convey in the four cardinal directions and connected by a serial connection, which is used to exchange all necessary information between adjacent modules. Seibold et al. (2013) analyse how powerful the FlexConveyor is in sorting of goods and what kind of layouts are most suitable. An extension of the FlexConveyor system is KARIS project developed by the Institute for Conveying Technology and Logistics (IFL) at the Karlsruhe Institute of Technology. It consists of several vehicles which are able to drive at the floor or stand at the floor while acting as a conveyor (Ommen et al. 2009). KARIS is a decentralised controlled system that enables quick installation and reaction to changing requirements by autonomous reconfiguration of the KARIS vehicles (Trenkle et al. 2013). It provides a wider range of scalability and reconfigurability. Based on the idea of puzzle-movement, a high-density storage system for physical goods called GridStore is introduced by Gue (2006). Some recent work has addressed autonomous vehicle based storage and retrieval systems (AVS/RS) as an alternative to crane-based AS/RS (Roy and Krishnamurthy 2011; Heragu et al. 2009; Ekren and Heragu 2011). Instead of aisle captive cranes, AVS/RSs employ a number of autonomous vehicles that are used to carry palletised loads and are capable of providing horizontal movement within a tier. More recently, a new automated material handling technology called the Cellular Transport System (CTS) is developed by Fraunhofer Institute for Material Flow and Logistics (IML) for the storage and order picking of small items. In order to cope with rigid design limitations, a group of dynamic, flexible mobile vehicles called Multishuttle Move (MSM) replace with the inflexible continuous conveyor system. The Multishuttle Move (MSM) is a novel fusion of conventional shuttle and automated guided vehicle system (Kamagaew et al. 2011). In contrast to automated guided vehicles (AGVs), MSMs can move on rack levels as well as freely within the warehouse. In other words, all transports in the rack and the surrounding area are covered with an autonomous vehicle swarm.

Since analytical methods have some limitations to handle the degree of complexity inherent to real-world problems, simulation method is today recognised as among the most promising tool for detailed investigations and reliable problem solving of complex systems (Chen et al. 2013). Most of the research concerning warehouse simulation (Liong and Careen 2009; Colla and Nastasi 2010; Ekren and Heragu 2011) has concentrate on the discrete-event simulation technique. However, using this technique, it is difficult to model the control of autonomous components in a system. Hence, researchers have shifted their attention on the agent-based modelling and simulation (ABMS) approach. Agent-based simulation, which is composed of different interacting computing entities called agents, offers an alternative way to design and implement complex behaviours of system components (Wooldridge 2002). Güller et al. (2013) have recently developed an agent-based simulation model of CTS for an experimental hall. In this paper, we extend the previous research through evaluating a novel performance measure called “performance availability” and the average cycle time with different time components.

## 2 Definition and Overview of the Performance Availability

Various performance indicators have been proposed and applied to warehouse operations in the literature, such as the utilisation which denotes the fraction of time in which the server is occupied and the system throughput which is defined as the number of customers served in a single time unit. Other major performance indicator is the performance availability that is defined in VDI-Guideline 4486 as “the degree of fulfilment of processes agreed between contract parties in accordance with the requirements and deadlines and in compliance with the agreed basic conditions” (VDI10 2010). Two measures are used to quantify the degree of fulfilment of the performance availability; the waiting time and the running time. The term "performance availability" was first introduced by Wittenstein (2007). It is defined as the state of a system in which a process is carried out according to requirement and the required result can be completed on time. Four essential steps are defined to reach the performance availability (Maier 2012):

1. Formulation of the business objective: The new system has the task of the operator to facilitate the achievement of its business objectives or facilitate. Therefore it is necessary that these goals are concretely defined.
2. Formulation of logistics processes: The business objectives are achieved by various logistics processes that are carried out successfully on the system. These processes must also be defined and quantified.
3. Formulation of boundary conditions: In order to measure and evaluate the performance in a meaningful way, reliable boundary conditions must be defined, based on which the necessary resources can be scheduled.
4. The difference between consequences when process disturbances occur:

If undesirable waiting times occur at the considered workplace due to a disturbance, the performance availability $\eta_W$ of this workplace is calculated as follow ($T_B$ is the observed time and $T_W$ is the waiting time in observed period):

$$\eta_W = \frac{T_B - T_W}{T_B} \tag{1}$$

If the process is not completed at a certain time due to the lack of availability, the performance availability $\eta_L$ is calculated as follow (N is the total load and n is the delayed loads in observed time):

$$\eta_L = \frac{N - n}{N} \tag{2}$$

Nevertheless, the above definition is not utilised directly for the assessment of performance of entire logistic systems. According to Klaus and Krieger (2009), a logistic process consists of a number of activities that is comprised of a measurable input, which is converted by a transformation into a measurable output. To meet business objectives, output of processes must be controlled by performance indicators, which usually involve efficiency and effectiveness metrics (Schmelzer and Sesselmann 2008). Implementing an appropriate performance measurement in increasingly dynamic and complex environments is critical to ensure that the agreement between the provider and customer is aligned with performance requirements of the system. However, defining, setting up and implementing an

effective performance management system for a highly turbulent business environment are challenges to be reached in order to adjust to fluctuating conditions of customer needs (Azevedo and Francisco 2007). These fluctuating conditions in internal and external environment must be already considered in the planning phase (Schuh et al. 2012). The system has to be ensured that these changes can be realised within a pre-defined and limited scope of action called flexibility corridor as shown in Figure 1. Until recently, little attention has been paid to the relationship between the performance availability and flexibility corridors.

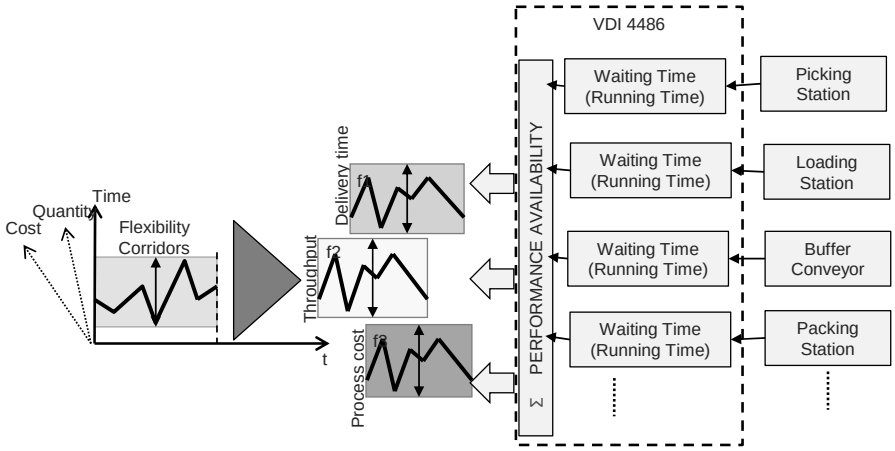



*Figure 1: Performance availability with flexibility corridors of performance dimensions*

## 3 Agent-Based Modelling for the Cellular Transport System

In order to manage the complexity of autonomous control of the Cellular Transport System, we have developed a simulation environment using agent-based modelling. Agent-based modelling (ABM) can be performed either by using programming languages or by using readily available simulation software packages that address the special requirements of agent modelling. Our simulation has been implemented in AnyLogic™, which is Java-based simulation software (AnyLogic 2013). One major advantage of AnyLogic is that it allows the user to create a simulation model combining differential equations; discrete events and agent based modelling. Another reason to use AnyLogic is the possibility to use Java code at any place of the program and an object oriented structure. This allows high freedom to expand the standard feature set of the tool.

There are several issues which must be considered during the simulation of mobile shuttles such as the number of shuttles, the path determination, and the control logic for dispatching shuttles. However, using traditional simulation approaches it is difficult to model the autonomous control of shuttles in a warehouse. Besides, they are insufficient to model complex behavioural patterns. Collision detection and deadlocks contribute to the complexity of autonomous shuttles' behaviour in an

open path manner in which shuttles are not guided by a floor path. In this context, ABM provides good background for solving collision detection and avoidance issues in distributed and decentralised approaches. Another contribution of ABM are decision-making capabilities, based on perception of the environment and knowledge obtained by interaction between agents. The developed simulation model is composed of a set of agents that communicate with one another by asynchronous message passing. The different developed agents that are captured to model consist of MSM agents, Lift agents, Enter/Exit agents and Workstation agents as illustrated in Figure 2. All these agents have their own characteristics and logic of behaviour.

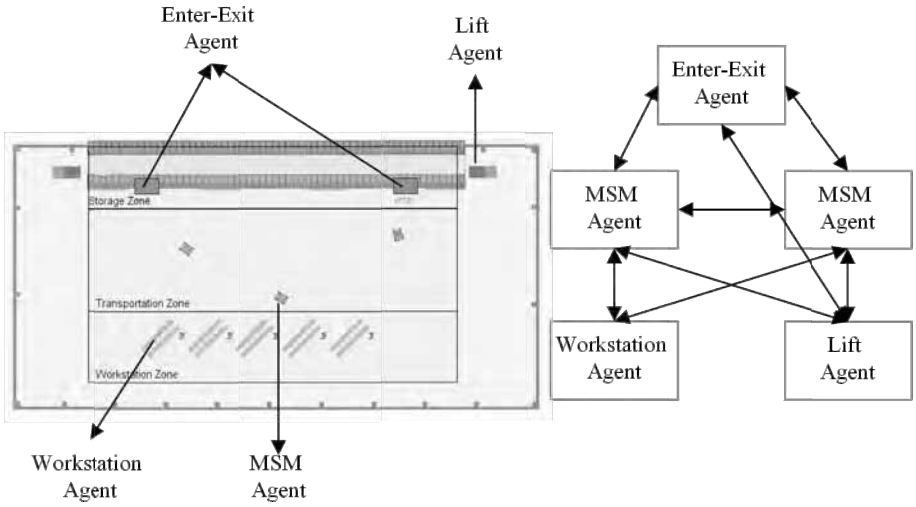



*Figure 2: Agent-based simulation concept of the Cellular Transport System*

*MSM Agent*: each MSM in the system is controlled by an MSM agent that is responsible for transporting bins to their destination. An MSM agent sends and receives information from the lift agent, workstation agent, enter/exit agents, and other MSM agents in the same environment. It receives information from the order pool regarding location of pick-up and drop-off workstations. Furthermore, agents can be programmed to move in x, y, and z directions. *moveTo(x, y, z)* is an example of the built-in function of the agent that starts a straight movement to a given destination point (x,y,z) in 3D space.

*Enter/Exit Agents*: there are two main reasons for adding these agents. One is to perform the rotation duration of the MSM at storage entry and exit points. When an MSM is moving inside/outside the storage area, it spends time within the enter/exit zone to perform the rotation of the MSM. The other reason is to avoid deadlock at storage entry/exit point. When an MSM agent arrives at the enter zone, the enter agent communicates with the lift agent to check the queue length between the entry point and the lift. Depending on the situation of the free space between the enter point and the lift, the MSM agent receives the permission to enter the storage area. On the other hand, when an MSM agent is ready to leave the storage area, the lift

agent communicates with the exit agent. If there is enough space between the lift and exit point, the MSM agent receives the permission to continue the movement.

*Workstation Agent*: each one of the order picking stations in the system is represented by a workstation agent that is responsible for the execution of tasks for corresponding orders at the picking station. Order pickers collect the orders from Eurobins and pack them into custom bins. After one bin is processed by the picker, it can be taken back by the MSM agent to its home position in the rack.

*Lift Agents:* These lift agents are introduced to control the vertical movement of MSMs in the rack system. Whenever an MSM agent arrives at the lift, it informs the lift agent about the target tier and the lift agent performs the vertical movement.

## 4 Simulation Experiments and Results

The goal of this section is to analyse the performance availability of the Cellular Transport System under different scenarios by using the developed Agent-based Simulation model. Simulation scenarios consist of different environmental factors, such as varying the number of order lines, stochastic demand, and number of Multishuttle Moves (MSM) in the system. The system is triggered by orders that enter the system at any time. Customer orders can have one or more order lines (product line), where each order line consists of a particular item type and quantity of the requested orders. Each of these items can be created as a separate line item, which rolls up into one order. In order to show the effect of different order structures on the performance of the system, the number of order lines per order is varied from 1 to 4. Other assumptions used in the simulation model are summarised below:

- The sample problem is considered with approximately 600 storage positions.
- The model is run for 15 independent replications and one day shift (8 hours length).
- Customer order arrivals in the system follow an exponential distribution with the arrival rates 60 per hour.
- The order size varies between 1 and 6 units and is generated from a uniform distribution.
- The transactions are served by MSMs on a first-come first-served (FCFS) rule.
- The system uses pure random storage policy. According to this policy, the probability that a retrieval transaction is required in a certain storage point is identical for each point.
- Vertical velocity of lift is assumed 2 m/s. The MSM move their loads with a speed of 1m/s on the floor and 2m/s in the rack system. Acceleration and deceleration delays for the MSM and lift are not considered.

In particular, we are interested in throughput of goods (the mean number of order that can be handled per hour by the system) and average cycle time (the time from arrival of an order until the time the order is disposed). As throughput may be estimated as the inverse of the average cycle time in conventional automated systems (i.e. AS/RS), AVS/RS this approach is not applicable due to the use of different resources (i.e. lifts and vehicles). Furthermore, throughput is a function of bottlenecks (i.e. of vehicles or lift). The creation of bottlenecks is a function of the rack configuration and kinematic features of vehicles and lifts. As Equation (1) shows, the cycle time ($\tau$) is the sum of four time components, namely the time spent

in the transportation zone, the time for vertical movements, the waiting time for lifts and the travel time in the rack.

$$\tau = \tau_1 + \tau_2 + \tau_3 + \tau_4 \quad (3)$$

where:

$\tau_1$: is the time spent in the transportation zone (time-transport);

$\tau_2$: is the time for vertical movement (time-lift);

$\tau_3$: is the waiting time for lifts and the entry point of storage (time-waiting);

$\tau_4$: is the time spent in the rack (time-storage).

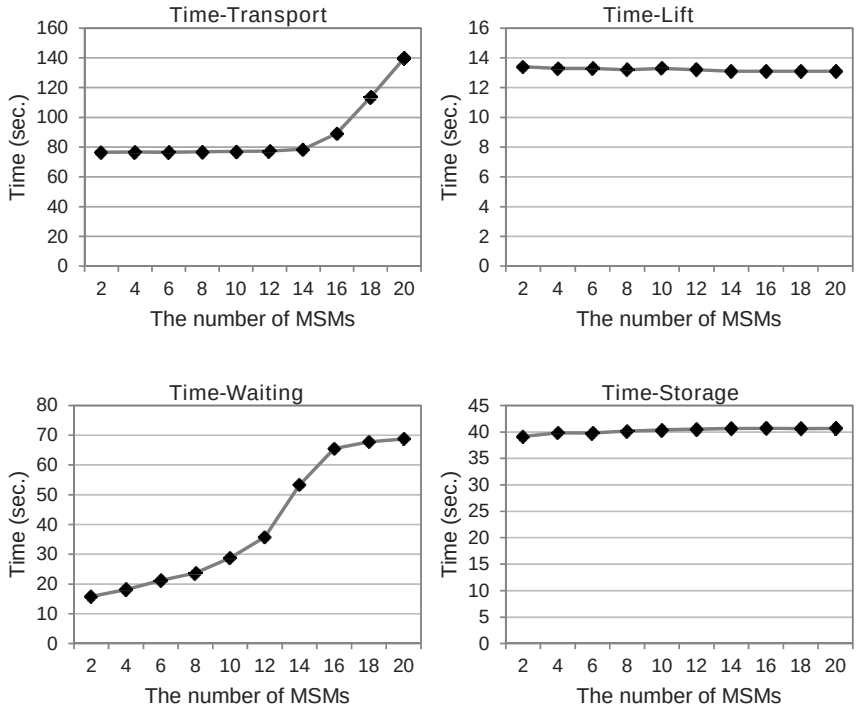



*Figure 3: Illustration of the impact of the number of MSM on the cycle time component*

Figure 3 shows the relationship between the cycle time components and the MSM number under the scenario I. Clearly, the MSM number does not significantly affect the travel time of lift ($\tau_2$) and storage ($\tau_4$) whereas it has a high impact on the waiting time for lifts ($\tau_3$). Furthermore, it is interesting to observe that as the number of MSMs increases, time spent in transportation area ($\tau_1$) remains nearly same. However, the transport time increases significantly with the respect to the increase of the number of MSMs, when it is higher than 14. This is due to the fact that

collision effects at the transportation area and congestion at the entrance of storage start to become significant when the number of MSMs increases to a large number. As Figure 3 shows, the waiting time for lifts increases as the number of MSMs in the system is increased, but when it is higher than 14, the slope of the diagram is lower, because the enter agent limits the entry capacity of storage.

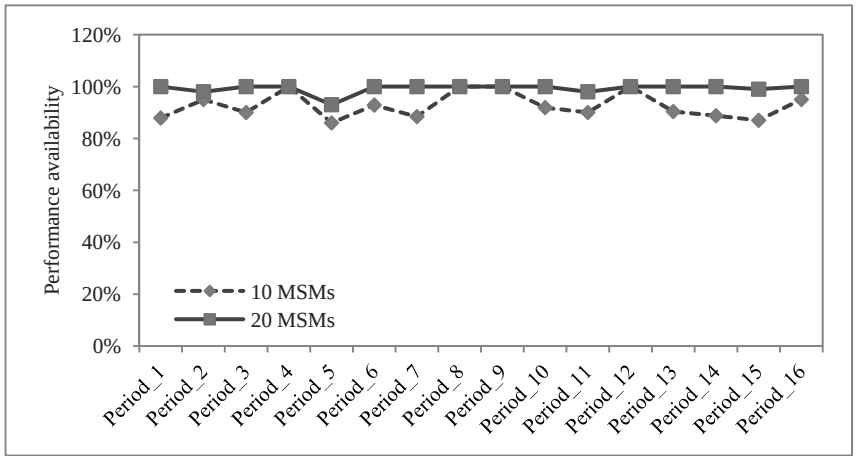



*Figure 4: Performance availability of the CTS and order picking process related to the number of MSMs*

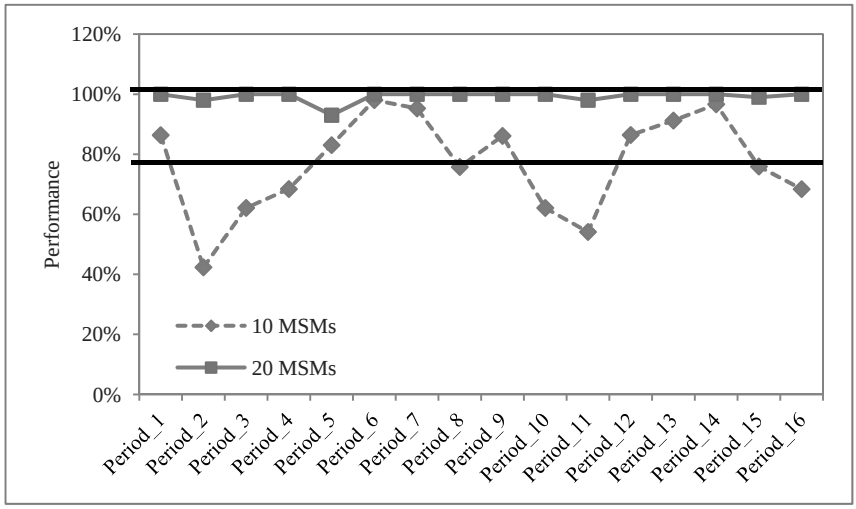



*Figure 5: Performance of the CTS based on on-time delivery*

The required performance of 60 item order/hour at the order picking stations corresponds to a cycle time of 60 s. The order picking stations are considered not available as long as the order picker must wait for the next item box. The performance availability of order picking process is calculated based on Equation 2. Figure 4 shows

the performance availability of the system in different time intervals of the observed period related to the number of MSMs. The throughput capacity and performance availability of the order picking process can be increased by increasing the number of MSMs. Performance of logistic processes is often measured from dimensions such as, time, quality, quantity, product, and cost. Figure 5 illustrates the performance of CTS based on on-time delivery dimension for different periods. The flexibility corridor of service level is assumed between 80% and 100%. As it can be seen, the system is out of the performance corridor for most of the period under 10 MSMs.

## 5 Conclusion

The evolution in intralogistics systems put forward new challenging requirements. Today, flexibility, reconfigurability, and high availability are more important than level of automation, cost effectiveness, and maximum throughput. Due to dynamic changes and uncertain environment, such as order variations, product diversity, and load variations, vehicles must be able to adapt their behaviour with changing circumstances. The Multishuttle Move (MSM) allows for a more efficient usage of available floor space and does not require any prior path definition. In this paper, we have presented an agent-based simulation approach to investigate the performance availability of Cellular Transport System that aims to cope with these new requirements. The simulation platform presented provides a basis to collisions with other vehicles in motion. Experiments have demonstrated that there is a relationship between throughput of the system, the performance availability and the efficient number of MSMs in the decentralised architecture.

The presented simulation model is today limited to the laboratory or trial hall installed at Fraunhofer IML. However, it is possible to extend the developed model to the simulation of an entire warehouse using an agent-based modelling architecture. Another limitation of this case study is that collisions are always solved by the prioritisation algorithm and the right-of-way rule. A possible solution would be the application of reducing its speed. In the near future, we are about to add architectural models to cope with order assignment and deadlock avoidance. Remaining challenge in Cellular Transport System is to optimise the physical layout of a distribution centre. In the long run, further experiments that are backed up by real-world validations are needed to be able to derive a model for a general design of Cellular Transport System.

## Acknowledgments

The findings illustrated in this paper are results of the research projects “Paketantrag 672 Logistics on Demand- Subproject C4 – Simulation-based Anticipatory Change Planning of Intralogistics Systems” funded by the German Research Foundation (Deutsche Forschungsgemeinschaft (DFG)).

## References

AnyLogic (2013), February, [Online], Available: http://www.anylogic.com/.

Azevedo, A. and Francisco, R.d.P.: Dynamic performance management in business networks environment. In: Cunha, P.F.; Maropoulos, P.G. (eds.) Digital Enterprise Technology. New York: Springer 2007.

Chen, X.; Ong, Y.S.; Tan, P.S.; Zhang, N.; Li, Z.: Agent-based modeling and simulation for supply chain risk management – A survey of the state-of the-art. Proceedings of 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester (UK), 13.-16. October 2013, pp. 1294-1299.

Colla, V. and Nastasi, G.: Modelling and simulation of an automated warehouse for the comparison of storage strategies. In: Rey, G.R.; Muneta, L.M. (eds.): Modelling Simulation and Optimization: InTech 2010, Available from: http://www.intechopen.com/books/modelling-simulation-and-optimization/modelling-and-simulation-of-anautomated-warehouse-for-the-comparison-of-storage-strategies

Ekren, B.Y.; Heragu, S.S.: Simulation based performance analysis of an autonomous vehicle storage and retrieval system. Simulation Modelling Practice and Theory 19 (2011) 7, pp. 1640–1650.

Furmans, K.; Schönung, F.; Gue, K.R.: Plug-and-work material handling systems. International Material Handling Research Colloquium 2010, Milwaukee, USA, 2010.

Gue, K.R.: Very high density storage systems. IIE Transactions 38 (2006) 1, pp. 79–90.

Güller, M. and Hegmanns, T.: Simulation-based performance analysis of a miniload multishuttle order picking system. In: ElMaraghy, H. (ed.): Variety Management in Manufacturing. Proceedings of 47th CIRP Conference on Manufacturing Systems, Windsor, Canada, 2014, pp. 475-480.

Güller, M.; Uygun, Y.; Karakaya, E.: Multi-agent simulation for the concept of cellular transport system in intralogistics. In: Azevedo, A. (ed.): Advances in Sustainable and Competitive Manufacturing Systems: Lecture Notes in Mechanical Engineering, Porto, Portogual, 2013, pp. 233-244.

Heragu, S.S.; Cai, X.; Krishnamurthy, A., Malmborg, C.J.: Analysis of autonomous vehicle storage and retrieval system by open queueing network. 5th Annual IEEE Conference on Automation Science and Engineering, Bangalore (India) 2009, S. 455-459.

Kamagaew, A.; Stenzel, J.; Nettsträter, A.; ten Hompel, M.: Concept of cellular transport systems in facility logistics. 5th International Conference on Automation, Robots and Applications, Wellington, New Zealand, 6.-8. December 2011, pp. 40-45.

Klaus, P. and Krieger, W.: Gabler Lexikon Logistik. Management logistischer Netzwerke und Flüsse , 3rd edition. Wiesbaden: Gabler 2009.

Liong , C.-Y. and Careen, S.E.L.: A simulation study of warehouse loading and unloading systems using Arena. Journal of Quality Measurement and Analysis 5 (2009) 2, pp. 45-56.

Maier, M.M.; Meurer, C.; Guido, F.: Die neue Richtlinie VDI 4486: Leistungsverfügbarkeit – Endlich kann Verfügbarkeit gemessen werden!. VDI-Gesellschaft Produktion und Logistik (Hrsg.), 21. Materialflusskongress, München 2012, pp. 15-27.

Marin, L.F.; Carrasco-Gallego, R.: A design framework for Flexible Automated Warehouses. 7th International Conference on Industrial Engineering and Industrial Management, Valladolid (Spain), 2013, pp. 463-471.

Ommen, D.; Beth, C.; Kamenik, J.: A system architecture for robotic movement of goods-approaches towards a cognitive material flow system. International Conference on Information in Control, Automation and Robotics, Milan (Italy), 2009, pp. 342-347.

Roy, D.; Krishnamurthy, A.: Performance analysis of unit-load transfer systems in multi-tier warehouses with autonomous vehicles. 44th CIRP Conference on Manufacturing Systems, Madison (WI), May 31- June 3 2011.

Schmelzer, H.J. and Sesselmann, W.: Geschäftsprozessmanagement in der Praxis, 6th edition. München: Carl Hanser 2008.

Schuh, G.; Potente, T.; Fuchs, S.; Thomas, C.: Quantitative approach for evaluation of company's changeability based on feedback data. POMS 23rd Annual Conference - Proceedings of the 23rd Annual Conference of the Production and Operations Management Society "Socially Responsible Operations", Chicago, Illinois (U.S.A), 2012.

Seibold, Z.; Stoll, T.; Furmans, K.: Layout-optimized sorting of goods with decentralized controlled conveying modules. IEEE International Systems Conference (SysCon 2013) , Orlando (FL), 15.-18. April 2013, pp. 628 - 633.

Trenkle, A.; Seibold, Z.; Stoll, T.: Safety requirements and safety functions for decentralized controlled autonomous systems. 24th International Conference on Information, Communication and Automation Technologies, Bosnia and Herzegovina, 30. October-1. November 2013, pp.1-6.

VDI Richtlinie 4486: Zuverlässigkeit in der Intralogistik: Leistungsverfügbarkeit. Berlin: Beuth 2010.

Wittenstein, Anna-Katharina: Bedarfs-synchrone Leistungsverfügbarkeit der kundenspezifischen Produktentwicklung. Dissertation, Heimsheim: Jost-Jetter 2007.

Wooldridge, M.: An introduction to Multiagent Systems. Chichester: John Wiley & Sons 2002.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Ein generisches Standardmodell zur Simulation spurgebundener Transportsysteme in der Halbleiterindustrie

## *An Adjustable Standard Simulation Model of Path Mover Transportation Systems in the Semiconductor Industry*

Sebastian Rank, Christian Hammel, Thorsten Schmidt, Technische Universität Dresden, Dresden (Germany), sebastian.rank@tu-dresden.de

Germar Schneider, Infineon Technologies Dresden GmbH, Dresden (Germany), germar.schneider@infineon.com

Abstract: Usually simulation studies of automated material handling systems in the semiconductor domain are time consuming because quite often highly detailed models are applied. These models only provide poor possibilities for customising and are costly in terms of runtime and build-up. Therefore a supplier-independent base-model will be presented. It allows easily building, adjusting and running simulation models of path-based systems without deeper programming knowledge. A use case of Infineon's Dresden fab revealed simulation results in the same range of accuracy as got from validated supplier provided models even when a comparatively higher level of abstraction was chosen for the base model. In some deeper research the connection of the base model's abstraction level and the result accuracy is reflected.

## 1 Einleitung

Die effiziente Planung und Steuerung automatisierter innerbetrieblicher Transportsysteme (engl. AMHS – Automated Material Handling System) verlangt Methoden, um deren Verhalten vorauszusagen. Insbesondere auf dem Gebiet der kapitalintensivsten Industrie überhaupt – der Halbleiterindustrie (Peng et al. 2012) – ist eine Einschätzung des Systemverhaltens von Bedeutung. Hier führen bereits kleinste Ineffizienzen zu erheblichen Opportunitätskosten (Terwiesch et al. 1999). Aufgrund der im Allgemeinen vorherrschenden Systemkomplexität von Halbleiterfabriken hat sich daher für deren Analyse die ereignisdiskrete Simulation etabliert.

Bei Infineon Dresden wird das AMHS der weltweit ersten 300-mm-Pilotlinie für Leistungshalbleiter simulativ untersucht. Bei der Betrachtung z. B. verschiedener Layouts, Ramp-Levels oder Steuerstrategien zeigen sich die Defizite der üblicherweise zum Einsatz kommenden Herstellermodelle. Diese bilden das System häufig sehr detailliert ab, sind in der Regel aber sehr schlecht konfigurierbar und langsam.

Gleichzeitig bestehen nur unbefriedigende Ansätze zur einfachen Erstellung, Parametrierung und performanten Simulation von Modellen von Waferfabriken (kurz, „Fabs“). Die Motivation der dem Beitrag zugrunde liegenden Arbeit bestand somit in der Entwicklung eines generischen Simulationsmodells, welches ähnlich genaue Ergebnisse wie Herstellermodelle liefert, aber die genannten Nachteile aufwiegt. Der Ansatz erfolgt über die Wahl und Variation des Abstraktionsgrades.

Der Beitrag ist wie folgt strukturiert: Im Kapitel 2 erfolgt zunächst eine Sichtung relevanter Literatur. Das betrifft die Beleuchtung bisheriger Modellierungsansätze sowie Diskussionen um die Wahl der „richtigen“ Modellkomplexität. Im Kapitel 3 wird ein neues Standardmodell vorgestellt. Kapitel 4 und 5 erläutern kurz Hintergründe zum Anwendungsfall und stellen das Standardmodell mit verschiedenen Abstraktionsgraden einem validierten Modell gegenüber. Der Beitrag schließt mit einer Zusammenfassung und dem Ausblick in Kapitel 6.

# 2 Stand der Technik

## 2.1 Simulation in der Halbleiterindustrie

Mit Vergrößerung des Waferdurchmessers von 200 auf 300 mm und der damit einhergehenden Einführung neuer Transport- Automatisierungskonzepte hat die Simulation im Halbleiterumfeld nochmals an Bedeutung gewonnen (Brain und Lin 2007; Han et al. 2005; Agrawal und Heragu 2006). Umfangreiche Literaturübersichten über die Anwendung von AMHS-Simulationen können in Tyan et al. (2004) und Montoya-Torres (2006) nachvollzogen werden. Es wird unter anderem beschrieben, wie ein Übergang von statischen Scheduling- und Dispatching-Problemstellungen hinzu Modellen hochdynamischer, automatischer Fabriken stattfindet. Jüngere Untersuchungen legen den Fokus z. B. auf die Bestimmung der optimalen Fahrzeuganzahl, die Reduzierung von Maschinen-Stillstandszeiten oder die Lastbalancierung (siehe z. B. Kiba et al. 2009; Hammel et al. 2012; Lin und Huang 2014).

Nahezu alle Waferfabriken ähneln sich in ihrer Struktur und ihren grundlegenden Prozessen, womit die Anwendung generischer Simulationsmodelle naheliegt. So können Entwicklungszeit eingespart und die Ergebnisgenauigkeit erhöht werden (Mackulak et al. 1998). In diesem Zusammenhang berichten Tyan et al. (2004) von einem integrierten Modell zur Simulation automatisierter Fabs. Es erlaubt die Abbildung des Prozesssystems und segregierter Transportsysteme. Ähnliche Ansätze und Untersuchungen zur Generizität von Simulationsmodellen sind Pillai et al. (2004), Nazzal und Bodner (2003), Mackulak et al. (1998), Schulz und Stanley (2000) oder Kim et al. (2009) zu entnehmen. Ohne die Mängel im Detail zu erläutern, bleibt festzustellen, dass alle Ansätze der zitierten Quellen Defizite aufweisen. Dies betrifft vor allem die Reduzierung des Modells auf einen Einzelfall, Einschränkungen beim Design des Layouts, die Notwendigkeit des Einsatzes proprietärer Software oder das separierte, nicht integrative Abbilden von Prozess- und Transportsystem.

## 2.2 Grad der Abstraktion von Simulationsmodellen

Die Diskussion und die Wahl des „richtigen“ Abstraktionsgrades ist für vorliegende Arbeit vor allem aus zweierlei Gründen wichtig: Zum einen wird die schlechte Lauf-

zeit bisheriger Modelle kritisiert und auf deren zu hohen Detaillierungsgrad zurückgeführt. Zum anderen verlangt der Ansatz, ein generisches, universell einsetzbares Standardmodell zu entwickeln, die Verständigung auf den „kleinsten, gemeinsamen Nenner“ – was zwangsläufig zu Simplifizierung und Abstraktion führt.

Erfahrungsgemäß werden detaillierte Simulationsmodelle sehr häufig von Auftraggebern nachgefragt. Der Forderung liegt die Annahme zu Grunde, dass mit niedrigeren Abstraktionsgraden eine höhere Ergebnisgenauigkeit erreicht wird (siehe dazu auch Kiba et al. 2009, Pillai et al. 2004 oder Astrup et al. 2008). Untersuchungen zeigen aber, dass ab einem bestimmten Punkt der Zuwachs an Genauigkeit abnimmt, bzw. sogar negativ wird (vgl. Abb. 1 – in Anlehnung an Benjamin et al. 1998; Chwif et al. 2000; Robinson 2007; Astrup et al. 2008).

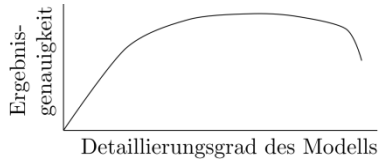



*Abbildung 1: Abhängigkeit des Abstraktionsgrades und der Ergebnisgenauigkeit*

Auch wenn aus wissenschaftlicher Sicht Konsens besteht, Modelle „so einfach wie möglich und so genau wie nötig“ zu erstellen, ist ein überproportionaler Anstieg der Modellkomplexität zu beobachten (Thesen und Travis 1989; Astrup et al. 2008) und wird meist mit Verweis auf leistungsfähige Computer gerechtfertigt (Chwif et al. 2000). Zusammenfassend gibt es keine allgemeingültige Definition, ab wann ein Modell zu detailliert ist. In der Literatur werden aber Methoden beschrieben, welche die Wahl eines geeigneten Abstraktionsgrades unterstützen bzw. diesen klassifizieren (z. B. Thesen und Travis 1989; Robinson 2007; Jimenez et al. 2008). Für den vorliegenden Fall soll daher vor allem die Methode des „Trial and Error“, wie von Benjamin et al. (1998) vorgeschlagen, zum Einsatz kommen.

## 3 Standardmodell

Im vorliegen Fall wird auf den ereignisdiskreten Simulator AutoMod zurückgegriffen. Aufgrund Compiler-basierter Modelle und deren Laufzeitvorteil hat sich AutoMod gegenüber anderen Tools als De-facto-Standard in der Halbleiterindustrie und im Wissenschaftsumfeld etabliert (siehe Quellen im Abschnitt 2).

Den im vorherigen Kapitel aufgeworfenen Mängeln bisheriger Modelle soll begegnet werden. Das nachfolgend vorgestellte Standardmodell erfüllt daher folgende Anforderungen: (1) Neben AutoMod wird keine weitere, proprietäre Software eingesetzt. Das Modell ist (2) nicht auf einen Anbieter bzw. Anwendungsfall beschränkt sowie Layout- und Daten-getrieben. Damit ist die Modellierung, Konfiguration und Simulation eines Systems ohne tiefere Programmierkenntnisse möglich und verlangt lediglich die Parametrierung vordefinierter Dateien und das Zeichnen des Layouts im Grafikeditor von AutoMod. Daraus folgt (3) die Generizität des Modells und die Möglichkeit, verschiedene Konfigurationen ohne erneutes Kompilieren zu simulieren. Das Modell ist (4) hochflexibel und erlaubt die Abbildung segregierter und geschlossener (sog. „unified“) Pfadlayouts (vgl. Abb. 2), verschiedene Pufferkonzepte

(z. B. sog. Stocker oder Side-Track-Buffer) und die Batch-weise Prozessierung von Losen.

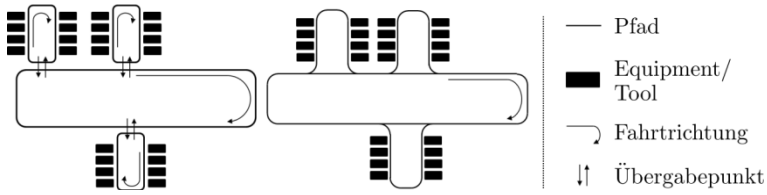



*Abbildung 2: Verschiedene Transportsystem-Paradigmen: segregiert mit drei Bays (links), geschlossenes System (unified system, rechts)*

Eine wesentliche Eigenschaft des Standardmodells ist (5) dessen Integration von Prozess- und Transportsystem. Somit werden ein prozessgetriebenes Verhalten des AMHS erreicht und entsprechend Interferenzen zwischen den Systemen abgebildet.

Wie bereits angedeutet ist eine exakte Spezifizierung des Abstraktionsgrades nicht trivial. Nach Jimenez et al. (2008) lässt sich das AMHS des hier vorgestellten Standardmodells ohne grundlegende Modifikation (vgl. Abschnitt 4) in *Klasse E*, das Prozesssystem etwa in *Klasse 5* einordnen. Übersetzt werden also nicht notwendiger Weise alle Komponenten abgebildet und Abläufe nicht in höchster Genauigkeit modelliert. Für das gesamte Modell ergibt sich eine Einstufung als *Fall d* im *Quadrant 4*, was eine Verflechtung der Tools und des AMHS ausdrückt.

Den prinzipiellen Aufbau des Standardmodells verdeutlicht Abbildung 3. Dieses macht sich die Möglichkeit zu Nutze, in AutoMod Systeme hierarchisch zu ordnen.

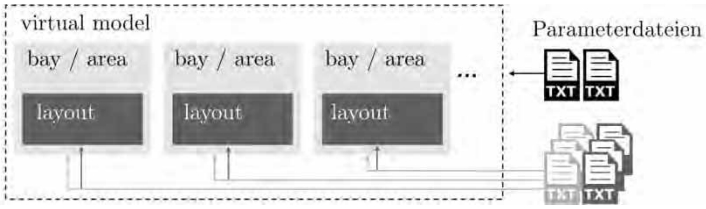



*Abbildung 3: Hierarchische Strukturierung und Parametrierung der Systeme*

Auf höchster Ebene im „virtual model“ sind alle Basisfunktionalitäten und die Grundlogik des Modells zusammengefasst. Dies beinhaltet bspw. den generellen Prozessablauf, die Lastgenerierung, das Lesen von Parameterdateien oder Regeln zum Dispatching, Routing und der Fahrzeugdisposition. Subsysteme (hellgrau dargestellt in Abb. 3) stellen physische oder logische Teile einer Fabe dar, also bspw. eine Lithographie-Bay (Bay: Fachbegriff für abgetrennten Bereich in einer Fab), ein Inter-Bay-Transportsystem oder Stocker. Subsysteme erben die genannte Funktionalität vom Elternmodel, sind aber jeweils in der Lage diese zu ersetzen bzw. zu erweitern, um spezifische Gegebenheiten abzubilden. Üblicherweise wird jedes Subsystem durch ein weiteres System (in Abb. 3 dunkelgrau dargestellt) visualisiert. Dies können bspw. Schienen, Equipment oder Puffer sein.

Nach Abbildung 3 werden die Systeme von Parameterdateien gespeist. So wird die beabsichtige Generizität und Flexibilität des Standardmodells sichergestellt und die Kompilierzeit des Programms minimiert. Dateien auf oberster Ebene enthalten z. B. Informationen zum Produktionsplan, Produktionsablaufplan samt Prozesszeiten der einzelnen Produkte oder über die Anzahl und Kapazität der Fahrzeuge in den Bereichen. Auf Submodellebene umfassen die Dateien u. a. Angaben zur Kapazität der platzierten Tools, der Größe der Puffer, zu möglichen Parkpositionen der Fahrzeuge sowie deren technischen Parametern (Geschwindigkeiten, Beschleunigungen...).

Um auf Basis des Standardmodells eine konkrete Fab zu modellieren müssen lediglich folgende Schritte durchgeführt werden: Zunächst sind (1) alle benötigten Subsysteme zu definieren, (2) deren Layouts zu zeichnen und (3) Übergabepunkte zu platzieren. Zur Visualisierung müssen (4) ggf. Tools, deren Loadports (Übergabeeinheit/-ort an den Tools) und Puffer angeordnet werden. Mit (5) der Konfiguration der vorgefertigten Parameterdateien ist das Modell bereits lauffähig. Soll das Modell in seiner Funktionalität erweitert werden, ist (6) der entsprechende Quellcodeteil zu ersetzen. In der Regel ist alle ersetzbare Funktionalität gekapselt, sodass der Änderungsaufwand gering und eine breite Individualisierung möglich ist.

Die Grundlogik des Basismodells ist verhältnismäßig einfach gestaltet – die übliche Definition von AutoMod-internen „work lists“ (Liste mit Orten, an denen Fahrzeuge nach Transportaufträgen suchen) und „park lists“ (Parkpositionen für Fahrzeuge im Ruhemodus) entfällt. Fahrzeuge fahren kürzeste Wege entsprechend des AutoMod-Algorithmus. Blockieren Lehrfahrzeuge die Strecke, werden diese zu den in den Parameterdateien angegebenen Positionen geschickt. Die Zuordnung der Transportaufträge zu Fahrzeugen erfolgt dezentral. Dazu werden permanent sowohl freie Transportkapazitäten als auch –forderungen gelistet und einander zugeordnet. Die Priorisierung beim Dispatching und Scheduling erfolgt nach örtlicher Nähe und dem Prozessfortschritt der Wafer im zu transportierenden FOUP. Noch einmal sei betont, dass auch diese Logik beliebig ersetzt oder erweitert werden kann – der Aufbau des Standardmodells sieht dies ausdrücklich vor.

## 4 Anwendungsfall

Aus Gründen der Geheimhaltung können zum folgenden Anwendungsfall nicht alle Details berichtet werden. Untersucht wird die 300 mm-Fab von Infineon Dresden und im Speziellen die Herstellung eines Produktes mit insgesamt 106 Bearbeitungsschritten an verschiedenen Tools. Diese sind je nach Prozessschritt zum Teil als sog. single-of-a-kind-Tool (Einzelmaschine) vorhanden oder stehen in bis zu 42-facher Redundanz bereit. Eine Prinzipskizze des Layouts ist in Abbildung 4 dargestellt. Insgesamt sind 11 verschiedene Bays mit je 1 bis 30 aktiven Fahrzeugen involviert.

Für das beschriebene System liegt ein vom Transportsystemlieferanten bereitgestelltes Simulationsmodell vor. In Voruntersuchungen konnte dessen Validität bestätigt werden. Das Modell (fortan MT-Modell) soll als Referenz für nachfolgende Untersuchungen dienen.

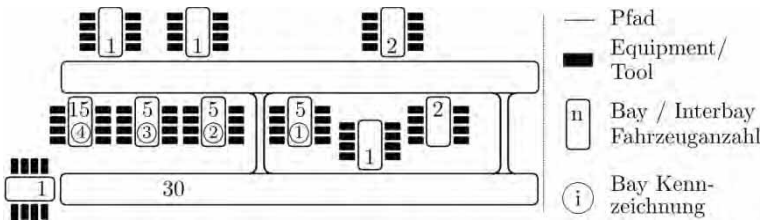



*Abbildung 4: Layout zum Anwendungsfall*

Das MT-Modell weist im Vergleich zum Standardmodell (vgl. Kapitel 3) einen erheblich niedrigeren Abstraktionsgrad auf. Das betrifft vor allem die deutlich komplexeren/realitätsnäheren Routing- und Dispatchingstrategien. Des Weiteren bildet es u. a. den Übergabeprozess am Loadport sehr detailliert ab und implementiert ausgefeilte Stauvermeidungsstrategien. Ein deutlicher Hinweis für den niedrigeren Abstraktionsgrad des MT-Modells ist dessen Quellcode. Dieser ist im Kern etwa dreimal so umfangreich wie der des hier eingeführten Standardmodells.

Das MT-Modell soll mit vier eigens erstellten Modellen verglichen werden. Diese basieren alle auf dem Standardmodell, unterscheiden sich aber im Abstraktionsgrad.

Ein erstes Modell hat exakt das Standardmodell, so wie in Kapitel 3 beschrieben, als Grundlage (fortan ID1). In einem zweiten Modell wird das Inter-Bay-System nicht physisch abgebildet, sondern dessen Transporte durch Zeitverzug nachgestellt (ID2). Ein drittes Abbild (ID3) stellt ein nicht-integriertes Modell dar, d. h. Prozess- und Transportsystem sind nicht verzahnt und die Aufträge werden anhand von Materialflussmatrizen generiert. Das vierte Modell (ID4) kombiniert die Vereinfachungen des ID2- und ID3-Modells. Erfahrungsgemäß sind das Abstrahieren des Inter-Bay-System und der Einsatz nicht-integrierter Modelle in der Praxis häufig diskutierte Modellierungsansätze – woraus sich die folgenden Untersuchungen begründen.

Im weiteren Verlauf sollen die Fragen geklärt werden, ob und welches der IDx-Modelle das Systemverhalten ähnlich genau abbildet wie das MT-Modell. Dazu wurde eine umfangreiche Simulationsstudie durchgeführt. Zur Anwendung kam ein Rechner mit 64 Bit Windows 7, i5 2500k Prozessor, 8 GB Ram und AutoMod 12.4.0.23.

Die Modelle werden identisch parametriert. Die Simulationszeit ist für alle Läufe konstant und „hinreichend lang“. Die Einschwingphase beträgt ein Viertel der Simulationszeit. Untersucht wurden unterschiedliche Lastszenarien.

Die Auswertungen in den nachfolgenden Kapiteln konzentrieren sich auf die Bays ①, ②, ③ und ④ (vgl. Abb. 4), weil für das MT-Modell die Analyse anderer Bays nicht ohne tiefgreifende Modifikation möglich ist.

## 5 Vergleich MT-Modell zu IDx-Modellen

Um Verwirrung zu vermeiden, werden nur die Ergebnisse eines Lastszenarios angegeben. Grundlegend andere Erkenntnisse ergeben sich zusammengefasst aus dessen Variation nicht. In einem ersten Schritt soll die Leistung des Transportsystems der Modelle verglichen werden. Dazu sind in Tabelle 1 für die einzelnen Bays Kennzahlen bzgl. fertiggestellter Transportaufträge aufgeführt.

*Tabelle 1: Anzahl fertiggestellter Transportaufträge – Grundlage sind je 10 Simulationsläufe; Intervall listet kleinsten und größten gemessenen Wert; Mean: Mittelwert; COV: Variationskoeffizient; Abweichung auf Basis des MT-Modells*

| | Bay ① | Bay ② | Bay ③ | Bay ④ | Bay ① bis ④ |
|---|---|---|---|---|---|
| MT (Intervall) | [957, 995] | [1368, 1565] | [1129, 1230] | [6462, 6945] | [9977, 10651] |
| MT (mean \| COV) | 981 \| 0,001 | 1499 \| 0,030 | 1186 \| 0,030 | 6867 \| 0,020 | 10534 \| 0,018 |
| MT (Abweichung) | – | – | – | – | – |
| ID1 (Intervall) | [1074, 1151] | [1344, 1432] | [1235, 1258] | [6997, 7132] | [10718, 10895] |
| ID1 (mean \| COV) | 1115 \| 0,020 | 1377 \| 0,020 | 1245 \| 0,007 | 7073 \| 0,005 | 10810 \| 0,005 |
| ID1 (Abweichung) | 14 % | 8 % | 5 % | 3 % | 3 % |
| ID2 (Intervall) | [1146, 1239] | [1226, 1287] | [1280, 1306] | [7285, 7345] | [10975, 11082] |
| ID2 (mean \| COV) | 1184 \| 0,019 | 1261 \| 0,017 | 1296 \| 0,007 | 7307 \| 0,003 | 11049 \| 0,003 |
| ID2 (Abweichung) | 21 % | 16 % | 9 % | 6 % | 5 % |
| ID3 (Intervall) | [849, 1102] | [1260, 1472] | [488, 570] | [4426, 5095] | [7027, 8191] |
| ID3 (mean \| COV) | 1014 \| 0,067 | 1379 \| 0,051 | 527 \| 0,055 | 4862 \| 0,048 | 7782 \| 0,050 |
| ID3 (Abweichung) | 3 % | 8 % | 56 % | 29 % | 26 % |
| ID4 (Intervall) | [944, 1113] | [1269, 1462] | [516, 592] | [4577, 5135] | [7344, 8243] |
| ID4 (mean \| COV) | 1019 \| 0,052 | 1389 \| 0,038 | 542 \| 0,035 | 4920 \| 0,035 | 7870 \| 0,036 |
| ID4 (Abweichung) | 4 % | 7 % | 54 % | 28 % | 25 % |

Zunächst sei festgestellt, dass sich der Variationskoeffizient aller Mittelwerte als kleiner 0,07 darstellt. Dies deutet jeweils auf robuste, wenig schwankende Simulationsergebnisse hin, womit sich eine Auswertung auf Basis des Mittelwertes anbietet. Dabei zeigt sich, dass nur die Modelle ID1 und ID2 für das in Kapitel 4 beschriebene System hinreichend genaue Simulationsergebnisse liefern. Mit dem MT-Modell als Referenz, ist das ID1-Modell am besten geeignet das Transportsystemverhalten vorherzusagen. Umgekehrt eignen sich ID3 und ID4 nicht. Zwar bilden diese die Bays ① und ② mit Abweichungen von 3 bis max. 8 % recht genau ab, zeigen aber mit Abweichungen von bis zu 56 % deutliche Ungenauigkeiten für Bay ③ und ④. Genau konträr werden mit den Modellen ID1 und ID2 die größten Ungenauigkeiten für Bay ① und ② beobachtet. Hierzu sei bemerkt, dass die Bays ③ und ④ im Vergleich ein deutlich komplexeres Pfadlayout aufweisen und jeweils bis zu 17 Tools aktiv sind – das Schienenlayout der Bays ① und ② ist dagegen durch einen einfachen Kreis mit Abkürzung festgelegt, bei teils nur 3 aktiven Tools.

Es sei wiederholt, dass ID1 und ID2 integrierte Modelle darstellen (vgl. Kapitel 3, Eigenschaft (5) des Standardmodells). Diese Eigenschaft scheint also wesentlich zur Ergebnisgenauigkeit beizutragen, womit von einer Abstraktion, wie in ID3 und ID4, abzuraten ist. Weniger kritisch stellt sich eine Vereinfachung des Inter-Bay-Systems dar (vergleiche hierzu ID1 zu ID2 und ID3 zu ID4). Insbesondere für die Bays ③ und ④ sowie für die Summe aller Bays erhöht sich im Vergleich zum MT-Modell die Abweichung aller fertiggestellten Transportaufträge um max. 4 Prozentpunkte.

Tabelle 2 lässt einen Vergleich des Prozesssystems bzw. der Tools in den Bays zu. Dazu wurde für drei zufällig gewählte Maschinen deren Durchsatz bzw. Auslastung untersucht. Außerdem sind die mittlere Durchlaufzeit der Lose durch das Gesamtsystem und der Durchsatz (Anzahl der fertiggestellten Lose) der Fab angegeben.

*Tabelle 2: mittlerer FOUP-Durchsatz – links gestrichelter Linie: zufällig gewählte Tools (Durchsatz | Toolauslastung | Abweichung zu MT); rechts gestrichelter Linie: Durchlaufzeit (DLZ) und Durchsatz des Gesamtsystems (normiert auf MT-Modell)*

| | | | | Gesamtsystem | |
|---|---|---|---|---|---|
| | Tool A | Tool B | Tool C | DLZ | Durchsatz |
| MT | 342 \| 0,63 \| – | 501 \| 0,28 \| – | 801 \| 0,45 \| – | 1,00 | 1,00 |
| ID1 | 359 \| 0,66 \| 5 % | 525 \| 0,29 \| 5 % | 722 \| 0,40 \| 10 % | 0,99 | 1,05 |
| ID2 | 419 \| 0,77 \| 23 % | 505 \| 0,28 \| 1 % | 802 \| 0,45 \| < 1 % | 0,98 | 1,15 |
| ID3 | 25 \| 0,09 \| 93 % | 507 \| 0,28 \| 1 % | 425 \| 0,24 \| 47 % | 0,97 | 0,69 |
| ID4 | 36 \| 0,13 \| 89 % | 471 \| 0,26 \| 6 % | 501 \| 0,28 \| 37 % | 0,98 | 0,61 |

Ähnlich der Auswertung von Tabelle 1 wird erneut deutlich, dass vor allem mit dem ID1-Modell realitätsnahe Ergebnisse erzielt werden können – im Vergleich zu MT beträgt bei den zufällig gewählten Tools die maximale Abweichung des Durchsatzes 10 % (Tool C). Gesamtsystemisch zeigen sich für ID1 ein leicht erhöhter Durchsatz (5 %) und eine minimale Abweichung in der Durchlaufzeit. Die Modelle ID2 bis ID4 zeigen dagegen alle bei je einer Kennzahl eine Abweichung von mind. 15 %. Besonders zu erwähnen ist, dass die nicht-integrierten Modelle ID3 und ID4 die DLZ recht genau wiedergeben, den Gesamtdurchsatz aber deutlich unterschätzen.

Wie einleitend beschrieben, soll mit dem hier vorgestellten Standardmodell u. a. eine Reduzierung der Simulationszeit im Vergleich zu (detaillierten) Herstellermodellen erreicht werden. Hierzu sind in Tabelle 3 die Ergebnisse einer Vergleichsstudie über die mittlere Laufzeit eines Simulationsexperiments gegeben.

*Tabelle 4: Laufzeit eines Simulationsexperiments – absolut | Verhältnis zu MT*

| | Lastszenario („wafer starts per week“) | | | |
|---|---|---|---|---|
| | niedrig | mittel | hoch | sehr hoch |
| MT | 1896 s \| – | 4353 s \| – | 6138 s \| – | 6487 s \| – |
| ID1 | 332 s \| 5,7 | 577 s \| 7,5 | 648 s \| 9,5 | 729 s \| 8,9 |
| ID2 | 257 s \| 7,4 | 471 s \| 9,2 | 591 s \| 10,4 | 683 s \| 9,5 |
| ID3 | 235 s \| 8,1 | 562 s \| 7,7 | 622 s \| 9,9 | 1464 s \| 4,4 |
| ID4 | 236 s \| 8,0 | 574 s \| 7,6 | 486 s \| 12,6 | 1445 s \| 4,5 |

Zusammengefasst zeigt sich: Das Standardmodell weist in jeder Konfiguration kürzere Laufzeiten als das MT-Modell auf. Für einzelne Szenarien werden Simulations-

ergebnisse in weniger als einem Zehntel der Zeit generiert. Deutlich wird ebenfalls, dass für das Standardmodell ein höherer Abstraktionsgrad nicht immer performantere Programme impliziert. Besonders die Verwendung von nicht-integrieren Modellen führt im vorliegenden Fall nicht zu den erwarteten Geschwindigkeitsvorteilen, wie z. B. in Jimenez et al. (2008) berichtet. Bemerkenswert ist außerdem, dass für sehr hohe Last bei den Modellen ID3 und ID4 ein deutlicher Anstieg der Laufzeit zu beobachten ist. Die Phänomene lassen sich erklären: Durch die Trennung von Transportsystem und Tools gibt es keinen Abgleich vorhandener Kapazitäten. Folglich werden an das Transportsystem Aufträge unreflektiert herangetragen. Dazu kommt eine Eigenheit von AutoMod, bei der hohe Aktivitäten in Fahrzeugsystemen zu langsamen Laufzeiten führen. Umgekehrt verhält es sich für ID1 und ID2. Ist keine Transportkapazität vorhanden, verharren die Aufträge an vorgesehenen Plätzen, warten auf ihren Abruf und verzögern die Laufzeit nicht.

Resümierend ist nur ID1 in der Lage, ähnlich genaue Ergebnisse wie das MT-Modell zu generieren. Das ID1-Modell entspricht dem Standardmodell ohne weitere Anpassungen. Der geringe Laufzeitvorteil von ID2 wiegt dessen Ergebnisungenauigkeiten nicht auf. Von einer Abstraktion der Integrität (ID3 und ID4) ist gänzlich abzuraten – zumal sich hieraus im vorliegenden Fall keine Laufzeitvorteile ergeben.

## 6 Zusammenfassung und Ausblick

Vorgestellt wurde ein Ansatz zur einfachen Erstellung, Parametrierung und Simulation von Modellen von Waferfabriken. Das entwickelte Standardmodell ermöglicht Vorhersagen zum Systemverhalten, deren Genauigkeit denen von Herstellermodellen entsprechen. Gleichzeitig weist es um bis zu Faktor 10 schnellere Laufzeiten auf.

Der Aufbau und die Eigenschaften des Standardmodells garantieren dessen Flexibilität und Generizität – es ist herstellerunabhängig und kann ohne Programmierkenntnisse genutzt werden. Die gesamte 300 mm-Fab von Infineon Dresden konnte in lediglich ca. 10 Arbeitsstunden modelliert werden („IDx-Modelle“).

Ferner bestätigt sich mit den Untersuchungen die aufgeworfene Diskussion um die Annahme, dass mit komplexen/niedrig abstrahierten Modellen die Ergebnissgenauigkeit nicht beliebig gesteigert werden kann, dies aber zu Laufzeitverlusten führt. Deutlich wird dies z. B. an Stauvermeidungsstrategien im MT-Modell, welche stets aktiv sind, obwohl das Layout in vielen Bays eine Vermeidung gar nicht ermöglicht.

Zukünftige Arbeiten werden sich auf das Testen des Standardmodells in weiteren Anwendungsfällen und die Variation des Abstraktionsgrades konzentrieren.

## Förderhinweis

The work has been performed in the project EPPL, co-funded by grants from Austria, Germany, The Netherlands, Italy, France, Portugal and the ENIAC JU.

## Literatur

Agrawal, G.K.; Heragu, S.S.: A survey of automated material handling systems in 300-mm SemiconductorFabs. In: Transactions on Semiconductor Manufacturing 19 (2006) 1, S. 112–120.

Astrup, R. et al.: Finding the appropriate level of complexity for a simulation model: An example with a forest growth model. Forest Ecology and Management 256 (2008) 10, S. 1659–1665.

Benjamin, P. et al.: Simulation modeling at multiple levels of abstraction. In: Proceedings of the Winter Simulation Conference 1998, S. 391-398.

Brain, M.D.; Lin, P.: Impact of speed on the performance of vehicle-free transport. In: International Symposium on Semiconductor Manufacturing 2007, S. 1-3.

Chwif, L. et al.: On simulation model complexity. In: Proceedings of the Winter Simulation Conference 2000, S. 449–455.

Hammel, C. et al.: Network optimization prior to dynamic simulation of AMHS. In: Proceedings of the Winter Simulation Conference 2012, S. 172:1-172:11.

Han, Y. et al.: Full Fabrication Simulation of 300mm Wafer Focused on AMHS. In: Baik D.-K. (Hrsg.): Systems Modeling and Simulation: Theory and Applications. Berlin u a.: Springer 2005, S. 514–520.

Jimenez, J.A. et al.: Levels of capacity and material handling system modeling for factory integration decision making in semiconductor wafer fabs. In: Transactions on Semiconductor Manufacturing 21 (2008) 4, S. 600–613.

Kiba, J.-E. et al.: Simulation of a full 300MM semiconductor manufacturing plant with material handling constraints. In: Proceedings of the Winter Simulation Conference 2009, S. 1601–1609.

Kim, B.-I. et al.: A layout- and data-driven generic simulation model for semicon. fabs. Transactions on Semiconductor Manufacturing 22 (2009), S. 225–231.

Lin, J.T.; Huang, C.-J.: A simulation-based optimization approach for a semiconductor photobay with automated material handling system. Simulation Modelling Practice and Theory 46 (2014), S. 76–100.

Mackulak, G.T. et al.: Effective simulation model reuse. In: Proceedings of the Winter Simulation Conference 1998, S. 979–984.

Montoya-Torres, J.R.: A literature survey on the design approaches and operational issues of automated wafer-transport systems for wafer fabs. Production Planning & Control 17 (2006) 7, S. 648–663.

Nazzal, D.; Bodner, D.A.: A simulation-based design framework for automated material handling systems in 300 mm fabrication facilities. In: Proceedings of the Winter Simulation Conference 2003, S. 1351–1359.

Peng, C. et al.: Capacity planning in the semiconductor industry. Manufacturing & Service Operations Management 14 (2012) 2, S. 170–185.

Pillai, D. et al.: 300mm full-factory simulations for 90- and 65-nm IC manufacturing. Transactions on Semiconductor Manufacturing 17 (2004) 3, S. 292–298.

Robinson, S.: Conceptual modelling for simulation Part I. Journal of the Operational Research Society 59 (2007) 3, S. 278–290.

Schulz, M.; Stanley, T. D.: Simulation based decision support for future 300mm automated material handling. In: Proceedings of the Winter Simulation Conference 2000, S. 1518–1522.

Terwiesch, C. et al.: An exploratory study of international product transfer and production ramp-up in the data storage industry. San Diego, UC 1999.

Thesen, A.; Travis, L. E.: Simulation for decision making: an introduction. In: Proceedings of the Winter Simulation Conference 1989, S. 9–18.

Tyan, J. C. et al.: Multiple response optimization in a fully automated FAB. Computers & Industrial Engineering 46 (2004) 1, S. 121–139.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulation-based Reduction of Traffic Jam Effects in a Distributed Automated Material Handling System

## *Simulationsbasierte Reduzierung der Störfallfolgen in einem verteilten Automated Material Handling System*

Robert Schmaler, Christian Hammel, Thorsten Schmidt, Technische Universität Dresden, Dresden (Germany), robert.schmaler@tu-dresden.de, christian.hammel@tu-dresden.de, thorsten.schmidt@tu-dresden.de

Matthias Schöps, Globalfoundries Module One LLC & Co. KG, Dresden (Germany), matthias.schoeps@globalfoundries.com

Abstract: Striving for automation in the semiconductor industry makes automated material handling systems (AMHS) a must in larger fabs. The demand for increasing throughput yields more complex requirements for the transportation system. Especially for transport systems which span several areas connected only by bridges, a system failure at the transportation system can have significant negative effects onto the manufacturing line, too. A new system logic has been developed, which keeps a minimum amount of transportation capacity per area to fulfil transport requirements during failure scenarios. This logic was developed with the help of simulation studies and is based on consideration of transport related restrictions inside the Manufacturing Execution System (MES). Since testing with the real system is not an option simulation is also a suitable method to get first impressions about functionality and performance of new system solutions for AMHS in semiconductor manufacturing facilities.

## 1 Problem Description

Automation can be found in every area of semiconductor industry. Especially larger semiconductor manufacturing facilities (fabs) cannot fulfil complex transport requirements without using automated material handling systems (AMHS). Mostly Overhead Hoist Transport (OHT) systems are used to save expensive cleanroom space. Modern versions of these systems are not segregated into single bays anymore and allow the travel of vehicles along the whole track network. As long as the system can cover e.g. differences in ceiling height by ramps the Front Opening Unified Pods (FOUPs) only have to be handled at source and destination ports. Transportation in between can be done without any further transitions among dif-

ferent systems. Therefore such unified systems can span several areas and even buildings connected only by bridges. For a part of GLOBALFOUNDRIES' Dresden campus this is depicted in figure 1.

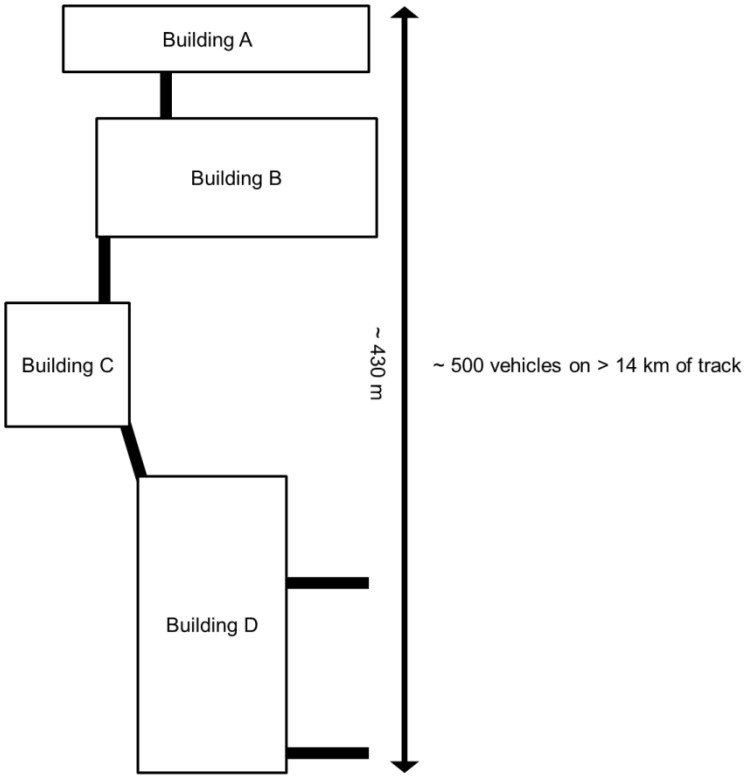



*Figure 1: Schematic of buildings served by connected OHT as AMHS*

The availability of a complex transport network brings specific challenges along, too. The following example gives an impression about how a single traffic jam can affect the whole transport system. In case of a significant traffic jam due to a system failure in a particular production area, many vehicles will pile up quickly because they are supposed to fulfil transport requests in this area. This can lead to other areas running short of vehicles because the affected area still keeps requesting vehicles. As a result, areas even not directly affected do not have enough vehicles to fulfil transport requests any longer. In that case the production of the whole fab can be endangered. To prevent this behaviour the Vehicle Area Balancer (VAB) which is presented in this paper has been developed. This logic ensures an appropriate distribution of the vehicles which represent the restricted resource of the transportation system. The VAB takes the overall situation of the transport system into account and keeps a minimum amount of vehicles in each specified area. This solution goes beyond any known watermark strategies as e.g. can be found in Jimenez et al. (2010). The goal is an at least limited functionality of areas which are not failure affected directly. Simulation studies were used to verify the different

necessary vehicle amounts. Since there is a need to check incoming transport requests for feasibility also further boundary conditions of fab automation have to be considered besides OHT regards: dependencies as well as interdependencies with the dispatching system in terms of transport request generation (Nishi and Tanaka 2012). Through integration of a replication of the developed solution into an existing simulation model it was possible to test different failure scenarios in advance and investigate functionality and benefit of the VAB.

The remainder of this paper is organised as follows: In section 2 analyses of layout, transport flow and vehicle quantities are presented, followed by section 3 where a detailed description of the VAB is provided. In section 4 details of the simulation model and the simulation experiments are presented accompanied by results. The paper ends with a summary in section 5.

# 2 Pre-Consideration

## 2.1 Layout and Transport Analysis

The bigger the connected transport system, the more tools and storages are available as possible destinations for a lot which has just finished a process step and has to be picked up to clear the load port of the tool. But if the OHT connects different buildings within one AMHS these connections can lead to increased traffic jam risks, too. In some cases it only needs one vehicle in error mode causing a traffic jam just in front of an entrance or exit of a building and the whole exchange of carriers and vehicles between the buildings will be disrupted. These transport and vehicle exchange dynamics are crucial for a semiconductor facility with different buildings because the throughput goals can only be achieved by interaction of different buildings and production areas. To get a better understanding of the correlation between MES-triggered inter-building transports on the one hand and bottlenecks of the transport system on the other hand layout and transport analyses have been carried out.

The main focus during transport pattern analysis was put on the relation between intra- and inter-building transports. Results have shown that transports changing a building (inter-building) exceeded those with source and destination within the same building (intra-building). For the transport system this ends in highly utilised connections (thick, black coloured) links between the different buildings as shown in the track utilisation chart in figure 2.

The next step was to calculate the betweenness centrality which gives information about how often a link is part of a shortest path (Anthonisse, 1971). When analysing links within the transport system by calculating the betweenness centrality with the Floyd-Warshall-Algorithm (Floyd 1962 and Warshall 1962) the resulting picture was pretty much the same: In a comparison of all connections of possible sources and destinations in the transport system, the links connecting different buildings are used most. This is depicted in figure 2 where again black coloured links show high utilised areas and grey coloured links show low utilised ones.

*Figure 2: Comparison of Betweenness Centrality (left) and Track Utilization Chart (right) showing utilisation differences from low (grey, small) to high (black, thick) utilisation*

In summary system failures affecting the dynamic exchange of carriers and vehicles between the different buildings will have significant impact onto the whole transport system. Area connections are the most important parts of the OHT from both points of view: manufacturing side (transport job generation) and transport system side (shortest paths). Nevertheless there will be intra-building transports which can be carried out even during failure scenarios. For this reason transport capacity has to be kept within the buildings to be able to clear tool load ports and fulfil transport requirements.

## 2.2 Vehicle Quantities

Goal of the VAB was to prevent certain areas from running out of vehicles to ensure a minimum functionality of areas not directly affected during traffic jams. A measure for the minimum amount of necessary vehicles per area to fulfil limited transport requirements had to be found. To find a measure for the minimum necessary amount of vehicles different aspects as transport patterns, delivery times and vehicle utilisation have been included (Mackulak and Savory 2001). Based on this, different simulation studies have been executed to find most important parameters to be included in this measure. This leads to the following formula to calculate minimum amounts of vehicles VAB ($B_{i,j}$) for the different transport relations as input parameters for the VAB in general:

$$VAB(B_{i,j}) = \frac{tph(B_{i,j}) \cdot t(B_{i,j})}{U} \quad (1)$$

Let:

$i$ Source area index i = 1, 2, …, I (I as amount of different source areas)

$j$ Destination area index j = 1, 2, …, J (J as amount of different destination areas)

$B_{i,j}$ Transport relation from source area i to destination area j

$tph(B_{i,j})$ Average transports per hour for transport relation from source area i to destination area j

$t(B_{i,j})$ Average delivery time for transport relation from source area i to destination area j

$U$ Vehicle Utilization as rate of vehicle usage time with $0 \leq U \leq 1$

For i=j the critical value for each defined area can be calculated. Again simulation was used to validate these values regarding their target achievement of fulfilling transport requirements for each area.

In the following the resulting values for the different transport relations are used as the critical values within VAB as explained in the next section.

## 3 Vehicle Area Balancer

Although minimum functionality of an area is ensured by defined vehicle amounts the realisation cannot be done without an according adjustment of dispatching rules and the transport jobs generated by the manufacturing execution system (MES). Even during failure scenarios transport requests for local to-tool transports have to be generated while inter-area transports through effected areas have to be rejected. If the transport controller recognises a critical value for the vehicle quantity within a certain area communication of this exception has to reach the MES. MES then has to identify the transport related situation and to draw the right conclusion. In the following, non-affected areas will be considered as local. That means they don't have any available transport relations which are using jammed areas. An overview about the ways of communication between the different parts of the system architecture can be seen in figure 3.

Bay Control Units (BCUs) calculate amounts of vehicles within each BCU in a periodic process. This information about BCU borders and vehicle quantities leads to the possibility to define specified areas and their critical values. Critical values have to be monitored by a sub program of the Overhead Hoist Transport Controller (OHTC). If one of the critical values is reached there has to be a defined warning message arriving at the Material Control System (MCS). This message finally has to be transmitted to the MES to be recognised by the Real Time Dispatcher (RTD). This trigger for the critical area has to be communicated and interpreted in the right way to identify the transport system state and to adjust the transports in the appropriate way.

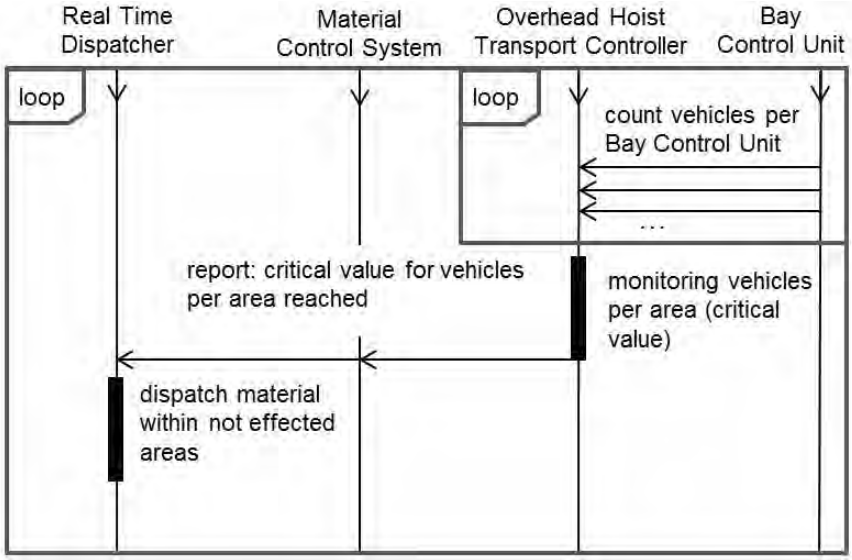



*Figure 3: VAB communication sequence diagram*

The innovation of the VAB is the consideration of transport related issues within the RTD decision level. RTD describes a software solution within the MES environment which dispatches available production material by considering the current fab situation (e.g. including tool utilisation and cycle time). Based on this, scheduling lists of production material are generated and finally certain material is assigned to a certain tool. To ensure an at least limited functionality of non-affected areas during traffic jams, consideration of this decision level is essential.

The described new system logic intervenes significantly into transport job generation and subsequently the control of the whole fab. Transport jobs are rejected or adjusted by taking into account not only parameters of the manufacturing line but also transport related issues. Rejecting transport requests due to AMHS problems is not state of the art. Since it could help to increase transport capacity and reduce the time until the traffic jam is cleared, negative consequences of system failures can be reduced. However, testing the new logic by implementing it into the real system is impossible for a bunch of reasons. Simulation is a suitable method to get a first idea about functionality and benefit of the VAB, if it is worth initiating further development steps and for optimisation of parameters.

# 4 Simulation

## 4.1 Simulation Model

The simulation model was developed in AutoMod. It combines four different buildings as depicted is figure 1. The simulation model uses the same layout as the real fab and includes the following system attributes:

- 825 process tools, direct accessible by AMHS
- about 20.000 storage locations in total

- more than 14 km of track
- more than 500 vehicles

The model logic also uses the same vehicle control strategies as the real system. Empty vehicle balancing dispatches empty vehicles according to common water-mark strategies. Secondary the vehicle routing is based on optimised cost settings instead of travel length or time only. Thus, not only shortest paths are taken into account but also system depending parameters (Hammel et al., 2012). This leads to beneficial paths for transporting vehicles in order to avoid bottlenecks whenever possible. Furthermore, the real system vehicle speed characteristics are implemented in the simulation model. Vehicle speed is set depending on whether the vehicle is using or approaching a curve, merge, diverge or a straight piece of track. By using even more details and logic of the real system than the described features the simulation model reaches a mean deviation in performance measures compared to the real system of less than 5%. This measured deviation makes it suitable to gather reliable results and conclusions from this reproduction of the real system.

The above mentioned features are essential for reflecting failure free simulation scenarios. Additionally, there are other features which make simulation the most suitable approach to get a first idea regarding VAB logic functionality. It is possible to stop vehicles within the simulation model at a certain place for a certain amount of time and therefore to reproduce system failures. Besides deadlock escape settings there are different vehicle amount restrictions, too. These are very important to reflect real system traffic jam behaviour which would be the use case for the VAB. For example there are parameters which define maximum vehicle quantities for each BCU or even ceiling load restrictions which have to be considered by the ceiling mounted OHT. Finally, there is the possibility to use and manipulate real fab transportation data as input for the simulation model. The VAB rejects and adjusts transport jobs depending on where the system failure takes place and which area is endangered to lose minimum transport capacity. Therefore the adjustment of the transport data sets is essential to be able to test VAB logic by using the simulation approach.

## 4.2 Simulation Experiment Description

To get a large set of information about VAB functionality and system performance effects different scenarios for simulation runs have been developed.

As already mentioned in section 2.1 the main bottlenecks regarding the transport system have been found along the connections between different buildings. According to these results the top 3 bottlenecks have been chosen for the simulation studies.

The impact of traffic jams for the transport system not only depends on the place but also on the duration. If a vehicle stops due to a failure and it can be released within a few minutes it will not have fab-wide impact. The longer it blocks a route the more vehicles will pile up within the traffic jam. This leads to higher transportation times, waiting times and lower throughput as well. To see the influence of the newly developed VAB logic as a function of the failure duration, three different intervals (described as short, medium, long) have been considered for the simulation scenarios.

Additionally, three different sets of transport data generated by the fab have been used to cover a higher variety of different system states. Therefore, a methodology to implement real fab data into the simulation model had to be developed. Combination of different denotations between real system and simulation model could be achieved by using a self-developed data base approach. Subsequently another step where the data sets can be adjusted according to the VAB logic has been implemented into synchronisation: depending on which bottleneck was blocked in each scenario transport jobs have been adjusted or rejected according to source and destination type.

An example of transport job adjustment can be found in Table 1. It refers to one of the scenarios where the entrance of area A (fig. 1) is blocked and other areas are not affected. After a certain period of time the pre-defined critical value for the minimum vehicle quantity of area A is reached and the VAB logic starts working. Hence all outgoing inter-building transports will be adjusted while ingoing inter-building as well as intra-building transports will not be affected. The way of adjustment depends on the source and destination type. The simplified logic is as follows:

- If source type is tool, the load port has to be cleared and the FOUP will be transported into a storage next to the tool
- If source type is storage, the FOUP stays at the port and the transport job is rejected

*Table 1: Transport Adjustment according to VAB rules (Src = Source, Dst = Destination)*

| SrcType | SrcArea | DstType | DstArea | Adjusted DstType | Adjusted DstArea |
|---|---|---|---|---|---|
| Tool | A | Tool | Other | Storage | A |
| Tool | A | Storage | Other | Storage | A |
| Storage | A | Tool | Other | Rejected | Rejected |
| Storage | A | Storage | Other | Rejected | Rejected |

Furthermore, each simulation run has been subdivided into a warm-up, failure and relaxation period to cover all different system states (normal, traffic jam with VAB logic and return into normal state).

## 4.3 Results

The evaluation of the simulation experiments was carried out with statistical analysis regarding transports per hour (TPH), vehicle utilisation, mean and 95%-quantile of delivery time. The results depicted in figure 4 reveal a significant reduction of delivery time measures for each simulated failure duration interval. Referring to the close approximation of simulation to real system the results show that the VAB improves the performance of the transport system during failure scenarios. For the short and medium failure durations the expected reduction in TPH due to transport job adjustment can be recognised. Although the VAB logic rejects a certain amount of transports, the TPH for the long failure durations with using the VAB logic is

even higher than in the original state where a certain amount of requests cannot be transported. By adjusting and rejecting transports according to the new logic, this transportation capacity can be increasing compared to the original state. This effect is even more obvious when looking at intra-building transports only because these should be the most affected ones. As depicted in figure 4 the increase of transport capacity is even reached at medium failure duration.

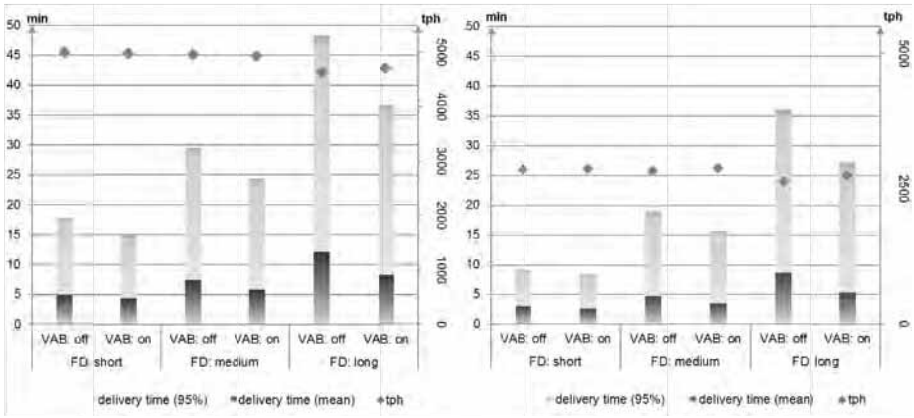



*Figure 4: Summary of simulation experiments for all transports (left) and for intra-building transports only (right)*

With this result in mind, simulation delivered the assumption that the VAB not only improves system performance. The developed solution also effectively increases transport capacity during failure scenarios to support production as good as possible.

By using simulation as a validation tool of the developed logic, first positive results regarding functionality and operational readiness have been achieved. Especially, in semiconductor industry simulation is a strong analysis tool which is indispensable because of the system complexity. Testing transport related changes within the real system would always involve high risks. The failure of the new method very easily could cause a loss for production in the amount of several millions of Euros.

## 5 Conclusion

Even in fully automated semiconductor manufacturing facilities the MES dispatches production material according to tool utilisation and shortest cycle times. Transport related restrictions have not been involved into these decisions upon today. The VAB is an innovative system solution which combines these issues. It was developed to reduce negative effects of failure scenarios in a distributed AMHS.

The presented approach goes beyond common techniques but the transport system as well as the manufacturing line will profit from this innovation. It offers new opportunities to reduce transport related risks for production especially for persons responsible for the transport system. Simulation is a suitable tool to bring this kind of innovative approaches forward and emphasise their usability.

Simulation was used during different parts of the VAB development: Evaluation of important parameters for the calculation of the critical values for each area, validation of these values and finally to test the VAB logic for the first time. Especially the possibility to manipulate real fab data regarding different dispatching rules and testing these rules with the simulation model in advance is beneficial. Simulation studies showed the expected behaviour and that the VAB is a possible tool to improve system performance and increase transport capacity during failure scenarios by keeping a minimum amount of vehicles in each predefined area.

## Acknowledgement

The work has been performed in the project “Entwicklung und Implementierung neuer Steuerungskonzepte zur Unterstützung des Transportsystems in der Halbleiterfertigung”, co-funded by grants from the European Union (EFRE) and the Free State of Saxony (SAB).

## References

Anthonisse, J.M.: The rush in a directed graph. Stitching Mathematisch Centrum. Mathematische Besliskunde (1971) 9, pp. 1-10.

Floyd, R. W.: Algorithm 97 (SHORTEST PATH). Communications of the ACM 5 (1962) 6, p. 345.

Hammel, C., Schmidt, T. and Schöps, M., 2012. Network optimization prior to dynamic simulation of AMHS. In: Laroque, C.; Himmelspach, J; Pasupathy, R.; Rose, O,; Uhrmacher, A.M. (eds.): Proceedings of the 2012 Winter Simulation Conference, Berlin (Germany), 9-12 December 2012, pp. 4673-4781.

Jimenez, J. A.; et al.: AMHS factors enabling small wafer lot manufacturing in semiconductor wafer fabs. In: Johansson, B.; Jain, S.; Montoya-Torres, J.; Hugan, J.; Yücesan, E. (eds.): Proceedings of the 2010 Winter Simulation Conference, Baltimore (USA), 5-8 December 2010, pp. 2575-2585.

Mackulak, G. T. and Savory, P.: A simulation-based experiment for comparing AMHS performance in a semiconductor fabrication facility. IEEE Transactions on Semiconductor Manufacturing 14 (2001) 3, pp. 273-280.

Nishi, T. and Tanaka, Y.: Petri net decomposition approach for dispatching and conflict-free routing of bidirectional automated guided vehicle systems. Systems, IEEE Transactions on Man and Cybernetics, Part A: Systems and Humans 42 (2012) 5, pp. 1230-1243.

Warshall, S.: A Theorem on Boolean Matrices. Journal of the ACM 9 (1962) 1, pp. 11–12.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Konfiguration von Kanallagern auf der Basis dynamischer Kanaltiefen mit Hilfe der Simulation

## *Simulation-based Configuration of Channel Storages by Dynamic Channel Depth*

Thomas Dietz, Lars-Boris Böttcher, Jochen Bernhard, SSI Schäfer Noell GmbH,
Giebelstadt (Germany),
thomas.dietz@ssi-schaefer.com, lars.boettcher@ssi-schaefer.com,
jochen.bernhard@ssi-schaefer.com

Abstract: This document presents the use of simulation for the configuration of multi-deep storage systems. These so called channel storages are especially of interest for the food and beverages industry. Due to the high complexity of influencing variables and the conflict between storage density and performance SSI Schäfer implemented a tool based on a simulation model to support the configuration process. The following chapters will describe constraints, determining factors, target variables, and the process of adapting the simulation results to the configuration problem. The simulation framework is not a fully automated configurator but a tool for target-oriented decision support during the planning process.

## 1 Einführung

Die Nahrungsmittel- und Getränkeindustrie ist i.d.R. charakterisiert durch eine geringe bis mittlere Sortimentstiefe bei einem stark schwankenden Absatz. Insbesondere zur Abdeckung saisonaler Spitzen wie z. B. an Ostern oder Weihnachten werden im Vorfeld hohe Bestände der einzelnen Artikel aufgebaut, die anschließend zu diesen Ereignissen, zeitlich oft sehr konzentriert und mit hohem Umschlagsvolumen an die Zwischenhändler und Kunden verladen und versendet werden müssen. Hierfür wurde in der Vergangenheit immer größere Lagerflächen für die Pufferung der Produkte im Bodenblocklager sowie eine temporäre Leistungserhöhung durch den zusätzlichen Einsatz von Stapler- und Verladepersonal installiert.

Die damit verbundene Komplexitätserhöhung der gezielten Steuerung dieser Prozesse zusammen mit der sinkenden Stapelmöglichkeit aufgrund von Einwegware in der Getränkeindustrie und instabilen Verpackungsformen im Allgemeinen, hat die Nachfrage nach automatischen Lagerlösungen forciert.

Die dargestellten Anforderungen legen den Einsatz einer Kanallagerlösung nahe, insbesondere vor dem Hintergrund des hohen Raumnutzungsgrades und der daraus

resultierenden effizienten und platzsparenden Lagerung. Dem gegenüber stehen die systemischen Nachteile eines Kanallagers wie LIFO-Lagerung (Gudehus 2008, S. 618 f.), längere Zykluszeiten, sowie der gegenüber einfach- oder doppelttiefen Lagern häufig niedrigere Füllgrad (Gudehus 2008, S 630 f.).

Ein Ansatz, um auf die Anforderung von temporär sehr hohen Auslagerleistungen zu reagieren, stellen die Satellitensysteme vieler Hersteller dar (Schrüfer und Frank 2015). Insbesondere mit der Installation von 1-Ebenen-Satellitensystemen ergeben sich in Bezug auf die Ein- und Auslagerung hohe Leistungspotentiale, die jedoch zumeist in ihrem Vollausbau durch die systemisch notwendige Peripherie (vertikaler Transport und Fördertechnikanbindung) limitiert bleiben. Als Alternative kann mit Mehr-Ebenen-Satellitenlösungen sehr individuell auf die Anforderungen der Kunden eingegangen werden.

Um dem niedrigen Füllgrad bei Kanallagerlösungen entgegenzuwirken hat die Firma SSI Schäfer auf der Basis ihres Mehr-Ebenen-Satellitensystems eine flexible und skalierbare Lösung geschaffen, die mittels einer konfigurierbaren Kanaltiefentrennung die ideale Verbindung von füllgradoptimierter Einlagerung und leistungsoptimierter Auslagerung schafft. Das Grundprinzip dieses Ansatzes besteht darin, dass in Lagern mit durchgängigen Kanälen (keine bauliche Trennung) zwischen zwei Fahrgassen für jeden einzelnen Lagerkanal individuell das Teilungsverhältnis festgelegt werden kann. Dadurch können abhängig von Artikelspektrum, Versand- und Produktionsstrategie unterschiedliche Kanaltiefen definiert werden, um einen erhöhten Füllgrad bei maximierter Auslagerleistung zu erreichen (Abb. 1).

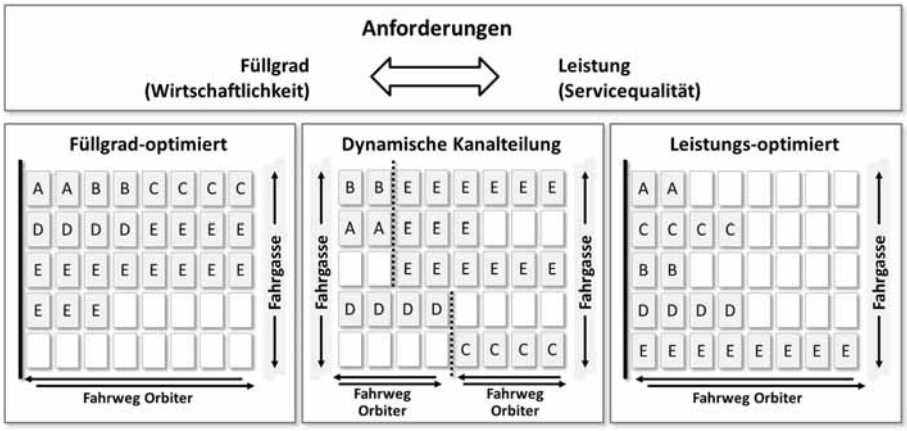



*Abbildung 1: Lagerstrategien*

Beispielsweise können Produkte mit hohen Beständen (A-Artikel) in lange Kanäle gestellt werden, während man Produkte mit niedrigeren Beständen bevorzugt in Kanäle mit weniger Stellplätzen lagert. Eine Reduktion des Lagerfüllgrades durch Artikel mit geringem Bestand oder geringen Umschlagsmengen in teilgefüllten Kanälen kann dadurch gemindert werden.

Fundierte Aussagen bezüglich der Leistung und des erreichbaren Füllgrades eines Kanallagers sind mit Hilfe statischer Berechnungen (VDI 2009) oder Heuristiken

nur schwer möglich. Aufgrund der hohen Dynamik, den vielfältigen Wechselbeziehungen und der Komplexität der Problemstellung, stellt die Simulation ein geeignetes Mittel zur Analyse dar. Aufgrund der in den Projekten variierenden Anforderungen und Einflussfaktoren, sinkt jedoch die Wiederverwendbarkeit der Simulationsmodelle, was wiederum den Aufwand für die Modellerstellung erhöht.

Durch die Wahl eines geeigneten Abstraktionsgrades und aufgabenbezogener Modellgrenzen muss ein Kompromiss zwischen Aufwand und Nutzen im Rahmen von Simulationsstudien gefunden werden.

Der Beitrag beschreibt wie mit Hilfe der Simulation verschiedene Lagerkonfigurationen bzgl. der Kanaltiefe im Zusammenspiel mit füllgradverbessernden Einlagerstrategien und leistungsverbessernden Auslagerstrategien aufwandsreduziert untersucht werden können.

Dazu erläutert der Beitrag ausgehend von der Analyse der Problemstellung das Konzept bezüglich Abstraktionsgrad, Systemgrenzen und Schnittstellen. Weiterhin werden die Stellgrößen und Konfigurationsparameter vorgestellt. Abschließend widmet sich der Beitrag dem Ablauf und der Interpretation der Ergebnisse dieser Simulationsstudien und veranschaulicht diese anhand von Beispielen.

## 2 Rahmenbedingungen und Einflussgrößen

Grob sind zwei Gruppen von Faktoren zu unterscheiden, die Einfluss auf die Konfiguration eines Kanallagers haben und aus denen sich auch die Zielgrößen ableiten. Die zweckgebundenen Faktoren beinhalten die Restriktionen und Anforderungen des Kunden bzw. der Branche. Die lösungsbedingten Faktoren ergeben sich aus der Konfiguration des Lagers und sollten bestmöglich mit den zweckgebundenen Faktoren korrelieren.

### 2.1 Zweckgebundene Faktoren

Kanallager werden in der Nahrungsmittel- und Getränkeindustrie zur Pufferung von Produktionslosen und der (sequenzierten) Versandbereitstellung verwendet. Eine erweiterte Nutzung sieht häufig die Lagerung von konfektionierten Waren oder das Zwischenpuffern von kommissionierten Ladungsträgern vor. Im Rahmen der dargestellten Betrachtungen werden über diese Anwendungsbereiche hinausgehende Funktionen nicht untersucht.

Basierend auf den Erfahrungen aus der Getränkeindustrie ist eine SKU-reine (eindeutige Kennzeichnung durch Artikelnummer und MHD) Befüllung der Kanäle zielführend. Hierzu wird leistungsabhängig eine definierte Anzahl Kanäle in unterschiedlichen Gassen für die Einlagerung disponiert. Nach vollständiger Befüllung eines Kanals wird für die SKU möglichst innerhalb einer bisher nicht berücksichtigten Gasse ein neuer, leerer Kanal ausgewählt und für die weitere Einlagerung verwendet. Mittels dieser Reglementierung über die maximale Anzahl gleichzeitig disponierter Einlagerkanäle sowie dem Gassenwechsel bei der Auswahl neu zu befüllender Kanäle wird sowohl ein hoher Füllgrad als auch eine maximale Verteilung der SKU über das Lager erreicht. Insbesondere Letzteres ermöglicht in Bezug auf die Auslagerung eine hohe Verfügbarkeit und eine maximale Zugriffsanzahl einer SKU, was für eine hohe Auslagerleistung elementar ist.

Die Konfiguration eines Kanallagers und insbesondere der Kanaltiefe ist somit abhängig vom Produktions- und Versandprogramm. Die wesentlichen Faktoren, die sich aus dem Produktionsprogramm ergeben sind:

- Anzahl Artikel
- Produktionsmenge pro Zeiteinheit
- Produktionslosgröße

Für den Versand sind folgende Faktoren ausschlaggebend:

- Anzahl Verladeeinheiten (bspw. LKW) pro Zeiteinheit
- Sequenzgruppen pro Ladeeinheit
- Ladungsträger pro Sequenzgruppe

## 2.2 Lösungsbezogene Faktoren

Bei der Konzeption von Kanallagern gibt es eine Vielzahl verschiedener Konfigurationsmöglichkeiten. Durch den Einsatz einer effizienteren 3D-MATRIX®-Konfiguration mit dem SSI Schäfer Lift&Run-System (SLR) erhöht sich die Komplexität zudem. Bei 3D-MATRIX®-Konfigurationen wird die zu- und abführende Fördertechnik nicht stirnseitig an das Lager angebunden, sondern verläuft im oder unter dem Lager quer zur Gassenausrichtung. Hierdurch wird eine dreidimensionale Lagersteuerung möglich und notwendig. Diese charakterisiert sich insbesondere dadurch, dass die SKU-Disposition für Ein- und Auslagerung nicht wie üblich von dem vorgelagerten Transportsystem, sondern systemimmanent und über das gesamte Lager auf allen Ebenen verteilt durch die Shuttlefahrzeuge (SLR) erfolgt.

SLR sind Regalbediengeräte mit jeweils einem Kanalfahrzeugen (Orbiter) die in der Regel 2 bis 5 Lagerebenen bedienen können. Mehrere SLR-Geräte bedienen, übereinander angeordnet, eine Lagergasse und sind über Lifte an die Fördertechnik angebunden. Bei Anordnung der Fördertechnik unter oder im Lagerblock sind die Lifte entsprechend mittig in der Gasse (im Regal) angeordnet (Abb. 2).

Laut Definition fahren die Regalbediengeräte in x-Richtung, die Lifte in y-Richtung und die Fördertechnik verläuft in z-Richtung im oder unter dem Lager.

Zentrale Parameter für die Lagerkonfiguration sind:

- Anzahl Gassen
- Anzahl Regalbediengeräte pro Gasse (übereinander)
- Gassentopologie
  - Anzahl Fächer pro Lagerebene (x-Koordinaten)
  - Anzahl Lagerebenen (y-Koordinaten)
  - Kanaltiefe
  - Anzahl und Position der Lifte

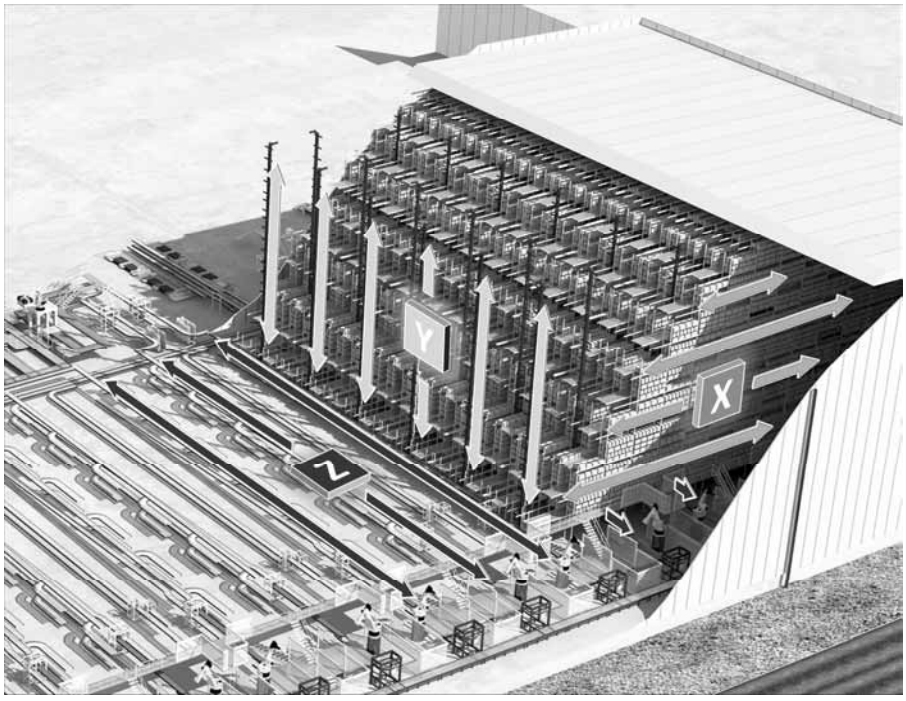



*Abbildung 2: 3D-MATRIX®*

## 3 Modellbeschreibung

Um die Aufgabenstellung beherrschbar und allgemeingültig zu gestalten, musste zunächst ein hinreichend genauer Abstraktionsgrad und geeignete Systemgrenzen (Betrachtungsgrenzen) definiert werden. Ziel der Untersuchung ist es, die Entwicklung des Lagerbestandes und der Lagerfragmentierung über einen längeren Zeitraum, idealerweise länger als 30 Tage, zu analysieren.

Im Fokus des Modells liegt die Bewertung der Leistung und des Verhaltens des Lagerblocks. Für diesen Bereich wurde daher ein sehr hoher Detaillierungsgrad gewählt. Im Gegensatz dazu wird vor- und nachgelagerte Fördertechnik abstrakt (als Blackbox) betrachtet. Schnittstellen bilden die Ein- und Auslagerpositionen der Regalbediengeräte. Für ein SLR-System bilden die Lifte die Schnittstelle zur Fördertechnik und werden als Modellgrenzen festgelegt.

Einlagerungsseitig erzeugt eine Quelle mit einem definierten Produktions- bzw. Wareneingangsprogramm entsprechende Ladungseinheiten (Paletten), die durch einen Logikbaustein auf die einzelnen Einlagerpositionen verteilt werden. Auslageraufträge werden anhand eines Versandprogrammes im Modell generiert.

Abbildung 3 verdeutlicht die unterschiedlichen Detaillierungsgrade und die grobe Struktur des Simulationsmodells.

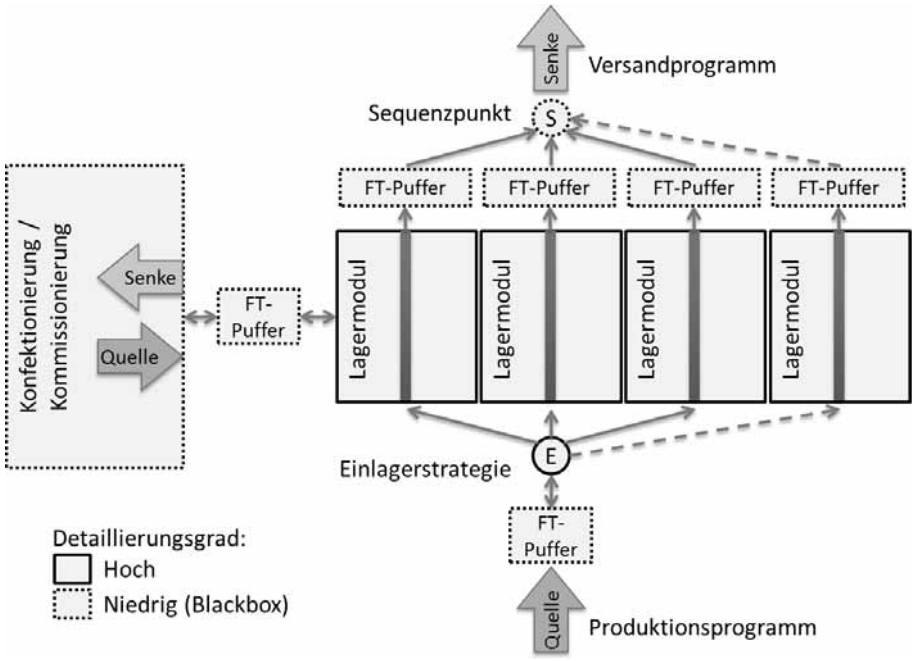



*Abbildung 3: Schema Modellkonfiguration*

## 3.1 Eingangsdaten

Eine valide und realistische Datenbasis ist die Grundlage für aussagekräftige Ergebnisse. Deswegen arbeitet SSI Schäfer soweit möglich mit (extrapolierten) Echtdaten. Eine besondere Herausforderung stellt dabei die Bereitstellung einer konsistenten Datenbasis dar, die in einer frühen Projektphase oder aufgrund der IT-Infrastruktur des Betreibers jedoch nicht immer zur Verfügung steht. Zu einem definierten Warenausgangsprogramm muss ein passendes und realistisches Wareneingangsprogramm erzeugt werden. Diese Aufgabe wird in der Realität durch entsprechendes Fachpersonal in der Produktionsplanung ausgeführt.

Aus diesem Grund wird momentan an einem Datengenerator gearbeitet, der über eine parametrierbare Versandauftragsstruktur verfügen soll und anhand von Rahmenparametern ein realistisches Einlagerprogramm erzeugt. Die wesentlichen Parameter werden im Folgenden beschrieben. Alternativ zu automatisch generierten Daten sollen auch weiterhin entsprechende (Echt-)Daten bspw. aus einem Tabellenkalkulationsprogramm ausgelesen werden können.

### *3.1.1 Definition Auslagerungen*

Ein Versandprogramm wird durch folgende Parameter definiert:

- Anzahl Versandgruppen pro Stunde (nach Typ)
- Versandgruppen-Typ:
  - Anzahl Aufträge pro Versandgruppe (bspw. Kunden, Sequenzanforderung)
  - Anzahl Paletten pro Auftrag
  - Anzahl SKU pro Auftrag
  - Anzahl kommissionierte Paletten pro Auftrag

### *3.1.2 Definition Einlagerungen*

Das Einlagerprogramm ist durch die folgenden Rahmenparameter gekennzeichnet:

- Anzahl SKU
- Losgrößen (jeweils für A-, B- und C-Artikel)
- Anzahl Quellen (z.B. Produktionslinien)
- Mittlere Leistung pro Quelle
- Rüstzeiten

Alle Parameter außer der Anzahl SKU können als Konstante oder statistische Verteilung definiert werden.

## 3.2 Lagerkonfiguration und Modellgenerierung

Die Lagerkonfiguration erfolgt detailliert über Parameter sowie eine MS Excel-basierten Tabelle, in der die physikalische Topologie der einzelnen Lagergassen definiert wird. Dazu wird für jedes Fach (x/y-Position) in der Gasse die Kanaltiefe auf der linken und rechten Seite des Regalbediengerätes angegeben (Abb. 4).

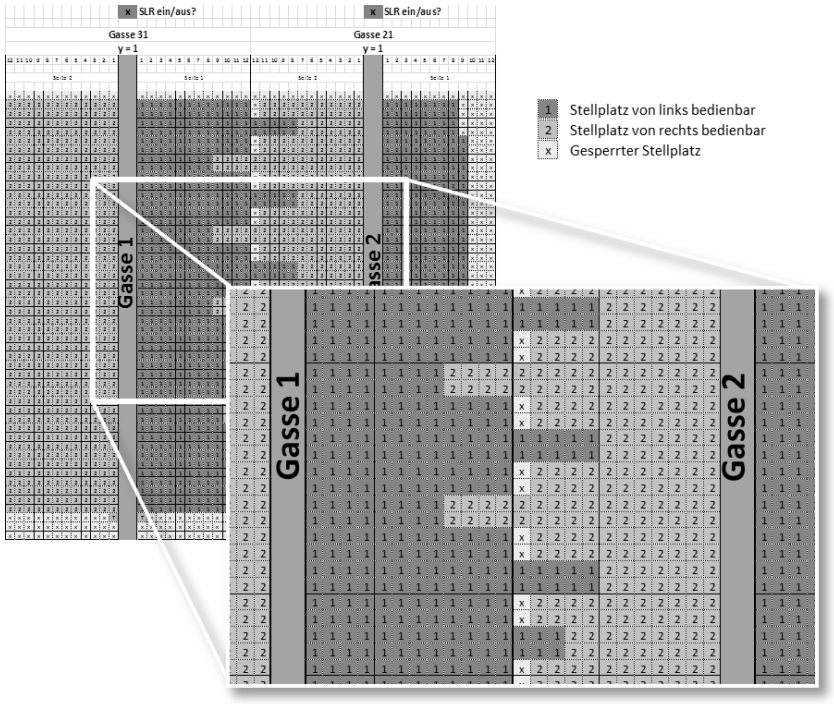



*Abbildung 4: Beispielkonfiguration Kanalraster*

Nach dem Einlesen der Topologie in die Simulationssoftware wird das Simulationsmodell automatisch generiert. Für die Lager- und Bestandsverwaltung des Lagerblockes wird eine entsprechende Datenbankstruktur aufgebaut. Die Fördertechnik wird abstrakt durch Puffer abgebildet.

Durch die Überwachung und Auswertung der Pufferfüllstände kann ausgewertet werden, ob es bei der Einlagerung oder bei der Sequenzbildung zu Engpässen gekommen ist. Die Einlagerstrategie definiert die Verteilung der Artikel auf die einzelnen Gassen. Der Sequenzpunkt stellt sicher, dass Schieflasten bei der Auslagerung erkannt werden. Stellt bspw. eine bestimmte Gasse den Engpass dar, füllt sich der Puffer der anderen Gassen, da Paletten zurückgehalten werden müssen.

## 4 Ergebnisse und Zielfunktion

Eine Besonderheit stellt die Definition der Initialbestände im Lager dar. Die Befüllung des Lages mit realistischen Beständen zu Simulationsbeginn führt zu einer wesentlich verkürzten Einschwingphase des Modells. So können bereits beim Start eines Simulationslaufes Auslageraufträge eingelastet werden ohne, dass diese zu Fehlbeständen im Lager führen. Zu Bestimmung eines geeigneten Initialbestandes wird zunächst ein Simulationslauf durchgeführt, der auch negative Bestände zulässt. Der maximal aufgetretene Negativbestand jeder SKU im Simulationszeitraum wird als Mindestmenge festgelegt, die zu Simulationsstart eingebucht sein muss (Abb. 5).

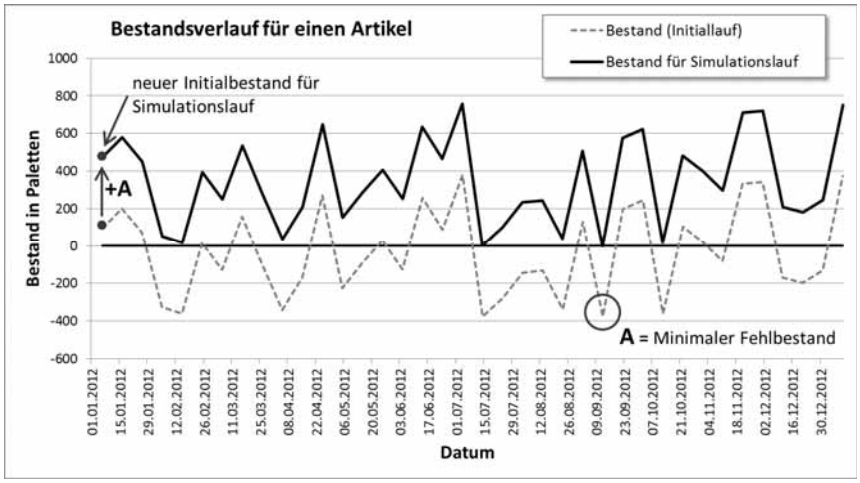



*Abbildung 5: Beispiel Initialbestandsermittlung*

Bereits nach diesem Lauf kann erkennbar sein, dass das Lager keine ausreichende Kapazität hat, um den Initialbestand aufzunehmen. Nach der Validierung und eventuellen Konfigurationsanpassungen werden die weiteren Läufe mit validem Initialbestand durchgeführt.

Am Ende jedes weiteren Simulationslaufes werden folgende Kennzahlen betrachtet:

- Anzahl freier, teilgefüllter und voller Kanäle im zeitlichen Verlauf (Abb. 6)
- Füllstände der Fördertechnik-Puffer im zeitlichen Verlauf
- Anzahl ein- bzw. ausgelagerter Paletten pro Stunde
- Anzahl abgeschlossener Versandgruppen pro Stunde
- Evtl. aufgetretene Fehlmengen

Die entsprechenden Kennzahlen stehen jeweils für das gesamte Lager oder definierte Teilbereiche (Gassen, SLR-Fahrebenen, etc.) zur Verfügung.

Die Qualität einer Lösung bemisst sich im Wesentlichen an:

- Lagerfüllgrad
- Liefertreue

Die Simulation stellt keine automatische Bewertung, bzw. Zielwert zur Verfügung. Es ist Aufgabe des Nutzers, anhand der Simulationsergebnisse die richtigen Schlüsse zu ziehen und geeignete Alternativen zu konfigurieren. Es kann auch sinnvoll sein eine Diskussion mit dem Betreiber über die Parameter, wie bspw. Losgrößen anzuregen. Die Auswirkungen können anhand einer Experimentserie anschaulich gemacht werden.

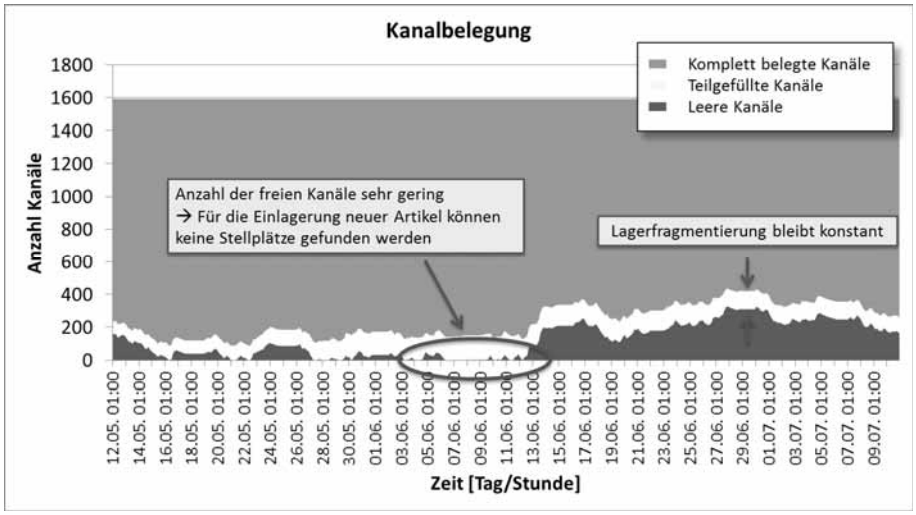



*Abbildung 6: Beispiel Kanalbelegung*

Das Tool soll zukünftig um eine Zielfunktion ergänzt werden. Das Konzept der Zielfunktion stellt sich wie folgt dar:

Parameter:

- $FG_{Ziel}$ – durchschnittlicher Zielfüllgrad (bspw. 90%)
- $G_{FG}$ – Gewichtungsfaktor bei Zielfüllgradabweichung pro Prozentpunkt
- $G_{Frag}$ – Gewichtungsfaktor für Lagerfragmentierung
- $G_{BZ}$ – Gewichtungsfaktor für Bereitstellzeit

Messwerte:

- $Frag_{avg}$ – durchschnittliche Lagerfragmentierung = (Anzahl freier Stellplätze in teilgefüllten Kanälen/ Gesamtzahl Stellplätze) * 100
- $BZ_{avg}$ – durchschnittliche Bereitstellzeit pro Palette = Zeitspanne von Anforderung der Palette durch Auslagerprogramm bis Ankunft am Sequenzpunkt in Minuten

- $FG_{avg}$ – Durchschnittlicher Füllgrad über die Simulationsdauer. Füllgrad = (Anzahl freie Stellplätze/ Gesamtzahl Stellplätze) * 100

Folgender Zielwert (Z) soll minimiert werden:

$$Z = G_{FG} \cdot \left| FG_{Ziel} - FG_{avg} \right| + G_{Frag} \cdot F_{avg} + G_{BZ} \cdot BZ_{avg} \quad (1)$$

Die Anzahl ein- und ausgelagerter Paletten fließt nicht in die Zielfunktion ein, da sie von den Eingangsdaten abhängig ist und das Ergebnis verfälschen könnte, wenn nicht zu jedem Zeitpunkt ausreichend Last auf dem System ist. Wird für eine einzulagernde Palette kein freier Stellplatz gefunden, ist dies ein Abbruchkriterium für die Simulation und der Zielwert wird auf das Maximum gesetzt. Die jederzeitige Verfügbarkeit von Lagerplätzen für die Einlagerung stellt ein KO-Kriterium für die Konfiguration dar.

Weitere mögliche Abbruchkriterien sind:

- $P_{max}$ - Maximale Anzahl Paletten im einlagerseitigen FT-Puffer. Wird dieser Wert überschritten, ist dies ein Abbruchkriterium, da ein Rückstau bspw. in die Produktion auftreten würde.
- $BZ_{max}$ - Maximale Bereitstellzeit. Wird dieser Wert überschritten, ist dies ein Abbruchkriterium, da bspw. die Lieferbereitschaft nicht eingehalten werden würde.

## 5 Zusammenfassung

Mit Hilfe des entwickelten Simulationsframeworks ist es möglich in kurzer Zeit durch Parametrierung ein generisches Simulationsmodell zu erzeugen. So können unterschiedliche Konfigurationen mit geringem Aufwand getestet werden. Die Simulation ist das geeignete Werkzeug, da die Entwicklung des Lagerfüllgrades und der Lagerfragmentierung im zeitlichen Verlauf und die dadurch bedingten Wechselwirkungen mit den vor- und nachgeschalteten Bereichen des Kanallagers mit Hilfe rechnerischer Methoden nicht oder nur mit hohem Aufwand ermittelt werden können.

Da nicht immer geeignete Eingangsdaten zur Verfügung stehen, wäre ein automatischer Datengenerator eine sinnvolle Erweiterung. Die Qualität der Zielfunktion muss in weiteren Projekten bestätigt und gegebenenfalls durch entsprechende Anpassungen verbessert werden.

## Literatur

Gudehus, T.: „Logistik: Grundlagen, Strategien, Anwendungen“. Berlin, Heidelberg: Springer 2005.

Schrüfer, M.; Frank, S.: „Drängt das Shuttle das RBG in die Nische?“. JM Fachmedien GmbH: Materialfluss (2015) 6, S. 18-23.

VDI: VDI-Richtlinie 3561 Blatt 4 „Spielzeitermittlung von automatischen Kanallager-Systemen“. Berlin: Beuth 2009.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Teilautomatisierte Softwaretests mithilfe einer mechanisch plausiblen Robotersimulation

## *Semi-automatic Software Tests Using a Mechanically Plausible Robot Simulation*

Fabian Aichele, Bernd Eckstein, Björn Schenke, TruPhysics GmbH, Stuttgart (Germany), aichele@truphysics.com, eckstein@truphysics.com, schenke@truphysics.com

Abstract: This article describes the simulation tool TruPhysics and how it can be used to test robot control programs. TruPhysics simulates an industrial robot in a mechanically plausible way and in real-time. This is achieved by using a newly developed, GPU-based collision-detection algorithm. The simulation directly connects to a robot controller and reacts to its signals like a real robot would. Additionally, TruPhysics provides methods for semi-automatic software testing that allow the definition of test cases. Based on this functionality, the simulation can monitor the execution of robot control programs under varying conditions and determine if these control programs behave as intended.

## 1 Einleitung

Zum Inbetriebnahme-Prozess eines Industrieroboters gehört die Erstellung eines Steuerungsprogramms. Bei manchen Anwendungsfällen (beispielsweise Schweiß- oder Lötprozessen) kann dies bereits erfolgen, noch bevor ein Roboter als Bestandteil einer Fertigungsanlage aufgebaut wird, indem ein Steuerungsprogramm geschrieben und anhand einer visuellen Simulation des Roboters und seiner Arbeitsumgebung getestet wird.

Aufgrund der Beschränkungen gängiger Simulationsumgebungen ist die Voraussetzung für dieses Vorgehen, dass die Bewegungsabläufe eines Roboters bereits im Voraus bekannt sein müssen. Daher ist es beispielsweise nicht möglich, eine Positionierungsstrategie in das Programm zu integrieren, um ein Werkstück an eine nicht exakt vorhersehbare Position (bedingt etwa durch Fertigungs- oder Positionierungstoleranzen) zu bringen. Wenn ein bestimmter Prozess eine derartige Variation im Steuerungsprogramm des Roboters erfordert, dann kann die Funktionsweise dieses Teils des Steuerungsprogramms nur am fertig konfigurierten realen Roboter getestet

werden. Dies birgt eine erhöhte Wahrscheinlichkeit von im späteren Produktionsbetrieb auftretenden software-bedingten Fehlern.

Eine Lösung für dieses Problem besteht darin, eine Simulation mit weniger Beschränkungen zu entwickeln, die auch den Test von komplexer strukturierten Steuerungsprogrammen ermöglicht. Dieses Ziel wird mit TruPhysics verfolgt.

Bei TruPhysics wird zur Kollisionserkennung ein Verfahren verwendet, welches sich die Rechenleistung von GPGPU-Prozessoren („General Purpose Graphics Processing Unit“, massiv parallele Grafikkarten-Prozessoren) zunutze macht. Die Steuerung simulierter Roboter erfolgt direkt durch die Verwendung realer Steuerungsrechner. Dies ermöglicht es, ein Steuerungsprogramm auf dem Steuerungsrechner zu erstellen und es mit Hilfe von TruPhysics zu testen.

Durch diese Vorgehensweise wird es möglich, wesentlich umfassendere Softwaretests durchzuführen, als es bei Steuerungsprogrammen für Industrieroboter bisher möglich ist. Einen systematischen Ansatz zum Testen von Steuerungsprogrammen liefert der Einsatz von teil-automatisierten Testfällen, die anhand vom Anwender spezifizierter Erfolgs-Kriterien und dem Einsatz von Überwachungsbausteinen realisiert werden.

Dieser Beitrag ist wie folgt strukturiert: Kapitel 2 beschreibt den aktuellen Stand der Technik in der virtuellen Inbetriebnahme. Kapitel 3 beschreibt das Simulationswerkzeug TruPhysics, insbesondere dessen Mechanik-Simulation und Kollisionserkennung. Kapitel 4 beschreibt den Softwaretest mithilfe von TruPhysics durch eine Beschreibung der allgemeinen Vorgehensweise, gefolgt von einem konkreten Beispielszenario.

## 2 Stand der Technik

Existierende Software-Lösungen mit ähnlichen Zielsetzungen wie TruPhysics sind in zwei verschiedenen Bereichen zu finden: Einerseits in der Produktionsplanung, speziell in der virtuellen Inbetriebnahme, andererseits im Bereich echtzeitfähiger Mechanik-Simulationen, die beispielsweise in Simulations-Lösungen für die Service-Robotik verwendet werden (Abb. 1 zeigt eine Übersicht verfügbarer Software).

### 2.1 Virtuelle Inbetriebnahme

In diesem Bereich sind Lösungen unterschiedlicher Hersteller verfügbar, die die Programmierung von Industrie-Robotern und deren Integration in automatisierte Fertigungsprozesse erlauben. Eine beispielhafte, nicht repräsentative Auswahl umfasst etwa Tecnomatix RobotExpert (Siemens 2015b) und NX Mechatronics Concept Designer (Siemens 2015a) der Firma Siemens, 3DS Robotics on Premise (Dassault Systèmes 2015) der Firma Dassault Systèmes, Kuka SimPro (KUKA 2015) und ABB RobotStudio (ABB 2015).

Diese Programme erlauben die Festlegung von Bewegungsbahnen von Robotern und deren Aktorik anhand vorliegender Konstruktionsdaten. Die vom Konstrukteur festgelegten Bahndaten werden in steuerungs-spezifischen Programm-Code überführt, der auf einen Steuerungsrechner eingespielt werden kann („Offline-Programmierung“). So können deterministische Montageschritte, die nicht auf dynamische Veränderungen während eines Ablaufs reagieren müssen, bereits komplett während

der Konstruktionsphase umgesetzt werden. Beispiele hierfür sind etwa Schweiß- oder Lackier-Arbeiten an Karosserieteilen. Dabei erfolgt die Interaktion mit Werkstücken zu fest definierten Zeitpunkten, und es sind keine Positionierungstoleranzen auszugleichen. Hinsichtlich des Funktionsumfangs verfügen diese Werkzeuge allerdings nicht über integrierte Mechanik-Simulationen. Abläufe werden daher nur kinematisch plausibel oder animiert dargestellt. Die Simulation dynamischen oder mechanisch plausiblen Verhaltens fehlt ebenso wie die Fähigkeit zur Erzeugung von Sensor-Informationen aus einer virtuell modellierten Umgebung. Einige dieser Werkzeuge verfügen zwar über integrierte Kollisionserkennungs-Komponenten, diese visualisieren mögliche Kollisionen allerdings nur und sind zudem nicht zum echtzeitfähigen Betrieb in Kombination mit einem Steuerungsgerät fähig.

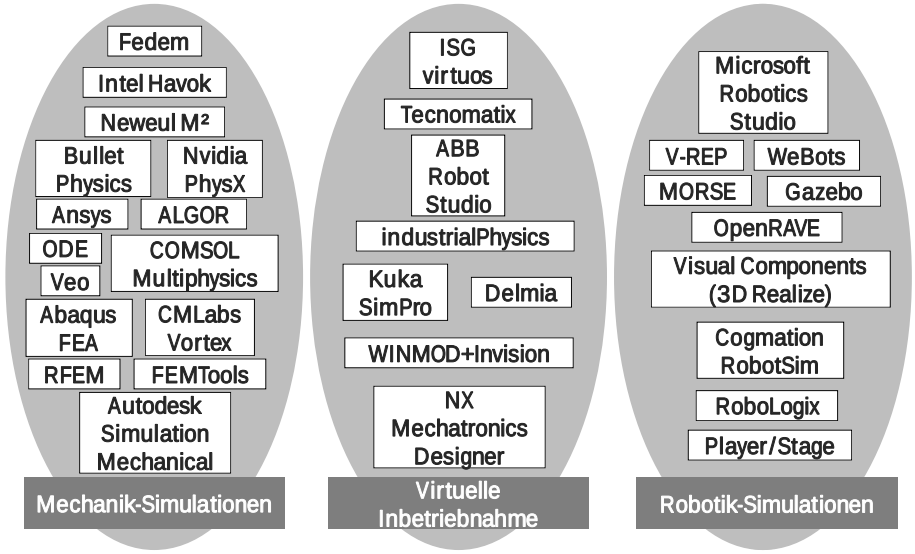



*Abbildung 1: Übersicht über verfügbare Software-Lösungen*

Eine weitere Klasse von Software-Werkzeugen für die virtuelle Inbetriebnahme sind Simulationsumgebungen für Werkzeugmaschinen, CNC-Fräsen oder Förderband-Systeme wie etwa industrialPhysics der Firma Machineering (Wünsch 2015), 3DRealize R der Firma Visual Components (Visual Components 2015) oder virtuos der Firma ISG (ISG 2015). Diese sind vergleichbar mit den zuvor genannten Lösungen ebenfalls für Offline-Programmierung konzipiert, verfügen jedoch teilweise über integrierte Mechanik-Simulationen: Die im folgenden Abschnitt skizzierten Physik-Engines. Dadurch ist zwar eine grundlegende Unterstützung für die Simulation von Prozessen mit dynamischem Objektverhalten gegeben, die jedoch wie im folgenden Abschnitt ausgeführt wesentlichen Einschränkungen unterliegt.

## 2.2 Echtzeitfähige Mechanik-Simulationen (Physik-Engines)

In diesem Bereich gibt eine große Bandbreite von Werkzeugen, die sich grob in zwei verschiedene Kategorien einteilen lassen. In die erste Kategorie fallen Physiksimulationen, die eine hohe physikalische Plausibilität aufweisen, dafür aber nicht

zum echtzeitfähigen Betrieb geeignet sind. Dies sind insbesondere Finite-Elemente-Methode (FEM) und Mehrkörper-Simulationen (MKS). Aufgrund der hohen Laufzeitanforderungen ist eine Eignung dieser Simulationslösungen für den Betrieb in Kombination mit realen Steuerungsrechnern nicht gegeben.

In die zweite Kategorie fallen echtzeitfähige Physik-Simulationen: Physik-Engines wie Bullet Physics (Coumans 2015), ODE (Smith 2015), DART (Georgia Tech 2015), SimBody (Sherman and Eastman 2015), MuJoCo (Todorov 2015), Intel Havok (Intel 2015), Nvidia PhysX (NVIDIA 2015) oder CMLabs Vortex (CM Labs 2015). Ein Vergleich einiger dieser Physik-Engines findet sich beispielsweise in (Erez et al.). Solche Physik-Engines finden oft in Computerspielen Anwendung.

Im Gegensatz zu FEM- und MKS-Systemen verzichten diese Physik-Engines auf die hoch-detaillierte Modellierung geometrischer und physikalischer Eigenschaften (z. B. Verformungsarbeit). Sie sind überwiegend für Szenarien mit einer großen Anzahl geometrisch einfach strukturierter beweglicher Objekte (etwa Kugeln, Quader oder konvexe Geometrien) konzipiert, die mit einer ansonsten statischen Umgebung interagieren. Auch die Plausibilität mechanischen Verhaltens ist vielfach nicht gegeben. So sind viele der genannten Lösungen nicht in der Lage, umschließende Kontakte zwischen geometrisch komplex strukturierten Objekten über längere Zeit stabil abzubilden (Aichele 2015). Es stellen sich oft bereits nach kurzer Zeit große Durchdringungen zwischen Objekten ein, die im Extremfall zum unkontrollierten Auseinanderfallen des umschlossenen und umschließenden Objekts führen.

## 3 Kurzbeschreibung von TruPhysics

Die Komponenten des Gesamtsystems sind in Abbildung 2 schematisch dargestellt.

Der zentrale Bestandteil von TruPhysics ist die Hauptanwendung, in der die Kollisionserkennungs-Komponente und die Mechanik-Simulation als Bestandteile des Simulationskerns zusammen mit einer in 3D visualisierten Darstellung der aktiven Simulation vereint sind.

Die Kontrolle von simulierten Robotern erfolgt über eine an die Hauptanwendung angebundene Robotersteuerung. Die Bewegungsdaten werden von der Robotersteuerung an die Hauptanwendung übermittelt und durch die Mechanik-Simulation auf die virtuellen Modelle der entsprechenden Roboter übertragen. Unter anderem mit Hilfe der Kollisionserkennungs-Komponente werden in der aktiven Simulation Sensorwerte (z. B. Abstandsmessungen oder mechanische Kontakte) ermittelt, die von der Hauptanwendung wieder an angeschlossene Steuergeräte zurück übermittelt werden.

Im Verbund mit der Hauptanwendung wird eine Testumgebung ausgeführt. Diese hat sowohl Zugriff auf die zwischen Steuergerät und Hauptanwendung kommunizierten Daten als auch auf die Berechnungsergebnisse der Kollisionserkennung und der Mechanik-Simulation. Mit diesen Daten werden Test-Bausteine versorgt. Diese Testbausteine kombinieren logische und zeitliche Verknüpfungen zu Erfolgs- oder Misserfolgs-Kriterien, und setzen so bestimmte Testfälle um.

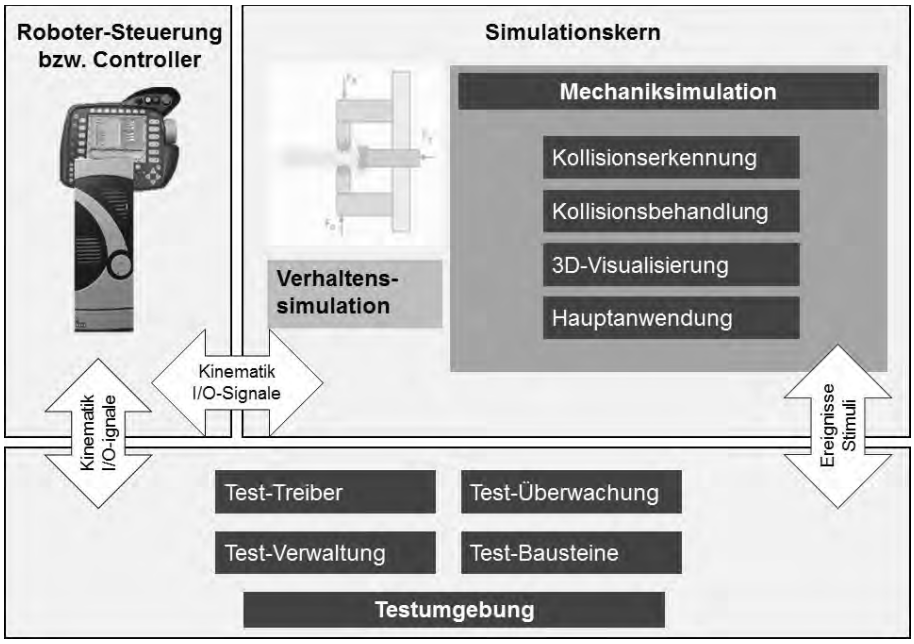



*Abbildung 2: Systemarchitektur von TruPhysics*

Besondere Relevanz hat bei der Implementierung die Mechanik-Simulation und insbesondere die Kollisions-Erkennung, also die Überprüfung komplex strukturierter dreidimensionaler Geometrien auf gegenseitige Berührung oder Überlappung. Die Kollisions-Erkennung ist der Teilschritt jeder Iteration, der den größten Laufzeitaufwand erfordert.

Zur Reduzierung des Laufzeitaufwands werden große Geometrien üblicherweise mit sogenannten Hüllkörper-Hierarchien umgeben. Ausgehend von der gesamten Geometrie werden sukzessive immer kleinere Teil-Geometrien mit Hüllkörpern umgeben (z. B. Hüll-Kugeln oder Hüll-Quadern). Diese Hüllkörper werden paarweise auf Überschneidung geprüft. Diese Überprüfung wird nur dann rekursiv für Kind-Geometrien fortgesetzt, wenn eine Überschneidung der Eltern-Hüllkörper vorliegt. Übliche Ansätze zur Kollisionserkennung (Eckstein 1999; Weller 2013; Erleben 2005) beschränken sich hierbei auf die Verwendung normaler Prozessoren. Das schränkt Möglichkeiten zur Parallelisierung von Kollisionsabfragen sowohl auf der Ebene von Objektpaaren als auch bei der Ausführung mehrerer Paar-Tests ein.

Um diesen Problemen entgegenzuwirken, verwendet TruPhysics eine speziell für die Verarbeitung CAD-konstruierter Geometrien konzipierte Kollisionserkennungs-Bibliothek (Aichele 2015). Diese verwendet GPU-Prozessoren (Graphics Processing Unit) zur massiv parallelen Ausführung von Kollisionsberechnungen. Die Bibliothek beruht auf einem hybriden CPU-/GPU-Verfahren. CPU-basiert werden nur die obersten Ebenen von Hüllkörper-Hierarchien ausgewertet. Die detaillierte Kollisionsberechnung wird GPU-basiert ausgeführt, mit Hilfe von raumpartitionierungs-basierten Verfahren, die sich besonders für massiv parallele Ausführung eignen. Dadurch wird eine enorme Leistungssteigerung gegenüber CPU-basierten Verfahren erreicht. Das ermöglicht die Simulation räumlich kompakter Arbeitsum-

gebungen (bis ca. 25 m²) mit komplex strukturierten Objekten im niedrigen dreistelligen Bereich. Im Vergleich zu anderen GPU-basierten Kollisionserkennungsverfahren (Lauterbach et al. 2010; Erbes et al. 2013) erreicht die Implementierung von TruPhysics eine Parallelisierung von Kollisionsberechnungen über mehrere Paartests hinweg, und ersetzt die vollständige Expansion von Hüllkörper-Hierarchien durch eine sehr effiziente Such-Operation.

## 4 Softwaretest mit TruPhysics

TruPhysics ist ein Werkzeug, welches spezifisch den sogenannten dynamischen Softwaretest unterstützt. Beim dynamischen Softwaretest wird das Programm ausgeführt, und es wird geprüft, welche Ausgaben es als Reaktion auf bestimmte Eingaben liefert.

Der dynamische Softwaretest lässt sich in zwei Kategorien einteilen, die struktur- und die funktionsorientierten Verfahren (Liggesmeyer 2009). Bei den strukturorientierten Verfahren, die zu den sogenannten „White-Box“-Techniken gehören, werden Testfälle anhand des Programmcodes erstellt. Testfälle werden dabei so formuliert, dass alle Teile des Programmcodes ausgeführt und auf korrekte Funktion getestet werden. Bei den funktionsorientierten Verfahren, die zu den sogenannten „Black-Box“-Techniken gehören, werden Testfälle anhand einer Spezifikation erstellt. Diese Spezifikation geht nicht auf die Funktionsweise des Programms ein, sondern legt fest, welche Ausgabe auf eine bestimmte Eingabe folgen muss, wenn das Programm korrekt abläuft.

Für beide Arten des dynamischen Softwaretests ist es nötig, das Programm auszuführen und die Ausgabe auszuwerten. Bei Anwendersoftware ist dies in den meisten Fällen kein Problem, da die Entwicklung des Programms, dessen Ausführung und die Ausgabe der Ergebnisse alle auf demselben Gerät erfolgen.

Dies ist anders beim Test eines Steuerungsprogramms für einen Roboter. Das Programm kann zwar auf dem Steuergerät ausgeführt werden, aber ohne einen angeschlossenen Roboter kann nicht beurteilt werden, ob das Verhalten eines Steuerungsprogramms korrekt ist.

Bisherige Ansätze zum Test von Steuerungsprogrammen konzentrieren sich ebenfalls auf die Durchführung dynamischer Tests. Allerdings erfolgt bei diesen Lösungen die Auswertung des Tests nicht automatisiert, d.h. die Durchführung und die Beurteilung eines Testdurchlaufs müssen manuell durch den Benutzer erfolgen (Sossenheimer 1989; Kohring 1993; Hengel 1994). Andere Lösungen beziehen das zu testende Steuergerät nicht direkt in den Testablauf mit ein (Strassacker 1997).

Die Anweisungen, aus denen ein Steuerungsprogramm besteht, lassen sich unterteilen in statische Bahnbefehle, Syntax-Anweisungen (z. B. Deklarationen, Funktionsaufrufe) und Programmbausteine für Montage-Aufgaben (z. B. Such-, Fehlerbehebungs- und Positionierungs-Strategien). Die beiden letzten Gruppen, die bei einer Analyse (Roßdeutscher 2011) der Steuerungsprogramme einer Montage-Linie für Hinterachsen mit einem Gesamtumfang von etwa 115000 Zeilen Quellcode insgesamt 78% des gesamten Quellcodes ausmachten, können beim derzeitigen Stand der Technik nur an fertig aufgebauten Anlagen entwickelt und erprobt werden.

## 4.1 Implementierung mit TruPhysics

Ein Testfall wird bei TruPhysics durch ein Ablaufdiagramm formuliert. Dieses Diagramm enthält bestimmte Bedingungen, die mit ihrem Eintreten über Erfolg oder Misserfolg eines Test-Durchlaufs entscheiden.

Diese Bedingungen können sich auf unterschiedliche Eigenschaften von einzelnen Objekten, oder generell auf den Zustand der Simulation beziehen. So kann etwa überwacht werden, ob ein einzelnes Werkstück nach einer bestimmten Laufzeit eine Zielposition erreicht oder nicht. Auch kann als Bedingung festgelegt werden, dass ein Verbund aus Objekten (etwa ein Roboterarm und das daran angebrachte Werkzeug) während des gesamten Ablaufs mit keinem anderen Objekt in der Simulation (außer einem Werkstück) in Berührung kommen darf.

In einem Szenario, in dem ein Objekt mit einem Roboterarm gegriffen und an einem bestimmten Ort abgelegt werden soll, könnte beispielsweise das Vorhandensein des Objektes am Zielort als Erfolg gewertet werden. Im selben Szenario könnte eine Kollision des Roboterarms mit einem beliebigen Teil der Szene als Misserfolg gewertet werden.

Neben Positionen und Kontaktbedingungen können auch weitere mechanische Eigenschaften (Geschwindigkeit, anliegende Kräfte) als Kriterien für die Bewertung von Testbedingungen herangezogen werden. Das ermöglicht eine sehr flexible und vielseitige Modellierung von Testfällen anhand von Daten, die aus dem Simulationskern abgerufen werden können.

Als Teil der Testumgebung fungieren zudem sogenannte „virtuelle Sensoren“. Dies sind in der Simulation vorhandene Objekte, die nicht der simulierten Physik unterliegen, und die für andere simulierte Objekte durchlässig sind. Bei Berührung oder Überlappung mit anderen Objekten wird ein entsprechendes Kollisionsereignis ausgelöst, allerdings ohne den Bewegungsablauf der anderen Objekte zu ändern.

Das parallel zur Simulation ablaufende Softwaremodul zum teil-automatisierten Test überwacht den Ablauf der Simulation. Dabei werden alle auftretenden Kollisionsereignisse erfasst: Sowohl Kollisionen regulärer Geometrien miteinander, als auch solche, die durch Kollisionen mit virtuellen Sensoren entstehen. Neben der Überwachung von Kollisionsereignissen sind weitere Ereignisse als Kriterien für Testdurchläufe möglich. Beispiele dafür sind etwa das Eintreten bestimmter Messwerte bei simulierten Sensoren, das Erreichen einer bestimmten Position durch überwachte Objekte, oder die Erzeugung eines bestimmten Statussignals durch ein an die Simulation angebundenes Steuergerät.

Die überwachten Werte werden dann solange mit dem Ablaufdiagramm eines Testfalls abgeglichen, bis entweder ein "Erfolg" oder "Misserfolg"-Endzustand erreicht wird.

## 4.2 Beispielszenario

Dieses Beispiel zeigt das Zusammenspiel zwischen virtueller Umgebung, der realen Robotersteuerung und dem darauf laufenden Programm.

Abbildung 3 zeigt exemplarisch eine Komponente der Testumgebung, das Verhaltensmodell eines simulierten Reflextasters (Roßdeutscher 2011). Ein Reflextaster ist ein Sensor, der Licht oder Schall aussendet und die entsprechende Reflektion

auswertet. Reflextaster werden beispielsweise verwendet, um das Vorhandensein eines Objekts auf einem Fließband oder Füllhöhen in einem Behälter zu bestimmen. Das Verhaltensmodell modelliert das Verhalten des Sensors während der Simulation. Zusammen mit anderen Bausteinen kann es als Basis für einen Testfall dienen.

Der Betrieb des Reflextasters wird dabei so unterteilt, dass er sich in einem von zwei Zuständen befindet. Der Reflextaster befindet sich im Zustand „Keine Störung“, solange er korrekte Ergebnisse liefert. Dieser Zustand wird noch einmal unterteilt in drei weitere Zustände: Es wird zwischen den Zuständen „Abstand_zu_klein“, „Innerhalb_Messbereich“ und „Abstand_zu_gross“ unterschieden. Im Zustand „Innerhalb_Messbereich“ erfasst der Reflextaster ein Objekt innerhalb seines Messbereichs, und er liefert die dazu passende Ausgabe zurück. In den beiden anderen Zuständen befindet sich ebenfalls ein Objekt innerhalb des Messbereichs, aber es ist zu nah, bzw. zu weit weg um (gemäß der Spezifikation des Sensors) korrekt erfasst zu werden.

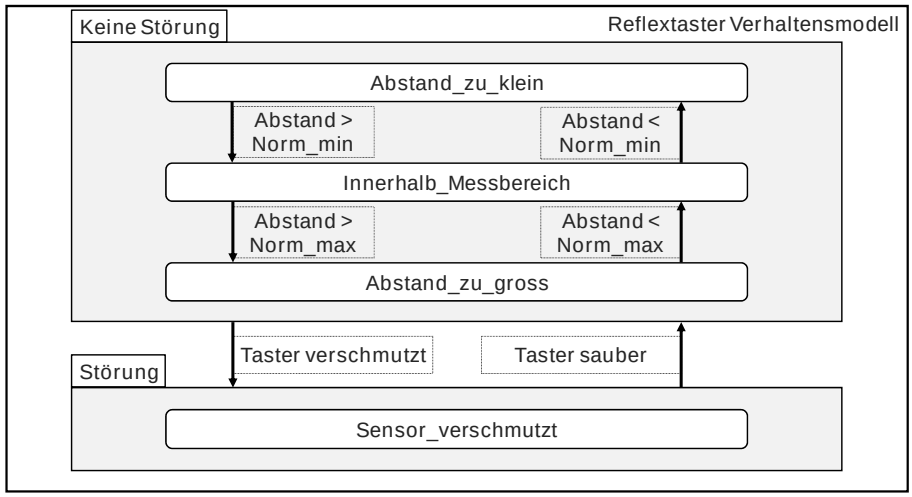



*Abbildung 3: Verhaltensmodell für einen Reflextaster*

Zu dem Zustand „Keine Störung“ kommt ein exemplarischer Zustand „Störung“ hinzu, der eine Störung des Sensors modelliert. In einem realen Testablauf können hier verschiedenste Fehlerszenarien erfasst werden. In diesem Beispiel wird eine Verschmutzung des Sensors modelliert, beispielsweise indem die simulierten Messergebnisse durch zufällige Abweichungen verfälscht werden. Der Zustand „Störung“ kann während eines Testablaufs hinzugeschaltet werden. Während der Simulation werden dann die verfälschten Messergebnisse an die angeschlossene Robotersteuerung übermittelt. Auf diese Weise kann getestet werden, ob das auf der Steuerung ausgeführte Programm die inkorrekte Funktion des Sensors erkennen und korrekt damit umgehen kann.

## 5 Zusammenfassung

Dieser Artikel beschreibt das Simulationswerkzeug TruPhysics und wie es benutzt werden kann, um teil-automatisierte Softwaretests durchzuführen. TruPhysics ist ein Werkzeug, welches einen Industrieroboter simuliert und Steuersignale einer Robotersteuerung entgegennimmt. Die Simulation ist dabei physikalisch plausibel und läuft in Echtzeit. Dies wird ermöglicht durch die Verwendung eines neu entwickelten Verfahrens zur Kollisionserkennung, welches sich die Rechenleistung von GPUs zunutze macht. Testfälle werden als Ablaufdiagramme formuliert, deren Zustandsübergänge unter anderem durch Kollisionen simulierter Objekte ausgelöst werden. Zusätzlich können in einer simulierten Szene „virtuelle Sensoren“ platziert werden, die für andere Objekte durchlässig sind, und deren Kollisionen auch Zustandsübergänge auslösen können.

## Literaturverzeichnis

ABB, 2015: RobotStudio: ABB Automation. Online verfügbar unter http://new.abb.com/products/robotics/de/robotstudio.

Aichele, F.: Kollisionserkennung für echtzeitfähige Starrkörpersimulationen in der Industrie- und Service-Robotik. Dissertation Universität Stuttgart, 2015.

CM Labs, 2015: Vortex: CM Labs Simulations. Online verfügbar unter http://www.cm-labs.com/market/robotics/products/vortex-dynamics-software.

Coumans, E., 2015: Bullet Physics. Online verfügbar unter http://bulletphysics.org.

Dassault Systèmes, 2015: 3DS Robotics on premise: Dassault Systèmes. Online verfügbar unter http://www.3ds.com/products-services/3dexperience/on-premise/robotics-on-premise/robot-simulation/.

Eckstein, J.: Echtzeitfähige Kollisionserkennung für Virtual reality Anwendungen. München: Utz Wiss 1999.

Erbes, R.; Mantel, A.; Schömer, E.; Wolpert, N.: Parallel collision queries on the GPU. In: Keller, R.; Kramer, D.; Weiss, J.-P. (Hrsg.): Facing the Multicore-Challenge III: Berlin Heidelberg: Springer 2013, S. 84–95.

Erez, T.; Tassa, Y.; Todorov, E.: Simulation tools for model-based robotics: Comparison of Bullet, Havok, MuJoCo, ODE and PhysX. In: 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, S. 4397–4404.

Erleben, K.: Stable, robust, and versatile multibody dynamics animation. Copenhagen, University of Copenhagen (DIKU), Department of Computer Science, Ph.D. thesis, 2005.

Georgia Tech, 2015: Dynamic animation and robotics toolkit: Georgia Tech Graphics Lab and Humanoid Robotics Lab. Online verfügbar unter http://dartsim.github.io/.

Hengel, K.: Softwareentwicklung für speicherprogrammierbare Steuerungen im integrierten, rechnergestützten Konstruktionsprozeß. Berlin, Heidelberg: Springer 1994.

Intel, 2015: Havok: Intel. Online verfügbar unter http://www.havok.com/physics/.

ISG, 2015: ISG-virtuos: ISG Industrielle Steuerungstechnik. Online verfügbar unter http://www.isg-stuttgart.de/de/isg-virtuos/virtuos.html.

Kohring, A.: Systematisches Projektieren und Testen von Steuerungssoftware für Werkzeugmaschinen. Aachen: Shaker 1993.

KUKA, 2015: KUKA.Sim Pro: KUKA Roboter. Online verfügbar unter http://www.kuka-robotics.com/de/products/software/simulation/.

Lauterbach, C.; Mo, Q.; Manocha, D.: gProximity. Computer Graphics Forum 29 (2010) 2, S. 419–428.

Liggesmeyer, P.: Software-Qualität. Heidelberg: Spektrum Akademischer Verlag 2009.

NVIDIA, 2015: PhysX: NVIDIA Corporation. Online verfügbar unter http://www.nvidia.de/object/nvidia-physx-de.html.

Roßdeutscher, M.: Entwicklung eines Verfahrens zum Programmtest in der robotergestützten Montage. Aachen: Shaker 2011.

Sherman, M.; Eastman, P., 2015: Simbody. Online verfügbar unter https://simtk.org/home/simbody/.

Siemens, 2015a: NX Mechatronics Concept Designer: Siemens PLM Software. Online verfügbar unter http://www.siemens.com/nx.

Siemens, 2015b: Tecnomatix RobotExpert: Siemens PLM Software. Online verfügbar unter www.siemens.com/tecnomatix.

Smith, R., 2015: Open Dynamics Engine. Online verfügbar unter http://www.ode.org/.

Sossenheimer, K.: Entwickeln von Instrumentarien zur rationellen Planung und Steuerung der Inbetriebnahme komplexer Produkte des Werkzeugmaschinenbaus. Dissertation, RWTH Aachen 1989.

Strassacker, D.: Testumgebung für die Implementierung und Inbetriebnahme eines adaptierbaren Leitsteuerungssystems. Berlin, Heidelberg: Springer 1997.

Todorov, E., 2015: MuJoCo: Roboti LLC. Online verfügbar unter http://www.mujoco.org/.

Visual Components, 2015: 3DRealize R: Visual Components. Online verfügbar unter http://www.visualcomponents.com/products/3drealizer/.

Weller, R.: New geometric data structures for collision detection and haptics. Cham, Switzerland, New York: Springer 2013.

Wünsch, G., 2015: industrialPhysics: machineering. Online verfügbar unter http://www.machineering.de/.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Verteilte Simulation und Emulation von Materialflusssystemen mit dezentraler Steuerung

## *Distributed Simulation and Emulation of Material Flow Systems with Decentralised Control*

Damian Daniluk, Michael ten Hompel, Fraunhofer-Institut für Materialfluss und Logistik IML, Dortmund (Germany), damian.daniluk@iml.fraunhofer.de, michael.ten.hompel@iml.fraunhofer.de

Abstract: In this paper a framework is presented which allows the distributed simulation and emulation of material flow systems with decentralised control using a commercial event-driven and field-proven material flow simulator. The term framework describes a system of interacting software components that support distributed simulation and emulation, without prescribing the details of the simulation model and its structure. In this paper the basic elements for framework implementation are motivated and described. These elements comprise reference algorithms for decentralised control of material flow systems, an optimised synchronisation algorithm, a partitioning methodology and a communication infrastructure for distributed simulation. Based on the framework it is shown, which factors have main influence on the achievable speedup.

## 1 Einführung

Seit einiger Zeit stellt die Dezentralisierung von Materialflusssteuerungen einen vitalen Forschungsgegenstand dar, der oft unter dem Begriff *Internet der Dinge* genannt wird (Günthner und ten Hompel 2010). Ein Ansatz zur Realisierung dezentraler Materialflusssteuerungen stellen Multiagentensysteme (MAS) dar (Jennings und Bussmann 2003). Zur Erforschung der Funktions- und Leistungsfähigkeit einer dezentralen, multiagentenbasierten Steuerung von Materialflusssystemen ist aufgrund der hohen Systemkomplexität der Einsatz der computergestützten Simulation unverzichtbar (Daniluk und Chisu 2010). Den Stand der Technik zur Modellierung von intralogistischen Abläufen bilden leistungsfähige kommerzielle Simulationswerkzeuge. Die *verteilte Simulation*, d. h. die nebenläufige Ausführung eines Simulationsmodells auf vernetzten Arbeitsplatzrechnern, bietet sich prinzipiell als Ausweg an, um die Ausführung aufwändiger Materialflussmodelle zu beschleunigen und die Leistungsfähigkeit der Hardware mehrerer Rechner für die Simulationsausführung zu bündeln (Lendermann 2006). Im Hinblick auf die Integration von dezentralen Steuerungsmechanismen in Simulationsmodelle

von ereignisgesteuerten Materialflusssimulatoren existieren noch keine Untersuchungen zum Einsatz der verteilten Simulation und deren Leistungsfähigkeit in diesem Anwendungskontext.

Ebenso fehlen Untersuchungen über Voraussetzungen und Realisierungs-möglichkeiten einer beschleunigten Ausführung von Emulationen unter Betrachtung von multiagentenbasierten Steuerungssystemen. Bei der *Emulation* erfolgt eine Kopplung der dezentralen Steuerungssoftware, die auch bei dem realen Betrieb einer Anlage zum Einsatz kommt, an den Materialflusssimulator. Der besondere Vorteil der Emulation liegt darin begründet, dass das Steuerungssystem als Anwendung ohne weitere Veränderungen für den Betrieb in einer realen Anlage übernommen werden kann. Der Vorteil der Simulation liegt dem gegenüber in der Reproduzierbarkeit von Simulationsläufen, die bei der Emulation nicht gegeben ist (Daniluk 2012).

Wesentliche Fragestellungen betreffen die mit verteilter Simulation gegenüber der Simulation auf einem Rechner erzielbare Beschleunigung (auch: *Speedup*) der Simulationsausführung sowie die Realisierungsaufwände für eine solche verteilte Ausführung.

Im vorliegenden Beitrag wird ein Framework vorgestellt, das die verteilte Simulation und Emulation von dezentral gesteuerten Materialflusssystemen unter Verwendung eines kommerziellen ereignisgesteuerten und in der Praxis bewährten Materialflusssimulators erlaubt. Der Beitrag bildet eine komprimierte Übersicht über das in der gleichnamigen Dissertation von Daniluk (2014) entwickelte Framework und die ermittelten Ergebnisse.

## 2 Framework für verteilte Simulation und Emulation von multiagentenbasierten Materialflusssystemen

Die Vorgehensweise zur Beantwortung der Forschungsfragen lässt sich in drei Themenfelder gliedern, in denen verschiedene Lösungselemente erarbeitet wurden, welche dann die Implementierung des Frameworks sowie die Durchführung der empirischen Studie, d. h. die Evaluation erlaubten. Diese drei Themenfelder sind:

- *Dezentrale Steuerung:* Das Ziel liegt hier in der Entwicklung konkreter Verfahren für die dezentrale Materialflusssteuerung, die im Rahmen der Evaluation genutzt werden.
- *Verteilte Simulation:* Es sind effiziente Verfahren für die Partitionierung und Synchronisation von Simulationsmodellen erforderlich. Darüber hinaus wird eine effiziente Kommunikationsinfrastruktur zum Austausch der Informationen zwischen den an der verteilten Simulation beteiligten Rechnern benötigt.
- *Beschleunigte Emulation:* Die o. g. Kommunikationsinfrastruktur ist auch für die beschleunigte Emulation wichtig, da zum Austausch von Informationen zwischen Steuerungssystem und der Anlagensimulation auch ein Kommunikationskonzept benötigt wird. Auch bei der Emulation ist ein Synchronisationsverfahren erforderlich, das bei beschleunigter Ausführung der Anlagensimulation eine korrekte Interpretation der Ergebnisse der Emulation erlaubt.

In Abbildung 1 sind die Vorgehensweise zur Durchführung der Evaluation sowie die Lösungselemente dargestellt.

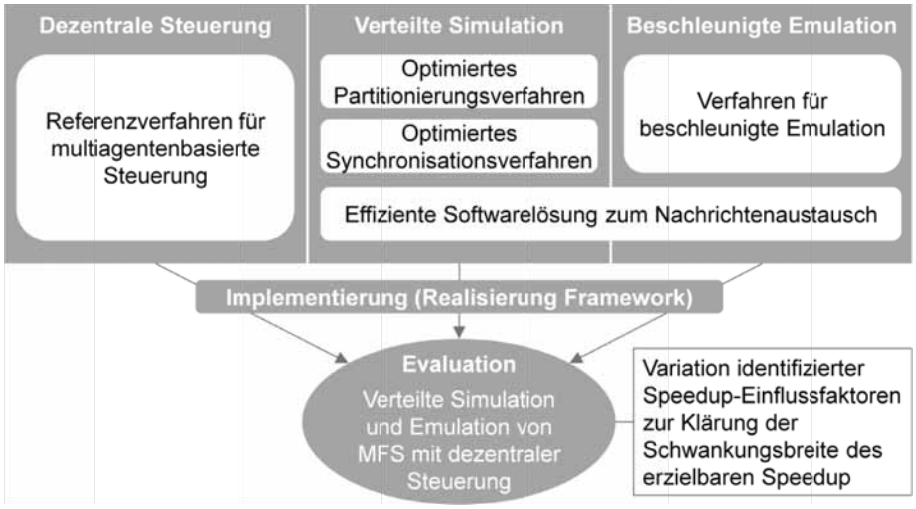



*Abbildung 1: Lösungselemente und Vorgehensweise zur Durchführung der Evaluation*

Das entwickelte Framework für verteilte Simulation und Emulation besteht aus mehreren Software-Komponenten, die in Abbildung 2 dargestellt sind. Ausgangspunkt ist ein Simulationsmodell ohne Verhaltenslogik. Im ersten Schritt wird die Modelldatei durch das Werkzeug eingelesen und geparst. Dabei werden die einzelnen Strukturelemente des Simulationsmodells, d. h. Quellen, Senken, Verzweigungen, Zusammenführungen usw. identifiziert. Danach ist die Topologie des Simulationsmodells bekannt, so dass auf dieser Basis ein agentenbasiertes Routing stattfinden kann. Die Hauptaufgabe des Werkzeugs besteht darin, das Simulationsmodell zu partitionieren und mit Agentenlogik anzureichern.

Im unteren Bereich von Abbildung 2 sind das Steuerungssystem, die Kommunikationsinfrastruktur sowie der Materialflusssimulator dargestellt. Diese Komponenten kommen bei der Simulations- bzw. Emulationsausführung zum Einsatz. Das Steuerungssystem mit der multiagentenbasierten Steuerung wird ausschließlich im Fall der Emulation eingesetzt. Immer zum Einsatz gelangt der Materialflusssimulator, der das Modell im Fall der verteilten Simulation inkl. multiagentenbasierter Steuerungslogik ausführt und im Fall der Emulation das technische System ohne Steuerungslogik simuliert.

Eine zentrale Komponente bildet die Kommunikationsinfrastruktur, welche die Vermittlung der Nachrichten zwischen den Teilmodellen und im Fall der Emulation auch zwischen dem Steuerungssystem sicherstellt. Darüber hinaus sammelt die Kommunikationsinfrastruktur statistische Daten, wie die Anzahl ausgetauschter Ereignis- und Synchronisationsnachrichten. Auch der Materialflusssimulator generiert Logdateien, in welchen z. B. Länge der Simulationsdauer des Teilmodells gespeichert wird. Die Logdateien von Kommunikationsinfrastruktur und Materialflusssimulator werden mittels eines Skripts in einer zentralen Datei zusammengeführt. Diese Datei enthält alle für die Evaluation erforderlichen Daten und Statistiken. Bestandteil des Frameworks ist ein weiteres Skript, welches zum Starten eines Simulations- oder Emulationslaufs alle Teilmodelle auf die verfügbaren

Rechnerknoten kopiert und die erforderlichen Komponenten (Steuerungssystem, Kommunikationsinfrastruktur, Materialflusssimulator) startet.

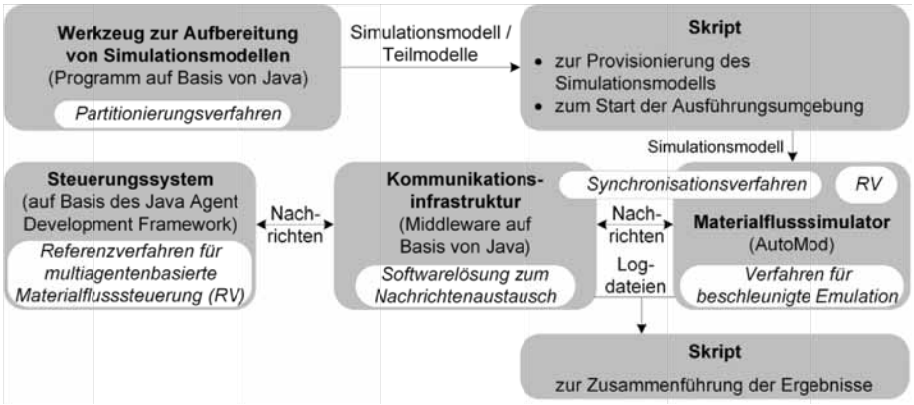



*Abbildung 2: Grobarchitektur und Einordnung der Lösungselemente in das Framework für verteilte Simulation und Emulation*

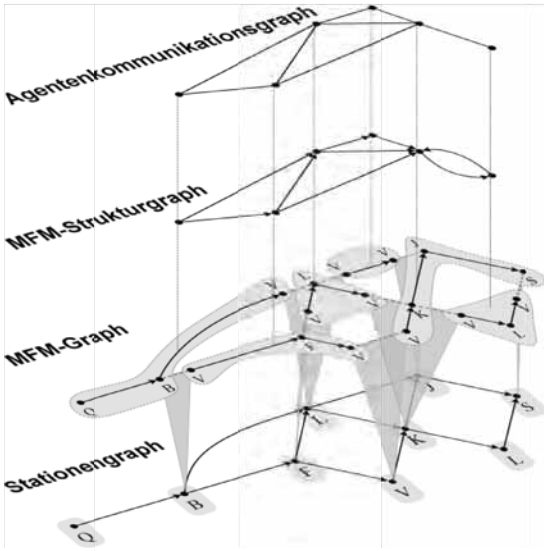
In Abbildung 2 sind den einzelnen Komponenten die getätigten neuen Entwicklungen bzw. Lösungselemente zugeordnet, wie sie bereits in Abbildung 1 im Kontext der grundsätzlichen Vorgehensweise zur Durchführung der Evaluation benannt wurden. Auf die einzelnen Lösungselemente wird in den folgenden Abschnitten näher eingegangen.

## 2.1 Referenzverfahren für die multiagentenbasierte Materialflusssteuerung

Ein Materialflusssystem kann als ein Netzwerk von Quellen, Senken, (Bearbeitungs-)Stationen und Lagermitteln angesehen werden, die durch die Verbindungselemente miteinander verknüpft sind (Gudehus 2005). Die *Stationen* eines Materiaflusssystems umfassen *Elementarstationen* und *Transportelemente*. Zu den Elementarstationen gehören alle Ursprungspunkte (Quellen), an denen die Arbeitsgegenstände, im Folgenden auch *Lastobjekte* genannt, in das Simulationsmodell eingesteuert werden sowie alle Zielpunkte (Senken, Bearbeitungsstationen, Lagermittel), die von den Lastobjekten als (Zwischen-)Ziele adressiert werden können. Die Transportelemente (Verbindungselemente, Verzweigungen, Zusammenführungen, Kreuzungen) ergeben sich durch die Verknüpfung von Förderstrecken.

Die im vorliegenden Beitrag betrachteten Simulationsmodelle setzen sich aus einer Menge von Materialflussmodulen (MFM) zusammen, die ein Stetigfördersystem abbilden. Die einzelnen MFM agieren dabei als Agenten. Die betrachtete multiagentenbasierte Steuerung fokussiert auf das Routing Lastobjekte von einer Quelle zu einem Ziel. Die Zielinformationen erhält der mit einem MFM assoziierte Agent von dem beförderten Lastobjekt selbst. Dazu speichert das Lastobjekt sein Bestimmungsziel z. B. auf einem RFID-Chip, welcher über entsprechende Lesegeräte am MFM ausgelesen werden kann.

Die in Abbildung 3 grau gestrichelten Linien zwischen den MFM im *MFM-Graphen* geben eine mögliche Ausprägung von Verknüpfungen zwischen den MFM an und beschreiben damit eine mögliche Struktur eines aus MFM zusammengesetzten Materialflusssystems. Legt man diese Strukturinformationen zu Grunde, so ist eine Überführung des MFM-Graphen in einen *Stationengraphen* möglich, auf dessen Basis das Routing der Lastobjekte erfolgen kann. Ebenso können der *MFM-Strukturgraph*, welcher die Flussbeziehungen zwischen den MFM wiedergibt und der *Agentenkommunikationsgraph*, der dem Datennetzwerk für das Routing entspricht, abgeleitet werden.




| | Station | Symbol |
|---|---|---|
| Elementarstation | Quelle | Q |
| | Senke | S |
| | Bearbeitungsstation | B |
| | Lagermittel | L |
| Transportelement | Verbindungselement | V |
| | Verzweigung (Weiche) | F |
| | Zusammenführung (Einmündung) | J |
| | Kreuzung | K |

*Abbildung 3: Ableitung verschiedener Graphendarstellungen des im Simulationsmodell betrachteten Materialflusssystems*

Für die Evaluation wurden auf Basis des aus Datennetzwerken bekannten Link-Statte Routing (Leon-Garcia und Widjaja 2000, S. 594 ff.) und des Dynamic Source Routing (Johnson und Maltz 1996) zwei in ihrer Funktionsweise verschiedene Algorithmen als Referenzverfahren entwickelt, die sich insbesondere in der zwischen den Agenten generierten Kommunikationslast signifikant voneinander unterscheiden. Diese Algorithmen spannen damit ein breites Spektrum im Hinblick auf den erzielbaren Speedup bei der verteilten Simulation auf.

## 2.2 Synchronisationsverfahren

Eine unabdingbare Anforderung an die Ausführung eines verteilten Simulationsmodells ist, dass die verteilte Simulation die gleichen Ergebnisse liefert wie die sequenzielle Simulation des entsprechenden sequenziellen Simulationsmodells. Hierzu ist eine Synchronisation der Teilmodelle erforderlich.

Eines der ersten entwickelten Synchronisationsverfahren wird in der Literatur nach den unabhängigen Veröffentlichungen der Autoren Chandy und Misra sowie Bryant als CMB-Verfahren bezeichnet (Chandy und Misra 1979). Die meisten existierenden

konservativen Synchronisationsverfahren basieren auf diesem Verfahren. Für den Einsatz im Framework wurde ein optimiertes Verfahren entwickelt, das ebenfalls auf dem CMB-Verfahren basiert. Dabei werden zur Verringerung der Anzahl im Rahmen der Synchronisation auszutauschender *Nullnachrichten* analog zu dem Vorschlag der „Loop Optimization“ in (Porras et al. 1997) pro Iteration der Simulationsschleife gleich mehrere Nachrichten empfangen und anschließend verarbeitet. Zur Vermeidung der Versendung vieler Zyklen von Nullnachrichten werden ähnlich wie in (Wood und Turner 1994) *erweiterte Nullnachrichten* verwendet.

Bei der Emulation ist zusätzlich eine Synchronisation zwischen Steuerungssystem und dem Simulator des technischen Systems notwendig, falls der Simulator schneller als Realzeit simuliert, da die zeitlichen Abweichungen der Simulationsereignisse gegenüber dem Realzeitverlauf zu Fehlern in der Steuerung führen können. Für die Evaluation wurde ein für das eingesetzte Steuerungssystem zugeschnittenes Synchronisationsverfahren verwendet, bei dem die Ausführungsgeschwindigkeit der Simulation abhängig vom Kommunikationszustand von Steuerungssystem und Simulator zwischen Realzeit und „schnellstmöglich“ variiert wird (Daniluk 2014).

## 2.3 Partitionierungsverfahren

Neben der Synchronisation gehört die Partitionierung zu den wesentlichen Aufgaben der verteilten Simulation. Die Partitionierung eines Simulationsmodells bestimmt die Zerlegung des Modells in möglichst unabhängige Teilmodelle zur Verteilung des Berechnungsaufwandes des Gesamtmodells auf mehrere Rechner. Die allgemeine Herausforderung bei der Partitionierung besteht darin, die prinzipiell parallel ausführbaren Teile des Simulationsmodells zu identifizieren und derart auf Partitionen aufzuteilen, dass (Mehl 1994)

- die *Kommunikationslast* zwischen den Teilmodellen *minimiert* und
- ein *ausgeglichenes Lastverhalten* der Teilmodelle erzielt werden.

Typischerweise wird dazu das Simulationsmodell als Graph abgebildet (Fujimoto 1998), wobei die Knoten eine durch ein Simulationsobjekt generierte Rechenlast und die Kanten die Kommunikationslast zwischen den Simulationsobjekten repräsentieren. Die Minimierung der Kommunikationslast entspricht dann der Minimierung der Schnittkosten der Partitionierung des Graphen. Die Herstellung eines ausgeglichenen Lastverhaltens kann durch die Ausgewogenheit der Graphenpartitionierung erzielt werden. Im Rahmen der Entwicklung des Frameworks wurde dieser Ansatz weiter optimiert, indem zur Minimierung der Kommunikationslast zusätzlich auch die Minimierung der Konnektivität des Teilmodell-Graphen, der die Kommunikationsverbindungen zwischen den Teilmodellen repräsentieren, berücksichtigt wurde. Darüber hinaus wurde bei der Maximierung der Güte der Lastverteilung nicht nur ein räumlicher Lastausgleich, sondern auch ein zeitlicher Lastausgleich angestrebt. Der räumliche Lastausgleich entspricht einer gleichmäßigen Verteilung der kumulierten von den Teilmodellen während eines Simulationslaufs generierten Rechenlast. Bei dem zeitlichen Lastausgleich wird zusätzlich gefordert, dass die Teilmodelle zu jedem Zeitpunkt des verteilten Simulationslaufs eine vergleichbare Last generieren, wodurch erst die eigentliche parallele Verarbeitung erzielt wird.

Mit METIS (Karypis 2011) konnte eine existierende Software-Bibliothek gefunden werden, welche einen Algorithmus anbietet, der so genutzt werden kann, dass sich eine Partitionierung gemäß der entwickelten Zielfunktion realisieren lässt.

Da das Partitionierungsergebnis nur dann nutzbare Ergebnisse liefern kann, wenn die Parameter für das jeweilige Partitionierungsverfahren hinreichend realistische Daten darstellen, liegt die zweite wesentliche Fragestellung bei der Partitionierung in der Ermittlung der Eingangsdaten für das Partitionierungsverfahren. Dazu wurden eine *statische Partitionierung*, bei der die Gewichtung des Partitionierungsgraphen auf Basis von Strukturinformationen des Simulationsmodells erfolgt, und eine *statische verhaltensbasierte Partitionierung*, bei der zusätzlich Informationen über die Lastdaten und die Ausfälle von MFM im Simulationsverlauf berücksichtigt werden, entwickelt.

### 2.4 Nachrichtenbasierte Middleware zur Kommunikation

Damit die an der verteilten Simulation und Emulation beteiligten Kommunikationsteilnehmer Daten untereinander austauschen können, ist eine entsprechende Kommunikationsinfrastruktur bereitzustellen. Im Framework erfolgt die Kommunikation über eine *Middleware*. Diese besteht aus einzelnen Middleware- Prozessen, die auf jedem Rechner gestartet werden, auf dem ein Teilmodell bzw. im Fall der Emulation das Steuerungssystem als Kommunikationsteilnehmer ausgeführt wird. Die einzelnen Middleware-Prozesse sind für den Verbindungsaufbau, die Durchführung der Datenübertragung zwischen den einzelnen Rechnern sowie den Verbindungsabbau verantwortlich. Darüber hinaus enthalten sie ein Modul, welches die Zählung von Ereignis- und Synchronisationsnachrichten pro Kommunikationsverbindung erlaubt und eine Ausgabe der entsprechenden Statistiken realisiert. Zudem enthält ein Middleware-Prozess einen Teil der Logik zur Synchronisation der Kommunikationsteilnehmer, um deren Anpassungsaufwände für eine verteilte Simulation bzw. Emulation zu reduzieren.

## 3 Evaluation und Ergebnisse

Übergeordnetes Ziel der Evaluation ist es, Erkenntnisse über den erreichbaren Speedup von verteilten Materialflussmodellen mit dezentraler Steuerung zu erhalten. Eine Testumgebung erlaubt es, gezielt den Einfluss bestimmter Faktoren zu untersuchen, die die Ausführungszeit der verteilten Simulation bzw. Emulation beeinflussen. Dies erfolgt dadurch, dass bei den Tests bestimmte Modelleigenschaften variiert werden, während die übrigen Modelleigenschaften unverändert bleiben.

Davon ausgehend wurden die Testdaten in Form der bei der Evaluation zu verwendenden Referenzmodelle sowie deren Eingabedaten definiert. Die Topologie der Referenzmodelle entspricht zur Wahrung des Realitätsbezugs der Struktur eines in der Industrie umgesetzten Materialflusssystems aus dem Bereich der Gepäckförderung. Es wurden drei unterschiedliche Topologien mit 716, 1.663 und 3.074 Agenten untersucht. Die mit den Agenten assoziierten MFM entsprachen dabei einzelnen Fördertechnikabschnitten (so genannten *Sections*) im verwendeten Materialflusssimulator *AutoMod*. Insgesamt wurden 28 Modellklassen erstellt. Zu einer Modellklasse gehört ein Simulationsmodell mit einer konkreten Ausprägung von Modelleigenschaften. Da die Testumgebung eine Verteilung des Simulationsmodells

auf bis zu fünf Teilmodelle erlaubt, die jeweils auf einem eigenen Rechner ausgeführt werden, erfolgten 5 Testläufe pro Modellklasse.

Die erzielten Ergebnisse lassen sich aufgrund der vielen Einflussfaktoren der verteilten Simulation nicht direkt auf andere Simulationsumgebungen übertragen, zeigen aber deutlich die Spannweite des erzielbaren Speedup bei der verteilten Simulation multiagentenbasierter Materialflusssysteme auf. Sie reicht von einer deutlichen Reduzierung der Ausführungsgeschwindigkeit gegenüber der sequenziellen Simulation bis zu einer mehrfachen Beschleunigung.

Im ungünstigsten Fall konnte eine 10-fache Verlangsamung der verteilten Simulation auf fünf Rechnerknoten gegenüber der sequenziellen Simulation auf einem Rechner beobachtet werden. Dies ist hauptsächlich auf die hohe kausale Abhängigkeit zwischen den in den unterschiedlichen Teilmodellen auftretenden Ereignissen dieser Modellklasse, einen damit einhergehend niedrigen *Lookahead* (ein Lookahead-Wert $l$ bezeichnet das Wissen zum Simulationszeitpunkt $t$ darüber, dass ein Teilmodell in Zukunft nur Ereignisnachrichten verschicken wird, die mindestens einen Zeitstempel von $t + l$ haben) und die dadurch implizierte hohe Kommunikationslast zurückzuführen. Pro ausgetauschter Ereignisnachricht wurden bei der o. g. verteilten Simulation durchschnittlich 182 Synchronisationsnachrichten ausgetauscht. Dies macht den starken Kommunikationsoverhead der Simulation deutlich.

Im günstigsten Fall konnte ein Speedup von 3,2 bei der verteilten Simulation auf 5 Rechnern erzielt werden. Die Eigenschaften der dazugehörigen Modellklasse waren dabei durch eine relativ hohe Ereignisgranularität, die Auswahl eines Routingverfahrens, welches einen relativ hohen Lookahead zulässt und die Nutzung des verhaltensbasierten Partitionierungsverfahrens geprägt.

Generell führt eine steigende Anzahl von Teilmodellen zu einer Erhöhung der Kommunikationslast. Bei den für die Experimente herangezogenen Modellklassen war häufig bei fünf an der verteilten Simulation beteiligten Teilmodellen ein Abfall des Speedup oder zumindest kein Anstieg des Speedup mehr zu beobachten. Daraus lässt sich schlussfolgern, dass unter Verwendung der aktuellen State of the Art-Verfahren eine weitere Verteilung des Simulationsmodells in der betrachteten Anwendungsdomäne nicht sinnvoll ist.

Bezogen auf die einzelnen Untersuchungsgegenstände, konnten folgende Beobachtungen gemacht werden:

- *Synchronisationsverfahren*: Das optimierte Synchronisationsverfahren führte in den Simulationsexperimenten zu einem besseren Speedup als das CMB-Verfahren. Insbesondere bei Modellen mit kleinem Lookahead konnte durch Einsatz des optimierten Synchronisationsverfahrens ein um den Faktor 2,7 höherer Speedup erzielt werden.
- *Ereignisgranularität*: Eine erhöhte Ereignisgranularität wirkt sich generell positiv auf den erzielbaren Speedup aus.
- *Partitionierungsverfahren*: Die statisch verhaltensbasierte Partitionierung bietet vor allem bei Simulationsmodellen Vorteile hinsichtlich des erzielbaren Speedup, bei denen die Lastverteilung im Modell während der Simulationsausführung ungleichmäßig ist.
- *Steuerungsverfahren*: Das verwendete Routingverfahren hat wesentlichen Einfluss auf den verfügbaren Lookahead und damit auf den erzielbaren Speedup.

- *Strukturelle Komplexität*: Bei Simulationsmodellen mit einer hohen Anzahl statischer Elemente (Anzahl der MFM) ist tendenziell die durchschnittliche Anzahl gleichzeig aktiver Lastobjekte pro Teilmodell und somit die darin generierte Rechenlast höher, was einen ähnlichen Effekt wie eine erhöhte Ereignisgranularität zur Folge und damit eine positive Wirkung auf den erzielbaren Speedup hat.
- *Externe dynamische Komplexität*: Kurze Zeitabstände zwischen der Einsteuerung von Lastobjekten führen ebenfalls zu einer durchschnittlich hohen Anzahl gleichzeitig im Simulationsmodell aktiver Lastobjekte und haben damit einen positiven Effekt auf den erzielbaren Speedup (s. o.).
- *Emulation*: Bei der Emulation können insbesondere dann gute Speedups erreicht werden, wenn die Verlagerung von Steuerungslogik in das Steuerungssystem auf Simulationsseite zu einem höheren Lookahead führt.

Aus den im Rahmen der Evaluation gemachten Beobachtungen lässt sich für ein multiagentenbasiertes Materialflussmodell mit den folgenden Eigenschaften in der Praxis ein guter Speedup prognostizieren:

- Hohe Ereignisgranularität (numerischer Aufwand pro Ereignis)
- Hohe Anzahl von statischen Elementen (Anzahl der Materialflussmodule bzw. Fördertechnikabschnitte, welche jeweils als Agent modelliert werden)
- Hohe dynamische Komplexität (Anzahl der gleichzeitig im Simulationsmodell aktiven Lastobjekte)
- Hoher Lookahead

Je mehr dieser Eigenschaften ein Simulationsmodell erfüllt, desto höher ist die Wahrscheinlichkeit, einen guten Speedup zu erzielen. Im Umkehrschluss sind Modelle, welche wenige oder keine der genannten Eigenschaften erfüllen, eher ungünstig. In der Praxis können diese Eigenschaften herangezogen werden, um das Beschleunigungspotenzial eines bestehenden oder zu entwerfenden Simulationsmodells mittels verteilter Simulation zu bewerten.

## 4 Fazit

Insgesamt zeigen die Ergebnisse, dass die Geschwindigkeitsvorteile einer verteilten Simulation von multiagentenbasierten Materialflusssystemen unter bestimmten Bedingungen auftreten, aber nicht für alle bzw. einen Großteil der Simulationsmodelle vorliegen. Sind häufig Versuchsreihen mit Simulationsmodellen durchzuführen, deren Eigenschaften einen potenziell hohen Speedup versprechen, kann die Schaffung eines entsprechenden Frameworks für die verteilte Simulation und Emulation lohnenswert sein. Nachdem die Hürde der erheblichen Aufwände zur Entwicklung eines solchen Frameworks genommen worden ist, können verteilte Simulationsexperimente aufgrund der weitgehenden Automatisierung mit geringem Aufwand durchgeführt werden.

## Danksagung

Die diesem Beitrag zugrunde liegende gleichnamige Dissertationsschrift wurde im Rahmen des DFG-Projektes „Verteilte Simulation und Emulation von dezentral gesteuerten Materialflusssystemen“ gefördert.

## Literatur

Chandy, K. M.; Misra, J.: Distributed Simulation: A Case Study in Design and Verification of Distributed Programs. IEEE Transactions on Software Engineering SE-5 (1979) 5, S. 440-452.

Daniluk, D.; Chisu, R.: Simulation und Emulation im Internet der Dinge. In: Günthner, W.; ten Hompel, M. (Hrsg.): Internet der Dinge in der Intralogistik. Berlin: Springer 2005, S. 149-166.

Daniluk, D.: Emulation mit der High Level Architecture: Ein Konzept zur Kopplung von multiagentenbasierten Steuerungssystemen und Materialflusssimulatoren. Saarbrücken: AV Akademikerverlag GmbH & Co. KG 2012.

Daniluk, D.: Verteilte Simulation und Emulation von Materialflusssystemen mit dezentraler Steuerung. Dortmund: Verlag Praxiswissen 2014.

Fjällström, P.-O.: Algorithms for Graph Partitioning: A Survey. Linköping Electronic Articles in Computer and Information Science 3 (1998) 10, S. 1-37.

Gudehus, T.: Logistik: Grundlagen, Strategien, Anwendungen. Berlin: Springer 2005.

Günthner, W.; ten Hompel, M. (Hrsg.): Internet der Dinge in der Intralogistik. Berlin, Heidelberg: Springer 2010.

Jennings, N. R.; Bussmann, S.: Agent-Based Control Systems. IEEE Control Systems Magazine 23 (2003) 3, S. 61-74.

Johnson, D. B.; Maltz, D. A.: Dynamic Source Routing in Ad Hoc Wireless Networks. In: Imielinski, T; Korth, H. (Hrsg.): Mobile Computing. Norwell, MA: Kluwer Academic Publishers 1996, S. 153-181.

Karypis, G.: METIS: A Software Package for Partitioning Unstructured Graphs, Partitioning Meshes, and Computing Fill-Reducing Orderings of Sparse Matrices. Manual, Version 5.0, 2011.

Lendermann, P.: About the Need for Distributed Simulation Technology for the Resolution of Real-World Manufacturing and Logistics Problems. In: Proceedings of the 2006 Winter Simulation Conference (WSC), 2006, S. 1119-1128.

Leon-Garcia, A.; Widjaja, I.: Communication Networks: Fundamental Concepts and Key Architectures. Boston, Massachusetts: McGraw-Hill 2000.

Mehl, H.: Methoden verteilter Simulation. Braunschweig, Wiesbaden: Vieweg & Sohn Verlagsgesellschaft 1994.

Porras, J.; Hara, V.; Harju, J.; Ikonen, J.: Improving the Performance of the Chandy-Misra Parallel Simulation Algorithm in a Distributed Workstation Environment. In: Summer Computer Simulation Conference, 1997, S. 657-662.

Wood, K. R.; Turner, S. J.: A Generalized Carrier-Null Method for Conservative Parallel Simulation. In: Proceedings of the Eighth Workshop on Parallel and Distributed Simulation, 1994, S. 50-57.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# VR-basierte Simulation zur Ermittlung der Leistungsverfügbarkeit zellularer intralogistischer Systeme

## *VR-based Simulation for Calculating the Performance Availability of Cellular Intralogistics Facilities*

Kevin Eilers, RIF Institut für Forschung und Transfer e.V., Dortmund (Germany), kevin.eilers@rt.rif-ev.de

Jürgen Roßmann, Institut für Mensch-Maschine-Interaktion (MMI), RWTH Aachen University, Aachen (Germany), rossmann@mmi.rwth-aachen.de

Abstract: Performance availability is a measurement for the performance of material flow systems. Its value is calculated based on data obtained while the system is in operation, which impedes predicting or planning the performance availability of a projected system. We developed a simulation model with the goal to calculate the performance availability of a cellular facility logistics system containing several AGV (Automated Guided Vehicles) to remedy this situation. The visual representation of a real world facility was obtained by performing a 3-D laser scan on site. Mechanical components as well as a control system were incorporated into the simulation to create a detailed representation of the system's behaviour. The simulation can be run in a VR environment where intuitive metaphors for the visualisation of inner simulation states and results as well as for the manipulation of simulation components and data processing algorithms help to communicate the simulation results.

## 1 Einführung

Leistungsverfügbarkeit ist nach der VDI-Richtlinie 4486 (VDI 2012) eine Kennzahl zur Bewertung der Leistung eines Materialflusssystems, die auf Basis von Messungen auftretender Warte- bzw. Laufzeiten ermittelt wird. Aufgrund der Tatsache, dass diese Daten erst während des laufenden Betriebs erhoben werden können, ist es schwierig, ein System auf eine bestimmte Leistungsverfügbarkeit hin zu planen und im Vorfeld Design-Entscheidungen zu treffen, die einen Einfluss auf die Leistungsverfügbarkeit haben können. Insbesondere bei der Betrachtung von hoch dynamischen, agentenbasierten Systemen, wie zellularen intralogistischen Anlagen, bei denen das gesamte Systemverhalten von Änderungen an einzelnen Komponenten beeinflusst werden kann, wird dieses Problem deutlich. Der Wunsch, logistische

Systeme an zukünftig auftretende Szenarien anzupassen, um auf interne oder externe Veränderungen zu reagieren, erfordern zudem gegebenenfalls weitere Planungsphasen während diese Systeme bereits in Betrieb sind.

Ein solches System sowohl schon vor der Inbetriebnahme als auch während des laufenden Betriebs zu simulieren und somit Daten zu erheben, um die Leistungsverfügbarkeit unter verschiedenen Bedingungen zu ermitteln, ermöglicht es der planenden Instanz, neue Steuerungs- und Sortierstrategien zu testen und Veränderungen am Layout oder den Komponenten zu planen, ohne Zugriff auf ein fertiges System zu haben oder ein schon laufendes System modifizieren zu müssen.

Unser Ziel ist es daher, ein Simulationsmodell eines zellularen intralogistischen Systems zu konzipieren, welches die Daten ermittelt, die Einfluss auf die Leistungsverfügbarkeit haben, und den aktuellen Zustand des Systems auf intuitive Art und Weise dem Benutzer zugänglich macht. Die Simulation soll dabei sowohl auf einfachen Desktop-PCs als auch in VR-Umgebungen wie einer CAVE (siehe Abbildung 1) lauffähig sein. Neu zu entwickelnde Interaktionsmethoden und -metaphern sollen den Zugriff auf die erhobenen Daten erleichtern und ein tiefes Verständnis der internen Vorgänge des Systems fördern, um in Zukunft die Entwicklung und die Wartung ähnlicher Systeme zu unterstützen.

Als Referenzsystem für das Simulationsmodell dient eine existierende Testanlage für zellulare Intralogistik, welche mit Hilfe eines 3-D-Laserscanners vermessen wurde. Die Scanergebnisse in Form von Punktwolken können direkt in der Simulation verwendet werden und bilden zudem die Basis für die Erstellung von Polygonmodellen und Texturen für die einzelnen Komponenten. Die mechanischen Bestandteile des Systems wurden mit den kinematischen Werkzeugen des Simu-lationssystems nachgebildet. Eine eigens entwickelte Steuerungslogik, oder aber eine über eine zusätzliche TCP/IP-Schnittstelle angeschlossene externen Steuerung, kontrollieren das System von der Auftragsvergabe bis hin zu der Ansteuerung der Motoren der einzelnen Komponenten.

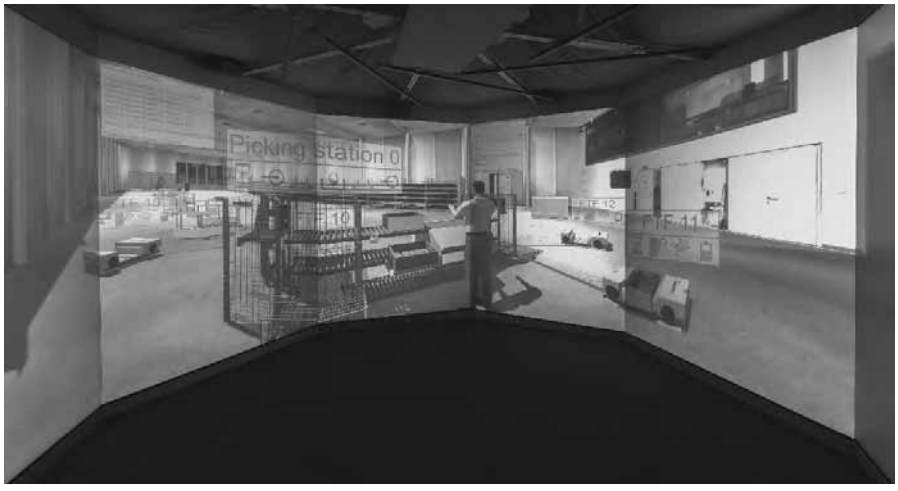



*Abbildung 1: Die Simulation in einer CAVE-Umgebung*

## 2 Stand der Technik

Da die Idee der Leistungsverfügbarkeit noch relativ neu ist, existieren momentan nur wenige Veröffentlichungen zu diesem Thema, insbesondere im Bereich der Simulation (Nagel 2010). Das Thema VR wird in keiner der Veröffentlichungen behandelt.

Viele Arbeiten behandeln den Bereich der Simulation von agentenbasierten Logistiksystemen, aber keine davon spannen den Bogen von einer echtzeitfähigen, VR-basierten Simulation bis hin zur Leistungsverfügbarkeit. Die Arbeiten von Günther et al. (2001), Raffel (2005), Raffel und Wagner (2011) sowie Roidl und Follert (2007) klammern den Bereich VR aus, während die Arbeiten von Günthner et al. (2010), Günthner und Tenerowicz (2010) und Günthner et al. (2012) die Simulation von autonomen Fahrzeugen und deren Sensorik in den Vordergrund stellen und weder Realzeitberechnung noch die VR betrachten.

Craighead et al. haben 2007 diverse kommerzielle Lösungen für die dynamische Simulation von autonomen Fahrzeugen untersucht, die aber aufgrund der Anforderungen an die Hardware nicht realzeit- und somit auch nicht VR-tauglich sind.

Im Bereich der Visualisierungs- und Interaktionsmetaphern für Intralogistiksysteme liegen Arbeiten von Poupyrev et al. (1998) und Steinicke et al. (2004) vor, die aber den Bereich der Leistungsverfügbarkeit noch nicht mit abdecken.

Das in dieser Arbeit verwendete Simulationsframework bietet einen hohen Grad an Modularität und stand im Zentrum mehrerer Forschungsprojekte (Rossmann und Alves 2009; Freund et al. 2003; Rossmann et al. 2009a; Rossmann et al. 2009b). Es ist VR-tauglich und bietet diverse Schnittstellen nach außen an, darunter Schnittstellen um verschiedene Eingabegeräte anzusteuern, sowie eine TCP/IP-Schnittstelle, die zum Anschluss externer Steuerungssysteme genutzt werden kann.

## 3 Leistungsverfügbarkeit

Die VDI Richtlinie 4486 definiert die Leistungsverfügbarkeit als den „anforderungs- und termingerechten Erfüllungsgrad von zwischen Vertragspartnern (Hersteller und Anwender) vereinbarten Prozessen unter Einhaltung der vereinbarten Rahmenbedingungen“. Das bedeutet, dass im Gegensatz zur technischen Verfügbarkeit, welche ein Maß für die Leistung und Zuverlässigkeit einer einzelnen Komponente darstellt, die Leistungsverfügbarkeit ein Maß für die Leistung des Gesamtsystems angibt.

Die Leistungsverfügbarkeit kann auf zwei Arten berechnet werden: Entweder auf Basis von Wartezeiten an Knotenpunkten des Systems oder auf Basis der Laufzeiten von Fördergütern. Da die Leistungsverfügbarkeit das gesamte System betrachtet, kann ihr tatsächlicher Wert sowohl von statischen Parametern wie der Verfügbarkeit einzelner Komponenten oder dem Design der Anlage als auch von dynamischen Größen wie der aktuellen Auftragslage oder Parametern der aktuell zu transportierenden Güter abhängen. Warte- und Laufzeiten können daher nur anhand von Messungen am sich schon im Betrieb befindlichen System erhoben werden. Dies macht die Ermittlung der Leistungsverfügbarkeit eines zukünftigen Systems unmöglich und steht somit dem Ziel der Richtlinie, eine verbindliche Metrik für die Leistung eines sich in der Planung befindlichen Systems und damit Rechtssicherheit für Hersteller und Anwender herzustellen, im Wege.

Um die Leistungsverfügbarkeit zu ermitteln, werden zunächst die Warte- bzw. Laufzeiten an den relevanten Schnittstellen bzw. Komponenten über einen oder mehrere definierte Zeiträume erhoben. Neben für die Berechnung der Leistungsverfügbarkeit notwendigen Parametern werden keine weiteren Daten betrachtet. Ausfälle einzelner Komponenten oder Verzögerungen gehen nur dann in die Berechnung mit ein, wenn Sie Einfluss auf die gemessenen Daten ausüben.

Soll die Leistungsverfügbarkeit auf Basis von Wartezeiten berechnet werden, kommt folgende Formel zum Einsatz:

$$\eta_W = \frac{T_B - T_W}{T_B}$$

Dabei steht $\eta_W$ für die Leistungsverfügbarkeit, während $T_B$ und $T_W$ die Länge der Beobachtungsperiode respektive die angefallenen Wartezeiten repräsentieren.

Soll die Leistungsverfügbarkeit hingegen auf Basis von Laufzeiten berechnet wer–den, kommt folgende Formel zum Einsatz:

$$\eta_L = \frac{N - n}{N}$$

Die Leistungsverfügbarkeit wird durch $\eta_L$ repräsentiert, wobei $N$ und $n$ für die erwartete Anzahl an Fördergütern bzw. die Anzahl verspäteter Fördergüter stehen.

In beiden Fällen ergibt sich die Leistungsverfügbarkeit $\eta$ zu einem Wert im Intervall [0; 1], wobei 1 den maximalen Erfüllungsgrad (100 %) darstellt.

## 4 Erstellung des Simulationsmodells

Das gewählte Referenzmodell, welches die Grundlage für unser Simulationsmodell darstellt, ist eine intralogistische Anlage des Fraunhofer Instituts für Materialfluss und Logistik (IML) in Dortmund. Die 17 x 55 m messende Halle beinhaltet ein Regallager mit einer Kapazität von 520 Behältern, die auf fünf Ebenen gelagert werden. Bis zu 50 autonome Fahrzeuge befahren den Hallenboden und das Regal, in welchem sie mit Hilfe zwei Fahrzeugheber an den Regalenden die Ebenen wechseln können. Die Fahrzeuge sind mit einem Differenzialantrieb ausgestattet, der sowohl das Vorwärts- als auch das Rückwärtsfahren sowie eine Drehung auf der Stelle erlaubt. Je ein Laserscanner am vorderen und einer am hinteren Ende eines Fahrzeugs erfasst dabei permanent die Umgebung, zum Zwecke der Selbstlokalisierung und Kollisionsvermeidung. Die Erfassungsbereiche der Scanner werden dabei von ihren eigentlichen 180° je nach Fahrtsituation angepasst (Kamagaew et al. 2011). Ein Teleskopgreifer erlaubt es den Fahrzeugen, Behälter aus dem Regal sowohl auf der linken als auch auf der rechten Seite aufzunehmen und in einer der sieben Kommissionierstationen gegenüber dem Regallager abzulegen. Eine achte mit einem Depalettierroboter ausgestatte Station fungiert als Wareneingang des Lagers.

Abbildung 2 zeigt eine Übersicht über das Hallenlayout. Der in der rechten oberen Ecke befindliche Depalettierroboter ist nicht mit abgebildet.

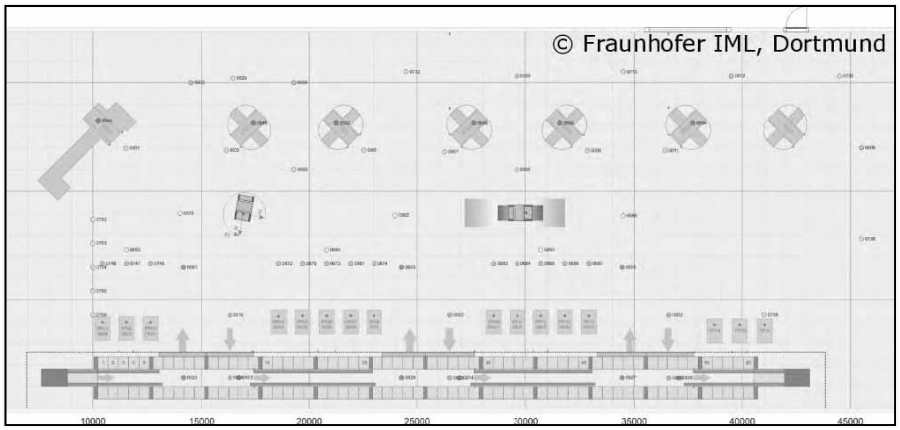



*Abbildung 2: Layout des Referenzsystems (© Fraunhofer IML, Dortmund)*

Um möglichst zeitnah ein exaktes Abbild der Halle zu erzeugen, wurde eine Kombination aus einem 3-D-Laserscanner und einer Digitalkamera eingesetzt. In einem Zeitraum von zweieinhalb Stunden wurden mit dieser Hardware vier Scans entlang der Achse der Halle und ein fünfter Scan im Regal durchgeführt. Das Ergebnis waren fünf Punktwolken vom Laserscanner (ein Ausschnitt ist in Abbildung 3 zu sehen) und fünf Panoramaaufnahmen von der Kamera. Diese Daten wurden bereinigt und miteinander kombiniert um die Punktwolken mit der Farbinformation der Kamerabilder anzureichern und somit eine durchgängige, eingefärbte und von Ausreißern befreite Punktwolke zu erzeugen, welche eine genaue Darstellung des Referenzsystems erlaubt (siehe Abb. 4).

Diese Punktwolke kann im Simulationssystem eingesetzt werden und bietet nicht nur dem menschlichen Betrachter eine dreidimensionale Ansicht der Halle, sondern ist auch für die Laserscanner der Fahrzeuge sichtbar, so dass die Fahrzeuge diese virtuelle Halle befahren können, ohne mit Hindernissen zu kollidieren.

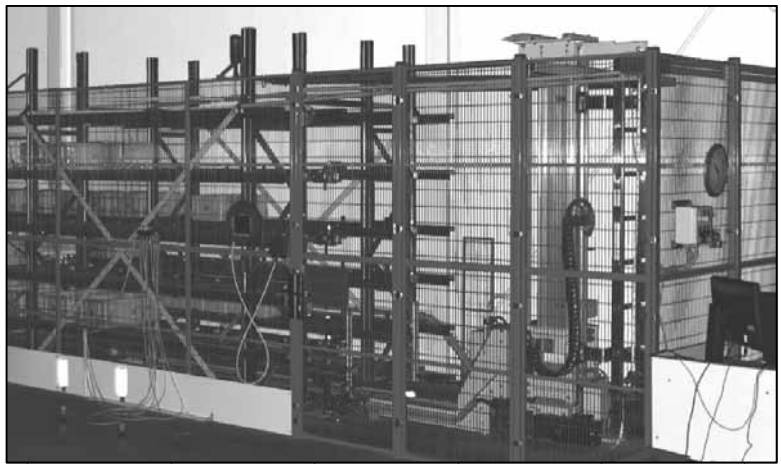


*Abbildung 3: Ausschnitt aus einer der fünf Punktwolken*

Diese rein geometrische Abbildung der Halle reicht natürlich nicht aus, um das gesamte System inklusive aller aktiven Komponenten, wie den Kommissionierstationen, den Fahrzeughebern und nicht zuletzt auch den Fahrzeugen zu simulieren.

Deshalb wurden auf Basis der Daten aus dem Laserscan Polygonmodelle aller aktiven Komponenten und derer beweglichen Teile erzeugt, die mit kinematischen und sensorischen Basiskomponenten aus dem Simulationssystem um ihre Funktionalität erweitert wurden (siehe Abb. 1 und 5).

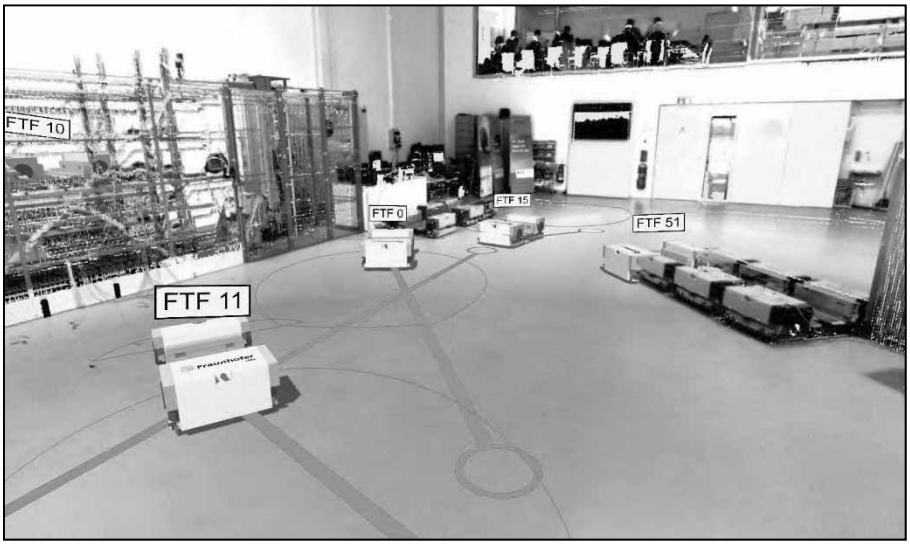



*Abbildung 4: Eingefärbte Punktwolken innerhalb des Simulationssystems*

## 5 Steuerung und Datenermittlung

Das Referenzsystem wird von einem Agentensystem gesteuert, dessen Software in drei Schichten zerfällt, die mittels TCP/IP miteinander kommunizieren. Das letzte Element in dieser Kette bildet das einzelne Fahrzeug, welches Zielkoordinaten empfängt, die aus der Interaktion der beiden übergeordneten Schichten (Auftragsgenerierung und agentenbasierte Auftragsvergabe) generiert werden.

Eine vereinfachte Abbildung dieser Kontrollarchitektur wurde von uns realisiert. Sie importiert Daten aus dem Materialflusssimulator OTD-NET (siehe Wagenitz 2007) und bereitet diese auf. Die daraus resultierenden Aufträge werden anschließend zu den am Auftrag hinterlegten Zeitpunkten in das System eingespeist und zentral an freie Fahrzeuge verteilt.

Da das Ziel des Simulationssystems die Ermittlung der Leistungsverfügbarkeit ist, muss die Simulation in der Lage sein, die für deren Berechnung notwendigen Daten zu erheben. Zu diesem Zweck ist jedes Objekt im Simulationssystem mit einer eigenen Logik ausgestattet, welche dafür zuständig ist, die relevanten Daten zu ermitteln. Kommissionierstationen protokollieren ihre Wartezeiten und messen den Durchsatz, während Behälter ihre Laufzeiten stoppen sobald sie Ziel eines Auftrags werden. Diese Daten werden in Echtzeit erhoben und können während der laufenden

Simulation dargestellt werden. Weitere für die Berechnung der Leistungsverfüg–barkeit notwendige Parameter, wie der erwartete Durchsatz, die Länge einer Beobachtungsperiode und die maximal erlaubte Laufzeit können jederzeit vom Benutzer angepasst werden, um die Einflüsse dieser Größen auf die Leistungsverfügbarkeit sofort sichtbar zu machen.

Simulationsläufe können jederzeit aufgenommen und wiedergegeben werden, um auch im Nachhinein eine detaillierte Analyse der Simulationsgeschehens zu ermöglichen.

## 6 Visualisierung, Interaktion und Evaluation

Soll eine Simulation in einer VR-Umgebung dargestellt werden, ist nicht nur die Echtzeitfähigkeit essentiell, sondern auch die Möglichkeit, dem Benutzer Einblicke in das Innere der Simulation und auf die ermittelten Daten zu gewähren und ihm die Möglichkeit zu eröffnen, intuitiv auf das System und die Datenerhebung einwirken zu können.

Zu diesem Zweck wurden eine Reihe neuer visueller Metaphern entwickelt, die diese Daten auf intuitive Art und Weise darstellen und zudem die Interaktion mit dem System erlauben. Abbildung 4 zeigt die adaptiven Erfassungsbereiche der Laserscanner und die Zielpositionen der Fahrzeuge zusammen mit dem Toleranzbereich des Ziels. In Abbildung 5 sind diverse Symbole zu sehen, die – auf den Boden projiziert – Auskunft darüber geben, welche Aktion ein Fahrzeug beim Erreichen der Zielposition durchführen soll (Warten auf einen Fahrzeugheber, Rotieren auf der Stelle, um eine bestimmte Orientierung einzunehmen, oder wie hier dargestellt: Ablegen eines Behälters). Symbole an den Fahrzeugen, Behältern, Kommissionierstationen und Fahrzeughebern geben Auskunft über deren Zustand.

An Kommissionierstationen können der Zustand der Auftragswarteschlange, die Belegung der Fließbänder und der Ein-/Ausgänge abgelesen werden, sowie der Zustand der Auftragsverarbeitung. Des Weiteren dienen die interaktiven Symbole dazu, Bestandteile der Kommissionierstation, wie die Fließbänder oder die Auftragsverarbeitung, zu deaktivieren oder aber detaillierte Informationen zu dem Einfluss der Station auf die Leistungsverfügbarkeit in Form von angefallenen Wartezeiten und deren Gewicht bei Bedarf einzublenden.

Symbole über den Fahrzeugen geben Auskunft über den aktuellen Status der Auftragswarteschlange, den aktuellen Bewegungszustand des Fahrzeugs, den Status des Kollisionsvermeidungssystems und den Ladestand der Batterie. Auch diese Symbole können vom Benutzer zur Interaktion genutzt werden, um die Auftrags-verarbeitung des Fahrzeugs zu deaktivieren, detaillierte Informationen zu den Motorparametern in Form eines Oszilloskops einzublenden, oder aber den Ladestand der Batterie zu manipulieren.

Symbole auf den Behältern geben Auskunft über eventuell erteilte Transportaufträge für diesen Behälter sowie dessen Füllstand und Gewicht, während die Symbole an den Fahrzeughebern Auskunft über deren aktuelle Position und Zielposition geben, sowie interaktiv das Bewegen des Fahrzeughebers zulassen.

Die dargestellten Interaktionsmethoden lassen sich nutzen, um dem System Störungen und Verzögerungen beizubringen, die potenziell die Leistungsverfügbarkeit be-

einflussen und somit einen Einblick gewähren, wie Robust das System gegenüber solchen Problemen ist, und wo noch Optimierungspotenzial besteht.

*Abbildung 5: Metaphern zur Visualisierung der inneren Systemzustände*

Ein zusätzlich einblendbares Pop-Up-Menü gibt dem Nutzer weitere Möglichkeiten, die von der Simulation erhobenen Daten zu visualisieren und Einfluss auf die Berechnung der Leistungsverfügbarkeit zu nehmen. So können die variablen Berechnungsparameter für die Leistungsverfügbarkeit, wie maximale Laufzeiten, Länge der Beobachtungsintervalle und erwarteter Durchsatz, in diesem Menü angepasst werden. Die Leistungsverfügbarkeit und der Einfluss einzelner Komponenten auf diese werden dabei in einem Overlay über das Simulationsgeschehen gelegt und in Echtzeit aktualisiert, so dass der Effekt der Parameteranpassung sofort evident wird. Des Weiteren können Zusatzinformationen zum Kollisionsvermeidungssystem der Fahrzeuge sowie Daten zur Flächennutzung auf den Hallenboden projiziert werden.

Ein Fahrzeug mit einer den realen Bedienmöglichkeiten nachempfundenen, manuellen Steuerung kann dem System bei Bedarf hinzugefügt werden, um z. B. die Wirksamkeit des Kollisionsvermeidungssystems zu testen, oder aber die Reaktion des Systems und der Werte für die Leistungsverfügbarkeit auf Verzögerungen und Staus.

## 7 Fazit und Ausblick

Wir haben ein Simulationsmodell einer zellularen Intralogistikanlage vorgestellt, dessen Zweck es ist, die Leistungsverfügbarkeit der simulierten Anlage unter verschiedenen Betriebs- und Berechnungsbedingungen zu ermitteln. Dabei wurde

ein genaues Abbild des realen Referenzsystems mit Hilfe von 3-D-Laserscannern erstellt und zusammen mit mechanischen Komponenten und Polygonmodellen in das Simulationsmodell integriert. Ein Datenimporter für Daten aus einem Materialflusssimulator, sowie ein dazu passendes Auftragsverwaltungsystem sorgen für eine realistische Auftragslage.

Neu entwickelte, intuitive und interaktive Metaphern stellen sowohl innere Zustände als auch Simulationsergebnisse intuitiv dar, und bieten dem Nutzer die Möglichkeit, auf die Simulation und die Berechnungen einzuwirken.

Momentan nutzt das Simulationsmodell noch seine Allwissenheit über das Geschehen, um die Position der Fahrzeuge exakt zu bestimmen. Um eine realistischere Simulation zu schaffen, ist es sinnvoll an dieser Stelle über ein heuristisches Modell oder besser noch, über eine exakte Simulation der Selbstlokalisierungsmechanismen und der an den Sensoren auftretenden Messfehler die Position ermitteln zu lassen. Die dafür notwendigen Komponenten sind bereits Teil des Simulationssystems und können an diesen Zweck angepasst werden (Rossman et al. 2010).

## Danksagung

Teile dieser Veröffentlichung entstanden im Rahmen des Forschungsvorhabens „Simulation zellularer Intralogistiksysteme zur Analyse der Leistungsverfügbarkeit in der Virtuellen Realität“, das von der Deutschen Forschungsgemeinschaft (DFG) unter dem Geschäftszeichen RO 4064/1-1 gefördert wird.

## Literatur

Craighead, J.; Murphy, R.; Burke, J.; Goldiez, B.: A survey of commercial & open source unmanned vehicle simulators. In: Proceedings of the 2007 International Conference on Robotics and Automation. Piscataway (NJ): IEEE 2007, S. 852-857.

Freund, E.; Rossmann, J.; Turner, C.: Application of robotic mechanisms to simulation of the International Space Station. In: Proceedings of the 2003 International Conference on Intelligent Robots and Systems, vol. 3. Piscataway (NJ) IEEE 2003, S. 3047-3052.

Günther, H.-O.; Lehmann, M.; Raffel, W.-U.; Wagner, G.: Modellierung und Simulation von Fahrerlosen Transportsystemen als Multiagentensysteme. 4. Kolloquium SPP Agenten, Freie Universität Berlin 2001.

Günthner, W. A.; ten Hompel, M.; Tenerowicz, P.; Büchter, H.: Auf dem Weg zur zellularen Fördertechnik – Fördern im Schwarm. Hebezeuge und Fördermittel (2010) 3, S. 78-79.

Günthner, W. A.; Tenerowicz, P.: Modularisierung und Dezentralisierung in der Intralogistik – Auf dem Weg zur zellularen Fördertechnik. Industrie Management (2011) 1, S. 25-29.Günthner, W. A.; ten Hompel, M.; Tenerowicz-Wirth, P.; Büchter, H.; Schipplick, M.: Forschungsbericht – Algorithmen und Kommunikationssysteme für die Zellulare Fördertechnik. München/Dortmund 2012.

Kamagaew, A.; Stenzel, J.; Nettsträter, A.; ten Hompel, M.: Concept of cellular transport systems in facility logistics. In: Proceedings of the 5th IEEE Internatio-

nal Conference on Automation, Robotics and Applications. Piscataway (NJ): IEEE 2011, S. 40-45.

Maier, M. M.: Prognose der Leistungsverfügbarkeit während der Planung. Logistics Journal: Proceedings vol. 6 (2010). ISSN 2192-9084.

Nagel, M.: Simulation der Leistungsverfügbarkeit der Modellfabrik. B.S. Thesis, Fakultät Maschinenbau, Technische Universität Ilmenau, 2010.

Poupyrev, I.; Billinghurst, M.; Weghorst, S.; Ischikawa, T.: Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques. Computer Graphics Forum 17 (1998) 3, S. 41-52.

Raffel, W.-U.; Wagner, G.: Agentenorientierte Modellierung von Fahrerlosen Transportsystemen. Technical Report B 01 05, Institut für Informatik, Freie Universität Berlin 2001.

Raffel, W.-U: Agentenbasierte Simulation als Verfeinerung der Diskreten-Ereignis-Simulation unter besonderer Berücksichtigung des Beispiels Fahrerloser Transportsysteme. Dissertation, Institut für Informatik, Freie Universität Berlin 2005.

Roidl, M.; Follert, G.: Simulation von multiagentenbasierten Materialflusssteuerungen. In: Informatik 2007 – Informatik trifft Logistik. Beiträge der 37. Jahrestagung der Gesellschaft für Informatik e.V. (GI), Proceedings 109, Band 1, Bremen 2007.

Rossmann, J.; Alves G.: A detailed timber harvest simulator coupled with 3-D visualization. In: Proceedings of the International Conference on Forestry Science and Technology, Amsterdam 2009, S. 319-324.

Rossmann, J.; Ruf, H.; Schlette, C. (2009a): Model based programming by demonstration – Fast Setup of robot System (ProDemo). In: Kröger, T.; Wahl, F.M. (Hrsg.): Advances in Robotics Research – Theory, Implementation, Application. Berlin, Heidelberg: Springer 2009, S. 159-268.

Rossmann, J.; Schluse, M.; Schlette, C. (2009b): The Virtual Forest: Robotics and simulation technology as the basis for new approaches to the biological and the technical production in the forest. In: Proceedings of the 13th World Multi-Conference on Systemics, Cybernetics and Informatics, vol. 2, Orlando (FL) 2009, S. 33-38.

Rossmann, J.; Schlette, C.; Emde, M.; Sondermann, B.: Discussion of a self-localization and navigation unit for mobile robots in extraterrestrial environments. In: Proceedings of the 10th International Symposium on Artificial Intelligence, Robotics and Automation in Space, Sapporo (Japan) 2010, S. 46-53.

Steinicke, F.; Ropinski, T.; Hinrichs, K. H.: Selektion von Objekten in Virtuellen Umgebungen mit der Improved Virtual Pointer Metapher. Tagungsband zum 1. Workshop Virtuelle und Erweiterte Realität der GI-Fachgruppe VR/AR, Chemnitz 2004, S. 59-67.

VDI 2012. VDI Richtlinie 4486: Zuverlässigkeit in der Intralogistik: Leistungsverfügbarkeit. Berlin: Beuth 2012.

Wagenitz, A.: Modellierungsmethode zur Auftragsabwicklung in der Automobilindustrie, Dissertation TU Dortmund, Fak. Maschinenbau 2007.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Mathematische Optimierung und Simulation von Paketsortieranlagen – Bessere Lösungen durch Kopplung der beiden komplementären Methoden

***Mathematical Optimisation and Simulation of Parcel Transshipment Terminals – Better Solutions by Linking These Two Complementary Methods***

Uwe Clausen, Daniel Diekmann, Jens Baudach, Moritz Pöting,
TU Dortmund – Institut für Transportlogistik, Dortmund (Germany),
clausen@itl.tu-dortmund.de, diekmann@itl.tu-dortmund.de,
baudach@itl.tu-dortmund.de, poeting@itl.tu-dortmund.de

Abstract: Mathematical optimisation and simulation represent two powerful methods that have been successfully used to solve a wide range of logistical problems. However, they are mainly applied separately. Simulation allows the modelling of logistics systems with almost unlimited complexity very close to reality. But finding the best system configuration is difficult and time-consuming since there are usually huge numbers of possible alternative scenarios. In contrast, mathematical optimisation has the ability to make complex decisions and find (near) optimal solutions. Real world logistic systems, however, can only be solved on a lower level of detail without stochastic behaviours. This paper presents a solution approach for the operational planning of a transshipment terminal that closely links both methods to make use of their complementary advantages. Computational results are presented for different system load scenarios using the example of two transshipment terminals.

## 1 Einleitung Motivation & Systembeschreibung

Das Marktsegment der Kurier-, Express- und Paketdienstleister (KEP) wächst kontinuierlich und stellt mittlerweile ein bedeutendes Segment der Transport-wirtschaft dar. Während vor einigen Jahren sowohl die steigende Nachfrage nach zeitkritischen Sendungen als auch der Trend hin zu kleineren Versandeinheiten („Atomisierung von Sendungen“) wesentliche Wachstumstreiber waren, sorgt gegenwärtig vor allem das boomende E-Commerce für steigende Sendungsmengen bei den Marktteilnehmern (Vahrenkamp et al. 2012).

Um Sortierengpässe zu vermeiden, welche unweigerlich zu Verzögerungen im Transportablauf und Laufzeitverlängerungen führen, errichten die Marktteilnehmer

neue Standorte bzw. erweitern bestehende Sortiersysteme. Die wachsenden Sendungsmengen und die fortschreitende Automatisierung des Sortierprozesses führen zu einer steigenden Komplexität der Umschlaganlagen, in denen vielfältige Zuordnungsentscheidungen getroffen werden müssen. Unterschiedliche auto-matische aber auch manuelle Umschlagtätigkeiten müssen daher aufeinander abgestimmt werden, um einen möglichst optimalen Betrieb zu gewährleisten.

Gegenstand und Ziel dieses Beitrags ist es, einen neuen methodischen Ansatz für eine optimale Steuerung sowie den Betrieb einer Paketsortieranlage mittels mathematischer Optimierung und Simulation vorzustellen. Im Beitrag wird gezeigt, wie die beiden Lösungsansätze der mathematischen Optimierung und der Simulation eng miteinander gekoppelt werden, um über die Ausnutzung der komplementären Eigenschaften der beiden Methoden Lösungen zu erzeugen, die zu besseren Ergebnissen führen als es eine der beiden Methoden für sich alleine könnte. Im Rahmen dieses Papers werden wir uns auf die Entscheidungen der eingehenden LKW-Entladetor-Anstellung sowie der ausgehenden Zuordnung von Zielrelation zu (Sortier)-Endstellen fokussieren.

## 2 Stand der Forschung

Lösungen zur Modellierung von Umschlaganlagen des Straßengüterverkehrs sind häufig im Rahmen von Forschungsarbeiten zu finden. Gleichwohl liegt der Anwendungsschwerpunkt dieser Beiträge überwiegend auf Cross-Docking oder Stückgutanlagen, deren innerbetrieblicher Umschlag im Gegensatz zu KEP-Sortieranlagen durch Stapler und nicht durch vollautomatische Sortieranlagen durchgeführt wird. Mathematische Modelle zur optimalen Steuerung der Torbelegung basieren dabei zumeist auf der Minimierung der innerbetrieblichen Transportwege und werden von verschiedenen Autoren mit unterschiedlichen Restriktionen betrachtet (z. B. Tsui und Chang 1992; Gue 1999; Yu und Egbelu 2008). Einen guten Überblick bestehender Arbeiten und eine Problemklassifikation bieten die Arbeiten von Boysen und Fliedner (2010) sowie von Van Belle et al. (2012).

Nur wenige wissenschaftliche Arbeiten besitzen einen spezifischen Anwendungsfokus auf Paketsortieranlagen. Existierende Arbeiten konzentrieren sich entweder auf Lösungen für die eingehende Torbelegungsplanung (Bereich Entladung) oder die Endstellenbelegungen für ausgehende Ziele (Bereich Beladung). Für die Endstellenbelegung stellen Werners et al. (2010) Optimierungsmodelle vor, welche die Zuordnung von Zustellgruppen zu Endstellen mittels eines binären linearen Optimierungsmodells (teilweise über hierarchische Dekomposition) mit dem Ziel der Minimierung innerbetrieblicher Wege optimieren (Werners et al. 2001; Werners und Wülfing 2010). Die Zuordnung der Endstellen zu den Beladetoren wird in Jarreh et al. (2014) durch Berücksichtigung des Personals und der Wechselzeit in einem mehrstufigen Optimierungsmodell erweitert. Die eingehende Torbelegungsplanung wird ausführlich in den Arbeiten von McWilliams et al. (2005), McWilliams et al. (2008) und McWilliams (2010) aufgegriffen. Hierfür wird ein simulationsbasierter Optimierungsansatz verfolgt, der je nach Veröffentlichung verschiedene Suchverfahren, wie z. B. generische Algorithmen, Simulated Annealing oder genetische Algorithmen verwendet. Ein Lösungsansatz, welcher auf einer gleichberechtigten Verknüpfung von komplexen Simulations- und Optimierungsmodellen für den Anwendungsfall einer KEP-Umschlaganlage basiert und

sowohl das Problem der eingehenden Tor- als auch der Endstellenbelegung adressiert, wurde erstmals im Artikel von Diekmann et al. (2014) vorgestellt. Die Idee der Kopplung wird in diesem Artikel aufgegriffen und im Folgenden weiterentwickelt.

# 3 Modellierung

In Kapitel 3 wird zunächst der methodische Kopplungsansatz beschrieben, bevor in einem zweiten Schritt die mathematische Modellierung und das Simulationsmodell erläutert werden.

## 3.1 Kopplung von Optimierung und Simulation

Abbildung 1 skizziert die Kopplung des Optimierungs- und Simulationsmodells. Im Gegensatz zu simulationsbasierten Optimierungsansätzen kommuniziert in dem vorliegenden Modell das eigenständige Optimierungsmodell mittels Interface mit dem detaillierten, ereignisorientierten Simulationsmodell.

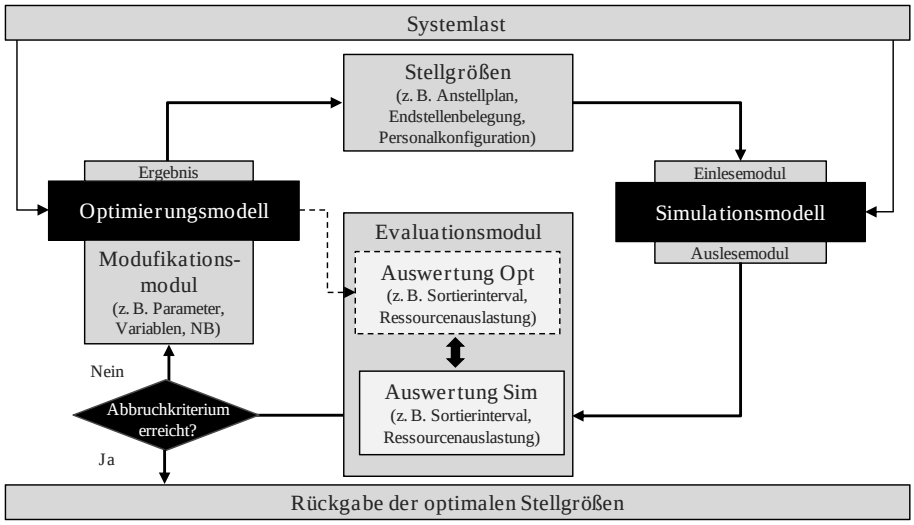



*Abbildung 1: Kopplung von mathematischer Optimierung und Simulation*

Im ersten Schritt des Kopplungskreislaufs wird die Systemlast, bestehend aus ankommenden LKWs sowie den dazugehörigen Sendungen, in beide Modelle geladen und zuerst vom mathematischen Modell bzgl. optimaler Entscheidungsvariablen analysiert. Die ermittelten Stellgrößen (Anstellplan und Endstellenbelegung) werden danach dem Simulationsmodell übergeben, welches die Simulationsläufe auf Basis der gegebenen Daten anstößt. Anschließend werden die vom Simulationsmodell ermittelten Kennzahlen an das Evaluationsmodul weitergeleitet. Das Evaluationsmodul analysiert den Output des Simulationsmodells und gleicht diesen mit dem Optimierungsmodell ab. Mittels des Modifikationsmoduls, das auf den Ergebnissen des Evaluationsmoduls basiert, können die Erkenntnisse des Systemverhaltens aus der Simulation somit (in vereinfachter Form) in das Optimierungsmodell integriert werden. Dies kann durch die Hinzunahme, Deaktivierung oder auch Anpassung von

Variablen, Parametern oder auch Nebenbedingungen erfolgen. Hierdurch wird das Optimierungsmodell schrittweise an die Umschlaganlage als auch Systemlast angepasst und weiterentwickelt. Auch die Simulation profitiert innerhalb des Kopplungszykluses von der Optimierung, da sie bereits optimierte Stellgrößen erhält, wodurch sich der Aufwand für die verschiedene Simulationsläufe maßgeblich – im Falle eine guten Übereinstimmung der Ergebnisse sogar auf einen Kopplungszyklus – reduziert.

## 3.2 Mathematische Modellierung

Der entwickelte Optimierungsansatz eignet sich für beide Referenzsysteme und erfordert lediglich kleinere Anpassungen aufgrund struktureller oder technischer Unterschiede. In einem ersten Schritt wird im Hinblick auf eine bestmögliche Ressourcenauslastung die *Minimierung des Sortierintervalls* (Makespan) aller Belade-, Entlade- und Sortierprozesse verfolgt. Anschließend wird in einem zweiten Schritt eine durchgehend *gleichmäßige Verteilung* der Sendungen auf die (logischen) Hauptsorter angestrebt, die entscheidend für die Grenzleistung eines Sortiersystems ist. Hierfür werden Zuordnungsentscheidungen sowohl im Entlade- als auch im Beladebereich einer Paketsortieranlage getroffen und mittels zwei Arten von binären Variablen modelliert: Die Variablen $x_i^t$ werden zu eins wenn ein LKW i ab der Zeitscheibe t (an einem bestimmten Tor) entladen wird und die Variablen $z_j^s$ zeigen an, ob eine ausgehende Relation j dem (logischen) Hauptsorter s zugeordnet ist (über ein entsprechendes Tor). Für die Diskretisierung der Zeit haben sich dabei für das Optimierungsmodell Zeitscheiben mit einer Länge von fünf Minuten als sehr guter Kompromiss zwischen Modellgröße, Rechenzeit und Detailierungsgrad erwiesen.

Um ein quadratisches Modell zu vermeiden, werden zusätzlich positive Flussvariablen $y_{i,t,s}^{\text{parcel/NC}}$ für Sendungen eingeführt, die in Zeitscheibe t aus LKW i über den Hauptsorter s laufen.

Die entwickelten linearen Optimierungsmodelle enthalten verschiedene Typen von Nebenbedingungen. Nebenbedingung (1) stellt sicher, dass jeder eingehende LKW genau einer Zeitscheibe zugeordnet wird, in der seine Entladung beginnt. Dabei ist zu beachten, dass immer nur eine begrenzte Zahl $|U|$ an Entladetoren zur Verfügung steht und ein LKW auch über mehrere Zeitscheiben ($dur_i$) hinweg entladen werden kann (2).

$$\sum_t x_i^t \quad = 1 \quad \forall i \tag{1}$$

$$\sum_i \sum_{t-dur_i<\hat{t}\leq t} x_i^{\hat{t}} \quad \leq |U| \quad \forall t \tag{2}$$

Die Restriktionen zur Beladung beinhalten sowohl die eindeutige Zuordnung jeder Zielrelation zu einem der Hauptsorter (3) als auch die begrenzte Zahl an verfügbaren Beladetoren (aufgeteilt in Nah- und Fernverkehrstore) an jedem Sorter (4).

$$\sum_s z_j^s \quad = 1 \quad \forall j \tag{3}$$

$$\sum_{j\in J_{long}} z_j^s \leq \left|L_{long}^s\right| \quad \forall s, \quad \sum_{j\in J_{short}} z_j^s \leq |L_{short}^s| \quad \forall s \tag{4}$$

Die Nebenbedingungen (5) und (6) garantieren in Kombination korrekte Sendungsflüsse. Mittels (5) wird sichergestellt, dass jeden LKW zu jeder Zeit (in Summe) die

korrekte Anzahl an Sendungen verlässt. Diese Sendungen werden dann durch (6) gemäß Zielrelationen auf die entsprechenden Sorter verteilt. Die Nebenbedingungen (5) und (6) gelten auch für NC-Sendungen.

$$\sum_{s} y_{i,t,s}^{parcel} = \sum_{j} \frac{ship_{i,j}^{parcel}}{dur_i} \sum_{t-dur_i<\hat{t}\leq t} x_i^{\hat{t}} \quad \forall(i,t) \tag{5}$$

$$y_{i,t,s}^{parcel} \leq \sum_{j} \frac{ship_{i,j}^{parcel}}{dur_i} z_j^s \quad \forall(i,t,s) \tag{6}$$

Die folgenden Nebenbedingungen (7) werden zur Bestimmung des Makespan aller Belade-, Entlade- und Sortierprozesse der Paketsortieranlage benötigt.

$$\sum_{t} n_t \, x_i^t \geq ms_{start} \quad \forall i \quad \text{and} \quad \sum_{t} (n_t + dur_i) \, x_i^t \leq ms_{end} \quad \forall i \tag{7}$$

Dabei bezeichnen $n_t$ die Nummer von Zeitscheibe t und $dur_i$ erneut die Entladedauer (in Zeitscheiben) von LKW $i$. Die ganzzahligen Variablen $ms_{start}$ und $ms_{end}$ repräsentieren den Start der Entladung und das Ende der Beladung. Andere praktische Restriktionen, zum Beispiel die maximal zulässige Wartezeit der eingehenden Fahrzeuge nach ihrer Ankunft an der Anlage, können durch Preprocessing und Fixierung von Variablen realisiert werden.

Das Optimierungsmodell zur Minimierung des Makespan aller Belade-, Entlade- und Sortierprozesse beinhaltet dann die Nebenbedingungen (1), (2) und (7), vervollständigt durch die Zielfunktion $\min(ms_{end} - ms_{start})$.

Das Optimierungsmodell zur gleichmäßigen Verteilung der Sendungen auf die Hauptsorter benötigt die Nebenbedingungen (1) – (6). Als geeignete Zielfunktion hat sich die Minimierung des maximalen Sendungsflusses (über alle Zeitscheiben und Sorter) herausgestellt:

$$\min \max_{(s,t)} \sum_{i} \left( y_{i,t,s}^{parcel} + y_{i,t,s}^{nc} \right)$$

Hinweis: Aufgrund der den Hauptsortern vorgeschalteten Entladebereiche erfolgt für Referenzsystem 2 eine weitere (optionale) Optimierung. Diese bestimmt für jeden LKW in Abhängigkeit seiner Sendungen den besten Entladebereich, an das dieser LKW (über ein entsprechendes Entladetor) angestellt werden kann.

Die entwickelten mathematischen Modelle wurden mit GAMS 24.3 implementiert, können mittels Cplex 12.6 in vertretbarer Zeit gelöst werden und führen zu guten Ergebnissen (vgl. Kapitel 4).

## 3.3 Simulationsmodell

Die Simulationsmodelle wurden mit der vom Institut für Transportlogistik entwickelten Simulationsumgebung ED Transport modelliert, welche auf dem Simulator Enterprise Dynamics (ED) basiert (Neumann und Deymann 2008). Diese ermöglicht die Darstellung und Simulation logistischer Knoten auf mikroskopischer Ebene anhand eines Katalogs eigens entwickelter Bausteine (Clausen und Diekmann 2012). In der Entladung wurde insbesondere der Einfluss stochastischer Prozesszeiten, die sequentielle Reihenfolge entladender Sendungen und manuelles Handling für nicht maschinell lesbare Sendungen abgebildet. Im Bereich der Sortierung wurde die Belegung einzelner Sorter-Belts sowie der Einfluss von Rezirkulierern (Sen-

dungen, die am Hauptsorter nicht nach dem ersten Umlauf ausgeschleust werden können) abgebildet und deren Einfluss auf das Gesamtsystem simulativ ermittelt. Dies ermöglicht es anhand der gemessenen Sortierleistung, den Einfluss auf vorgelagerte Prozesse zu bestimmen und detaillierte Rückschlüsse auf das tatsächliche Systemverhalten zu ziehen. Insgesamt wurden zwei Systeme (Abb- 2) modelliert und zur Untersuchung des entwickelten Kopplungsansatzes verwendet.

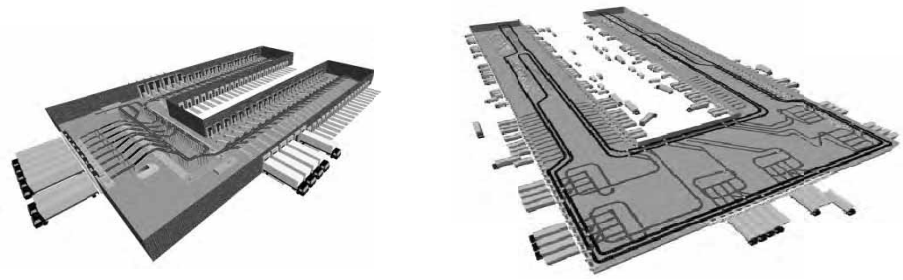


*Abbildung 2: Referenzsysteme (Depot- und Hubstandort, RS1 links und RS 2 rechts)*

Das erste modellierte Referenzsystem (RS1) ist ein Depotstandort mit U-Layout und etwa 6.000 m² Hallenfläche. Es teilt sich auf in einen Entladebereich an der Stirnseite der Anlage sowie zwei Beladebereiche, die sich in den Schenkeln der Halle befinden. Die Sortierung auf die Endstellen erfolgt durch ein automatisches Sortiersystem. Nicht automatisch sortierbare Sendungen (NC) werden von einem zweiten System aus Förderbändern manuell auf die Endstellen sortiert. Der Anteil der NC-Sendungen an der Gesamtsendungsmenge beträgt etwa 10%.

Das zweite Referenzsystem (RS2) repräsentiert exemplarisch einen großen Hub-Standort im U-Layout mit ca. 24.000 m² Hallenfläche. Dieser verfügt über sechs Entladeabschnitte, die sich im stirnseitigen Entladebereich der Halle befinden und die Sendungen auf einen von zwei Hauptsortern vorsortieren. Die Hauptsorter fördern die Sendungen zu den Endstellen der Beladebereiche.

Für die beiden entwickelten Referenzsysteme wurde durch Variation der Systemlast drei Szenarien entwickelt, implementiert und untersucht. Das Basisszenario (S1) beruht auf den Verteilungen von Realdaten. Aufbauend auf das Basisszenario wurde anschließend ein Verspätungsszenario (S2) aufgebaut. Ausgehend von dem Basisszenario wurden die Ankunftszeiten der Fahrzeuge modifiziert und eine Ballung des Ankunftsintervalls der Fahrzeuge erzeugt (durchschnittliche Verspätung ca. 30-40 Minuten). Ebenfalls aufbauend auf dem Basislastszenario wurde ein Schieflastszenario (S3) entwickelt. Im Gegensatz zu den relativ gleichmäßig verteilten Mengen in S1 wurden hier teils stärkere Schieflasten in der Verteilung von Zielrelationen der Sendungen einzelner Fahrzeuge generiert. Diese Schieflast führt zu einer ungleichmäßigen Inanspruchnahme der Fließbänder und Personal.

Der Lösungsansatz wird im Folgenden an den beiden oben beschrieben Referenzsystemen mit den drei generierten Systemlasten getestet.

## 4 Anwendung und Ergebnisse

Zunächst werden die Ergebnisse für den Depotstandort (RS1) vorgestellt und erläutert. Daran anschließend erfolgt die Analyse und Zusammenfassung der Ergebnisse des Hubstandorts (RS2).

Die Ergebnisse von RS1 sind in Abbildung 3 für die unterschiedlichen Systemlasten (S1 bis S3) visualisiert. Die erste Spalte zeigt die Ergebnisse der Simulationsläufe für die bestehende IST-Belegung (manuelle Planung) des Anstellplans und der Endstellenbelegung, die mit Hilfe des in der Praxis derzeit typischerweise angewendeten FIFO-Prinzip simuliert wird. Die zweite Spalte zeigt die Ergebnisse des Optimierungsmodells mit den neuen optimierten Stellgrößen (Anstellplan und Endstellenbelegung). Die dritte Spalte visualisiert die Ergebnisse des Simulationsmodells mit optimierten Stellgrößen. Die wichtigsten Kennzahlen der jeweiligen Läufe sind in den schwarzen Boxen zusammengefasst. Dies ist zum einen die Kennzahl für die gleichmäßige Auslastung der Hauptsorter (Workload Balance - WLB). Diese wird als Differenz der Auslastungslinien der beiden Sorter von der (als Linie skizzierten) Gleichauslastung über alle Zeitscheiben gemessen. RS1 besitzt z. B. zwei Hauptsorter, daher liegt die Gleichauslastung hier bei 50%. Die Auslastung der beiden Sorter ist prozentual auf der linken Y-Achse abgetragen. Die zweite Kennzahl ist die Länge des Sortierintervalls (MSP). Die Anzahl der Zeitscheiben, welche sich aus der Länge des Intervalls zwischen den Zeitscheiben der ersten und letzten sortierten Sendung ergeben, ist auf der X-Achse mit dem Doppelpfeil gekennzeichnet. Darüber hinaus ist in grau im Hintergrund die Anzahl der von dem Sortiersystem verteilten Sendungen pro Zeitscheibe visualisiert und auf der rechten Y-Achse abgetragen.

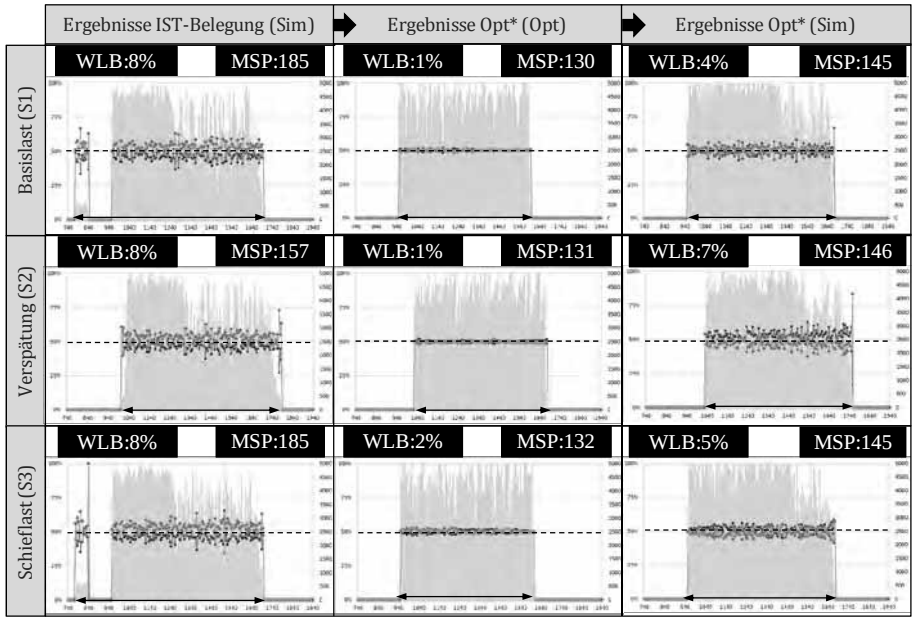



*Abbildung 3: Ergebnisse Referenzsystem 1*

Im RS1 konnte durch Anwendung des Kopplungsverfahrens in allen untersuchten Szenarien eine deutliche Verkürzung des MSP (22% in S1 und S3) und eine Verbesserung des WLB realisiert werden. Es zeigt sich, dass gegen Ende der abgebildeten Schicht die Ergebnisse von Simulation und Optimierung auseinanderlaufen. Gegenüber der Ist-Planung konnte die Grenzlast des Systems jedoch deutlich länger gehalten werden. In den Szenarien S2 und S3 führen im Vergleich zum Basisszenario die ungünstigen Parameterausprägungen der Systemlast, insbesondere in der Simulation, zu schlechteren Ergebnissen im WLB, welche jedoch keine signifikanten Auswirkungen auf den MSP haben.

Abbildung 4 zeigt die Ergebnisse des RS2. Hier weist das Ist-Szenario deutliche geringere Sortierleistung auf. Insbesondere S3 zeigt, dass eine hohe Abweichung des WLB zu Reduktionen in der Sortierleistung und damit einhergehend zu einer deutlichen Verlängerung des MSP führen. Die Übereinstimmung der Simulationsergebnisse mit den Optimierungsergebnissen ist sehr viel höher als in dem RS1. Dies lässt sich u.a. durch die höhere Systemlast in RS2 begründen, da dem Optimierungsverfahren von Beginn an mehr Kombinationsmöglichkeiten bei der Anstellung von Fahrzeugen zur Verfügung steht und sich die Anlage schon bei Schichtbeginn in einem hohen Lastbereich betreiben lässt.

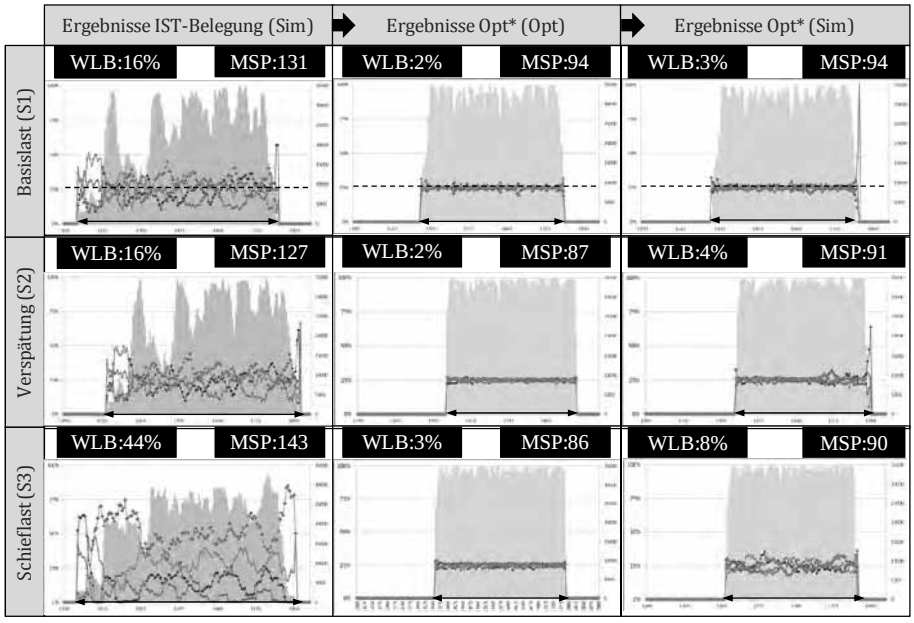



*Abbildung 4: Ergebnisse Referenzsystem 2*

Die Auswertung der Szenarien beider Systeme lässt Aussagen über die entwickelte Kopplungsmethode aus mathematischer Optimierung und Simulation zu. Die ereignisorientierte Simulation liefert im Gegensatz zur reinen Optimierung realitätsnahe Ergebnisse und zeigt u.a. stochastische Schwankungen des Systems auf. Durch eine gute iterative Abstimmung lassen sich die Ergebnisse beider Verfahren jedoch zu einem hohen Grad annähern. Dies zeigt, dass das Verfahren der Kopplung funk-

tioniert. Die zur Optimierung verwendeten Verfahren lassen sich so gut auf die Problemstellung kalibrieren und zur operativen Planung verwenden. Die Laufzeiten des mathematischen Optimierungsmodells (GAMS 24.3, Cplex 12.6) betragen in Abhängigkeit des RS und der Zielfunktion zwischen wenigen Minuten und mehreren Stunden. Die Simulationslaufzeit liegt je nach Referenzsystem und Systemlast im Minutenbereich, wodurch sich für die Vorplanung einer Schicht ausreichend schnelle Laufzeiten ergeben. Die ermittelten Kennzahlen aller Untersuchungsszenarien belegen, dass eine optimierte Planung sehr viel bessere Ergebnisse erzielen kann als eine manuelle Planung. Im RS2 konnten der ermittelte MSP durch annähernd vollständige Gleichauslastung des Sortiersystems um 38% reduziert werden (S3) und eine Grenzlast von 35.000 Sendungen pro Stunde über das gesamte Sortierintervall realisiert werden.

Der Vergleich der beiden RS zeigt, dass mit steigender Anzahl an Fahrzeugen und Sendungen sich zunehmend bessere Ergebnisse erzielen lassen. Gerade für Umschlaganlagen mit hohem Sendungsaufkommen bei gleichzeitig niedrigem NC-Anteil ist die Anwendung des Verfahrens daher sehr vielversprechend. Eine entsprechende Planung auf Basis aktueller Fahrzeug- und Sendungsdaten ließe sich bei gegebener Datenbasis aufgrund der Laufzeit in der operativen Betriebsplanung einsetzen.

## 5 Fazit

Der Beitrag zeigt, dass sich durch Kopplung von mathematischer Optimierung und Simulation für Paketsortieranlagen sehr gute Ergebnisse ableiten lassen und validiert diese anhand zwei repräsentativer Referenzsysteme. Im Vergleich zu einer manuellen Planung weist die Simulation der Anlagen auf Grundlage optimierter Parameter deutliche Verbesserungen auf. Das Verfahren liefert auch bei ungünstigen Bedingungen, wie z. B. Verspätungen (S2) oder stärkeren Schieflasten (S3) sehr vielversprechende Resultate.

Zukünftige Arbeiten werden die Kopplung der beiden Methoden weiterentwickeln und das Verfahren verbessern, z. B. durch eine verbesserte Abbildung der manuellen Sortierung (NC-Sendungen). Neben den beiden untersuchten RS soll ein weiteres RS aufgebaut werden und anhand dieser Systeme operationalisierbare Muster für die Praxis abgeleitet werden. Darüber hinaus gilt es weitere Planungsprobleme durch Anwendung des entwickelten Kopplungsverfahrens zu lösen.

## Danksagung

Der vorliegende Beitrag ist Teil des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekts „Entwicklung eines Verfahrens zur Kopplung von diskreter Optimierung und stochastischer Simulation für die Planung und Steuerung logistischer Knoten“.

## Literatur

Boysen, N.; Fliedner, M.: Cross dock scheduling: Classification, literature review and research agenda. Omega 38 (2010) 6, S. 413-422.

Clausen, U.; Diekmann, D.: 2012: Simulation of transshipment terminals in the parcel delivery industry. In: Proceedings of the European Simulation and Modelling Conference, Essen 2012, S. 386-398.

Diekmann, D.; Baudach, J.; Clausen, U.: A new approach of linking mathematical optimization and discrete-event simulation for operating transshipment terminals in the parcel delivery industry. In: Brito, A. C.; Tavares, J. M.; Braganca de Oliveira, C. (Hrsg.): Proceedings of the 2014 European Simulation and Modelling Conference, Porto 2014, S. 254-260.

Gue, K. R.: The effects of trailer scheduling on the layout of freight terminals. Transportation Science 33 (1999) 4, S. 419-428.

Jarrah, A.; Qi, X.; Bard, J.: The destination-loader-door assignment problem for automated package sorting centers. Transportation Science 48 (2014), published online.

McWilliams, D.; Stanfield, P.; Geiger, C.: The parcel hub scheduling problem: A simulation-based solution approach. Computer & Industrial Engineering 49 (2005) 3, S. 393-412.

McWilliams, D.; Stanfield, P.; Geiger, C.: Minimizing the completion time of the transfer operations in a central parcel consolidation terminal with unequal-batch-size inbound trailers. Computer & Industrial Engineering 54 (2008) 4, S. 709-702.

McWilliams, D.; A genetic based scheduling algorithm for the PHSP with unequal batch size inbound trailers. Journal of Industrial and System Engineering 4 (2010) 3, S. 167-182.

Neumann, L.; Deymann S.: Transsim-Node - A simulation tool for logistics nodes. In: Industrial Simulation Conference. Lyon: The European Simulation Society, 2008, S. 283-287.

Tsui, L. Y.; Chang, C.-H.: Optimal solution to a dock door assignment problem. Computer & Industrial Engineering 23 (1992) 1, S. 283-286.

Vahrenkamp, R.; Kotzab, H.; Siepermann, C.: Logistik – Management und Strategien. München: Oldenbourg Wissenschaftsverlag 2012.

Van Belle, J.; Valckenaers, P.; Cattrysee, D.: Cross-docking: State of the art. Omega 40 (2012) 6, S. 827-846.

Werners, B.; Thorn, J.; Freiwald, S.: Innerbetriebliche Transportoptimierung für ein Paketzentrum der Deutschen Post World Net. OR-Spektrum 23 (2001) 4, S. 507-523.

Werners, B.; Wülfing, T.: Robust optimization of internal transports at a parcel sorting center operated by Deutsche Post World Net. European Journal of Operational Research 201 (2010) 2, S. 419-426.

Yu, W.; Egbelu, P.: Scheduling of inbound and outbound trucks in cross docking systems with temporary storage. European Journal of Operational Research 184 (2008), S. 377-396.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Anwendungen von Simulation zur Verbesserung von Prozessabläufen in Industriewäschereien

## *Applications of Simulation to Improve Operations in Commercial Laundries*

Annegret Brandau, David Weigert, Otto-von-Guericke-Universität Magdeburg, Magdeburg (Germany), annegret.brandau@ovgu.de, david.weigert@ovgu.de

Juri Tolujew, Fraunhofer-Institut für Fabrikbetrieb und -automatisierung IFF Magdeburg, Magdeburg (Germany), Juri.Tolujew@iff.fraunhofer.de

Abstract: Insufficient basic data, increasingly complex customer and product structures, and a lack of transparency in the process structure increase the quantity of laundry items in circulation, lessen machine utilisation, delay deliveries, and increase the liability of errors in operations. Competition is making it essential for laundries to operate quickly, reliably and cost effectively. A simulation analysis of a commercial laundry is intended to identify and facilitate the reorganization of ineffective processes. This article describes two simulation studies. The first study aids in the reorganization of a commercial laundry's receiving area by means of automation and Auto-ID-technologies. The second study specifies a simulation model that is used to identify the best strategy for laundry order scheduling. Both of the studies demonstrate the feasibility of simulating complex laundry operations and of using such simulation studies to improve laundry operations.

## 1 Einleitung und Problemstellung

Industrielle Großwäschereien in Deutschland erbringen umfassende Dienstleistungen zur Bereitstellung, Reinigung und Aufbereitung waschbarer Textilien. Mit zunehmendem Wettbewerb und ausgeprägter Kundenorientierung steigen die Anforderungen an die Wäschereien. Individuelle Kundenwünsche im Gesundheitswesen und Hotelgewerbe fordern eine termingerechte und flexible Reinigung und Auslieferung von Wäsche zu niedrigen Preisen. Durch die historische Entwicklung vieler Wäschereien sind deren Planung, Steuerung und Überwachung vom Erfahrungswissen der Mitarbeiter geprägt. Der geschlossene Prozessablauf einer Wäscherei unterscheidet sich von einem produzierenden Unternehmen dadurch, dass das Produkt, die Wäsche, nicht aus vorhandenen Ressourcen hergestellt wird, sondern einen Wäschekreislauf durchläuft (Abb. 1).

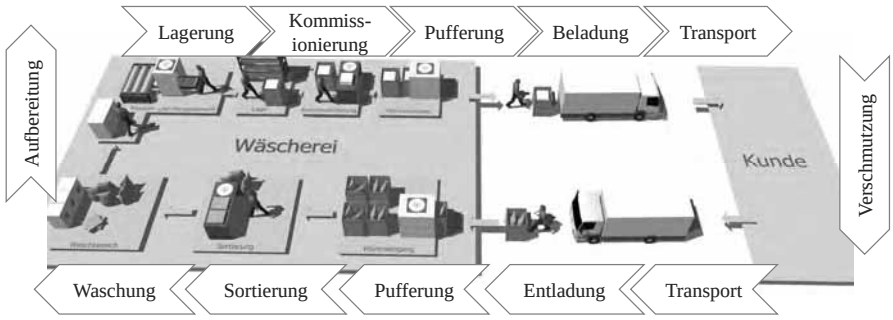



*Abbildung 1: Schematische Darstellung der Prozesse eines Wäschereikreislaufs*

Die Kernprozesse wie Waschen und Aufbereiten (Mangeln, Trocknen, Falten, Stapeln) sind hoch automatisierte Prozesse, deren Schnittstellen häufig noch manuell ausgeführt werden. Durch die gestiegene Komplexität in der Kunden- und Produktstruktur behandeln Wäschereien eine vielfältige Anzahl von Wäschestücken. Um eine Produktion vorausschauend und optimal planen zu können, ist es notwendig die Art und Anzahl der vorhandenen Ressourcen im Wareneingang zu kennen. Im Fall der Wäschereien ist diese Anforderung nur teilweise erfüllt. Die Art und Anzahl der eintreffenden Schmutzwäsche der Kunden ist vorab nicht bekannt. Der Inhalt der Container ist durch den geschlossenen Transport nur schwer einsehbar. Die Planung der Waschauftragsreihenfolge basiert somit auf unsicheren Daten, was zu aufwändigen Umplanungen im operativen Betrieb führt. Weiterhin ist in kleineren Wäschereien wenig Lagerfläche vorhanden, ein An- und Neubau ist in den seltensten Fällen möglich.

Diese Gründe führen zu hohen Umlaufbeständen an Wäsche, schlechter Maschinenauslastung, verspäteten Auslieferungen und einer erhöhten Störanfälligkeit im Prozessablauf. Aufgrund der aktuellen Markt- und Wettbewerbssituation ist es für Großwäschereien zukünftig von großer Bedeutung schnell, zuverlässig und kostengünstig zu Arbeiten. Dieser Problemstellung widmet sich das Kooperationsprojekt „Entwicklung eines Simulationsmodells zur Produktionsplanung und -steuerung von industriellen Großwäschereien“, das durch eine Zuwendung im Förderprogramm „Zentrales Innovationsprogramm Mittelstand“ des Bundesministeriums für Wirtschaft und Technologie (BMWi)–Fördermodul ZIM-KOOP/Fördersäule Kooperationsnetzwerke finanziert wird.

Die erste Anwendung stellt eine Simulationsstudie zur Umstrukturierung des Wareneingangs vor. Die zweite Anwendung umfasst die Entwicklung eines Simulationsmodells zur Bewertung von Waschauftragsreihenfolgen hinsichtlich Liefertreue unter Normalbetrieb sowie unter Störereignissen. Die Durchführung solcher Simulationsstudien wurden in der Literatur und Praxis für herkömmliche Produktionssysteme zahlreich dargestellt, vgl. (Kuhn 1998; Verein Deutscher Ingenieure 2000; Dangelmaier et al. 2013). Jedoch sind keine Anwendungen für den Industriezweig der Großwäschereien zu finden.

## 2 Umstrukturierung des Wareneingangs

Wenn durch den Einsatz von RFID-Technologie im Wareneingang einer Wäscherei bekannt ist, welche Wäschestücke sich in welchem Rollcontainer befinden, bietet diese Transparenz neue Möglichkeiten zum Einsatz von verschiedenen Ein- und Auslagerungsstrategien. Daher ist Ziel der Simulationsstudie die Bewertung verschiedener Lagerstrategien für einen transparenten Wareneingang. Mittels einer wirtschaftlichen und ganzheitlichen Betrachtung der eintreffenden Rollcontainer sowie der Produktvielfalt und Lagerhaltung, wurde eine kombinierte Lagerstrategie entwickelt und deren Praxistauglichkeit mit einer Simulationsstudie nachgewiesen. Es wurden zusätzlich die maximal notwendige Anzahl an Stellplätzen, die maximale Pufferkapazität vor und nach dem automatischen Identifikations- und Vereinnahmungssystem der Rollcontainer sowie die maximal notwendigen fahrerlosen Transportfahrzeuge ermittelt.

### 2.1 Kombinierte Lagerstrategie

Der erste Schritt zur Entwicklung einer kombinierten Lagerstrategie ist die Auswahl einer geeigneten Lagertechnik. Dazu wurde eine detaillierte Analyse der Begebenheiten im Wareneingang sowie der vorhandenen Lagerhilfsmittel durchgeführt. Es hat sich ergeben, dass sich im Wareneingang das Bodenzeilenlager am besten eignet, da die Wäsche in Rollcontainern gebündelt und transportiert wird.

Zur Bestimmung einer geeigneten Belegungsstrategie wurden Zusammenhänge zwischen der komplexen Kunden- und Produktstruktur sowie den Behandlungsprozessen mit Hilfe von Experteninterviews analysiert. Das Ergebnis der Analyse ist in Abbildung 2 dargestellt. Die verschiedenen Strukturen zeigen, dass sich eine Zonierung als Belegungsstrategie sehr gut eignet. Eine chaotische sowie Festplatzlagerung bedingt, dass jeder Stellplatz zu jeder Zeit direkt erreichbar ist. Wegen der Produkt- und Kundenstruktur ist dies aber in Wäschereien nicht unbedingt notwendig, außerdem reicht die vorhandene Lagerfläche nicht aus. Daher ist eine Zonierung nach den verschiedenen Strukturen eine sehr gute Lösung.

Aus den verschiedenen Zusammenhängen (Wirtschaftstyp, Kundentyp, Behandlungskreis und Kundenkreis) lassen sich 48 Kombinationen zur Bestimmung der Zonen ableiten. Um die Reihenfolge der Zuordnung zu bestimmen, wurden Prioritätsregeln aufgestellt. Zuerst wird die Wäsche nach dem jeweiligen Kundentyp, anschließend nach Kundenkreis, dann nach Wirtschaftstyp und anschließend nach dem Behandlungskreis zugeordnet.

### 2.2 Validierung der kombinierten Lagerstrategie und Bestimmung der Größe der Zonen

Um die praktische Anwendung der Lagerstrategie und Zonierung zu validieren, wurde eine Simulationsstudie durchgeführt. Als logistisches System wird der Wareneingang der Wäscherei betrachtet. Die Eingabedaten sind unterschiedliche Rollcontainer, die mit unterschiedlichem Inhalt als LKW–Lieferungen eintreffen. Demzufolge wird zur Abbildung der Containerflüsse der ereignisdiskrete Simulationsansatz gewählt. Durch stochastische Verteilungen werden die Ankunftsfrequenzen und Inhalte der LKW-Lieferungen sowie die Fahrt- und Ladezeiten der fahrerlosen Transportsysteme abgebildet. Über die Einflussgrößen werden die

Lagerstrategie mit der entsprechenden Zonierung, die Anzahl an Stell- und Pufferplätzen sowie die Frequenz und Anzahl der Rollcontainer parametrisiert. Die Ausgabedaten werden durch Bestand im Wareneingang, Auslastung, Still- und Wartezeiten des automatischen Vereinnahmungssystem und des fahrerlosen Transportsystems beschrieben (Abb. 3).

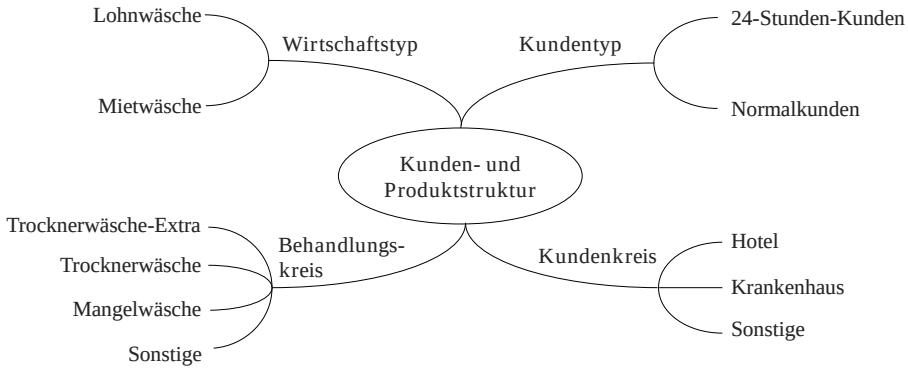



*Abbildung 2: Einflüsse der Kunden- und Produktstruktur*

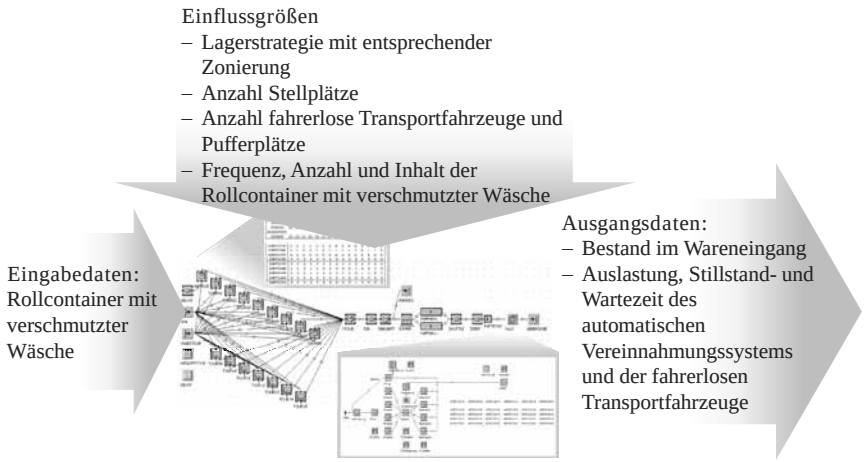



*Abbildung 3: Eingabe- und Ausgabedaten sowie Einflussgrößen des Simulationsmodells zur Umstrukturierung des Wareneingangs*

Das ereignisdiskrete Simulationsmodell wurde mit *Tecnomatix PlantSimulation* erstellt (Abb. 4). Die Eingangsgrößen wie Anzahl, Inhalt und Frequenzen der eingehenden Rollcontainer werden zufällig erzeugt. Die Grenzen der Zufallszahlen sind in Expertengesprächen realitätsnah ausgewählt worden.

Das Lager besteht aus 48 Zonen, die durch das fahrerlose Transportsystem angefahren werden. Hierzu wurde das Lager in unterschiedliche Streckenabschnitte gegliedert. Die Einlagerungsstrategie wird anhand der vorherigen Priorisierung der

Rollcontainer vorgenommen und entspricht der kombinierten Lagerstrategie. Die Auslagerungsstrategie entspricht der 1:3-Regel (auf eine Waschladung Trockenwäsche folgen drei Waschladungen Mangelwäsche) die in Wäschereien häufig Anwendung findet.

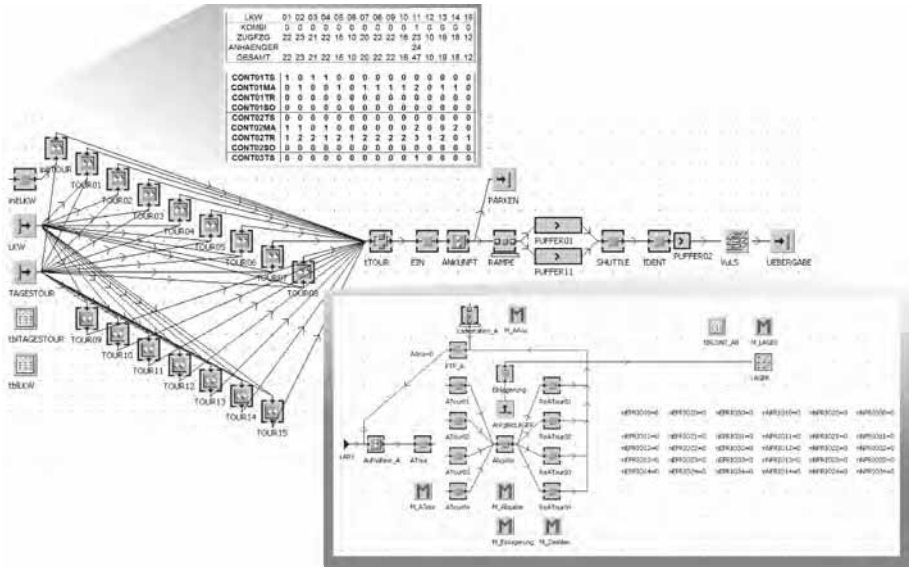



*Abbildung 4: Simulationsmodell des Wareneingangs der Wäscherei*

Durch die Beobachtung und Kontrolle einzelner Bearbeitungsschritte der eintreffenden Rollcontainer im Simulationsmodell (Trace-File-Analyse), wurde die Verifikation erfolgreich durchgeführt. Vorherige Überschlagsrechnungen unterstützten zusätzlich die Modellverifizierung.
Wegen einer unzureichenden Datengrundlage aus der realen Wäscherei zur Validierung des Simulationsmodells, wurden die Teilergebnisse durch Expertengespräche validiert. Diese wurden durch Begehungen und Aufnahme von Prozesszeiten und Ablaufbeschreibungen in direkter Zusammenarbeit mit den Wäschereien geführt.

Der Experimentenplan enthält die Auswahl der Szenarien, die getestet werden, um die Auslastung der Zonen im Wareneingang zu bestimmen. Die verschiedenen Testszenarien bestehen aus insgesamt 3750 zufällig erzeugten Rollcontainern, welche 50-mal das Simulationsmodell durchlaufen. Der betrachtete Zeitraum beträgt 24 Stunden mit einer Einschwingphase von 16 Stunden. Betrachtet werden die gemittelten Durchsätze der jeweils belegten Zonen über die 50 Umläufe. Die Auslastung der jeweiligen Zone ergibt die Zonierung der Lagerstellplätze.

## 2.3 Ergebnisse

Die Experimente zeigen, dass mit der kombinierten Lagerstrategie der Ablauf im Wareneingang nicht gestört wird und es zu keinen Verzögerungen bei der Auslagerung kommt.

Mit den Szenarien kann der gemittelte Durchsatz pro Zonierung ausgewertet werden. Wie in Abbildung 5 dargestellt, belegen die erzeugten Container 30 von bis zu 48 möglichen Zonen.

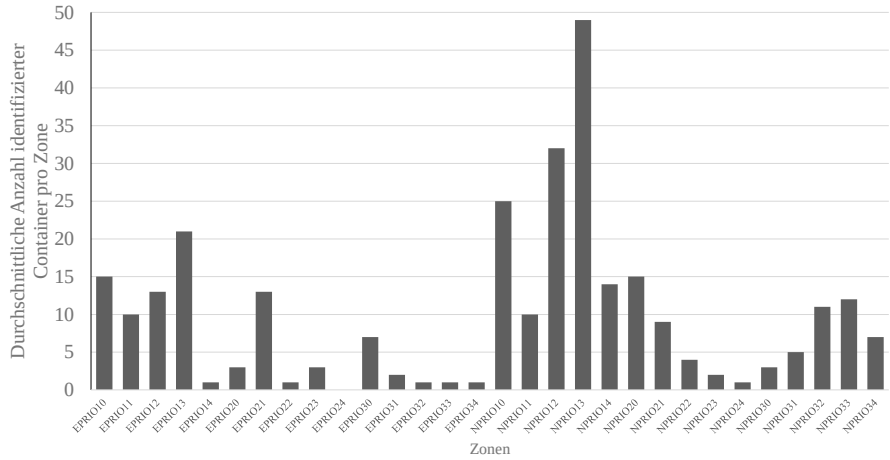



*Abbildung 5: Anzahl Container pro Zone*

Die Top 3–Zonen mit den höchsten gemittelten Durchsätzen, sind NPRIO13 mit 49 Containern, NPRIO12 mit 32 Containern sowie NPRIO10 mit 25 Containern. Die Zonen nehmen Container von Hotelkunden mit Miet- und Lohnwäsche auf, deren Behandlungskreis mehrheitlich aus Trocknerwäsche besteht. Die Zonen EPRIO14 und EPRIO32 bis EPRIO34 bilden mit jeweils einem Container die kleinsten Zonen ab. Diese Zonen beinhalten 24–Stunden-Kunden mit einem erhöhten Anteil an Mangelwäsche.

Die Gesamtergebnisse geben eine Empfehlung zur Verteilung der Zonen und die Anzahl der Stellplätze in der Zone. Die Kombination aus Zonierung und Prioritätsregel entspricht der kombinierten Vereinnahmungs- und Lagerstrategie.

## 2.4 Fazit

Die geringen Lagerflächen, steigende Kunden- und Produktstruktur und die Vielzahl an manuellen Tätigkeiten in den Großwäschereien verhindern einen transparenten Einlagerungsprozess. Erst wenn die Transparenz gegeben ist, kann eine adäquate Zonierung und Stellplatzzuweisung erfolgen.

Durch die Simulationsstudie sowie durch die Entwicklung einer kombinierten Vereinnahmungs- und Lagerstrategie konnte eine Empfehlung für die Zonierung und die Anzahl an Stellplätzen in der jeweiligen Zone gegeben werden. Mit bis zu 48 potentiellen Zonen sind Wäschereien in der Lage unterschiedliche Rollcontainer mit verschiedenen Artikelzusammensetzungen einzulagern.

# 3 Planung von Waschauftragsreihenfolgen

Durch die Einführung von RFID–Technologie sowie einem automatischen Vereinnahmungs- und Identifikationssystem im Wareneingang ist die notwendige Transparenz gegeben, um die Waschauftragsreihenfolge nicht mehr nur mit Erfahrungswissen zu planen sondern verschiedene bekannte Strategien aus anderen Produktionssystemen einsetzen zu können. Es wurden drei Strategien zur Planung der Waschauftragsreihenfolge ausgewählt und mit einem Simulationsmodell hinsichtlich unterschiedlicher Störszenarien nach dem Zielkriterium Liefertreue verglichen.

## 3.1 Strategien zur Planung der Waschauftragsreihenfolge

Aktuell werden die Reihenfolgen der Waschaufträge mit Erfahrungswissen der Mitarbeiter geplant. Je nach Tourenplan wird vorgeschrieben, welcher Kunde wann gewaschen wird. Die Waschaufträge eines Kunden werden anschließend nach Erfahrung der Mitarbeiter im Verhältnis 1:3 eingesteuert (vgl. Kapitel 2.2). Durch die gewonnene Transparenz im Wareneingang lassen sich nun weitere Strategien zur Reihenfolgeplanung umsetzen. Es wurden drei typische Strategien zur Planung ausgewählt: First–Come First–Served (FCFS), Shortest Processing Time (SPT) und Longest Processing Time (LPT).

## 3.2 Simulationsmodell zur Bewertung der Waschauftragsreihenfolgen

Zur Bewertung der drei ausgewählten Strategien zur Planung der Waschauftragsreihenfolge wurde die Produktion der Wäscherei als Simulationsmodell abgebildet. Dazu wird der Wäschefluss vom Wareneingang bis zum Warenausgang betrachtet. Einzelne Wäschestücke werden zu Waschaufträgen zusammengefasst und der Waschstraße zugeführt. Nach der Waschstraße werden die Aufträge teilweise getrennt und zu Aufbereitungsaufträgen neu zusammengesetzt. Im Lager und an den Kommissionierstationen werden die Wäschestücke wieder getrennt und zu Kundenaufträgen kommissioniert. Das Simulationsmodell soll somit den Fluss einzelner Wäschestücke abbilden. Jedoch kann dies wegen der Vielfalt und hohen Anzahl an Wäschestücken zu langen Simulationszeiten führen. Daher wurde der mesoskopische Simulationsansatz gewählt, vgl. (Reggelin 2011). Die verschiedenen Wäschearten werden als Produkttypen und die unterschiedlichen Aufträge als Produktportionen abgebildet. Mit dem mesoskopischen Simulationsansatz können sowohl Ereignisse als auch Flüsse der Produkttypen abgebildet werden. Die Prozesszeiten der manuellen Aufbereitungsprozesse wurden als stochastische Verteilungen abgebildet. Eingabe- und Ausgabedaten sowie Einflussgrößen des Simulationsmodells sind in Abbildung 6 dargestellt.

Das Simulationsmodell wurde mit der Simulationssoftware *ExtendSim* erstellt (Abb. 7). Die Waschaufträge werden als Impulse in das Simulationsmodell eingesteuert und als ereignisdiskrete Objekte behandelt. Nach dem Umschlag in die Sackcontainer werden die Waschaufträge in Flüsse umgewandelt und den Waschmaschinen zugeführt. Der Fluss enthält verschiedene Wäschearten, die nach der Waschmaschine in Produktflüsse getrennt werden. Diese Trennung bilden die Sortierstationen ab. Zur Aufbereitung im Trockner, in der Mangel und den Falt-/ Stapelmaschinen werden die Flüsse in Impulse umgewandelt und anschließend wieder als

Flüsse den einzelnen Kommissionierstation zugeführt. An der Kommissionierstation werden die Waschaufträge zusammengesetzt und der Senke des Systems (Warenausgang) übergegeben.

*Abbildung 6: Eingabe- und Ausgabedaten sowie Einflussgrößen des Simulationsmodells zur Planung der Waschauftragsreihenfolge*

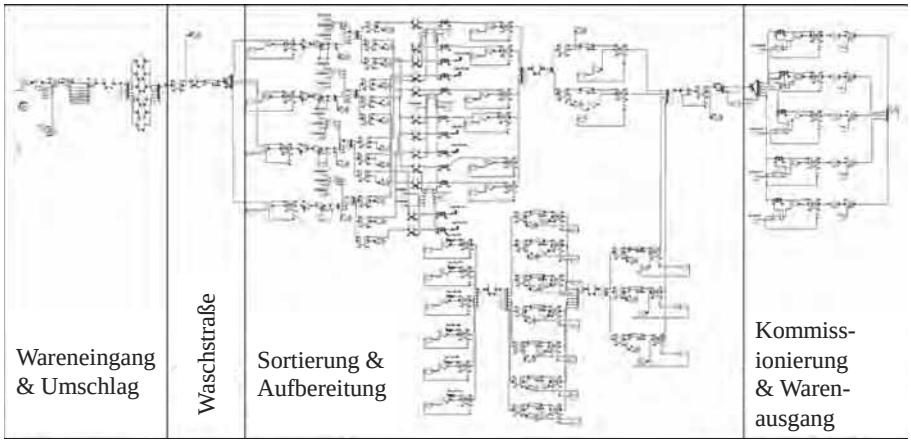



*Abbildung 7: Mesoskopisches Simulationsmodell zur Planung der Waschauftragsreihenfolge*

Die Verifizierung des Modells wurde mit einer Trace-File-Analyse durchgeführt. Einzelne Modellbausteine wurden hinsichtlich ihrer Richtigkeit überprüft und der Fluss der einzelnen Waschaufträge durch die Produktion verfolgt. Zur Validierung des Modells wurden die Ergebnisse aus der vorab durchgeführten Überschlagsrechnung mit den Ergebnissen aus den Testläufen des Simulationsmodells verglichen.

Zur Bewertung der Waschauftragsreihenfolge wurde ein Basisszenario erstellt, das die Anlieferung von Wäsche verschiedener Kunden sowie die Auslieferung von

Kundenbestellungen mit Lieferterminen enthält. Für dieses Szenario wurden die drei ausgewählten Strategien zur Reihenfolgeplanung eingesetzt und mit dem Simulationsmodell validiert. Außerdem wurde die Liefertreue unter den drei Reihenfolgestrategien sowie unter drei Störszenarien: Ausfall einer Waschmaschine, Ausfall einer Mangel und Auftreten eines Eilauftrags bewertet. Insgesamt wurden je Szenario und Strategie 3 Simulationsdurchläufe mit je 24 Stunden durchgeführt.

## 3.3 Ergebnisse

Für alle drei Strategien und den vier Szenarien wurden die Liefertermine der Durchläufe aufgezeichnet und deren Mittelwerte betrachtet (Tab. 1).

*Tabelle 1: Mittelwerte der Fertigstellungszeitpunkte der verschiedenen Szenarien und Reihenfolgestrategien*

| | Szenario ohne Störung | Szenario Störung Waschmaschine | Szenario Störung Mangel | Szenario Eilauftrag |
|---|---|---|---|---|
| FCFS | 14:02 | 14:20 | 16:34 | 14:20 |
| SPT | 14:13 | 13:46 | 16:06 | 14:15 |
| LPT | 13:48 | 13:47 | 16:45 | 14:15 |

Der geforderte Soll-Liefertermin lag bei 16:30 Uhr. Es ist erkennbar, dass die Strategien FCFS und SPT den Lieferzeitpunkt unter Normalbedingung sowie unter allen Störeinflüssen einhalten. Mit der Strategie LPT wird außer im Szenario Störung Mangel der geforderte Liefertermin eingehalten. Daraus lässt sich schließen, dass die Störung einer Mangel einen großen Einfluss auf die Prozesszeiten und den Ablauf des Wäschereikreislaufs hat. Die Fertigstellungstermine aller drei Strategien sind bei einer Störung der Mangel um zwei Stunden verzögert. Dieser Zusammenhang zeigt sich auch bei der Betrachtung der durchschnittlichen Durchlaufzeiten, siehe Abbildung 8.

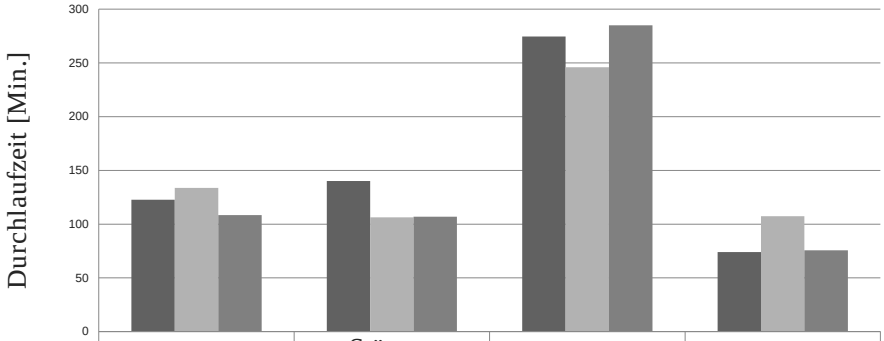
## 3.4 Fazit

Zusammenfassend lässt sich sagen, dass die Produktion einer Wäscherei mit einem mesoskopischen Simulationsmodell abgebildet werden kann. Der Vergleich und die Bewertung der verschiedenen Strategien zur Reihenfolgeplanung haben gezeigt, dass sich alle drei Strategien eignen, jedoch unter Störeinflüssen Verzögerungen auftreten können. Besonders der Ausfall einer Mangel ist in der betrachteten Wäscherei ein großes Problem und führt zu zeitlichen Verzögerungen. Daher sollte der Ausfall einer Mangel schnellst möglich beseitigt bzw. vorab schon verhindert werden.

# 4 Ausblick

In diesem Beitrag wurden zwei Simulationsstudien für Wäschereien vorgestellt. Zukünftig sollen die Ergebnisse der Umstrukturierung des Wareneingangslagers in einer realen Wäscherei umgesetzt werden. Die Simulation der Wäschereiprozesse wird zusätzlich zum mesoskopischen Simulationsansatz noch mit einem ereignis-

diskreten Ansatz durchgeführt. Es sollen außerdem weitere Strategien zur Reihenfolgeplanung getestet sowie weitere Störszenarien entwickelt und hinsichtlich ihrer Auswirkungen bewertet werden. Der Beitrag zeigt, dass der Wäschereikreislauf ein großes Potential zur Verbesserung der Prozessabläufe bietet, das mit Simulationsstudien ausgeschöpft werden kann.




| | Szenario 0 | Störung Waschmaschine | Störung Mangel | Eilauftrag |
|---|---|---|---|---|
| FCFS | 122,67 | 140 | 274,66 | 74 |
| SPT | 133,67 | 106,33 | 246 | 107,33 |
| LPT | 108,33 | 107 | 285 | 75,6 |

*Abbildung 8: Mittelwerte der Durchlaufzeiten der verschiedenen Szenarien und Reihenfolgestrategien*

## Danksagung

Die Inhalte des Beitrags wurden im Kooperationsprojekt „Entwicklung eines Simulationsmodells zur Produktionsplanung und -steuerung von industriellen Großwäschereien“ erarbeitet. Das Projekt wird durch das Förderprogramm „Zentrales Innovationsprogramm Mittelstand“ des Bundesministeriums für Wirtschaft und Technologie (BMWi)-Fördermodul ZIM-KOOP/Fördersäule Kooperationsnetzwerke finanziert. Außerdem danken wir Herrn B.Sc. Marcel Müller für die Erstellung des mesoskopischen Simulationsmodells im Rahmen seiner Bachelorarbeit an der Otto-von-Guericke-Universität Magdeburg.

## Literaturverzeichnis

Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik 2013: Heinz Nixdorf Institut 2013.

Kuhn, A.: Simulation in Produktion und Logistik: Fallbeispielsammlung. Berlin [u.a.]: Springer 1998.

Reggelin, T.: Mesoskopische Modellierung und Simulation logistischer Flusssysteme. Magdeburg: DocuPoint 2011.

Verein Deutscher Ingenieure: VDI-Richtlinie 3633. Simulation von Logistik-, Materialfluss- und Produktionssystemen. Begriffsdefinitionen. In: VDI-Handbuch Materialfluss und Fördertechnik. Berlin: Beuth Verlag 2000.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Prototyp eines simulationsbasierten Assistenzsystems zur Entscheidungsunterstützung bei der Pflege von Planungsparametern eines ERP-Systems im laufenden Betrieb

## *Prototype of a Simulation-based Assistance System to Support Decision Making Regarding the Update of Material Planning Parameters of an ERP System in Real Time*

Ulrike Stumvoll, Thorsten Claus, TU Dresden, Dresden (Germany),
ulrike.stumvoll@mailbox.tu-dresden.de, thorsten.claus@tu-dresden.de

Willi Ertl, Ostbayerische Technische Hochschule Regensburg, Regensburg (Germany), willi.ertl@oth-regensburg.de

Abstract: In an Enterprise Resource Planning (ERP) System production planning is influenced by a variety of parameters. Previous investigations have shown that the setting of parameters is of high relevance. The setting should be checked and adjusted, e. g. after a change of environmental factors. This is done by material planners. In practice it is difficult due to a large number of parameters, materials and other reasons. In this paper a simulation-based assistance system is proposed to update material planning parameters in real-time. The concept offers support to material planners during all phases of decision processes. A prototype of the system will be presented in this paper.

## 1 Einführung

Module zur Materialbedarfsplanung und Disposition sind wichtige Bestandteile von ERP-Systemen. Mit Hilfe einer Vielzahl von Planungsparametern können die in diesen Systemen hinterlegten Algorithmen an die spezifischen Gegebenheiten in einem Unternehmen angepasst werden. Pro Material können z. B. die Parameter Vorlaufzeit, Losgrößenheuristik und die Losgrößenmodifikatoren eingestellt werden, welche die Materialbedarfsplanung beeinflussen. Die korrekte Einstellung der Parameter hat auf die Kennzahlen Terminabweichung, Kapitalbindung und Durchsatz gravierende Auswirkungen. Dies zeigen die von Dittrich et al. (2009) wiedergegebenen und mit Hilfe von Simulation durchgeführten Untersuchungen.

Bei der Einstellung der Dispositionsparameter eines ERP-Systems wird zwischen der Parameterinitialeinstellung und der Parameterpflege oder -optimierung im laufenden Betrieb unterschieden. Die Parameter sind im laufenden Betrieb nach einer wesentlichen Änderung von Umweltfaktoren oder des Produktionssystems, mindestens jedoch einmal jährlich zu prüfen und ggfs. anzupassen (Jodlbauer 2008). Vor jeder Durchführung der Materialbedarfsplanung sollte somit ein Disponent die Entscheidung treffen, ob die Parameter im Hinblick auf das Zielsystem eines Unternehmens anzupassen sind. Das Treffen einer Entscheidung ist dabei kein punktueller Akt sondern ein Vorgang, der sich im Zeitablauf vollzieht (Schiemenz und Schönert 2005). Der Entscheidungsprozess besteht aus den folgenden fünf Phasen: *Anregung, Suche, Auswahl, Vollzug* und *Kontrolle* (Heinen 1985).

In diesem Beitrag wird das Konzept eines *s*imulationsbasierten *A*ssistenzsystems zur *E*ntscheidungsunterstützung bei der *P*flege von *P*lanungsparametern (SAEPP) eines ERP-Systems im laufenden Betrieb aufgezeigt. Ziel dieses Beitrags ist es erstmals den mit Plant Simulation Tecnomatix erstellten Prototypen sowie die Ergebnisse der Anwendung des Konzepts für eine Schneckengetriebeproduktion vorzustellen. Dies erfolgt in Kapitel 4 und 5. Zuvor werden in Kapitel 2 die Herausforderungen für Disponenten bei der Parameterpflege skizziert und in Kapitel 3 das Konzept des Assistenzsystems SAEPP beschrieben. Abgerundet wird der Beitrag durch einen Ausblick in Kapitel 6.

## 2 Herausforderungen für Disponenten

Die Frequenz, in der die Planungsparameter in der Industrie überprüft werden, ist relativ gering (Jonsson und Mattsson 2006). Dies liegt an einer Vielzahl an Herausforderungen. So ist z. B. allein für den Parameter Vorlaufzeit ein Wert im Bereich von 0 bis 999.999 im SAP-System möglich. Zudem wird die Suche nach Handlungsalternativen z. B. dadurch erschwert, dass ein Disponent in einem Betrieb üblicherweise für mehr als 500 Artikel verantwortlich ist (Gulyássy et al. 2009), aber auch dadurch, dass zwischen den einzelnen Parametern nichtlineare Effekte, positive Verbund- und schädliche Nebenwirkungen auftreten können (Mertens et al. 1991). Aufgrund der bestehenden Herausforderungen werden in vielen Unternehmen die Parameter, nachdem sie einmal eingestellt wurden, meist nicht mehr angepasst.

In der Literatur wurden bereits Systeme zur Unterstützung von Disponenten vorgestellt. Ein wissensbasiertes Verfahren wurde z. B. von Wedel (1990) und Hartinger (1995) entwickelt. Der Ansatz, dass PPS-Systeme selbst günstigere Einstellungen mit Hilfe eines genetischen Algorithmus und einer Proberechnung ermitteln, wurde von Kernler (1994) vorgestellt. Auf dem Markt vorhandene ERP-Systeme, wie z. B. das SAP-System, bieten jedoch keine Unterstützung bei der Parameteroptimierung (Gulyássy et al. 2009).

Die bestehenden Ansätze berücksichtigen im Rahmen der Willensbildung des Entscheidungsprozesses nur unzureichend die Kosten, die durch eine Änderung einer Parametereinstellung entstehen. Zudem wurde die Unterstützung der Phase Kontrolle, in welcher die Zielerreichung der getroffenen Entscheidung bestimmt wird, bisher nicht betrachtet. Durch das Assistenzsystem SAEPP wird ein Disponent in allen Phasen des betrieblichen Entscheidungsprozesses bei der Pflege der Planungsparameter eines ERP-Systems im laufenden Betrieb unterstützt.

## 3 Konzept des simulationsbasierten Assistenzsystems

Einen Überblick über die einzelnen Bestandteile des Assistenzsystems SAEPP, welches an ein ERP-System angeschlossen wird, gibt die nachstehende Abbildung 1. Das System wurde, insbesondere in der Phase Suche und Kontrolle, gegenüber Stumvoll et al. (2013) weiterentwickelt.

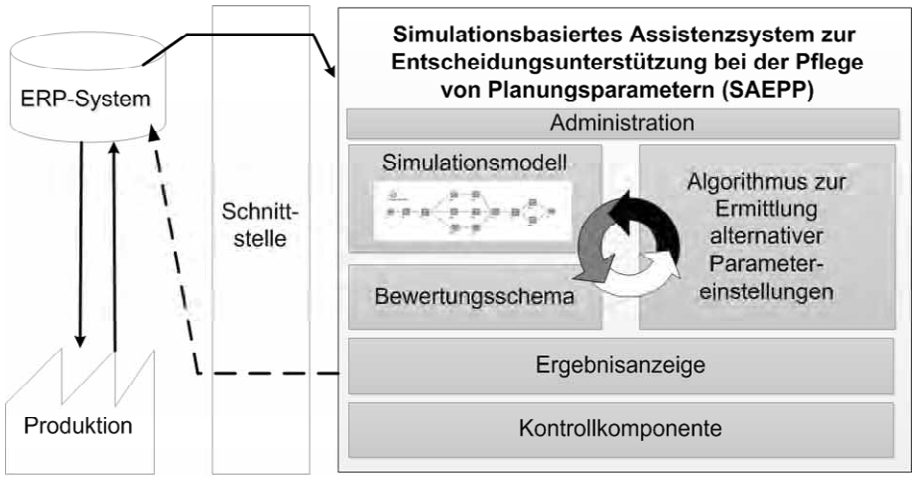



*Abbildung 1: Bestandteile des simulationsbasierten Assistenzsystems SAEPP*

In der Komponente Administration des Systems SAEPP ist das Zielsystem eines Unternehmens zu hinterlegen. Das Zielsystem ist eine zentrale Einflussgröße des Entscheidungsprozesses (Heinen 1985). Von Wiendahl (2008) stammt das Zielsystem der Produktionslogistik. Dieses beinhaltet die Größen hohe Termintreue, niedrige Durchlaufzeit, hohe Auslastung und niedrige Bestände sowie im Kern die Wirtschaftlichkeit. Aus diesem Grund können in der Administration des Systems SAEPP auch weitere finanzielle Größen, auf welche die Einstellung der Parameter einen Einfluss hat, wie z. B. die Anzahl an Transportvorgängen hinterlegt werden.

Durch das jeweilige Unternehmen ist ein glaubwürdiges Simulationsmodell zur Verfügung zu stellen. Im Rahmen der Phase Anregung des Entscheidungsprozesses wird die aktuelle Unternehmenssituation in das Simulationsmodell, welches in das System SAEPP eingebunden ist, übernommen. In diesem Modell wird anschließend ein Probebetrieb, unter Verwendung der aktuellen Parametereinstellung, durchgeführt. Dadurch werden die zukünftigen Auswirkungen auf die einzelnen Größen des Zielsystems ermittelt.

Daran anschließend werden die Auswirkungen für die Unterlassungsalternative, im sogenannten Bewertungsschema, bewertet. Für die Zielgrößen Auslastung und Bestand ist es möglich, vorausgesetzt dass entsprechende Kostensätze vorliegen, die Auswirkungen monetär zu bewerten. Die Umwandlung der Termintreue in Fehlmengenkosten ist jedoch hochbrisant. (Kernler 1994) Auch die monetäre Bewertung der Durchlaufzeit ist schwierig. Aus diesem Grund wird im Rahmen des Bewertungsschemas die Kosten-Wirksamkeits-Analyse eingesetzt. Bei dieser Methode wird die Nutzwertanalyse nur auf die nicht in Geldeinheiten (GE) quantifizierbaren

Kriterien angewendet, wodurch eine Aussage über die Wirksamkeit einer Lösung vorliegt. Kriterien, welche monetär bewertet werden können, gehen unverändert in die Bewertung ein. Durch Division werden die Kosten- und Wirtschaftlichkeitszahlen zueinander ins Verhältnis gesetzt, wodurch sich die Kosten-Wirksamkeits-Kennziffer der Alternative ergibt. (Schulte-Zurhausen 2014)

In der Phase Suche des Entscheidungsprozesses werden, ausgehend von der aktuellen Parametereinstellung, Handlungsalternativen erzeugt. Für das Assistenzsystem SAEPP wurde ein eigener Algorithmus zur Ermittlung alternativer Parametereinstellungen konzipiert. Dieser besteht aus zwei Teilen. Im regelbasierten Teil 1 wird für einen Parameter eine Startlösung erzeugt. In Teil 2 des Algorithmus wird, unter Verwendung der Heuristik Simulated Annealing, der Einstellungsvorschlag für den betrachteten Parameter weiter verbessert. Diese Vorgehensweise wurde bereits von Stumvoll et al. (2013) vorgestellt. Wird durch den Algorithmus eine alternative Einstellungskombination erzeugt, so werden mit Hilfe von Simulation wieder die zukünftigen Auswirkungen bestimmt.

Nach Durchführung des Simulationslaufes werden für jede Alternative, im Rahmen des Bewertungsschemas, Kosten- und Wirtschaftlichkeitszahlen ermittelt und die zugehörigen Kennziffern berechnet. Die Anzahl der Replikationen sowie die Vorgehensweise bei der Berechnung der einzelnen Kennzahlen, werden durch das jeweilige Unternehmen definiert. Damit die Nutzwertanalyse durchgeführt werden kann, ist eine Skala zur Bewertung der einzelnen Alternativen erforderlich. Im Rahmen des Assistenzsystems SAEPP wird die beste bisher gefundene Lösung für eine nicht monetär bewertbare Zielgröße mit 10 Punkten und die schlechteste Lösung mit 1 Punkt bewertet. Für Lösungen dazwischen wird ein anteiliger Punktwert bestimmt. Bei der Bestimmung der Kostenzahlen können auch einmalige Kosten, wie in Kapitel 4 näher ausgeführt, berücksichtigt werden. Durch die Berücksichtigung von Strafkosten, deren Höhe vom letzten Zeitpunkt der Änderung der Einstellung der Parameter abhängt, kann z. B. vermieden werden, dass eine Planungsnervosität eintritt.

Die Phase Suche ist beendet, wenn eine Abbruchbedingung des Algorithmus zur Ermittlung alternativer Parametereinstellungen erfüllt ist. Im Rahmen der anschliessenden Phase Auswahl wird die Alternative mit der kleinsten Kosten-Wirksamkeits-Kennziffer bestimmt. Hat die Unterlassungsalternative die kleinste Kennziffer, so liegt zum aktuellen Zeitpunkt kein Handlungsbedarf bei der Pflege der Planungsparameter vor. Andernfalls liegt ein Verbesserungspotenzial und damit Handlungsbedarf vor. Dies wird einem Disponenten in der Ergebnisanzeige des simulationsbasierten Assistenzsystems SAEPP angezeigt. Zur Visualisierung des Vorliegens eines Handlungsbedarfs wird eine Ampelgrafik eingesetzt.

Auf Basis der in der Ergebnisanzeige dargestellten Informationen trifft ein Disponent die Entscheidung, ob die vorgeschlagene Alternative in das ERP-System übernommen wird. Für ein Assistenzsystem ist es kennzeichnend, dass eine Bestätigung durch einen Anwender erforderlich ist, bevor eine Handlungsalternative automatisch ausgeführt wird (Hauß und Timpe 2002). Liegt diese vor, so wird die vorgeschlagene Alternative im Rahmen der Phase Vollzug in das ERP-System übernommen, wozu die in Abbildung 1 dargestellte Schnittstelle genutzt wird.

Das simulationsbasierte Assistenzsystems SAEPP wird eine definierte Zeitspanne vor jeder Ausführung der Planungsverfahren im ERP-System gestartet. Nach dem

Start wird sofort die Zielerreichung, der bei der letzten Ausführung getroffenen Entscheidung, überprüft. Dazu wird ein Soll-Ist-Vergleich unter Verwendung der Methode Validierung von Vorhersagen eingesetzt. Wird im Rahmen dieser Kontrollphase eine relevante Abweichung festgestellt, so sind zunächst die Ursachen mit Hilfe einer Abweichungsanalyse zu bestimmen.

## 4 Prototyp des Assistenzsystems SAEPP

Das Konzept des Assistenzsystems SAEPP wurde exemplarisch für die Parameter der programmorientierten Material- und Losgrößenplanung von eigengefertigtem Material, in einem Prototyp realisiert. Dieser Prototyp umfasst die in Abbildung 1 dargestellten Bestandteile.

Als Simulationsmodell wird in dem Prototyp das Werkstattsimulationsmodell des Innovationszentrums für Produktionslogistik und Fabrikplanung (IPF) der Ostbayerischen Technischen Hochschule verwendet. Das Simulationsmodell beinhaltet einerseits die Planungsverfahren der Phase der Mengenplanung eines ERP-Systems und eine Fertigungssteuerung, andererseits beinhaltet das Modell auch eine Nachbildung einer Produktion, sodass ein Probebetrieb ermöglicht wird. (Herrmann und Stumvoll 2010) Zu dem mit Plant Simulation realisiertem Modell wurden das Bewertungsschema und der Algorithmus zur Ermittlung alternativer Parametereinstellungen des Assistenzsystems SAEPP hinzugefügt. Die vorgenommene Weiterentwicklung umfasst ca. 2.000 Quellcodezeilen. Während der Entwicklung wurden u. a. die Verifikations- und Validierungstechniken, Test von Teilmodellen, Dimensionstest und Schreibtischtest eingesetzt. Zudem wurde der Ablauf in Aktivitätsdiagrammen dokumentiert.

Die Realisierung der Administration, Ergebnisanzeige und Kontrollkomponente ist nicht im Werkstattsimulationsmodell des IPF erfolgt. In der Software Plant Simulation wäre es sehr schwierig und zeitaufwändig gewesen, eine ansprechende Bedienoberfläche zu schaffen. Deshalb wurde für diese Bestandteile des Systems SAEPP Microsoft Excel verwendet. Da Plant Simulation Dynamic Data Exchange (DDE) unterstützt, ist die Anbindung von Dateien im .xlsx-Format an das Modell des IPF möglich. So kann das Simulationsmodell aus Excel heraus gesteuert werden. Zudem können aus dem Modell Daten in eine Excel Datei geschrieben oder von dieser ausgelesen werden. (Bangsow 2008)

Im nachstehenden Kapitel 5 wird die Anwendung des erstellten Prototyps für ein Fallbeispiel aus der Literatur beschrieben. In diesem Beispiel stehen somit kein reales ERP-System und keine Produktion zur Verfügung. Es wird daher stellvertretend wieder das Werkstattsimulationsmodell des IPF verwendet.

## 5 Ergebnisse der Anwendung des Systems SAEPP für eine Schneckengetriebeproduktion

In diesem Kapitel wird das verwendete Fallbeispiel beschrieben. Anschließend wird aufgezeigt, welche Einstellungen in der Administration des Systems SAEPP vorgenommen worden sind.

## 5.1 Beschreibung des verwendeten Fallbeispiels

Das in diesem Beitrag verwendete Fallbeispiel basiert auf einer Dokumentation von Prêt (2015), welcher wiederum Daten u. a. von der Firma G.U.N.T. Gerätebau GmbH Hamburg erhalten hat. Schneckengetriebe werden z. B. in Fahrantrieben von Kränen, Wischerantrieben von Kraftfahrzeugen oder Rolltreppen eingesetzt. Ein Schneckengetriebe umfasst im Beispiel von Prêt (2015), die in nachstehender Abbildung 2 dargestellten, eigengefertigten Materialien.

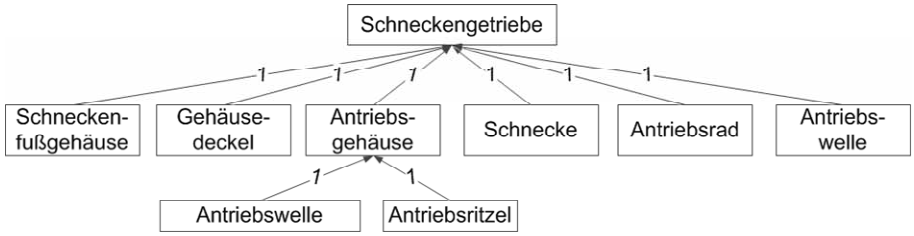



*Abbildung 2: Stückliste des Fallbeispiels Schneckengetriebeproduktion*

In der Ausarbeitung von Prêt (2015) ist, mit Ausnahme von dem Schneckengetriebe, auch dokumentiert, an welchen Arbeitsplätzen die einzelnen Produkte bearbeitet werden und wie dabei die Rüst- und Bearbeitungszeiten sind. Für die Montage des Getriebes wurden entsprechende Annahmen getroffen. Die betrachtete Fertigung besteht aus zehn unterschiedlichen Arten von Betriebsmitteln. In der Produktion sind die 19 vorhandenen Maschinen nach dem Werkstattprinzip angeordnet.

Ebenfalls sind in der Dokumentation von Prêt (2015) die Rüstkosten pro Material sowie der Zinssatz der Lagerung angegeben. Diese Kostendaten werden u. a. zur Bewertung der Auswirkungen einer Parametereinstellung im Assistenzsystem benötigt (siehe Kapitel 5.2). Zudem kann aus der Ausarbeitung entnommen werden, von welcher Fertigungslosgröße ausgegangen wird. Diese Information wird für die Einstellung der Planungsparameter in der Ausgangssituation verwendet. Da aus dem Beispiel von Prêt (2015) keine Information über die Einstellung des Parameters Vorlaufzeit entnommen werden kann, wird für jedes Material eine der Fertigungslosgröße angepasste Annahme getroffen.

Die von Prêt (2015) beschriebene Fertigung ist auf eine Produktion von 33.000 bis 35.000 Schneckengetrieben pro Jahr ausgelegt. In der Ausgangssituation werden Kundenaufträge erzeugt, durch welche die Produktion nicht vollständig ausgelastet ist. Zum Zeitpunkt $t_1$ wird die Höhe der eingehenden Kundenaufträge deutlich angehoben. Dies stellt eine Veränderung der Umweltfaktoren dar, welche dazu führen kann, dass die Einstellung der Parameter anzupassen ist (siehe Kapitel 1). Durch die Anwendung des simulationsbasierten Assistenzsystems SAEPP wird ein Disponent bei der Entscheidung, ob die Parameter zum aktuellen Zeitpunkt anzupassen sind, unterstützt.

## 5.2 Einstellungen in der Administration des Systems SAEPP

Vor der ersten Ausführung des simulationsbasierten Assistenzsystems SAEPP sind Informationen über den Zeitpunkt des Planungslaufes des ERP-Systems im jewei-

ligen Betrieb in der Administration zu hinterlegen. Zudem sollte die initiale Konfiguration des Algorithmus zur Ermittlung alternativer Parametereinstellungen kontrolliert werden. Ebenfalls sollten die Schwellenwerte der Ampelgrafik, d. h. ab wann eine rote bzw. gelbe Ampel angezeigt wird, geprüft werden. Im Fallbeispiel wird eine rote Ampel bei einem Verbesserungspotenzial von > 25 % angezeigt. Darüber hinaus ist in der Administration, wie in Kapitel 3 erwähnt, das Zielsystem zu hinterlegen. Die nachstehende Abbildung 3 zeigt die Einstellungen im Abschnitt Zielsystem.

| | | Gewichtung |
|---|---|---|
| Zielgrößen | Verspätung Mittelwert | 30 |
| | Verspätung Standardabweichung | 30 |
| | Durchlaufzeit Mittelwert | 20 |
| | Durchlaufzeit Standardabweichung | 20 |
| | | ∑ = 100 |
| finanzielle Zielgrößen | Bestand | 0,08 GE / pro Stück pro Periode gelagert |
| | Maschinenlaufzeit | 33,68 GE / pro Stunde |
| | Rüstkosten | 32,64 - 104,86 GE / Rüstvorgang (poduktspezifisch) |
| | Transportkosten | 2,00 GE / Transportvorgang |
| | Dokumentationskosten | 6,00 GE / Rüstvorgang |
| | Kosten für Qualitätsprüfungen | 42,00 GE / Rüstvorgang |
| | Transporthilfsmittel erforderlich | falls ja, 50 Stück für 5.000,00 GE |
| | Mitarbeiterinformation erforderlich | falls ja, einmalig 800,00 GE |
| | Lieferanteninformation erforderlich | falls ja, einmalig 500,00 GE |
| | Strafkosten (Planungsnervosität) | nach vier Ausführungszeitpunkten 2.500,00 GE |

*Abbildung 3: Hinterlegtes Zielsystem in der Administration im Fallbeispiel*

Die Auswirkungen auf die Größen Verspätung und Durchlaufzeit werden, aufgrund der problematischen Umwandlung, nicht monetär bewertet (siehe Kapitel 3). Für diese Kennzahlen wird, wie in Abbildung 3 dargestellt, jeweils der Mittelwert und die Standardabweichung bestimmt. In diesem Fallbeispiel wird angenommen, dass dem Unternehmen eine hohe Termintreue wichtiger als eine geringe Durchlaufzeit ist. Dies wird bei der Nutzwertanalyse durch unterschiedliche Gewichtungsfaktoren berücksichtigt.

Bei der Bewertung einer Alternative werden zudem die Auswirkungen auf finanzielle Zielgrößen, wie z. B. Bestand, Maschinenlaufzeit und Rüstkosten berücksichtigt. Die durch die Simulation gewonnenen Produktablaufinformationen werden dazu mit den hinterlegten Kostendaten verknüpft, d. h. es erfolgt eine nachgelagerte Kostensimulation. (VDI3633-7) Die in Abbildung 3 abgebildeten Kostensätze wurden, soweit vorhanden, aus Prêt (2015) übernommen. Andernfalls wurden entsprechende Annahmen getroffen. Es sei erwähnt, dass diese Sätze mit Hilfe der Maschinenstundensatz- bzw. Prozesskostenrechnung bestimmt werden können.

Zur Vermeidung einer Planungsnervosität können, wie in Kapitel 3 erwähnt, in der Administration auch Strafkosten hinterlegt werden. Im Fallbeispiel wurde darüber hinaus definiert, dass z. B. bei einer deutlich gestiegenen Anzahl an Transportvorgängen Kosten für die Beschaffung von weiteren Transporthilfsmitteln anfallen. Diese einmaligen Kosten werden bei der Kosten-Wirksamkeits-Analyse ebenfalls berücksichtigt.

## 5.3 Angezeigte Ergebnisse des Systems SAEPP

Für das in Kapitel 5.1 skizzierte Beispiel wurde das System SAEPP, mit dem in Kapitel 5.2 beschriebenem Zielsystem, angewendet. Die in den Phasen Anregung, Suche und Auswahl ermittelten Informationen werden einem Disponenten in der Ergebnisanzeige des Assistenzsystems angezeigt. Diese beinhaltet eine Ampelgrafik, die angibt, ob ein Handlungsbedarf bei der Pflege der Parameter zum aktuellen Zeitpunkt vorliegt. Die nachstehende Abbildung 4, welche ebenfalls Bestandteil der Ergebnisanzeige ist, zeigt das Bewertungsschema für die genannten Alternativen. Im oberen Bereich der Abbildung ist das Ergebnis der Nutzwertanalyse dargestellt. Das Ergebnis der Kostensimulation sowie die sich daraus ergebende Kosten-Wirksamkeits-Kennziffer sind im unteren Bereich dargestellt.

| | | Unterlassungsalternative | | | Günstigste Alternative | | |
|---|---|---|---|---|---|---|---|
| | | Ergebnis | Punkte | Gewichtet | Ergebnis | Punkte | Gewichtet |
| Zielgrößen | Verspätung Mittelwert | 26,63 | 3,81 | 114,43 | 9,70 | 10,00 | 300,00 |
| | Verspätung Standardabweichung | 4,36 | 1,00 | 30,00 | 1,38 | 9,47 | 284,20 |
| | Durchlaufzeit Mittelwert | 15,69 | 1,46 | 29,28 | 6,13 | 10,00 | 200,00 |
| | Durchlaufzeit Standardabweichung | 9,44 | 3,31 | 66,26 | 3,36 | 10,00 | 200,00 |
| | | 239,97 | | | 984,20 | | |
| finanzielle Zielgrößen | Bestand | 417.770 | 33421,60 GE | | 141.102 | 11288,16 GE | |
| | Maschinenlaufzeit | 11.741 | 395436,04 GE | | 11.090 | 373521,58 GE | |
| | Rüstkosten | 46 | 2298,99 GE | | 89 | 4747,05 GE | |
| | Transportkosten | 185 | 370,00 GE | | 402 | 804,00 GE | |
| | Dokumentationskosten | 46 | 276,00 GE | | 89 | 534,00 GE | |
| | Kosten für Qualitätsprüfungen | 46 | 1932, 00 GE | | 89 | 3738,00 GE | |
| | Transporthilfsmittel erforderlich | 0 | 0,00 | | 0 | 0,00 | |
| | Mitarbeiterinformation erforderlich | 0 | 0,00 | | 1 | 800,00 GE | |
| | Lieferanteninformation erforderlich | 0 | 0,00 | | 1 | 500,00 GE | |
| | Strafkosten (Planungsnervosität) | 0 | 0,00 | | 1 | 2500,00 GE | |
| | | 433.734,63 GE | | | 398.432,79 GE | | |
| | Kosten-Wirksamkeits-Kennziffer | 1.807,48 | | | 404,83 | | |

*Abbildung 4: Ergebnis der Anwendung des simulationsbasierten Assistenzsystems*

Durch den Algorithmus zur Ermittlung alternativer Parametereinstellungen wurden in diesem Beispiel 107 Alternativen erzeugt. Bei der günstigsten Alternative hat sich, mit einem Wert von 404,83 die kleinste Kosten-Wirksamkeit-Kennziffer ergeben. Dieser Wert ist deutlich geringer als der von der Unterlassungsalternative mit einem Wert von 1.807,48. Zum Zeitpunkt $t_1$ wird dem Disponenten somit ein Handlungsbedarf bei der Pflege der Parameter und damit eine rote Ampel angezeigt.

Der Anstieg der Kundenaufträge zum Zeitpunkt $t_1$ hat zur Folge, dass sich bei Beibehaltung der aktuellen Parametereinstellung im Vorhersagezeitraum eine Verspätung von 26,63 Tagen (Mittelwert) bei einer Standardabweichung von 4,36 Tagen ergeben würde. Bei der günstigsten Alternative ergibt sich ein Mittelwert von 9,70 Tagen und eine Standardabweichung von 1,38 Tagen. Der Mittelwert der Durchlaufzeit beträgt 6,13 Tage und die Standardabweichung 3,36 Tage. Mit Ausnahme von der Standardabweichung bei der Verspätung sind diese Ergebnisse die besten ermittelten Lösungen. Dies geht aus Abbildung 4 hervor. Wie in Kapitel 3 definiert, wird die beste Lösung jeweils mit 10 Punkten bewertet.

Aus Abbildung 4 ist u. a. auch ersichtlich, welche Laufzeit und Prozessmengen sich für die Zielgrößen, welche monetär bewertet werden, im Vorhersagezeitraum

ergeben haben. Bei der Laufzeit der Maschinen besteht zwischen der aktuellen Parametereinstellung und der günstigsten Alternative nur ein geringer Unterschied. Der Bestand ist hingegen bei der günstigsten Alternative im betrachteten Zeitraum insgesamt deutlich kleiner, wodurch sich die Lagerkosten sehr stark reduzieren. Dies ist darauf zurückzuführen, dass bei dieser Alternative deutlich kleinere Lose gebildet werden. Die Bildung von kleineren Losen hat jedoch auch zur Folge, dass fast doppelt so viele Rüstvorgänge durchgeführt werden müssen, wodurch diese Kosten steigen. Neben diesen haben sich auch die Kosten für Dokumentationen und Qualitätsprüfungen erhöht. Ebenfalls gestiegen ist die Anzahl an Transporten, was wiederum auf die kleinere Losgröße zurückzuführen ist. Obwohl zu den Kosten der günstigsten Alternative Kosten für die Information von Mitarbeitern und Lieferanten sowie Strafkosten addiert werden, sind die Kosten geringer als bei der aktuellen Parametereinstellung. Die Gesamtkosten der Unterlassungsalternative betragen 433.734,63 GE, wohingegen sich die der günstigsten Alternative nur auf 398.432,79 GE belaufen. Da die günstigste Alternative im Vorhersagezeitraum beim Nutzwert und bei den Kosten besser ist als die Unterlassungsalternative, ergibt sich ein kleinerer Wert für die Kosten-Wirksamkeits-Kennziffer. In diesem Beispiel beträgt das Verbesserungspotenzial 77,06 %.

In der Ergebnisanzeige des simulationsbasierten Assistenzsystems SAEPP wird zudem der zur günstigsten Alternative gehörende Einstellungsvorschlag angezeigt. Gegenüber der Unterlassungsalternative wurde vom Algorithmus zur Ermittlung alternativer Parametereinstellungen die Einstellung der Losgrößenheuristik geändert. Bei allen Produkten wird die Verwendung des Groff-Verfahrens empfohlen. Zudem sollte die Einstellung der Vorlaufzeit bei allen Produkten, mit Ausnahme des Gehäusedeckels, deutlich reduziert werden. Durch den Algorithmus zur Ermittlung alternativer Parametereinstellungen wurde auch die Einstellung des Parameters Minimale Losgröße betrachtet. Zum Zeitpunkt $t_1$ konnte durch diesen Parameter keine weitere Verbesserung der Kosten-Wirksamkeits-Kennziffer erreicht werden.

## 6 Ausblick

Der Prototyp, der in diesem Beitrag vorgestellten wurde, zeigt, dass es möglich ist, das Konzept des simulationsbasierten Assistenzsystems SAEPP in Software abzubilden. Auch wurde gezeigt, dass der Prototyp lauffähig ist und nachvollziehbare Ergebnisse liefert. Dies wird durch die Durchführung von weiteren Fallbeispielen weiter verifiziert. Durch das simulationsbasierte Assistenzsystem SAEPP werden Disponenten in allen Phasen des Entscheidungsprozesses bei der Pflege der Parameter im laufenden Betrieb unterstützt. Aufgrund der hohen wirtschaftlichen Relevanz der Parameter ist der Einsatz eines Hilfsmittels zur Entscheidungsunterstützung zielführend.

Weiterer Forschungsbedarf besteht bei multikriteriellen Optimierungsmodellen, insbesondere in Situationen, bei denen die eingeführte Meta-Zielfunktion monetär und nicht monetär bewertbare Zielgrößen umfasst.

## Literatur

Bangsow, S.: Fertigungssimulationen mit Plant Simulation und SimTalk. München: Hanser 2008.

Dittrich, J.; Mertens, P.; Hau, M.; Hufgard, A.: Dispositionsparameter in der Produktionsplanung mit SAP. 5. Auflage, Wiesbaden: Vieweg+Teubner 2009.

Gulyássy, F.; Hoppe, M.; Isermann, M.; Köhler, O.: Disposition mit SAP. Bonn: Galileo Press 2009.

Hartinger, M.: Die Pflege der Parameter von Standardsoftware. Wiesbaden 1995.

Hauß, Y.; Timpe, K.-P.: Automatisierung und Unterstützung im Mensch-Maschine-System. In: Timpe, K.-P.; Jürgensohn, T.; Kolrep, H. (Hrsg.): Mensch-Maschine-Systemtechnik. Düsseldorf: Symposion 2002, S. 41-62.

Heinen, E.: Einführung in die Betriebswirtschaftslehre. 9. Auflage, Wiesbaden: Gabler 1985.

Herrmann, F.; Stumvoll, U.: Überwachung und Verbesserung von Losgrößenmodifikatoren eines PPS-Systems im laufenden Betrieb. In: Claus, T.; Herrmann, F. (Hrsg.): Simulation als betriebliche Entscheidungshilfe. Zittau: Internationales Hochschulinstitut Zittau 2010, S. 45-64.

Jodlbauer, H.: Produktionsoptimierung. 2. Auflage, Wien: Springer 2008.

Jonsson, P.; Mattsson, S.-A.: A longitudinal study of material planning applications in manufacturing companies. In: International Journal of Operations & Production Management 26 (2006) 9, S. 971-995.

Kernler, H.: Programme werden selbstständig. Logistik Heute 10 (1994), S. 101-102.

Mertens, P.; Wedel, T.; Hartinger, M.: Management by Parameters?. Zeitschrift für Betriebswirtschaft 61 (1991) 5/6, S. 569-588.

Prêt, U.: Teilefertigungs- und Montageprojekt „Schneckengetriebeproduktion“, www.uwe-pret.de/fabrikplanung/skripte/stapel3/getriebe.pdf, 2015, Zuletzt geprüft am 7.3.2015.

Schiemenz, B.; Schönert, O.: Entscheidung und Produktion. 3. Auflage. München, Wien: Oldenbourg 2005.

Schulte-Zurhausen, M.: Organisation. 6. Auflage. München: Vahlen 2014.

Stumvoll, U.; Nehls, U.; Claus, T.: Ein simulationsbasiertes Assistenzsystem zur Pflege von Dispositionsparametern eines ERP-Systems im laufenden Betrieb. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik 2013. Paderborn: Heinz-Nixdorf-Inst. Univ. Paderborn 2013, S. 569-578.

VDI Norm 3633 Blatt 7: Simulation von Logistik-, Materialfluss und Produktionssystemen, Kostensimulation. Verein Deutscher Ingenieure: Düsseldorf, 2001.

Wedel, T.: Wissensbasierte dynamische Konfiguration der Parameter von Standardpaketen zur Produktionsplanung und -steuerung am Beispiel des IBM-Modularprogramms COPICS. Universität Erlangen-Nürnberg 1990.

Wiendahl, H.-P.: Betriebsorganisation für Ingenieure, 6. Auflage, München: Hanser 2008.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Bessere Prioritätsregeln für komplexe Produktionssysteme mittels multi-kriterieller simulationsbasierter Optimierung

## *Improved Dispatching Rules for Complex Manufacturing Systems Using Multi-objective Simulation-based Optimisation*

Torsten Hildebrandt, Michael Freitag, BIBA – Bremer Institut für Produktion und Logistik GmbH an der Universität Bremen, Bremen (Germany), hil@biba.uni-bremen.de, fre@biba.uni-bremen.de

Abstract: This document describes the use of simulation-based, multi-objective optimisation (multi-objective Genetic Programming) to automatically develop improved dispatching rules. As the result of the optimisation is a new scheduling heuristic this can be categorized as a hyper-heuristic approach. For a complex manufacturing scenario from semiconductor manufacturing we show that resulting rules clearly outperform state-of-the-art rules from the literature. As a further novelty we demonstrate that the method allows to explicitly and effectively search for rules that are both good and short. Using rule length as an additional objective dimension we are able to find rules that are easy to understand and can also quantify the trade-offs possible between rule quality and rule complexity.

## 1 Einleitung

Im Rahmen dieses Beitrags wird die multi-kriterielle simulationsbasierte Optimierung eingesetzt, um verbesserte Prioritätsregeln für die operative Reihenfolgeplanung für ein komplexes Szenario aus dem Bereich der Halbleiterfertigung zu erstellen. Dieses ist gekennzeichnet durch: über 200 Maschinen; bis zu 355 Operationen, um ein Produkt zu fertigen; reihenfolgeabhängige Rüstzeiten; Batchprozesse; zyklische Materialflüsse. Im Unterschied zu den meisten Vorarbeiten aus der Literatur, können mittels multi-kriterieller Optimierung nun Regeln gefunden werden, die sowohl Kriterien wie die Durchlaufzeit als auch Termintreue verbessern und es außerdem erlauben, die Regellänge als ein weiteres Zielkriterium zu betrachten. Somit wird es möglich, explizit nach guten und kurzen Regeln sowie den dazwischen möglichen Abstufungen zu suchen. Diese Erweiterung bietet einen deutlichen Mehrwert für den Nutzer, da als Optimierungsergebnis eine Lösungs*menge* resultiert, in der jede darin enthaltene Regel einen in sich optimalen Kompromiss bspw. zwischen den größtenteils konfliktären Zielgrößen Durchlaufzeit und Termintreue darstellt. An-

hand dieser Ergebnismenge kann dann durch den Nutzer auf Basis der Menge möglicher Kompromisslösungen eine fundierte Entscheidung getroffen werden, welche Regel am besten seinen Präferenzen entspricht und den für ihn besten Kompromiss in der Zielerreichung darstellt.

Zunächst wird kurz das Problem der Reihenfolgeplanung für komplexe Produktionssysteme vorgestellt, wie es beispielsweise in der Halbleiterfertigung auftritt. Darauf folgt eine Vorstellung des Lösungsansatzes zur Kopplung der multi-kriteriellen Genetischen Programmierung (GP) als verwendetem Optimierungsverfahren mit der Simulation. Nach einer knappen Beschreibung des betrachteten Szenarios und des genauen Experimentaufbaus erfolgt eine Darstellung und Diskussion der Ergebnisse.

# 2 Stand der Technik

## 2.1 Automatischer Entwurf von Prioritätsregeln

Prioritätsregeln zur operativen Reihenfolgeplanung (vgl. Haupt 1989) sind ein weit verbreitetes Verfahren in Softwarelösungen zur Fertigungsfeinplanung. Immer, wenn eine Ressource frei wird und noch Aufträge auf Bearbeitung warten, wird jeder Auftrag mittels einer Prioritätsregel bewertet. Der Auftrag mit dem höchsten Prioritätswert wird als nächstes gestartet. Zur Beliebtheit von Prioritätsregeln trägt bei, dass sie einfach zu verstehen und zu implementieren sind. Darüber hinaus stellen sie nur minimale Rechenzeitanforderungen, so dass sie auch als Echtzeit/Online-Verfahren eingesetzt werden können und somit ihre Entscheidungen auf Basis der jeweils aktuellsten verfügbaren Daten treffen können. Dies ist wichtig zur geeigneten Reaktion auf unvorhergesehene Störungen (z.B. Maschinenausfälle). Beispiele für einfache Standardregeln sind FIFO (First In (buffer), First Out) oder die Kürzeste-Operationszeit-Regel (engl. SPT, Shortest Processing Time first). Simulationssoftware als auch die meisten MES (Manufacturing Execution System) mit Feinplanungsfunktion enthalten meist eine Reihe an Standardregeln, aus denen man die für die jeweilige Problemstellung beste auswählen kann.

Um darüber hinaus gehende Verbesserungen zu erreichen, wird oft auf einen manuellen Entwurfsprozess zurückgegriffen, in dem ein Systemexperte sich auf Basis seiner Erfahrungen und den Simulationsergebnissen mit Standardregeln neue Regeln ausdenkt. Diese werden implementiert und mittels Simulation bewertet. Dies führt zu einem zyklischen Verbesserungsprozess, an dessen Ende eine verbesserte Regel steht. Um diesen sehr zeit- und kostenaufwendigen, größtenteils manuellen Entwurfsprozess zu verbessern, werden zunehmend Verfahren vorgeschlagen, die automatisiert Prioritätsregeln erstellen. Eine aktuelle Übersicht über diese Arbeiten findet sich in Branke et al. (2015). Die meisten dieser Verfahren verwenden die Genetische Programmierung (Poli et al. 2008) als sogenannte Hyper-Heuristik. Hyper-Heuristiken (Burke et al. 2013) sind ein relativ neues Forschungsfeld, insbesondere auch im Zusammenhang mit Simulation. Hierin werden Verfahren untersucht, bei denen eine Meta-Heuristik (wie die Genetische Programmierung) nicht eingesetzt wird, um direkt ein Problem zu lösen (z. B. ein Reihenfolgeplanungsproblem mit konkreten bekannten Aufträgen). Stattdessen stellt das Ergebnis der Hyper-Heuristik eine neue Heuristik dar, mit der die zugrunde liegende Problemstellung gelöst werden kann.

## 2.2 Multi-kriterielle Optimierung von Prioritätsregeln

Klassische Ansätze zur simulationsbasierten Optimierung gehen davon aus, dass sich die zu minimierende bzw. zu maximierende Zielgröße als ein einzelner skalarer Wert ausdrücken lässt. Für viele Fragestellungen stellt dies jedoch eine starke Vereinfachung dar. Die multi-kriterielle Optimierung (Marler und Arora 2004) ermöglicht auch in solchen Fällen eine Lösung. Ohne Beschränkung auf eine einzelne Zielgröße bzw. ohne die a-priori Festlegung von Gewichtungen für die einzelnen (Teil-)Zielgrößen, lassen sich hierzu insbesondere die auf dem Konzept der Pareto-Optimalität beruhenden Ansätze verwenden, die in den letzten Jahren zunehmend an Popularität gewinnen (Coello Coello 2006). Hierbei ist die Lösung eines Optimierungslaufs nicht eine einzelne „optimale“ Lösung, sondern stattdessen eine Pareto-Front, also eine Menge an Lösungen, die alle einen bestimmten erreichbaren Kompromiss zwischen den Teilzielgrößen darstellen. Form und Lage dieser Lösungsmenge aus jeweils nicht-dominierten Lösungen stellen für den Nutzer wichtige Informationen dar, um aus dieser Menge an möglichen Kompromiss-Lösungen die für seine Präferenzen beste auszuwählen.

In der Literatur konnten wir nur zwei Arbeiten identifizieren, die ebenfalls multi-kriterielle Optimierung im Kontext der Erstellung von Prioritätsregeln einsetzen. Tay und Ho (2008) betrachtet drei Zielkriterien für eine flexible Werkstattfertigung mit bis zu 15 Maschinen. Während der Optimierung werden die drei Zielkriterien jedoch mit festen Gewichtungen zu einer gemeinsamen Zielfunktion kombiniert. Somit wird letztlich eine uni-kriterielle Optimierung durchgeführt und das Ergebnis der Optimierungsläufe stellt somit keine Regelmenge dar. Deutlich näher an den in diesem Beitrag durchgeführten Untersuchungen ist Nguyen et al. (2013). Darin werden Prioritätsregeln für ein dynamisches Werkstattfertigungsszenario mit 10 Maschinen entwickelt. Es wird ein Pareto-Ansatz verwendet, d. h., das Ergebnis eines Optimierungslaufs ist tatsächlich eine Regelmenge wie in den hier vorgestellten Experimenten. Das von Nguyen et al. (2013) verwendete Szenario ist jedoch deutlich einfacher als das diesem Beitrag zugrunde liegende. Darüber hinaus verwendet der von uns entwickelte Ansatz zur multi-kriteriellen Genetischen Programmierung verschiedene Techniken, um die Konvergenzgeschwindigkeit der Optimierungsläufe zu verbessern (Normierung der Attributwerte, Erkennung und Entfernung von Duplikaten). Eine Betrachtung der Regellänge als weiterem Zielkriterium findet ebenfalls nicht statt.

# 3 Lösungsansatz

Grundsätzlich arbeitet das für dieses Paper verwendete Optimierungsverfahren der Genetischen Programmierung (GP) mit einer Baumrepräsentation für die Lösungen. Bei der hier vorliegenden Anwendung der Erstellung von Prioritätsregeln ist dieser Baum die Baumstruktur des arithmetischen Ausdrucks der Regel. Ein Bespiel hierzu ist in Abbildung 1 (links) dargestellt. Der dort dargestellte Ausdrucksbaum stellt die (hypothetische) Prioritätsregel „0-(PT+PT+S+B)“ dar, wobei PT, S und B jeweils Attribute des zu bewertenden Auftrags darstellen oder Größen zur Charakterisierung des aktuellen Systemzustands repräsentieren. Die möglichen Bestandteile eines solchen Baums, also die möglichen Rechenoperationen und die zur Verfügung stehenden Attribute, sind vor dem Optimierungslauf festzulegen. Im vorliegenden Beitrag wurden hierfür im Wesentlichen die häufig in Standardregeln verwendeten

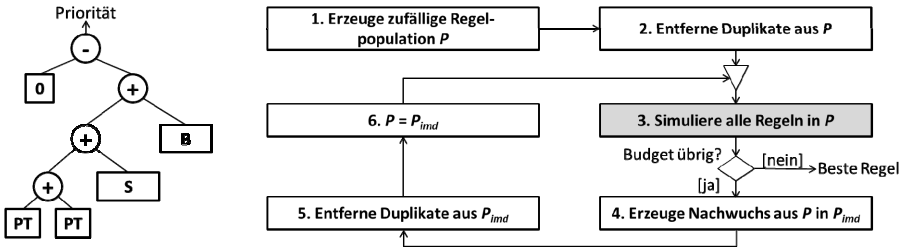



*Abbildung 1: Beispiel einer als Baum repräsentierten Prioritätsregel (links); Grafische Darstellung des verwendeten Optimierungsalgorithmus (rechts)*

Rechenoperationen und Attribute verwendet. Die Größe des Baumes muss bei GP nicht im Voraus festgelegt werden, sondern ergibt sich aus den Eigenschaften des zu lösenden Problems. Als Maß für die Länge einer Prioritätsregel wird im vorliegenden Beitrag die Anzahl an Knoten des entsprechenden Baumes betrachtet, im Beispiel in Abbildung 1 (links) ist sie also 9.

Der grundlegende Ablauf der Optimierung ist in Abbildung 1 (rechts) dargestellt. In Schritt 1 wird zunächst eine Menge an zufälligen Regeln erzeugt. Aus diesen lassen sich in Schritt 2 mithilfe des in Hildebrandt und Branke (2015) näher beschriebenen Vorgehens ohne rechenaufwändige Simulationsläufe syntaktisch unterschiedliche, in ihrem Verhalten jedoch gleiche Regeln erkennen und entfernen. Die Hauptiteration besteht aus den Schritten 3 bis 6. Im mit Abstand rechenaufwändigsten Schritt 3 werden nun alle Regeln der Regelpopulation P mittels Simulation bewertet. Auf Basis dieser Bewertung werden anschließend unter Nutzung der Grundprinzipien der Evolution (Selektion, Mutation, geschlechtliche Fortpflanzung/Cross-Over) eine neue Zwischenpopulation $P_{imd}$ erzeugt. Aus dieser werden dann wie für die Anfangspopulation in Schritt 2 eventuelle Duplikate erkannt und durch zusätzlich aus P erzeugte Regeln ersetzt. In Schritt 6 schließlich ersetzt die Zwischen-population $P_{imd}$ die aktuelle Population und eine neue Iteration beginnt. Als Abbruchbedingung wird nach Schritt 3 geprüft, ob eine bestimmte Anzahl an Simulationsläufen bereits erreicht wurde die Hauptschleife ggfs. verlassen.

Die Erweiterung zu dem in Hildebrandt et al. (2014) verwendeten Vorgehen besteht nun zum einen in der Verwendung eines anderen Selektionsoperators in Schritt 4. Hierzu wird das NSGA-II-Verfahren adaptiert (Non-dominated Sorting Genetic Algorithm-II, Deb et al. 2002). Dies ermöglicht es, mehreren Zielkriterien gerecht werden zu können und während eines einzelnen Optimierungslaufs eine Menge nichtdominierter Lösungen zu ermitteln.

Darüber hinaus erwies sich eine zusätzliche Feinauswahl nach Ende der Haupt-Iteration als sinnvoll zur Erzielung besserer Ergebnisse. Dies ist ebenfalls eine Erweiterung zum bisherigen Vorgehen, bei dem die beste Regel aus der letzten Iteration der Optimierung als Gesamtergebnis des Optimierungslaufs zurückgegeben wurde. Die für dieses Paper verwendete Feinauswahl berücksichtigt stattdessen die jeweils beste Regel der letzten 5 Iterationen und bewertet diese in jeweils 30 von der Optimierung unabhängigen Replikationen, um dann mit höherer Wahrscheinlichkeit die tatsächlich beste Regel zurückliefern zu können. Anstelle einer einzelnen Regel ergibt sich bei mehreren Zielgrößen am Ende jeder Iteration jeweils eine Menge nichtdominierter Lösungen, aus denen dann analog die endgültige Regelmenge ermittelt wird.

# 4 Szenariobeschreibung und Experimentaufbau

## 4.1 Szenario und Benchmark-Regeln

Das für diesen Beitrag betrachtete Produktionsszenario aus dem Bereich der Halbleiterfertigung ist das FAB6-Modell des „MIMAC (Measurement and Improvement of Manufacturing Capacities) testbed“ (Feigin et al. 1996). Dieses Modell ist durch die folgenden Charakteristika gekennzeichnet: 104 Maschinengruppen mit insgesamt 223 Maschinen, Batch-Maschinen, reihenfolgeabhängige Rüstzeiten, 9 verschiedene Produkte, die zwischen 234 und 355 Operationen erfordern, zyklischer Materialfluss, Maschinenausfälle und –Wartung.

Das Modell wird für 18 Monate simuliert (etwa 3600 fertiggestellte Aufträge/Lose), wobei die hier wiedergegebenen Ergebnisse nur die letzten 12 Monate dieses Zeitraums erfassen und die ersten 6 Monate zur Vermeidung von Einschwingeffekten für die Ergebnisauswertung ignoriert werden. Als stochastische Einflussgrößen sind Maschinenausfälle und -wartung modelliert, alle weiteren Größen (insbesondere auch Auftragsankünfte) sind deterministisch. Bzgl. Produktmix werden hier die in Hildebrandt et al. (2014, Table 1) mit „Setting 1“ bezeichneten Einstellungen verwendet, auf deren Wiedergabe hier aus Platzgründen verzichtet wird. Im Unterschied hierzu wird jedoch ein Flussfaktor von 2.12 statt 2.30 verwendet, da sich letzterer als zu einfach erwies.

Wenn nicht anders angegeben, sind die im Folgenden wiedergegebenen Ergebnisse zur Leistungsbewertung einer Regel die Mittelwerte aus 30 unabhängigen Replikationen. Diese Replikationen sind ebenfalls unterschiedlich zu den während der Optimierung und (Fein-)Auswahl der besten Regeln verwendeten. Als Simulationssoftware kommt die am BIBA entwickelte Simulationsbibliothek jasima (JAva SImulator for MAnufacturing and logistics, http://jasima.googlecode.com/) zum Einsatz.

Vorgelagert zu den Optimierungs-Experimenten wurde in einer konventionellen Simulationsstudie zunächst die beste Standard-Regel für dieses Szenario ermittelt. Hierzu wurde eine Menge von 7 Standard-Regeln betrachtet, vgl. Haupt (1989): FIFO (first in first out), ERD (earliest release time first), EDD (earliest due date first), ODD (operational due date), MOD (modified operational due date), SPT (shortest processing time first), sowie CR (critical ratio). Darüber hinaus wird die ODD+ Regel untersucht (siehe Zhou und Rose (2012), dort als IFD-Regel bezeichnet). Diese stellt eine Erweiterung von ODD dar, die Aufträgen, die bereits länger als 2 Tage vor einer Maschine warten die höchste Priorität zuweist, unabhängig von ihrem Due Date. Bei allen 8 Regeln wird ERD als „tie breaker“ verwendet, sollten zwei Aufträge anhand der Basisregel die gleiche Priorität zugewiesen bekommen. Alle Regeln werden darüber hinaus mit der „Kürzesten Rüstzeit“-Regel kombiniert. Auf Maschinen mit reihenfolgeabhängigen *Rüstzeiten* werden also zunächst Aufträge mit möglichst kurzer Rüstzeit bevorzugt und nur bei gleicher Rüstzeit wird die Prioritätsregel herangezogen. Für die *Batchmaschinen* wird ein „greedy batching“ durchgeführt. Hierzu wird zunächst der Auftrag mit der höchsten Priorität gemäß der Basis-Regel ausgewählt. Mit diesem Auftrag werden dann möglichst viele batchkompatible Aufträge kombiniert und der Batch gestartet. Als Parameter in dieser Heuristik kann eine bestimmte minimale Batch-Größe (MBS-Minimum Batch Size) eingestellt werden. Im Folgenden wird dies mit MBS(u) gekennzeichnet, wobei u eine Zahl zwischen 0 und 1 ist. Ein u von 0 startet einen Batch sofort, sobald

mindestens ein Auftrag vorhanden ist. U=1 wartet immer, bis ein Batch der maximalen Größe gebildet werden kann. In der Voruntersuchung wurden alle 8 verschiedenen Regeln mit jeweils 6 Einstellungen der Steuerung der Batch-Maschinen kombiniert (MBS(0), MBS(0,2), …, MBS(0,8), MBS(1,0)).

### 4.2 Multi-kriterielle Genetische Programmierung

Die Simulation mittels *jasima* wurde mit dem Framework ECJ (Evolutionary Computation in Java, http://cs.gmu.edu/~eclab/projects/ecj/, Version 20) zur Durchführung der Optimierungsläufe gekoppelt. Die verwendeten GP-Parameter entsprechen den in Hildebrandt et al. (2014) genutzten Einstellungen, eine detailliertere Beschreibung lässt sich dort nachlesen. Insgesamt werden in einem Optimierungslauf jeweils 30000 Simulationsläufe durchgeführt (60 Generationen mit jeweils 500 Regeln). Die möglichen Bestandteile von Prioritätsregeln sind hierbei vorgegeben. Neben den vier Grundrechenarten, max und if-then-else als mathematischen Operationen und vier verschiedenen Konstanten (0, 1, 2, 5) stehen insgesamt 10 verschiedene Attribute zur Verfügung. Diese sind allesamt leicht zu ermittelnde Größen, wie sie auch in Standardregeln verwendet werden: (A1) Bearbeitungszeit der aktuellen Operation, (A2) Durchschnittliche Bearbeitungszeit aller vor der Maschine wartenden Aufträge, (A3) Rüstzeit, (A4) Durchschnittliche Rüstzeit, (A5) Anzahl der verbleibenden Operationen eines Auftrags, (A6) Verweilzeit des Auftrags im System, (A7) Verweilzeit des Auftrags in aktueller Warteschlange, (A8) Größe der Batch-Familie (wie viele rüstkompatible Aufträge sind in der Warteschlange), (A9) Due Date der aktuellen Operation, (A10) Durchschnittliches Due Date aller wartenden Operationen. Darüber hinaus wurde ECJ zur Verwendung der NSGA-II-spezifischen Klassen entsprechend konfiguriert und erweitert. Als Zielfunktionen wurden die mittlere Durchlaufzeit (DLZ), mittlere Verspätung (Versp.) und die Länge der Regeln als Zielkriterien verwendet. Im Einzelnen wurden folgende Kombinationen untersucht:

1. mittlere Durchlaufzeit (uni-kriteriell)
2. mittlere Durchlaufzeit, Regellänge (2 Zielkriterien)
3. mittlere Durchlaufzeit, mittlere Verspätung (2 Zielkriterien)
4. mittlere Durchlaufzeit, mittlere Verspätung, Regellänge (3 Zielkriterien).

Um Aussagen über die mittlere Güte der zu erwartenden Optimierungsergebnisse machen zu können, wurden für jede der 4 Einstellungen jeweils 10 unabhängige Optimierungsläufe durchgeführt.

## 5 Ergebnisse und Diskussion

### 5.1 Primäre Zielfunktion: Durchlaufzeit

In den durchgeführten Experimenten zur Auswahl der besten Benchmark-Regel erweist sich die ODD+-Regel als am besten. Unabhängig von der verwendeten Regel ist im betrachteten Szenario die Variante MBS(0) die beste. Es lohnt sich (wahrscheinlich aufgrund der hohen Auslastung von 95%) hier nicht, auf Aufträge zu warten, um einen Batch besser füllen zu können. Die Kombination ODD+/MBS(0) erzielt als beste Benchmark-Regel ein Ergebnis von 27,37 (±0,19) Tagen mittlerer Durchlaufzeit bei einer mittleren Verspätung von 76,9 (±31,7) Minuten. Etwa 11% aller Aufträge sind verspätet.

*Tabelle 1: Mittlere Qualität der besten DLZ-Regeln für verschiedene Zielfunktionen*

| Zielkriterien | DLZ [d] | Versp. [min] | Regellänge |
|---|---|---|---|
| DLZ | 24,87 (±0,05) | 2.9 (±2,7) | 233,5 (±62,9) |
| DLZ, Länge | 24,89 (±0,06) | 3.7 (±3,1) | 48,0 (±11,9) |
| DLZ, Versp. | 24,83 (±0,03) | 0.1 (±0,1) | 197,5 (±50,5) |
| DLZ, Versp., Länge | 24,87 (±0,08) | 1.1 (±1,6) | 84,9 (±23,9) |

Um die Qualität der erreichbaren Regeln in den Optimierungs-Varianten 1-4 bewerten und einfacher miteinander vergleichen zu können, erfolgt eine Fokussierung auf das Zielkriterium „mittlere Durchlaufzeit“. Das heißt, bei den Varianten mit mehr als einem Zielkriterium, die also eine Regelmenge zum Ergebnis haben, erfolgt jeweils eine Auswahl der nach dem Zielkriterium „mittlere Durchlaufzeit“ besten Regel. Somit ergibt sich je Variante aus den jeweils 10 durchgeführten Optimierungsläufen eine Menge von 10 Regeln. Die sich hieraus ergebenden Resultate sind in Tabelle 1 aufgeführt. Jede Zeile beschreibt dabei eine der untersuchten Varianten und gibt die hierfür erzielten mittleren Ergebnisse (gemittelt über die jeweils 10 Optimierungsläufe) hinsichtlich Durchlaufzeit, mittlerer Verspätung der Aufträge und Regellänge an. Die in Klammern angegebenen Werte stellen jeweils den doppelten Standardfehler der Ergebnisse aus den 10 Optimierungsläufen dar. Wie zu erkennen ist, führen alle untersuchten Varianten bzgl. Durchlaufzeit zu sehr ähnlichen Ergebnissen. Zumindest nach 10 Optimierungsläufen ist keine signifikant besser oder schlechter. Dies ist ein wichtiges Ergebnis, zeigt sich doch, dass gegenüber einem uni-kriteriellen Vorgehen (Variante 1) keine Abstriche hinsichtlich Regelqualität zu machen sind. Im Vergleich mit der besten Benchmarkregel lässt sich die Durchlaufzeit sehr deutlich um ca. 2,5 Tage bzw. 9.0% reduzieren.

Was das Kriterium der mittleren Verspätung angeht, erzielen alle Varianten ebenfalls sehr gute Werte. Obwohl die Tabelle 1 zugrunde liegenden Regeln allesamt „Spezialisten“ für das Kriterium der mittleren DLZ sind, kommen allenfalls einzelne Aufträge zu spät, so dass sich insgesamt nur sehr kleine mittlere Verspätungen ergeben. Die Verbesserungen in Bezug auf die mittlere DLZ sind so groß, dass der verwendete Flussfaktor, der für die Benchmark-Regeln eine große Herausforderung darstellt, für die optimierten Regeln kein Problem ist. Die untersuchten Varianten, die während der Optimierung explizit die Verspätung mit berücksichtigen, erzielen ggfs. etwas geringere Verspätungen, die Differenzen sind aber statistisch ebenfalls nicht sehr aussagekräftig und sollten in weiteren Arbeiten in Versuchen mit engeren Due Dates näher untersucht werden.

Sehr deutlich ist der Effekt der Reduktion der Regellänge, wenn die Länge als zusätzliches Kriterium während der Optimierung mit berücksichtigt wird. Wird nur nach Durchlaufzeit minimierenden Regeln gesucht, schrumpft die Länge auf ein Viertel ihres ursprünglichen Werts. Wird nach Regeln gesucht, die sowohl nach DLZ als auch Verspätung gut sind, halbiert sich die mittlere Regellänge immerhin noch. Diese Betrachtung bezieht sich allerdings nur auf die Regeln maximaler Länge. Auf Basis der Lösungs*menge* als Ergebnis der multi-kriteriellen Optimierung lassen sich jedoch auch Aussagen treffen, welche Ergebnisse sich mit kürzeren Regeln erzielen lassen. Hierauf wird in Abschnitt 5.2 noch näher eingegangen.

*Tabelle 2: Ergebnisse der jeweils besten gefundenen Regeln*

| Zielfunktionen | DLZ [d] | Versp. [min] | Regellänge |
|---|---|---|---|
| DLZ | 24,79 (±0,10) | 0,08 (±0,06) | 388 |
| DLZ, Länge | 24,73 (±0,17) | 8,76 (±17,06) | 37 |
| DLZ, Versp. | 24,77 (±0,10) | 0,05 (±0,07) | 303 |
| DLZ, Versp., Länge | 24,55 (±0,37) | 0,80 (±1,12) | 69 |

In Tabelle 2 sind die Ergebnisse der jeweils besten Regel aus den jeweils 10 Optimierungsläufen einzeln aufgeführt. Die in Klammern aufgeführten Zahlen stellen den doppelten Standardfehler über die 30 genutzten Replikationen dar. Die Ergebnisse unterstreichen im Grunde die eben gemachten Beobachtungen auf Basis der gemittelten Werte. Die insgesamt beste Regel wird interessanterweise von der eigentlich komplexesten Optimierungs-Variante mit drei gleichzeitigen Zielkriterien gefunden. Diese erzielt eine mittlere DLZ von 24,55 Tagen und somit ein um 2,8 Tage bzw. 10,3% besseres Ergebnis als die beste Benchmark-Regel.

## 5.2 Abwägung Durchlaufzeit vs. Regellänge

Die explizite Berücksichtigung der Regellänge während der Optimierung von Prioritätsregeln stellt eine Neuheit dieses Beitrags dar. In den Experimenten unter Einbeziehung der Regellänge kann nun nicht nur nach guten, sondern nach guten und möglichst kurzen Regeln gesucht werden. Grundsätzlich ist eine Abwägung erforderlich: kurze Regeln sind gut interpretierbar und lassen sich besonders schnell berechnen. Andererseits ist zur Erreichung besonders guter Ergebnisse eine gewisse Regellänge notwendig. Dass die Berücksichtigung der Regellänge als Zielfunktion deutlich die Regellänge senkt, ohne dass hierdurch die Qualität bzgl. des Zielkriteriums mittlere Durchlaufzeit sinkt, konnte bereits in Abschnitt 5.1 gezeigt werden. Im Folgenden erfolgt eine detaillierte Betrachtung der Abwägung Regellänge vs. erreichbare Durchlaufzeit.

Hierzu wurden die Ergebnisse aller durchgeführten Optimierungsläufe ausgewertet und aus allen Regeln diejenigen ausgewählt, die jeweils einen in sich optimalen Kompromiss zwischen DLZ und Regellänge darstellen. Dies ergibt eine Menge an 11 Regeln mit Mindestgröße 5, welche in Abbildung 2 grafisch dargestellt sind. Wie man sieht, beträgt der Unterschied zwischen der kürzesten (Länge 5), aber schlechtesten Regel und der besten, dafür aber komplexesten Regel (Länge 69) etwa einen Tag an mittlerer Durchlaufzeit. 9 Abstufungen liegen jeweils dazwischen. Je nach Präferenz hin zu einfacheren oder aber möglichst guten Regeln kann nun eine fundierte Auswahl getroffen werden, um die letztlich am besten geeignete Regel auszuwählen und diese dann bspw. in operativen Planungs- und Steuerungssystemen einzusetzen.

# 6 Zusammenfassung und Ausblick

Dieser Beitrag beschreibt den Einsatz der multi-kriteriellen, simulationsbasierten Optimierung, um deutlich leistungsfähigere und kürzere Prioritätsregeln zur Reihen-

folgeplanung für ein komplexes Produktionsszenario aus dem Bereich der Halbleiterfertigung mit einem automatisierten Verfahren zu erstellen.

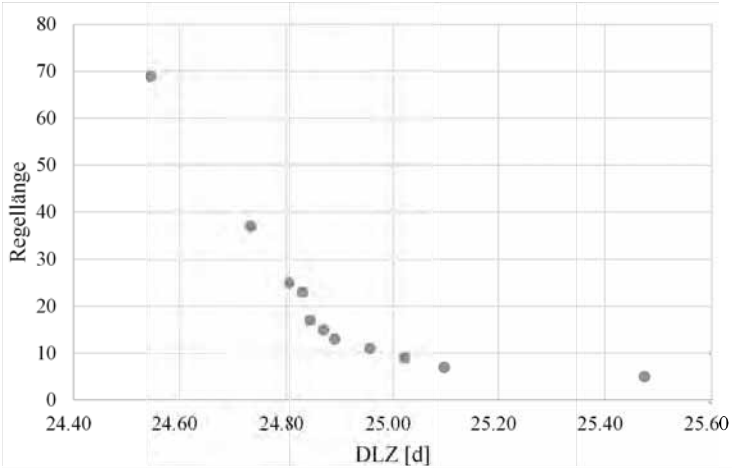



*Abbildung 2: Mögliche Kompromisslösungen zwischen Regellänge und erzielbarer Durchlaufzeit*

Der Beitrag beschreibt nach bestem Wissen der Autoren zum ersten Mal in der Literatur den Einsatz multi-kriterieller Optimierung für die Prioritätsregelentwicklung in einem Szenario der Komplexität wie das hier betrachtete. Mit den entwickelten Regeln konnte die mittlere Durchlaufzeit von Aufträgen um bis zu 9,4% gesenkt werden. Darüber hinaus ergaben die durchgeführten Experimente keine Hinweise auf Leistungseinbußen gegenüber einer uni-kriteriellen Vorgehensweise. Als weitere Neuerung gegenüber der bestehenden Literatur wird die multi-kriterielle Optimierung eingesetzt, um explizit nach *kurzen*, guten Prioritätsregeln zu suchen. Während diese Vorgehensweise in der GP-Literatur für verschiedene Problemstellungen bereits erfolgreich angewendet wurde (vgl. de Jong und Pollak 2003), stellt ihr Einsatz im Rahmen der Entwicklung von Prioritätsregeln eine Neuerung dar. In den durchgeführten Experimenten konnte hierdurch die Regellänge sehr deutlich reduziert werden. Hierdurch ergibt sich auf Basis eines einzelnen Optimierungslaufs eine fundierte Entscheidungsgrundlage, um zwischen kurzen, leicht interpretierbaren Regeln, die aber nur verhältnismäßig schlechte Ergebnisse erreichen, und sehr guten, dafür aber komplexen und schwer interpretierbaren, Regeln die geeignetste auszuwählen.

Im Ergebnis wird also eine Erweiterung der automatisierten Entwicklung von Prioritätsregeln vorgestellt, die die Anwendbarkeit und den Nutzen des Verfahrens durch den Einsatz der simulationsbasierten multi-kriteriellen Optimierung in mehrfacher Hinsicht verbessern. Das vorgestellte Verfahren fließt in ein entsprechendes Optimierungsmodul ein, das im Laufe der letzten Jahre bereits sehr erfolgreich zur unikriteriellen Entwicklung von Prioritätsregeln eingesetzt wurde. Im Rahmen einer aktuellen Ausgründung aus dem BIBA werden hierdurch interessierten Unternehmen diese neuen Möglichkeiten zur Effizienzsteigerung ihrer Prozesse eröffnet.

## Danksagung

Wir danken dem Bundesministerium für Wirtschaft und Energie für Unterstützung im Rahmen des EXIST-Gründerstipendiums „jasima solutions“.

## Literatur

Branke, J.; Nguyen, S.; Pickardt, C.; Zhang, M.: Automated design of production scheduling heuristics: A Review. IEEE Transactions on Evolutionary Computation 2015, doi:10.1109/TEVC.2015.2429314, to appear.

Burke, E.; M. Gendreau; M. Hyde; G. Kendall; G. Ochoa; E. Ozcan; Qu, R.: Hyper-heuristics: a survey of the state of the art. Journal of the Operational Research Society 64 (2013) 12, S. 1695–1724.

Coello Coello, C.: Evolutionary multi-objective optimization: A historical view of the field. IEEE Computational Intelligence Magazine 1 (2006) 1, S. 28–36.

Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6 (2002) 2, S. 182-197.

Feigin, G.; Fowler, J.; Leachman, R.: MASM test data sets. 1996. Available from http://masmlab.engineering.asu.edu/ftp.htm. Accessed June 06, 2014.

Haupt, R.: A survey of priority rule-based scheduling. OR Spektrum 11 (1989) 1, S. 3–16.

Hildebrandt, T.; Branke, J.: On using surrogates with Genetic Programming. Evolutionary Computation 2015, doi:10.1162/EVCO_a_00133, to appear.

Hildebrandt, T.; Goswami, D.; Freitag, M.: Large-scale simulation-based optimization of semiconductor dispatching rules. In: Proceedings of the 2014 Winter Simulation Conference, Savannah 2014, S. 2580–2590.

de Jong, E.; Pollak, J.: Multi-objective methods for tree size control. Genetic Programming and Evolvable Machines 4 (2003) 3, S. 211-233.

Marler, R.; Arora, J.: Survey of multi-objective optimization methods for engineering. Structural and multidisciplinary optim. 26 (2004) 6, S. 369–395.

Nguyen, S.; Zhang, M.; Johnston, M.; Tan, K.: Dynamic multi-objective job shop scheduling: A genetic programming approach. In: Uyar, A.; Ozcan, E.; Urquhart, N. (Hrsg.): Automated Scheduling and Planning. Heidelberg, New York, Dordrecht, London: Springer 2013, S. 251–282.

Poli, R.; Langdon, W., McPhee, N.: A field guide to genetic programming. 2008. http://www.gp-field-guide.org.uk. Accessed June 06, 2014.

Tay, J.; Ho, N.: Evolving dispatching rules using genetic programming for solving multi-objective flexible job-shop problems. Computers & Industrial Engineering 54 (2008) 3, S. 453–473.

Zhou, Z.; Rose, O.: WIP balance and due date control in a wafer fab with low and high volume products. In: Proceedings of the 2012 Winter Simulation Conference, Berlin 2012, S. 2019-2026.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulation zur Entwicklung und Dimensionierung von Schichtmodellen

## *Use of Simulation for Development and Capacity Determination of Shift Models*

Harald Mutzke, Simulation Expert Services, Nürnberg (Germany),
info@mutzke-ses.com

Gregor Blauermel, B416 GmbH & Co. KG, Cadolzburg (Germany),
info@B416.de

Abstract: In the growing market of internet pharmacies, optimal working-time models are of vital importance. The flexible arrangement of working hours of the employees in the picking area was identified as the most effective parameter and was systematically optimised within this project. For this application it was important to create a discrete event simulation model, which established the lab environment for optimisation studies outside real time operation. To achieve high speed run-time behaviour, it was necessary to apply a high level of abstraction to the models in the optimisation suite, using evolutionary strategies. For data input into the model and the subsequent analysis, a SQL database connection was established. Graphically generated results were used for the communication of the new strategies. Results indicated shift and break patterns as well as customised headcount, which are currently tested in real operation. A methodological view to the transferability of the approach completes the paper.

## 1 Ausgangssituation

Für Unternehmen im Apotheken-Internethandel – wie dem Partner im vorliegenden Projekt – werden bis 2020 zweistellige Wachstumsraten prognostiziert, u. a. durch Erhöhung des Anteils an rezeptpflichtigen Medikamenten (Nielsen 2012). Gleichzeitig erwarten die Kunden eine immer kleinere Zeitspanne zwischen Bestellen und Liefern. Lag die akzeptierte Versanddauer im Jahr 2013 bei drei bis fünf Arbeitstagen (BITKOM 2013), erwartet der Kunde heute den Paketeingang am nächsten Arbeitstag. Dies erfordert angepasste Arbeitszeitmodelle in der Bestellbearbeitung, die von den Mitarbeitern akzeptiert werden und gleichzeitig eine an die Menge und den Bestellzeitpunkt orientierte Schichtstärke (eingesetzte Mitarbeiter). Flexibilisierung findet bisher statt, indem kurzfristig (Leih-) Mitarbeiter eingestellt und Schichtmodelle wie gewohnt weiter genutzt werden. Damit wird Experimentieren

am Arbeitszeit-Modell vermieden, um den laufenden Betrieb und das zu bewältigende Wachstum nicht zu gefährden.

# 2 Problemstellung

Bei einer qualitativen Analyse der Bestellbearbeitung lassen sowohl die umgesetzte Artikelmenge als auch der Servicelevel Optimierungspotentiale erkennen, während Leerlaufzeiten des Personals die Kosten in die Höhe treiben. Um für die zukünftigen Anforderungen gerüstet zu sein, ist eine besser adaptierbare Wachstumsstrategie nach modernsten analytischen Methoden erforderlich.

# 3 Vorgehensweise

Um die Strategie zu entwickeln, wurde dieses Projekt aufgesetzt. An den Workshops waren je nach Bedarf unterschiedliche Experten beteiligt. Die Vorgehensweise orientiert sich an der VDI-Richtlinie 3633, Blatt 1.

# 4 Prozessmodellierung

Das Prozessmodell wurde in Zusammenarbeit mit Projektbeteiligten im Zielfindungs-Workshop entwickelt und online in MS VISIO 2013 Professional umgesetzt. Dabei wurde die für diesen Kunden relevante Prozesstiefe festgelegt und zeitunkritische Schritte eliminiert (Abb. 1).

## 4.1 Workshop Datenaufbereitung

### *4.1.1 Arbeitszeitmodell*

Mit Belegschaft und Betriebsrat wurden geeignete Schichtmodelle entwickelt. Wesentliche Fragestellungen dabei waren:

- Ist Sonntagsarbeit möglich?
- Welche Schichtmodelle sind denkbar?
- Wie werden Überstunden gehandhabt?
- Werden Saisonarbeitskräfte beschäftigt?
- Gibt es Nachtarbeit?

### *4.1.2 Auftragseingang*

Zugrunde gelegt wurden die Auftragsdaten eines halben Jahres (Abb. 2). Bei der Analyse zeigten sich Wochenmuster und Tagesmuster.

- Wochenmuster: Spitzenwerte am Wochenanfang, Täler Ende der Woche.
- Tagesmuster: 70 % vor 13:00 Uhr. Kein Auftragseingang während der Nacht.
- Diese Gesetzmäßigkeiten wurden später bei der Anwendung von Schichtclustern verwendet (übereinstimmende Wochentage bildeten ein Cluster). Das zweistellige Wachstum war ebenfalls klar ersichtlich und folgte den bereits bekannten Mustern.

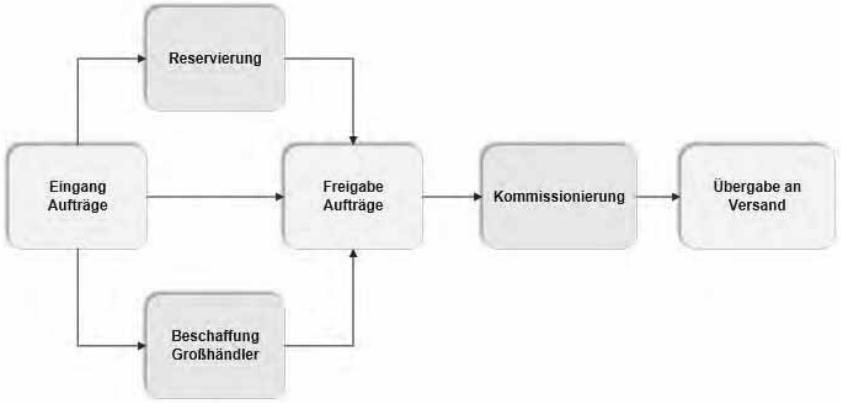



*Abbildung 1: Prozessmodell*

Um die Laufzeit für einen Simulationslauf zu reduzieren, wurden die einzelnen Aufträge zu Zeitscheiben verdichtet. Die Aufträge, die innerhalb einer Zeitspanne eingingen, wurden zu einem Paket zusammengefasst (vgl. Abschnitt Abstraktionsgrad).

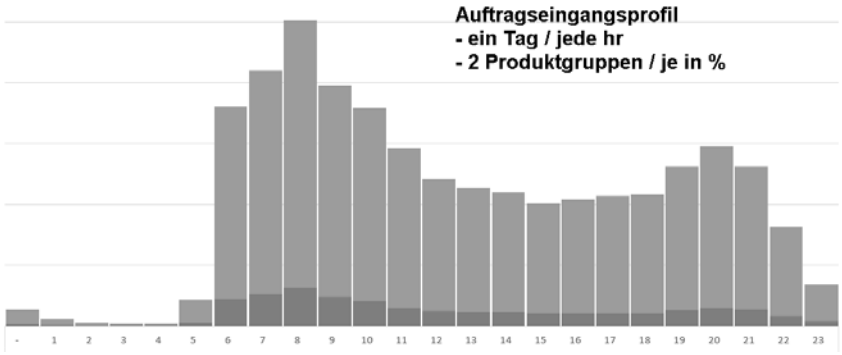



*Abbildung 2: Auftragseingang*

## 4.2 Modellierung

Bei der Erstellung des Simulationsmodells lag das Hauptaugenmerk auf folgenden Funktionalitäten:

- Unterschiedliche Mitarbeiterzahlen je Zeiteinheit sollen verglichen werden.
- Die Mitarbeiter sollen für unterschiedliche Zeitscheiben (z. B. Stunde, 4h-Teilschichten, …) zu Pools zusammengefasst werden können.
- In der späteren Optimierung sollen die Zeitscheiben an-/abschaltbar sein.
- Gemäß dem Betriebskalender wurden relevante Liefertermine festgelegt.
- Zugrunde gelegt wurde das Prozessmodell, das alle mitarbeiterrelevanten Zeiten berücksichtigt (s. o.).

### 4.2.1 Systemgrenzen

Das Modell beginnt mit dem Eintritt des Auftrags im Funktionsbereich „Kommissionierung“ und endet mit Übergabe an die Spedition.

### 4.2.2 Abstraktionsgrad

Der Abstraktionsgrad ist entscheidend für die spätere Laufzeit des Simulationsmodells. Jeder Prozessschritt und seine Komplexität kosten Zeit und verlängern bereits den einzelnen Simulationslauf. Angesichts der hohen Anzahl an einzelnen Experimenten summiert sich jeder überflüssige Zeitbedarf in einer Optimierungsstudie schnell zu Stunden oder gar Tagen. Folgende Entscheidungen und Annahmen wurden getroffen:

- Bündelung von Aufträgen zu Zeitscheiben: Welches Zeitintervall ist ausreichend und genau genug? (1 min, 10 min, 1-h, 2-h oder 4-h oder 6-h Schicht …)
- Bildung von Arbeitstag-Clustern: Wie lassen sich die Arbeitstage gruppieren? (Wochentage mit ähnlichem Auftragseingang)
- Für alle Mitarbeiter aller Schichten wird ein einheitlicher Produktivitätswert angenommen.
- Feiertage werden behandelt wie normale Arbeitstage.
- Die Kundenartikel stammen aus einer unendlichen Quelle, d. h. es gibt keine Versorgungsengpässe.

Der vorgenommene Abstraktionsgrad hilft bei der Fokussierung der Studie auf den zentralen Sachverhalt.

## 5 Simulator Entwicklung

Abbildung 3 zeigt das Zusammenspiel zwischen der Inputdatengenerierung aus dem SQL-DB-Export), dem Simulator und dessen Output. Das VISIO Modell des Prozesses wurde mittels Process Simulator 2014 Professional vollautomatisch in ein ProModel Optimization Suite 2014 Simulationsmodell überführt. Dieser Simulator Prototyp wurde dann weiter verfeinert und in die Simulations- und Optimierungsumgebung integriert.

### 5.1 Workshop zur Validierung

Bei der Validierung wurden die simulierten Kennzahlen Servicelevel und Ausbringung mit dem Kennzahlen-Reporting des Unternehmens verglichen. Die Gegenüberstellung von Maßnahmen z. B. „mit und ohne erweiterte Kapazität in Stoßzeiten“ und deren gemessene Wirkung auf den Servicelevel wurde im Modell auf plausible Verhaltensmusteränderungen intensiv ausgetestet (Sensitivitäts-analyse). In einem Kunden-Experten-Workshop wurden veränderte Arbeitszeitmodelle und Schichtstärken gegenüber gestellt (Plausibilitätskontrolle). Zusätzlich wurden exemplarisch einzelne Aufträge zu kritischen Tageszeiten im Detail datentechnisch analysiert, grafisch aufbereitet und geprüft.

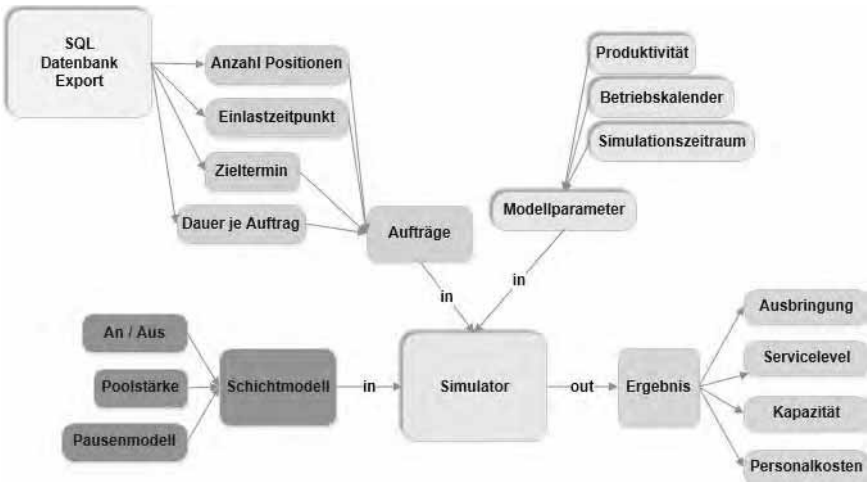



*Abbildung 3: Ablauf Simulator-Entwicklung*

## 5.2 Ergebnisse

Das Simulationssystem liefert am Ende jeder Simulation eine Datenbank mit allen gemessenen Kennzahlen des Geschäftsprozesses. Diese werden zur besseren Visualisierung in Histogrammen, Zeitverlaufskurven oder Tortendiagrammen dargestellt. Zusätzlich können Daten exportiert und in gängigen Statistikpaketen wie „R“ aufbereitet werden. Darüber hinaus werden sogenannte Log-Dateien vom Simulator geschrieben, um einzelne Auftragspositionen konkret zu analysieren. Damit steht ein umfassend geprüftes Simulationssystem für die Optimierung zur Verfügung.

# 6 Optimierer Entwicklung

Zweck einer Optimierung ist es, mit möglichst wenig zeitlichem Aufwand eine optimale Lösung für ein komplexes System zu finden. Im vorliegenden Fall ist das komplexe System im Simulationsmodell abgebildet (i. W. Simulator). Für die Lösungssuche wird eine Zielfunktion (s. u.) benötigt und der Lösungsraum (s. u.) für die Optimierung festgelegt, wie in Abbildung 4 skizziert ist.

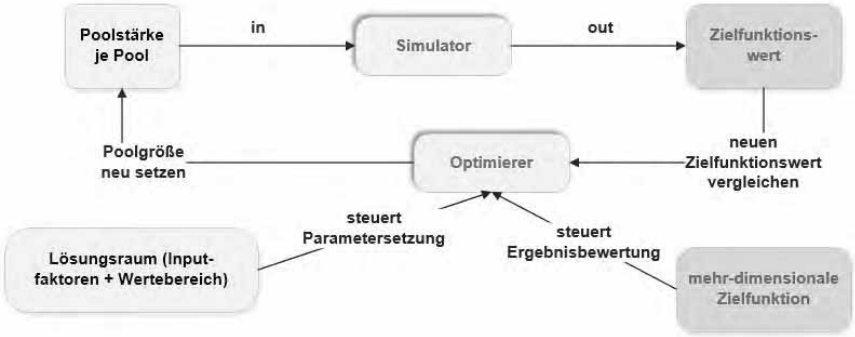



*Abbildung 4: Ablauf Optimierer-Entwicklung*

Die Optimierungssoftware (i. W. Optimierer), die alle notwendigen Informationen erhalten hat, ermittelt auf dieser Basis einen Startsatz der Mitarbeiterzahlen (Poolstärken). Mit diesen Werten wird das Simulationsmodell parametriert („gefüttert“) und nach einem ersten Simulationslauf der Wert der Zielfunktion berechnet und ausgegeben. Damit beginnt ein iterativer Prozess, in dem der Optimierer versucht, mithilfe des implementierten Algorithmus mit möglichst wenig Experimenten die optimale Kombination der Poolstärken und zusätzlich eine An- oder Abschaltung von Schichten zu ermitteln.

## 6.1 Eckdaten

Unterschiedliche Zielfunktionen verkürzen oder verlängern die Dauer der Optimierung, d. h. die Anzahl der durchzuführenden und zu analysierenden Optimierungsexperimente oder Simulationsläufe. Folgende Einstellungen sind vorzunehmen:

- Input Faktoren und deren Intervall
- Zielfunktion und deren Komplexität
- Warm up, Laufzeit des Modells und Anzahl Wiederholungen
- Optimierungsprofil und Konfidenzlevel

## 6.2 Zielfunktion

Eine Zielfunktion dient dazu, quantitativ auszudrücken, wie sich ein System oder das Modell dieses Systems verhält (Performance). Dazu werden Faktoren ermittelten, die auf diese Performance Einfluss haben. Im vorliegenden Projekt wurden vier Faktoren ermittelt, die ausdrücken, welche Leistung das System erbringt.

- *Mengenindex:* Steht für das Verhältnis von tatsächlich ausgebrachter Menge zur gewünschten Menge. Der Index liegt im Intervall von null bis eins und soll maximiert werden.
- *Servicelevel:* Drückt die Zufriedenheit des Kunden aus. Bei Aufträgen, die innerhalb des vorgegebenen Zieltermins versendet werden, ist der Servicelevel erfüllt und wird mit 1 bewertet. Benötigt die Bearbeitung länger, so liegt der Servicelevel bei 0. Dieser Wert sollte für die Summe aller Aufträge möglichst hoch liegen.
- *Kapazitäten:* Je mehr Wochenstunden die Mitarbeiter eingesetzt werden müssen, umso teurer ist die Kommissionierung.
- *Kostenindex:* Steht für die gesamten Personalkosten, die durch das gewählte Schichtmodell verursacht werden, in Relation zur Hochrechnung der Produktivitätsrate.

Kapazitäten erlauben eine direkte Einschätzung des Simulationsexperiments wie z. B. „200 Stunden mehr Kapazität“ eingesetzt. Der Kostenindex liefert dagegen eine relative Einschätzung des verwendeten Schichtmodells, z.B. 50 % Mehrkosten. Eingebunden wurden Kosten gemäß Vorgehensweise zur Kostensimulation (Mutzke 1997).

Für jeden Einzelfaktor wird bestimmt:

- Wie genau lautet die *Definition* (z. B. Mittelwert, …)?
- Beeinflusst er das Ergebnis positiv oder negativ? (Suche nach *Maximum oder Minimum*)
- Im welchem Intervall bewegt sich der *Erwartungswert*?

- Wie stark beeinflusst er das Gesamtergebnis im Vergleich zu den anderen Faktoren? (*Gewichtung*)

Die Zielfunktion wurde ebenfalls mithilfe einer Sensitivitätsanalyse untersucht, Stichproben zum Realitätsabgleich wurden durchgeführt und gemeinsam mit den Experten und deren Erfahrungswissen plausibilisiert.

## 6.3 Lösungsraum

Der Lösungsraum umspannt alle möglichen Kombinationen der Inputparameter (28 Poolstärken). Für jeden Wochentag sind vier Schichten möglich, das bedeutet 28 mögliche Schichten pro Woche (7 Arbeitstage). Jede dieser Schichten kann einen Wert zwischen 0 und 35 annehmen. Die Spitze jeder Pyramide steht für den ausgewählten Wert für diese Schicht. In diesem Fall handelte es sich z.B. um *4 741 632* mögliche Kombinationen. Durch Einsatz des Optimierers konnte die Anzahl der Simulationsläufe gegenüber der vollständigen Berechnung aller Kombinationsmöglichkeiten auf *12 908* reduziert werden. Das entspricht einer Einsparung *von 99,7%*.

## 6.4 Workshop Zielfunktion

Bei der Entwicklung der Zielfunktion sind einige Themen zu beachten:

- Welche Kennzahl beschreibt die „bessere“ Lösung und sortiert die Experimente?
  - Ein größerer Servicelevel beschreibt ein „besseres“ Modell.
- Welche Randbedingung ist erforderlich?
  - Es sollen nur Lösungen verglichen werden, bei denen alle Kunden auch bedient wurden.
- Welche Abbruchbedingungen sind sinnvoll?
  - Wird der Bestand im Modell sehr groß, so ist davon auszugehen dass das Modell keine verwertbaren Ergebnisse liefert und der Optimierungslauf kann vorzeitig abgebrochen werden.
- Wie könne triviale Lösungen verhindert werden? Wie kann eine Vergeudung von eingesetzten Kapazitäten bei gleichzeitigem erforderlichem Servicelevel gefunden werden?
  - Eine mehrdimensionale Zielfunktion unter Berücksichtigung minimaler Kosten verhindert Verschwendung von Kapazität.

## 6.5 Algorithmus

Die verwendeten *Evolutionären Algorithmen* (EA) sind eine Klasse von stochastischen, metaheuristischen Optimierungsverfahren, deren Funktionsweise von der Evolution natürlicher Lebewesen inspiriert ist. In Anlehnung an die Natur werden Lösungskandidaten für ein bestimmtes Problem künstlich evolviert, EA sind also naturanaloge Optimierungsverfahren. (SIMRUNNER 2014, S. 7).

## 6.6 Struktur der Optimierung

Die Gesamtoptimierung kann gedanklich zur Veranschaulichung in drei Abschnitte unterteilt werden. Dadurch entsteht eine hierarchische Vorgehensweise.

*I Fokus Schichtmodell: Arbeitszeit,* verglichen wurden die verschiedenen Arbeitszeitmodelle: Samstag und Sonntag ja / nein? Nachtschicht? Schichtlänge?

*II Fokus Mitarbeiter Anzahl,* gesucht wurde hier die optimale Mitarbeiteranzahl für die kostengünstige Realisierung des Prozessablaufs.

*III Fokus Arbeitszeit-Details,* welchen Effekt hatten verschiedene Pausenregelungen (Position und Dauer)?

## 7 Ergebnis-Workshop

Bei diesem Workshop wurden die optimalen Parameterkonstellationen vorgestellt und diskutiert. Anhand der Grafiken erhielten alle Teilnehmer ein vertieftes Verständnis für die Einflüsse der Parameter auf die einzelnen Terme der Zielfunktion und damit auf die Kennzahlen der Kommissionierung und deren Zusammenhänge. Aufgetragen ist beispielhaft in Abbildung 5 der Einfluss der Kapazitäten auf den Servicelevel. Für einen Servicelevel von x müssen Kapazitäten von y eingesetzt werden. Eine falsche Verteilung der Kapazität führt hingegen wieder zu einer Verschlechterung des damit möglichen Servicelevels.

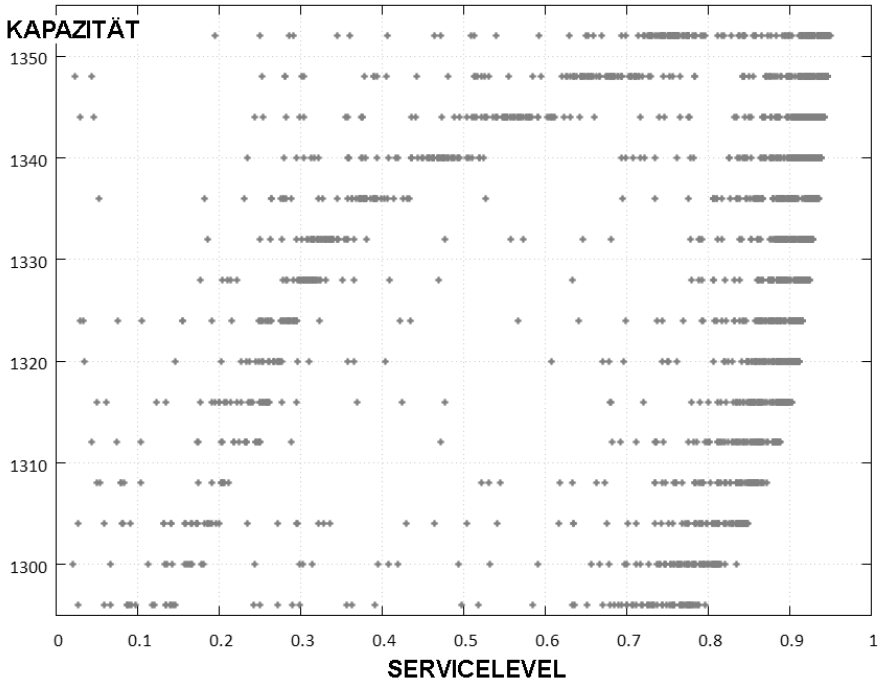



*Abbildung 5: Unterschiedlicher Servicelevel bei gleicher Kapazität*

Auch lassen sich die bekannten klassischen Betriebspunkte aus der Praxis darin nachvollziehen. Eine Gesamtsicht auf die Zielfunktion und den zu maximierenden Servicelevel bei zu minimierenden Kosten zeigt Abbildung 6.

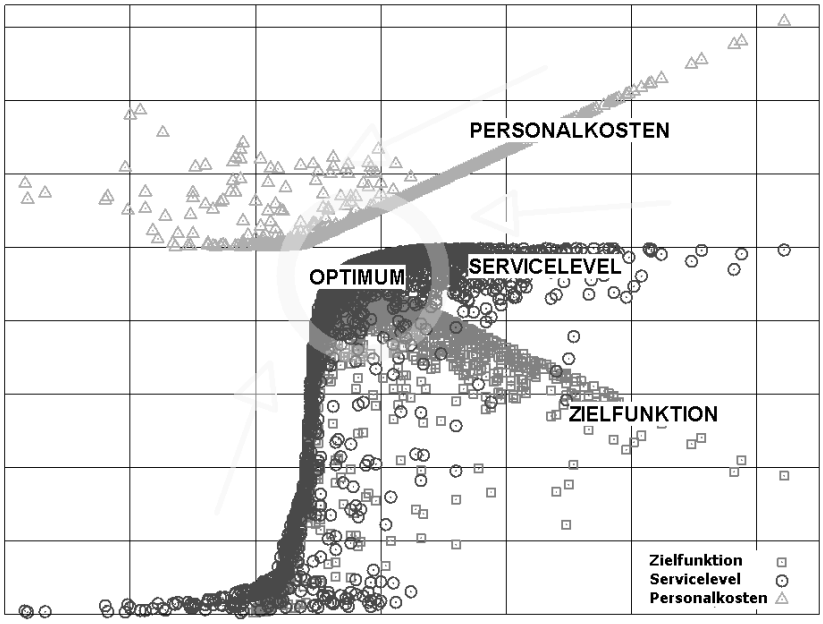



*Abbildung 6: Zusammenfassung Ergebnisse*

## 8 Realisierung

Derzeit werden die gefundenen optimalen Parameter im Unternehmen implementiert und die Kennzahlen auch erzielt. Die entwickelte Vorgehensweise wird in Zukunft regelmäßig eingesetzt, um die Schichten zu überprüfen und kontinuierlich an das Wachstum anzupassen. Die Konzentration auf Wachstum und nicht auf Organisation und Fehlversuche steht damit im Fokus. Fazit: O-Ton „Das hätten wir uns weder ausdenken noch im laufenden Betrieb ausprobieren können!“ … und … „Die Vorhersagegenauigkeit und Qualität der prognostizierten Kennzahlen übertraf unsere Erwartungen!“

Grundsätzlich eignet sich die Methode für jedes Unternehmen, das ähnliche Herausforderungen bewältigen muss. Die Weiterentwicklung des projektspezifischen Simulators zu einem generisch anwendbaren Produkt-Methoden-Baukasten zur kostenintegrierten Optimierung von Schichten und Poolstärken steht aktuell auf dem Fahrplan der beteiligten Unternehmen für das Jahr 2015/2016. Die Ergebnisse überzeugten durch verständliche Visualisierung, zeit- und kostensparende Bewertung von Alternativ-Szenarien im Simulationsmodell sowie einen nachvollziehbaren Verlauf der Optimierung. Für die Praktiker waren diese Effekte wichtig:

- bedarfsgerechte Anpassungen des Simulationsmodells kurzfristig möglich
- keine Fehlversuche beim Umbau und Testen der Schichten
- Ausschluss nicht hilfreicher Schichten
- Platzierung der Ressourcen am richtigen Ort zur richtigen Zeit
- deutliche zweistellige Steigerung des Servicelevels

Übereinstimmende Erkenntnis: „10000 Versuche am Rechner weisen den Weg!“

## 9 Ausblick

Aktuelle Studien zum Thema eCommerce stellen übereinstimmend fest (Bonsai 2013, S. 26; Xu 2014):

- Der Kunde erwartet Lieferung schnellst möglich (Servicelevel!) und zum bestmöglichen Preis.
- Die zunehmende Internationalisierung, die in den nächsten Jahren den Einzelhandel hier zu Lande zunehmend betreffen wird (der Apotheken Versandhandel nimmt eine Sonderstellung ein aufgrund der bestehenden gesetzlichen Regelungen BMG 2014), erfordert weitere Anstrengungen bei der Einsparung von Zeitverlusten.

Der Prozess Kommissionierung steht branchenübergreifend im Fokus. Die wesentlichen Eigenschaften des optimierten Prozesses sind:

- Personal- oder kapazitätsintensiv
- stark schwankende Nachfrage
- Trendmodell (Wachstum, Schrumpfung)
- viele Lösungsmöglichkeiten möglich!
- Berechnung unmöglich oder zu zeitaufwendig!

Sollten diese Eigenschaften zutreffen, so kann die beschriebene Vorgehensweise adaptiert und angewendet werden. Einsatzgebiete können ebenfalls sein:

- personalintensive Montage
- Abläufe im Gesundheitswesen wie z. B. Behandlungspfade, OP-Saal-Auslegung
- Projektmanagement Prozesse (Fähigkeiten + Qualifikationen)
- Serienanlauf / Hochlaufmodelle

## Literatur

BITKOM e. V.: Online-Medikamentenkauf in Deutschland immer beliebter. www.bitkom.org, 12.02.2013.

Bonsai Market Research, Bundesverband Deutscher Versandapotheken: Versandhandelsstudie 2013, www.bdva.de, 28.05.2013.

Bundesministerium für Gesundheit (BMG): Versandhandel mit Arzneimitteln 2014. www.bmg.bund.de, 18.11.2014.

Bundesverband Deutscher Versandapotheken: Daten und Fakten zum Arzneimittelversandhandel in Deutschland. www.bdvda.de, 08.09.2014.

Nielsen Company: QUO VADIS Apotheken-Versandhandel? Vortrag beim 5. BVDVA Kongress, Berlin Mai 2012.

Mutzke, H.: Integration der Kostensicht in die Simulation. In: 11. Symposium Simulationstechnik Dortmund, November 1997, S. 232-236.

Xu, Annie: 7 E-commerce Trends Small Businesses Need to Know in 2014. www.huffingtonpost.com, updated 9.3.2014.

PROMODEL 2014 Users Guide Version 9.2, ProModel Corporation, 2015.

Process Simulator Professional 2014 Users Guide Version 9.2, ProModel Corporation, 2014.

SIMRUNNER 2014 Users Guide, Version 3.0, ProModel Corporation, 2010.

VDI 3633 Blatt 1: Simulation von Logistik-, Materialfluss- und Produktionssystemen – Grundlagen. Beuth Verlag, Berlin 2014.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Material Flow Analysis at BASF – A Huge Variety in Complexity

## *Materialflussanalyse bei BASF – Eine große Bandbreite an Komplexität*

Markus Feist, Ana Maria Fernández Alcalde, BASF SE, Ludwigshafen (Germany),
markus.feist@basf.com, ana-maria.fernandez-alcalde@basf.com

Abstract: Material flow analysis (MFA) is a common method for the investigation of production and logistic processes. Its objective is finding, quantifying and (if possible) avoiding bottlenecks and saving potentials. Other applications of MFA are the designing of capacities of production parts or logistic devices such as loading stations or storages. Usually, the complexity of the process and the questions to be answered determine the required level of detail of the MFA. The huge variety in complexity starts from simple estimations and calculations, going up to building detailed and customised simulation models. One common simulation method is the so called discrete event simulation. In this contribution, three examples are presented for using MFA in the BASF SE, a company in the field of chemical industry. An overview is given on how established methods are used in an industrial setting to solve different tasks with various data basis and purposes.

## 1 Introduction

In industry, material flow analysis (MFA) is often used to gain knowledge about existing production and logistic processes, or to determine capacities for plants or logistic devices for investment projects, Jahangirian et al. (2013) and references cited therein. In either case, it must be determined whether the benefit is worth the effort of setting up a complex and detailed simulation or if a less detailed calculation is sufficient. The required level of detail is determined by several boundary conditions. For example, the complexity of the underling processes and the respective tasks which should be answered are as important as the available data basis. In that way it could happen that a MFA covers only a data analysis of production data, followed by a straight-forward and not very complex calculation (notation: *static MFA*). But it can also cover a complex dynamic model for discrete event simulations (notation: *dynamic MFA*).

Independent of the level of detail, MFA takes into account the plants and logistic behaviors as well as the product output and delivery reliability with respect to capacity, product recipes and logistic setup. In case of static calculations, many

assumptions, generalisations and simplifications usually have to be made. This process of decision is accompanied by a deep discussion on the expected impact on the calculation results. For dynamic modeling a detailed process model including fewer simplifications is built, aiming to approximate the real process behavior more closely. Using the method of discrete event simulations (Jahangirian et al. 2013), even unexpected breakdowns or planned events as well as simultaneous usage of equipment and resources can easily be considered.

Independent on the level of complexity, it is possible to investigate the influence of production order, delays in transport, lack of raw materials or buffer capacity.
Upon the completion of the model, the verification (usually accomplished by comparison with historical data) is a milestone in the MFA process. Based on the historical data, results are evaluated to check whether the simulation captures the influence of the key parameters on the underling process qualitatively or even quantitatively. Afterwards, future scenarios can be investigated, for instance, variations in the production amount or changes of the product portfolio. Finally, with the help of these results, proposals for an increase of production efficiency such as the elimination of bottlenecks, new production strategies as product order or better resource utilisation can be developed and evaluated.
Following a short general introduction, three representative examples of MFA at BASF SE illustrating typical applications are discussed. Since the applied methods are of common use, the presentation focusses on the underling questions, the data basis and boundary conditions as well as the effort to gain robust results.
In the order of increasing complexity and level of model detail, the MFA examples focus on the following topics:

- Example 1: Interwoven production of two products – static MFA
- Example 2: Raw material delivery and its influence on the production – dynamic MFA
- Example 3: Simulation of an entire batch production plant – static and dynamic MFA

# 2 Examples

## 2.1 Example 1: Interwoven Production of two Products – Static MFA

For MFA of this interwoven production three main questions should be answered:

- Which production order is required to be reliable for the customers demand?
- What is the maximal annual production of products?
- Is it possible to make the calculations with a model/tool that is as complex as needed yet easy to operate?

In addition, it was requested to make the calculation model adaptable for certain changes in boundary conditions and for further and repeated use during planning. Keeping static and time discrete calculation to be as easy as possible, MS Excel was chosen as an established tool.

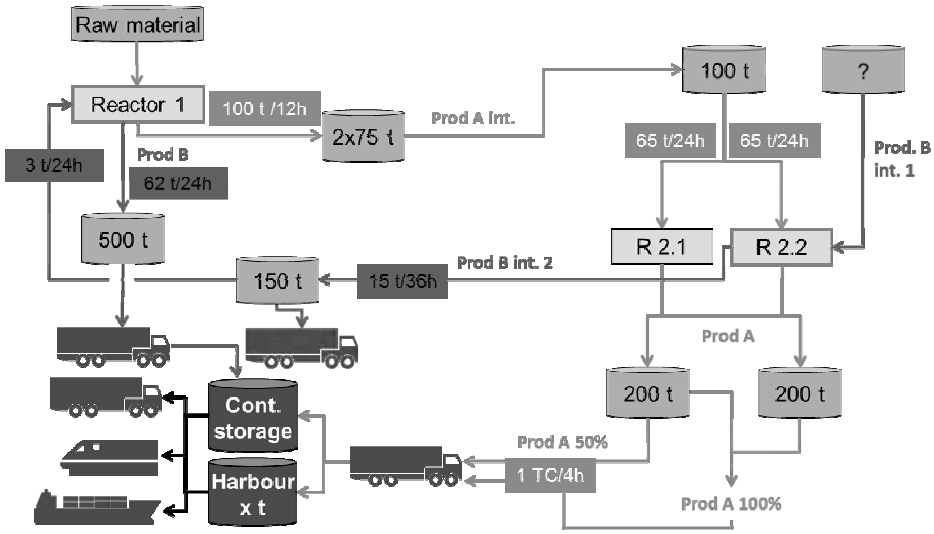



*Figure 1: Outline – Interwoven production chain*

Figure 1 shows the interwoven production. The main challenges result from the different capacities of parts in the production chain, the limited buffer and storage capacities and the parallel usage of certain reactors. The main goal is to continuously maintain reliability for the customers demand independent of its seasonality.

Necessary assumptions and simplifications result from several discussions during the analysis process. One main assumption is the time discretisation of 12 h derived from the minimal batch time of one reactor. That way, especially a discrete event behavior or stochastic breakdowns in the process are neglected. However, the key parameters and their influences on the production chain could still be considered. The out coming result is an MS Excel-tool that automatically generates a rough production plan depending on the forecast of the customers demand and planned shutdowns of the plant. As a generalised rule, a certain production order, variable customer demands and a limited capacity of the production capacity and buffer tanks are taken into account. Further, non-generalised knowledge can be specified manually.

## 2.2 Example 2: Raw Material Delivery and its Influence on the Production – Dynamic MFA

For MFA of this process three main questions should be answered:

- Which is the maximal annual production volume?
- What is the order of bottlenecks by raising the output of the products?
- Which improvements in the production chain are possible to raise the actual maximum?

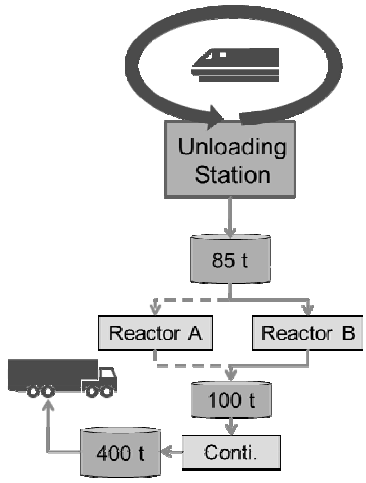



*Figure 2: Outline of the production chain*

The starting point of this production chain is a rail car supply circle. Here, the raw material is delivered across several intermediate stops to a loading station at the plant. The loading station cannot be used exclusively for supplying this raw material. Further a maneuvering of the rail cars is possible only twice a day during fixed time slots. The station is connected to a raw material tank of approximately one and a half day full production quantity in capacity. In the reactors, two products are produced in campaigns having different raw material utilisation factors. The downstream process started again with a buffer tank followed by a continuous process, another buffer tank and a filling station, see figure 2.

At some points, the complexity of this problem is greater than the complexity of the first example. This is mainly attributed to the many uncertainties across the production chain and to the more detailed questions. Here, the focus is not laid on delivery reliability but on the raw material supply and the maximal possible outcome of the plant. Because of the low buffer capacity and the present uncertainties, a time discretisation would lead to a relatively high estimation error - or it has to be very fine and hence would not be suitable. The stochastic behavior mainly occurs at the supply of raw materials by rail cars. One usually has to account for variable cycle times resulting from unexpected delays. The data basis for this analysis is gathered by GPS tracking of all rail cars, collected during the last half year. With this information the time for the delivery circle was modeled using the same distribution, hence covering behavior in the past. The loading station has defined time slots to maneuver the rail cars, and the unloading process shows variable durations. These delays can cause the missing of a maneuver slot. Consequently, the rail car will block the loading station, until the next free slot for maneuvering. The time variations during the unloading process result from few breakdowns and also from different loading times caused by the outer temperature and the waiting times in the supply circle. Finally, the production contains unforeseeable breakdowns that, due to low buffer capacities, lead to a blockage of the loading station.

To sum up: all uncertainties have a significant influence on the whole production chain, on the storage capacity and in the end on the maximal outcome of the

products. Consequently, a detailed dynamic model was required. Such a model was implemented using the software INOSIM, (Hellenkamp and Balling 2012). It is used to investigate the influence of an additional loading station, a change in the campaign production, more buffer capacity etc. on the product chain. In a final step, the results of the scenarios were evaluated with respect to the benefit for the production and the best investment strategy.

## 2.3 Example 3: Simulation of an Entire Batch Production Plant – Static and Dynamic MFA

For MFA of the whole plant three main questions should be answered:

- Where are the bottlenecks of the production?
- Can additional buffer tanks reduce waiting times?
- What improvements could be done in the logistics to get a higher overall capacity?

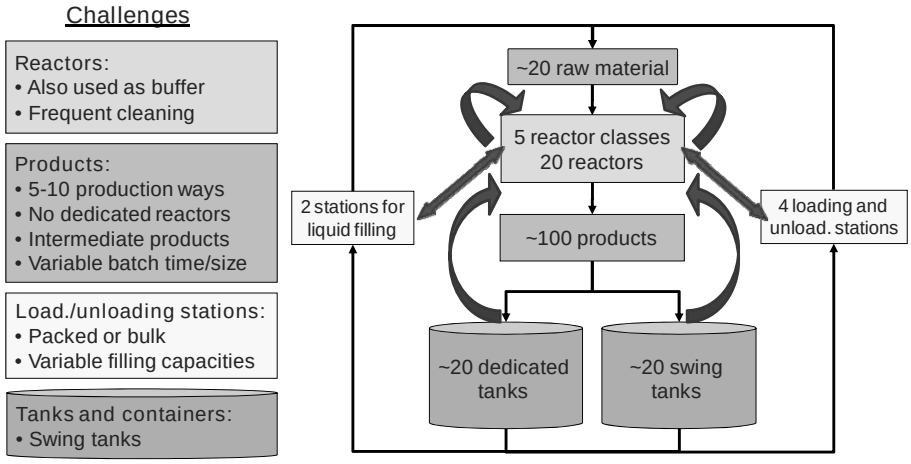



*Figure 3: Illustration of the production with main boundary conditions*

The entire production plant consists of 20 batch reactors of different sizes, clustered in 5 classes. In these, approximately 100 products are produced from 20 major raw materials. The products are stored in several dedicated or swing tanks and containers and are delivered to customers via filling stations as packed or via loading station as bulk ware, respectively. The production is not linear for all products, i. e., some products are produced from intermediate products that are also stored in product tanks. Each product can be produced in several reactors and even parallel, if needed. In addition, they can have different batch sizes and batch times.

The initial step of the MFA was a detailed data analysis. In order to gain a deep understanding of the plant and the production rules as well as of the planning process, regularly interviews with plant operators and planning personnel were conducted. The batch times and batch sizes for each product were analysed by using process data from a PCS system. Here, a huge variance in the batch times was discovered, related to waiting times due to full storage tanks or blocked loading

stations. Therefore, one main task was to separate waiting times from production times to define the ”golden batch” for each product. Due to the high flexibility in the production and since the whole complexity of the production had to be covered, the generalisation of the production rules was another challenge. Finally, the product demands and seasonality were analysed by using historical selling data. Altogether lead to the basis of a high detailed discrete event model of the whole production and loading process.

The simulation model was able to manage over 100 products, up to 10 different production paths, variable batch times and sizes as well as dedicated and swing tanks. This high level of detail was necessary in order to account for the high flexibility and complexity of the real production process. By validation it has been proved that the model was as detailed as necessary to approximate the historical production sufficiently and to correctly describe the influence of the key parameters.

In a final step, the model was used to calculate scenarios to work out and quantify possible improvements. As a main result it turned out that a low filling capacity during the weekend leads to a product accumulation after approximately one day. Consequently, the reactors are used as buffers and have to wait. In addition, the production during the first days of the week is affected catching up these product amounts. Finally, a higher filling capacity during a weekend was assumed and its improvement was quantified with respect to the overall production amount.

## 3 Summary

In summary, this presentation gives an overview on how different ‘Material Flow Analysis’ topics at BASF SE really are. Selected examples and results illustrate the variety of questions and constraints, like complexity, level of detail and data basis. Special attention was paid to stress the different approaches with respect to the questions and the final usage of the MFA results.
The examples show that a static MFA is used for not very complex problems or as basis for validation of a complex and dynamic MFA. In the first example straightforward calculations lead to appropriate results. The main advantage is the relatively low effort to gain suitable statements as basis for best decisions. The use of a dynamic MFA is shown in the second example where unexpected events can cause big influence on the final results. There a static MFA would lead to an over- or underestimation of the capacity of the considered parts. The combination of static and dynamic MFA is used if the underlying conditions are very complex. A static MFA is needed to achieve a well analysed data basis. In addition static calculation results are used for validation of simulation results to be sure that the model is appropriate to cover the main influencing parameters.

Overall, the discussed case studies illustrate that the MFA provides a deep insight and understanding in the status quo of the production, and it embodies a comprehensive analysis basis for future decisions and investments.

## References

Hellenkamp, T.; Balling, P.: Production Logistics Analysis with INOSIM Professional In: Laroque, C.; Himmelspach, J.; Pasupathy, R.; Rose, O..; Uhrmacher, A.M. eds.: Proceedings of the 2012 Winter Simulation Conference (WSC), Berlin (Germany), 09.-12. December 2012.

Jahangirian, M; Eldabi, T.; Naseer, A. Stergioulas, A. N.; Young, T.: Simulation in manufacturing and business: A Review. European Journal of Operational Research 203 (2010) 1-13

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Joint Management of Material and Cash Flows

## *Gemeinsames Management von Material und Cash Flows*

Enver Yücesan, INSEAD, Fontainebleau (France), enver.yucesan@insead.edu

Illana Bendavid, ORT Braude College of Engineering, Karmiel (Israel), illana@tx.technion.ac.il

Yale T. Herer, Israel Institute of Technology, Haifa (Israel), yale@technion.ac.il

Abstract: The objective of inventory management is to design effective policies for managing the trade-off between customer satisfaction and the cost of service. Over the years, these models have become increasingly sophisticated, incorporating many complicating factors that are relevant in practice. Curiously absent from these models are the financial constraints on the working capital. In this paper, we analyse the materials management practices of a self-financing firm whose replenishment decisions are constrained by cash flows driven by purchases and sales. Our study shows that arbitrarily imposed constraints on the working capital not only fail to prevent any violations, but also significantly distort operational decisions. We propose less myopic policies to eliminate these distortions.

## 1 Introduction

Working capital refers to the difference between current assets, which include cash, inventory and accounts receivable (A/R), and current liabilities, which include accounts payable (A/P) and short-term loans. The top 1000 European public companies have approximately 900 Billion Euros (9.4% of EU GDP) tied up in working capital (REL Consultancy 2012). Working capital management, whose objective is to increase the profitability of a firm by investing in long-term, high-return assets while ensuring sufficient liquidity during the operating cycle (Pass and Pike 1984), therefore represents a huge economic opportunity, not only for the economy as a whole, but for multitudes of individual firms as well. In particular, efficient working capital management is necessary to manage the inherent trade-off between profitability and liquidity since keeping too much cash for operations may lead to a smaller return from long-term investments.

Against this backdrop, our investigation into the management of inventory under working capital constraints was motivated by a distributor of industrial equipment that purchases heavy goods such as generators and earthmoving equipment from a manufacturer and sells them to various customers. The distributor grants trade credit

to its customers while receiving trade credit from the manufacturer. The latter, which has a finite production capacity, ships orders requiring a non-negligible delivery lead time and requires full payment after a fixed number of periods following delivery. This type of trade credit represents a one-part (net term) contract. Customer demand is random as it is driven by major construction projects. Once a delivery is made to a customer, the customer must pay for the invoice after a fixed number of periods. The distributor's challenge is to determine an inventory management policy that balances its inventory and shortage costs. With trade credits, inventory decisions have a direct impact on working capital levels since deferred inventory payments are recorded as A/P while delayed sales collection is captured as A/R. Typically, firms adopt a working capital policy to achieve the optimal trade-off between profitability of net current assets and the ability to pay for current liabilities under a clearly defined risk framework. Such a policy designates the portion of the sales revenue to be retained for working capital (Pass and Pike 1984). The inventory management problem thus becomes further complicated by the working capital policy. In this particular case, the distributor has an aggressive policy, choosing to keep a smaller portion of sales revenue as working capital. As a result, the working capital for the distributor cannot exceed a maximum value that is set by the distributor's parent company in an effort to limit inventory and collection risk. Throughout its operating cycle, which spans procurement, inventory holding, sales and cash collection, the distributor must therefore ensure adequate customer service in a self-financing fashion, that is, generating sufficient funds through sales to procure additional inventory. Note that this is in contrast with most of the recent work in the literature that focuses on payment default –hence, promotes a lower bound.

The recent spotlight on micro-retailing is another key motivator for our study. With little or no access to external financing, small, typically family-owned firms must carefully manage their operating cycle, generating sufficient cash from sales to replenish their inventories for business continuity. Pierson and van Ryzin (2010) describe such a micro-retail operation where the owning family needs to keep an eye on its working capital not only to effectively manage its operating cycle, but also to generate a sufficient cash surplus to make a living.

In this setting, cash flows generated by a firm`s operations have to be explicitly incorporated into the operational decision making process beyond the common practice of imposing an exogenous budget constraint on activities such as procurement, production, and sales. In the literature, however, until recently there has been little investigation of the relationship between finance and operations, and, in particular, of the interactions between material flows and cash flows. As a result, these decisions are often made separately without a unifying perspective on how these trade-offs affect the performance of the firm. This is also reflected in practice whereby a hard WCR constraint may be imposed on procurement while sales decisions are made without considering this constraint. The objective of this paper is therefore to construct a model to depict the interactions between material and cash flows as well as to quantify the impact of these interactions on performance. Through this study, we identify some unintended consequences of financial constraints; in particular, our study shows that arbitrarily imposed constraints on the working capital not only fail to prevent any undesired violations, but also sig–

nificantly distort operational decisions. We therefore propose an approach to elimi–nate these side effects.

## 2 Literature Review

Due to the fundamental Modigliani and Miller (1958) result in corporate finance theory whereby, in perfect and competitive markets, financing decisions can be made independently of investment and production decisions, there is traditionally little overlap between the extensive literature on cash management and the vast literature on production and inventory management. This separation is further reinforced in practice as the two functions normally reside in separate “silos" within the company, the former being part of treasury while the latter being part of operations. The realization that market imperfections such as accounting and bankruptcy costs, tax advantages of debt, and information asymmetry violate the key assumptions underlying the Modigliani-Miller result recently triggered a significant amount of work on the interaction between financial and operational decisions.

Dada and Hu (2008) consider the single-period inventory procurement problem of a capital-constrained newsvendor (CCNV). Given the procurement and selling price along with the distribution of demand, the CCNV determines the quantity that satisfies the newsvendor fractile, but may not have the financial means to order the optimal quantity. It must therefore determine how much to borrow at a given interest rate to finance additional procurement. In a Stackelberg game between the lender (the leader) and the CCNV (the follower), the authors show that even if the borrowing cost is not prohibitively high, the CCNV borrows funds to procure a quantity that is less than the channel optimal quantity. They further show that the lender charges an interest rate that decreases in CCNV's equity.

Li et al. (2013) characterize the optimal production, borrowing, and dividend decisions of a firm that aims to maximize the expected present value of the infinite-horizon flow of the dividends under uncertain demand, capacity, borrowing, and liquidity constraints.

Babich and Sobel (2004) focus on the coordination of financial and operational de-cisions to maximize the expected discounted proceeds from an initial public offe-ring, modelled as a stopping time in an infinite-horizon Markov decision process.

Chao et al. (2008) analyse the replenishment decisions of a self-financing firm with the objective of maximizing its expected wealth at the end of a finite horizon. At the beginning of each period, the firm places an order with its cash on hand and deposits its surplus cash into a savings account to earn interest. During the period, demand is realized. At the end of the period, the firm collects cash generated from sales and interest earned from the savings account. Their model, however, assumes negligible replenishment lead times and does not consider the operating cycle, ignoring collection and payment periods.

In the presence of trade credits, Luo and Shang (2013) consider non-negligible payment and collection periods, and develop base stock policies that minimize the discounted total inventory and default costs over a finite horizon. When the collection period is longer than or equal to the payment period, they show the optimality of a myopic base stock policy. In the reverse setting, they propose a heuristic policy. We also examine the impact of trade credit on the operating cycle in

a working capital constrained environment where borrowing from a third party is not an option. Our analysis, however, differs from theirs in several aspects: first, we consider an infinite-horizon setting. Second, while we impose a hard constraint on the working capital, they consider a penalty for violating this constraint that reflects the risk of payment default. Finally, our model, which will be introduced in the next section, incorporates non-negligible replenishment lead times.

# 3 The Model

Consider a firm that procures goods from a manufacturer at a fixed purchase price to satisfy random demand. In particular, the firm faces an infinite-horizon newsvendor setting with backorders. In addition, the manufacturer has uncertain capacity that is independent of the order quantity. Mean demand is assumed to be smaller than mean capacity. The manufacturer delivers the product after a fixed order lead time. The firm has to pay for the received goods after a given payment period following delivery. When the firm sells a unit at a given selling price, the customer has to pay for the product after a fixed collection period. The goal is to minimize the long-run average cost of demand-supply mismatch while taking into account the uncertainty in customer demand and manufacturer capacity. In other words, the firm must determine an optimal inventory policy to minimize its (long-run average) cost for a specific combination of demand distribution, capacity distribution, order lead time, and cost parameters (holding and shortage costs). This challenge is further complicated by a constraint on the firm's working capital during its operating cycle. In this setting, two factors may limit the firm's ability to procure goods. First, the manufacturer may not be able to ship the requested replenishment quantity due its own capacity constraint. Second, due to the constraint on the firm's working capital, the firm may not have the funds to finance the totality of the desired replenishment quantity.

## 3.1 Notation and Problem Formulation

Throughout the paper, we will use the following notation:

*Random variables*

$D_t$ is a random variable describing the I.I.D. demand in period $t$, with cumulative distribution function $F_D$, mean $\mu_D$, finite variance, and realization $d_t$.

$K_t$ random variable describing the I.I.D. manufacturing capacity available in period $t$, with cumulative distribution function $F_K$, mean $\mu_K$, having a finite variance, and realization $k_t$.

$R_t$ is a random variable describing the I.I.D. shortfalls/residuals at the end of period $t$, with cumulative distribution function $F_R$ and realization $r_t$.

*Problem parameters*

$c$ is the per unit purchase price from the manufacturer.

$L_O$ is the order delivery lead time from the manufacturer.

$PP$ is the duration of the payment period, i.e., the number of periods during which the firm carries accounts payable (A/P).

$V$ reflects the selling price per unit.

*CP* is the duration of the collection period, i.e., the number of periods during which the firm carries accounts receivable (A/R).

*h* is the holding cost per unit per period.

*p* is the shortage cost per unit per period.

*CR* represents the newsvendor critical ratio given by $p/(p + h)$.

*W* is the maximum allowable value of the Working Capital Requirements (WCR), henceforth referred to as the WCR limit.

ρ is the capacity utilization rate or the system load on the manufacturer, defined as $\mu_D / \mu_K$.

*Auxiliary and Decision Variables*

*S* represents the order-up-to level.

$Q_t$ is the quantity ordered by the firm in period *t* taking into account the order-up-to level, the available capacity, and the WCR limit.

$I_t$ is the net inventory at the end of period *t*.

$I_t^+$ represents the on-hand inventory at the end of period *t* (= $\max\{0, I_t\}$).

$I_t^-$ captures the backorder level at the end of period *t* (= $\max\{0,-I_t\}$).

$WCR_t$ reflects the requirements for working capital at the end of period *t*.

We wish to minimize the average system cost per period calculated as the sum of the cost of excess inventory and the cost of backorders, i.e.,

$$\frac{1}{n} \sum_{t=1}^{n} (hI_t^+ + pI_t^-) ,$$

Subject to the following constraints:

$$Q_t = \max\left\{0, \min\left\{S - I_t - \sum_{m=t-L_O+1}^{t-1} Q_m, K_t, \frac{W - \mathrm{WCR}_t}{c}\right\}\right\} \quad \text{for all } t > 0$$

$$\mathrm{WCR}_t = \mathrm{WCR}_{t-1} + c\left(I_t^+ - I_{t-1}^+\right) + V\left(\min\{I_{t-1}^+ + Q_{t-L_O}, D_t + I_{t-1}^-\} - \min\{I_{t-CP-1}^+ + Q_{t-CP-L_O}, D_{t-CP} + I_{t-CP-1}^-\}\right) - c\left(Q_{t-L_O} - Q_{t-PP-L_O}\right) \quad \text{for all } t > 0$$

$$I_t = I_t^+ = I_t^- = 0 \quad \text{for all } t \le 0$$

$$\mathrm{WCR}_t = Q_t = 0 \quad \text{for all } t \le 0$$

The first equation defines the quantity ordered in period *t*, as the minimum of the quantity needed to attain the order-up-to level *S*, the realized manufacturer capacity, and the quantity corresponding to the monetary value of the working capital

available to the firm. The last element of this minimum highlights the difference between $W$ and $WCR_t$, where the former is the amount of wealth to which the firm has access and the latter is the amount of this wealth that the firm is using at time $t$. The next equation updates the *WCR*, where we add the change in the monetary value of the inventory and the change to the A/R, adding the value of sales in period $t$ and subtracting the value of sales in period $t - CP$. Finally, we subtract the change in the A/P, the difference in the value of items delivered in period $t$ and the items delivered in period $t - PP$. The last two equations initialize the system with boundary conditions. Note that the optimal order-up-to level, $S$, is computed numerically following the approach described in Güllü (1998). We refer to this formulation as the *Base Model*.

Unfortunately, an arbitrarily imposed constraint on the working capital not only fails to prevent any undesired WCR violations, but also significantly distorts operational decisions. This phenomenon is driven by the fact that while a hard constraint on working capital is strictly enforced in procurement, sales decisions are typically made without taking it into consideration. To eliminate these distortions, instead of limiting our order quantity by the current *WCR* value, we order no more than the quantity, which ensures that the probability of violating the constraint on the working capital in the future, due to this order, will be less than or equal to $\beta$ ($0 < \beta < 1$), where $\beta$ is a model parameter that reflects the firm's risk appetite. More specifically,

$$Q_t = \min\left\{ S - I_t - \sum_{m=t-L_O+1}^{t-1} Q_m, K_t, \arg\max_{Q}\left\{ \text{Prob}\left( \max_{\tau > t+L_O} wcr_\tau(Q) > W \right) \leq \beta \right\} \right\}$$

$$\text{for all } t > 0$$

Since the only two time epochs at which WCR increases are when we pay for the items we ordered (i.e., $L_O + PP$ time units after placing the order) and when we sell an item, it suffices to evaluate $wcr_\tau$ for $\tau$ from $t + L_O$ up to $t + L_O + PP$, or whenever the inventory runs out, whichever occurs later. We begin the evaluation at time $t + L_O$ because this is the first time the effect of an order placed at time $t$ is felt. To evaluate the term, we use simulation. That is, we independently replicate the system starting in period $t$ and, for each replication, we find the maximal order quantity in period $t$ that does not violate the constraint on the working capital.

The manufacturer then ships the minimum of the replenishment quantity ordered by the firm and its realized capacity. Let us refer to this modified version as the *Spend-at-Risk Model*.

## 3.2 Observations

As anticipated, the system with no constraints on the working capital has the lowest average cost (AC) and the highest percentage of violations (PV). The effect of $\beta$ was also consistent with our expectations. Lower values of $\beta$ (tighter constraints) result in higher AC and lower PV values. Of the highest interest is the comparison of the Base Model with the Spend-at-Risk Model. The AC of the Base Model is higher than even the most constrained Spend-at-Risk Model ($\beta = 0$). In and of itself, this result is not surprising; one would expect a trade-off between AC and PV. However, combined with the fact that the Base Model had a higher PV value than

the Spend-at-Risk Model for both $\beta = 0$ and $\beta = 0.25$, this finding means that both of these policies dominate the Base Model. This is best illustrated graphically in Figure 1, where we plot the performance of various policies based on their 90th percentile AC and PV values. Note that the Base Model does not even lie on the efficient frontier.

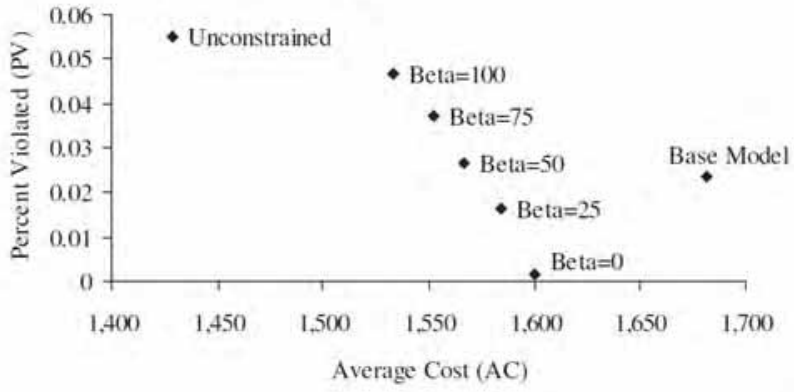



*Figure 1: Average cost versus WCR violations*

Having observed that the Spend-at-Risk Model is superior to the Base Model, we now examine its behavior. We start by focusing on the newsvendor parameters. For $\beta = 0$, there is no qualitative difference in the impact these parameters have on the AC between the Base Model and the Spend-at-Risk Model. The less myopic perspective of the latter model, however, is noticeable in the percentage of time the order quantity is constrained by WCR, PM, and PV. For instance, a higher *CR* leads to a higher PM 41% of the time in the Base Model and 30% of the time in the Spend-at-Risk Model. Similarly, a longer $L_O$ leads to a higher PV 22% (14%) of the time in the Base (Spend-at-Risk) Model.

With respect to the financial parameters, the AC is significantly lower in the Spend-at-Risk Model. An increase in *V* and *CP* leads to an increase in the AC in a smaller percentage of the cases compared with the Base Model. Notable is the behavior of the parameter *c*. Whereas in the Base Model an increase in the unit purchase price causes an increase in AC in more cases than it causes a decrease (21% versus 16%), this trend is reversed in the Spend-at-Risk Model where AC decreases in more cases than it increases (14% versus 12%). The reason for this switch lies in how this parameter is treated in the two models. While the Base Model incorporates this parameter in three ways (accounts payable, inventory evaluation, and directly in the ordering decision), the Spend-at-Risk Model incorporates the parameter only in two ways (accounts payable and inventory evaluation). In other words, an increase in *c* could very well force the Base Model to order less, i.e., further from the average cost minimizing order-up-to level. There is, however, no such effect in the Spend-at-Risk Model. We also note that the 'lethal' interaction between *V* and *CP*, which consumes a larger portion of WCR for a longer period of time, has a more mitigated impact on both AC and PM under the Spend-at-Risk Model. We can therefore conclude that

this model not only eliminates this instability, but also lowers the average total cost per period. Unfortunately, it does not eliminate all the violations on the working capital constraint.

## 4 Conclusions

The objective of inventory management models is to determine effective policies for managing the trade-off between customer satisfaction and the cost of service. Over the years, these models have become increasingly sophisticated, incorporating many complicating factors that are relevant in practice such as demand uncertainty, finite supplier capacity, and yield losses. Curiously absent from these models are the financial constraints imposed by working capital requirements (WCR). In practice, however, many firms are self-financing, i.e., their ability to replenish their own inventories is directly affected not only by their current inventory levels, but also by their receivables (trade credit they have extended to their customers) and payables (trade credit they have received from their supplier). Such constraints have gained added significance during the lingering economic crisis, which has made external financing increasingly difficult to secure. In this paper, we analyzed the materials management practices of a self-financing firm whose replenishment decisions are constrained by cash flows, which are updated periodically following purchases and sales in each period. In particular, we investigated the interaction between the financial and operational parameters as well as the impact of the constraints on the working capital on the long-run average cost.

The key messages of our study can be summarized as follows:

1. The imposition of a myopic constraint on WCR renders an otherwise stable operating cycle unstable by leading to shortfalls that become increasingly difficult to eliminate.
2. An arbitrarily imposed constraint on the working capital not only fails to prevent any undesired violations, but also significantly distorts operational decisions. This phenomenon is driven by the fact that while a hard constraint on working capital may be strictly enforced in procurement, sales decisions are typically made without considering the constraint.
3. The unintended consequences of the myopic WCR constraint mentioned before lead us to define a more flexible and forward-looking approach, which is implemented in the form of a spend-at-risk model. This model not only eliminates this instability, but also lowers the average total cost per period. Unfortunately, it does not eliminate all the violations on the working capital constraints.

## References

Babich, V; Sobel, M.J.: Pre-IPO operational and financial decisions. Management Science 50 (2004) 7, pp. 935-948.

Chao, X.; Chen, J.; Wang, S.: Dynamic inventory management with cash flow constraints. Naval Research Logistics 55 (2008) 8, pp. 758-768.

Dada, M; Hu, Q.: Financing newsvendor inventory. Operations Research Letters 36 (2008) 5, pp. 569-573.

Güllü, R.: Base stock policies for production/inventory problems with uncertain capacity levels. European Journal of Operational Research 105 (1998) 1, pp. 43-51.

Li, L.; Shubik, M; Sobel, M.: Control of dividends, capital subscriptions, and physical inventories. Management Science 59 (2013) 5, pp. 1107-1124.

Luo, W.; Shang, K.: Managing inventory for entrepreneurial firms with trade credits. Technical Report, Working Paper, Fuqua School of Business, 2013.

Modigliani, F. ; Miller, M.H.: The cost of capital, corporation finance and the theory of investment. American Economic Review 48 (1958) 3, pp. 261-297.

Pass, C.; Pike, R.: An overview of working capital management and corporate financing. Management Finance 10 (1984) 3, pp. 1-11.

Pierson, M.; van Ryzin, G.: *Frogtek: Mobile technology for micro-retailing*. Columbia Business School Case Study, 2010.

REL Consultancy. *Europe Working Capital Survey*. London, 2012.

Simulation in Production
and Logistics 2015
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Inventory Management and Markup Pricing in the Presence of Price Fluctuations

## *Bestandsmanagement und Preisgestaltung bei Vorliegen von Preisschwankungen*

Caner Canyakmaz, Süleyman Özekici, Fikri Karaesmen, Koç University, Istanbul (Turkey), ccanyakmaz@ku.edu.tr, sozekici@ku.edu.tr, fkaraesmen@ku.edu.tr

Abstract: We consider a single-item, multi-period inventory model where for a retailer purchase and sales prices fluctuate randomly. We assume that a time-homogeneous and Markovian stochastic market price process represents purchase prices for the retailer and determines the sales prices through a markup rate. Customers arrive according to a doubly-stochastic Poisson process where stochastic arrival rates are affected by the sales prices. Upon observing the inventory level and market price, the retailer decides on an order quantity in each period to maximise the expected profits. We provide conditions on the price process that ensure the concavity of the profit functions and show that a price-dependent base-stock policy is optimal. In a numerical study, we implement a simulation model to estimate the expected profit functions and to analyse the effect of price volatility on the optimal expected profits and base-stock levels. We observe that a more volatile price process leads to lower optimal expected profits.

## 1 Introduction

Price uncertainties are among the most critical challenges that retailers and manufacturers have to face. Companies whose operations require procuring from commodity markets are exposed to commodity price fluctuations which in fact experienced sharp movements in the near past. Successful inventory management is an effective approach to mitigate risks due to input or even sales prices. Besides its importance in managing the usual trade-off between holding, shortage and purchase costs, it can create additional value in fluctuating price environments by adjusting the order sizes in response to price. Even more value can be created by integrating pricing decisions with inventory planning as pricing is one of the most effective tools a firm has to control the demand (Chan et al. 2004).

Despite their practical significance, price fluctuations are usually not directly addressed in inventory management literature. Classical inventory models usually take purchase and selling prices as constants. It is clear that for some products, input

prices are volatile. Moreover, it is common that selling prices in some industries may be difficult to predict as well. For instance, a wholesaler that sells in a different currency will bear an exchange rate risk in sales prices. Some industries such as apparel and high technology face the problem of variable selling price due to rapid product substitution and short life cycles. As another example, one can think of jewellery store managers who buy and sell products that are made of precious metals such as gold, silver and precious gemstones such as diamond. Sales prices of these products will be randomly varying over time and customers arriving at different times may be charged different prices.

In this paper, we investigate the inventory management problem of a retailer that sells a commodity-based product whose inherent value is changing according to a stochastic price process. In a multi-period, single-item, periodic review setting, we explicitly model a stochastic market price process which determines both purchase and sales prices and also stochastically describes the customer demand. To incorporate the effect of stochastic sales prices, we model the customer arrival process which is modulated by the market price process such that times of customer arrivals become important in determination of sales prices. In a numerical study, we use Monte Carlo simulation to compute expected profits by first generating random price paths from a continuous price process and then the conditional customer arrival process. We generate insights on the effect of price volatilities on retailer's optimal expected profits and inventory decisions.

The rest of the paper is organised as follows. In Section 2, we briefly review the relevant literature. In Section 3, we present the inventory control model involving price fluctuations. In Section 4, we use simulation to conduct a sensitivity analysis on the optimal values of the models developed before. Lastly, in Section 5, we give our concluding remarks and further research ideas.

## 2 Literature

Several researchers studied the effect of volatile purchase prices on inventory control problems. Kalymon (1971) studies a multi-period inventory model by incorporating the random purchase prices which are determined by a Markov process. He proves that a state-dependent ($s$, $S$) policy is optimal. Berling and Martinez de Albeniz (2011) investigate a Poisson demand system where purchase price is a Markov process to see the effect of price evolution on the optimal policy and its parameters. There is also a wide range of models that consider the use of volatile spot markets where the firms can buy and sell. Goel and Gutierrez (2006) consider a multi-period stochastic inventory model where a firm may purchase from both spot and future markets. Haksoz and Seshadri (2007) review existing models that incorporate spot market procurements with volatile prices in several supply chain operations. Secomandi (2010) studies the warehouse problem of a merchant that involves in commodity-trading activities. He assumes that at each period, the spot price of the commodity evolves as a Markov process. In the presence of both space and buy-sell limits, he shows that optimal policy is characterised by two-stage price dependent base-stock targets.

Despite numerous papers that consider stochastic input prices, few papers investigate the effect of changing sales prices which are mostly due to inflation or a deterministic continuous price decrease. Erel (1992) demonstrates the sensitivity of

the basic model to continuous price changes and concludes that relationship between price change rate and holding cost rate is highly important. Hariga (1995) studies EOQ models with linearly increasing price and demand. Khouja and Park (1993) consider EOQ models with decreasing purchase price motivated by the statistical finding that in high technology industry, component prices decline at a rate of 1% per week. Banerjee and Meitei (2009) consider a single period stochastic demand inventory model with random lead time and continuously decreasing selling price.

The model presented in this paper differs from existing models in the sense that it incorporates the effect of random sales prices by explicitly modelling a continuous-state stochastic price process. It also relates random purchase and sales prices through a proportional retail markup unlike classical models that take sales price as constants or a random variable unrelated to the purchase costs. This case is applicable to situations where a firm is selling in a foreign currency or selling a commodity-based item such that any price fluctuations in the material cost also pass to customers. In addition, besides using random variables for random demands to be realised at the end of sales periods, we model the customer demand also as a stochastic process that depends on the prevailing stochastic prices. The next section presents the details of our model.

## 3 The Model

We consider a multi-period and periodic-review inventory model with fluctuating prices. We assume that there are $M$ periods whose lengths are equal to $T$ units of time where at the beginning of each sales period the retailer places an order. For the model in consideration, we assume the existence of a stochastic market price that follows a continuous-time and time-homogeneous Markov process $P = \{P_t; t \geq 0\}$ with state space $\mathbb{R}_+ = [0, \infty)$. We assume that these random market prices actually represent the purchase prices of the retailer in the analysis, i.e., the value of the market price at review times (i.e., $P_T$, $P_{2T}$, ..) are the purchase prices for the retailer. Unlike most of the inventory models that model the customer demand as a random variable to be realised at the end of each review period, we assume that there is a customer arrival process and it is modulated by the market price process that we consider. More specifically, we assume that the unit customer demand process is a modulated Poisson process where stochastic arrival rate at time $t$ is $\Lambda_t = \lambda(P_t)$ and $\lambda(.)$ is a nonnegative function of the random prices. Customers arrive according to this process and at each arrival they demand one unit of the item. We here note that all the following analysis holds if each customer demands a random amount of the item provided that they are independent and identically distributed. In our setting, since arrival rate is a function of stochastic price, we have a stochastic arrival rate process $\Lambda = \{\Lambda_t; t \geq 0\}$ which modulates the customer arrival process. These types of models are referred as doubly stochastic Poisson processes introduced by Cox (1955) or shortly, Cox processes. If $\Lambda$ is a deterministic function rather than a stochastic process, we have a nonhomogeneous Poisson process. Since we assume that $P$ is a Markov process, $\Lambda$ is also Markovian. We denote the customer purchase process by $N = \{N_t; t \geq 0\}$ where $N_t$ denotes the number of sales by time $t$ and $N_0 = 0$. The arrival times of the customers form a random sequence $S = \{T_n; n \geq 1\}$ where $S$ and $N$ are related as $\{T_n \leq t\} = \{N_t \geq n\}$.

Sales prices are assumed to be proportional to the market price with a fixed markup ratio of $\alpha \geq 1$. In other words, if a customer arrives at time $t$, the sales price is $\alpha P_t$ where $P_t$ is the current random market price. This is appropriate for products that have constantly changing underlying commodity prices like diamond rings, gold bracelets etc. For such products, it is common practice to set a proportional markup rather than an incremental markup. We assume that $\alpha$ is an exogenous economic parameter to the retailer which is plausible for environments with high competition. The type of the product that the retailer in question sells is also relevant. For a gold retailer, for example, setting a retail markup freely will be very difficult for gold coins which are non-differentiated products and easily be substituted by customers at other retailers. On the other hand, for some exclusive products like designer bracelets, necklaces, etc., the retailer may have more freedom in setting his markup since an arriving customer may not find exactly the same product elsewhere. Although we assume a fixed markup rate $\alpha$ in the analysis of this section, in the numerical analysis of Section 4 we analyse the case where the retailer also sets the markup together with the inventory decision.

At the beginning of each period, the decision maker observes the current price and inventory level to make an ordering decision. There is no lead time and the entire order is received immediately at the beginning of each period. Customers that arrive in the case of shortage are assumed to be lost. We assume that for each period, unit inventory holding is $h \geq 0$ and unit lost-sale cost is $b \geq 0$. Moreover an interest rate parameter $r \geq 0$ is used for discounting the revenues. The following analysis also holds if these parameters are functions of initial price at each period.

The objective is to maximise the immediate and subsequent discounted profits. Since we assume that the unit demands arrive according to a Cox process which is modulated by a Markovian price process and the length of intervals are the same, the probability distribution of number of sales in any interval only depends on the initial price at the beginning of that period. In other words they are conditionally independent. We denote the total demand in a period by $N_T$, whose distribution depends on the initial market price $P_0$. We also do not put any restriction on the price process except the Markov property. Since we assume that each demand is of size one, the state space for inventory level is nonnegative integers.

We can write the total expected revenue during a period as a function of initial price p and order-up-to decision $y$ as

$$r(y;p) = E\left[\sum_{n=1}^{\min(N_T,y)} e^{-rT_n}\alpha P_{T_n} \mid P_0 = p\right] = \sum_{n=1}^{y} E\left[e^{-rT_n}\alpha P_{T_n} 1_{\{T_n \leq T\}} \mid P_0 = p\right] \qquad (1)$$

where $N_T$ denotes the number of arrivals during the period and $T_n$ is the time of $n$th arrival since the beginning of the period. The summation inside the expectation in 1 is the total discounted revenue in which the item is sold, if available, at the current sales price $\alpha P_{T_n}$ upon arrival of the $n$th customer since the beginning of the period. Summation is performed until the last sale and every time revenue is collected, it is discounted to the beginning of the period. The expectation in 1 is taken with respect to the random components; number of arrivals, the time of arrivals and the market price. A similar approach appears in Grubbstorm (2007) who considers a single-period problem where demand is modelled as a compound renewal process. He

assumes that sales price is constant and customers that arrive according to a renewal process demand a random amount of the product. There is no fixed sales period and items are sold until the inventory is depleted. Although in our model we sum all individual revenues from each arriving customer, our model construction is different in the sense that we have a finite selling season and sales price is a stochastic process that also modulates the customer arrival process.

The total expected discounted one-period profit can be written similarly as

$$g(y;p) = -py + r(y;p) - E\left[b(N_T - y)^+ + h(y - N_T)^+ \mid P_0 = p\right] \tag{2}$$

The first term in 2 is the total purchase cost for $y$ units at the initial price $p$ and the last term is the one-period expected lost-sale and inventory holding costs. Note that the one-period expected profit is independent of the period. This is due to the fact that conditional random prices $P_{kT+T_n} \mid P_{kT}$ and $P_{T_n} \mid P_0$ have the same distribution for any period $k$ since stochastic market price process is assumed to be Markovian and time-homogenous.

We use dynamic programming to solve the inventory problem to optimality. We define the value function $V_k(x, p)$ as the maximum total expected discounted profit for periods from $k$ to $M$ if the initial inventory is $x$ and market price is $p$. We write the dynamic programming equation as

$$V_k(x,p) = \max_{y \geq x} \{g(y;p) + \gamma \Psi_k(y;p)\} + px \tag{3}$$

where

$$\Psi_k(y;p) = E\left[V_{k+1}\left((y - N_T)^+, P_T\right) \mid P_0 = p\right] \tag{4}$$

denotes the total expected future profits from period $k + 1$ to $M$. Note that the discount factor for a sales period is $\gamma = e^{-rT}$. Since we do not allow backorders, the inventory level can not be negative in the next period. It should be zero if the demand turns out to be more than the total inventory in the current period. We assume that the salvage price is zero. Therefore, the terminal value function is

$$V_{M+1}(x,p) = 0\,. \tag{5}$$

*Lemma 1:* A necessary condition for $g(y;p)$ to be concave in $y$ is

$$E\left[1_{\{T_n \leq T\}}\left(e^{-rT_n}\alpha P_{T_n} + b + h\right) \mid P_0 = p\right] \tag{6}$$

is decreasing in $n$. Moreover, if the expected discounted prices are nonincreasing in time, the condition given in 6 is always satisfied.

*Proof:* Note that forward difference of $g(y; p)$ can be written as

$$\begin{aligned} g(y+1;p) - g(y;p) &= -p + E\left[e^{-rT_n}\alpha P_{T_{y+1}} 1_{\{T_{y+1} \leq T\}} \mid P_0 = p\right] \\ &\quad + bP\{N_T \geq y+1 \mid P_0 = p\} - hP\{N_T \leq y \mid P_0 = p\} \\ &= -(p+h) + E\left[1_{\{T_{y+1} \leq T\}}\left(e^{-rT_n}\alpha P_{T_{y+1}} + b + h\right) \mid P_0 = p\right] \end{aligned} \tag{7}$$

Note that under the condition given in 6, the forward difference of the expected one-period profit function is decreasing in y which ensures that it is integer concave. □

*Theorem 1:* Assume that the necessary condition is satisfied. Then $V_k(x, p)$ is concave in $x$ for every $p$ and a base-stock policy is optimal, i.e., there exists a base-stock level $S_k(p)$ for each period $k$ such that if the inventory level is less than the base-stock level, it is optimal to raise the inventory up to $S_k(p)$. Otherwise, it is optimal to order nothing. Moreover, optimal base-stock level for period $k$ is given by

$$S_k(p) = \inf\left\{ y \geq 0 : \begin{array}{l} P\{N_T \leq y \mid P_0 = p\} \\ \qquad \geq \dfrac{-p + b + E\left[1_{(T_y \leq T)} e^{-rT_y} \alpha P_{T_y} \mid P_0 = p\right] + \gamma \Delta \Psi_k(y; p)}{b + h} \end{array} \right\} \tag{8}$$

*Proof:* We prove the result by induction. First note that terminal value function is trivially concave. Assuming that that $V_{k+1}(x, p)$ is concave for any $k < M$. This makes $\Psi_k(y; p)$ concave. By lemma 1, since $g(y; p)$ is concave in $y$, $g(y; p) + \gamma\Psi_k(y; p)$ is concave in $y$ which also makes $V_k(x, p)$ concave in $x$. This suggests that a price-dependent base-stock policy is optimal. □

## 4 Numerical Analysis

In this section, we conduct a sensitivity analysis on the model presented in Section 3 and aim to investigate the effect of price volatility on the optimal profits and optimal base-stock levels. We focus on the single period model and through a simulation study, we generate insights on the sensitivity of the optimal profits to price volatility by comparing them for various volatility levels.

In our analysis, we assume that the sales period is $T = 1$ unit of time and market prices follow a geometric Brownian motion process with drift $\mu$ and volatility $\sigma$. We also assume that the arrival process is a doubly-stochastic Poisson process with the following piecewise linear rate function

$$\Lambda_t = (A - B\alpha P_t)^+. \tag{9}$$

A direct analytical approach to the above model is challenging. Therefore, we employ a simulation approach to estimate the performance measures. To generate the market prices, we use $n = 100$ equally-spaced discretisation of the interval [0,1] to simulate a random path for a Wiener process $W = \{W_t; 0 \leq t \leq T\}$ where it is well-known that $W_{tk} - W_{tk-1} \sim$ Normal $(0, 1/n)$. We then use these Wiener realisations in the analytical Ito' solution for the geometric Brownian motion which is

$$P_t = P_0 e^{(\mu - \sigma^2/2)t + \sigma W_t}. \tag{10}$$

(Baxter and Rennie 1996). This way we generate a sample path for a geometric Brownian motion with drift $\mu$ and volatility $\sigma$. Then, conditional on the price path, we generate a nonhomogeneous Poisson process stream for the customer arrivals using the thinning algorithm of Lewis and Shedler (1979). According to the thinning algorithm, for each price path, we set $\lambda_{max} = A$, and generate a Poisson process with rate $\lambda_{max}$. At each arrival time, we generate an independent $U =$ Uniform (0, 1) random variable and accept the arrival if $U \leq \Lambda_U / \lambda max$ where

$\Lambda_U = (A - B\alpha P_U)^+$. This way, we generate a nonhomogeneous Poisson process for each realised price path.

We take the initial price $P_0$ = 20 and assume that h = b = r = 0 for simplicity. To investigate the effect of price volatility, we take the drift of the price process $\mu$ = 0. This makes the price process in 10 a martingale, i.e., the theoretical expectation is $E[P_t] = P_0$ which means that the process remains the same in expectation. Although increasing the value of price volatility $\sigma$ does not change the theoretical expected future values of prices, the variability increases. We use 10000 replications to estimate the expected profit function

$$g(y, \alpha; p) = -py + E\left[\sum_{n=1}^{\min(N_T, y)} \alpha P_{T_n} \mid P_0 = p\right]. \tag{11}$$

Note that 11 is the same as 2 when h = b = r = 0 and in addition, the markup $\alpha$ is also a decision variable. We use volatility values of [0, 0.02, .. , 0.8] for the analysis and at first we fix the markup $\alpha$ = 1.8. As seen in Figure 1, although optimal base-stock levels that maximise 11 exhibit nonmonotone behaviour, they have a tendency to decrease as price volatility increases.

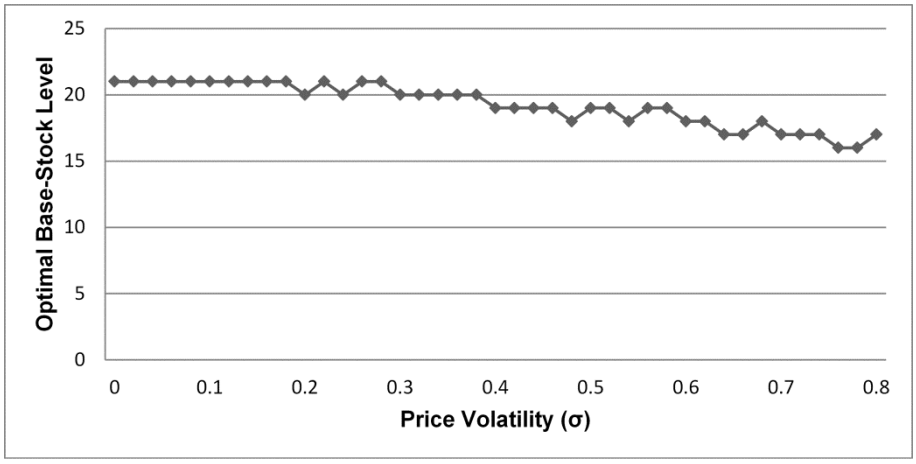



*Figure 1: Sensitivity of the optimal base-stock level to price volatility*

The bottom curve in Figure 2, on the other hand, shows the optimal expected profits for the same price volatility levels when $\alpha$ = 1.8 is used as the sales markup. The top curve titled g(y*, α*) shows the optimal expected profits when the expected profit function in 11 is optimised over order quantity $y$ and markup $\alpha$. We observe that whether the markup is exogenous to the retailer or not, the optimal expected profits decrease as prices get more volatile while their expected future values remain the same. This suggests that price volatility is undesirable for a retailer that faces the risk of random sales price although it has the power of adjusting the order size and setting the retail markup.

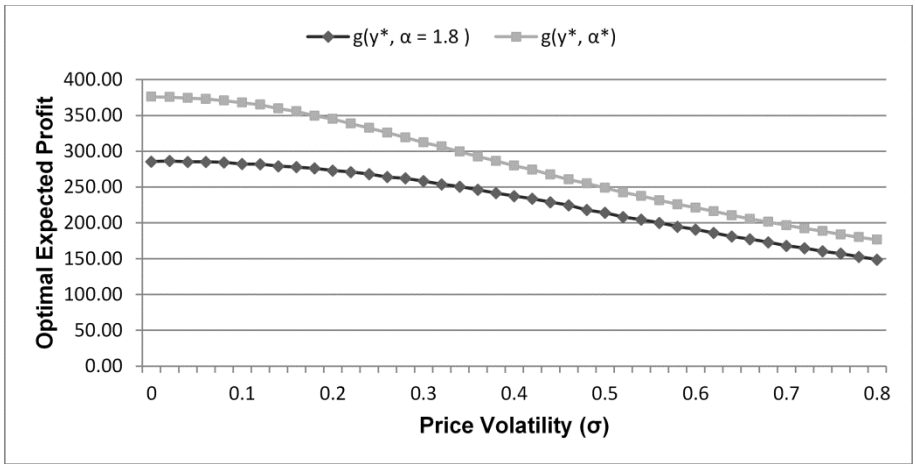



*Figure 2: Sensitivity of the optimal expected profits to price volatility*

## 5 Conclusion

This paper considers a multi-period inventory management problem of a retailer that faces both stochastic input and sales prices. From a modelling point of view, we assume that inherent value of the product in question is fluctuating due to exchange rates or commodity content. We refer to these fluctuating prices as the market price and assume that they follow a time-homogeneous Markov process. We also model the customer arrival process as a doubly stochastic Poisson process that is governed by a stochastic intensity process where the stochastic intensity at any time is determined by the random market price and the sales markup. We model this inventory management problem using dynamic programming with states being the current inventory level and market price at the review periods. We show that a price-dependent base-stock policy is optimal under a condition on the price process. Moreover, it is sufficient to have non-increasing expected discounted market prices to ensure the concavity of the expected profit functions.

In the last part, we conduct a sensitivity analysis to observe the effect of volatile prices on the optimal decisions and optimal profits. To compute the value of the expected profit functions, we use Monte Carlo simulation as it is rather difficult and time-consuming to compute it analytically for continuous-time, continuous-state price processes. Using the fact that a doubly stochastic Poisson process reduces to a nonhomogeneous Poisson process conditioned on the stochastic arrival rates, we first generate market prices at each replication as they define the arrival rates. We assume that market prices follow a geometric Brownian motion having no drift. For a fixed volatility level, we generate the prices at equidistant time points in the sales season. Then for each realised price path, we generate a nonhomogeneous Poisson stream using a thinning algorithm and estimate the single-period profit function. We observe that as the prices get more volatile, optimal expected profits decrease when the sales markup is a decision variable or not.

## References

Banerjee, S.; Meitei, N.: Effect of declining selling price: Profit analysis for a single period inventory model with stochastic demand and lead time. Journal of the Operational Research Society 61 (2009) 4, pp. 696-704.

Baxter, M; Rennie, A.: Financial Calculus An introduction to derivative pricing. Cambridge: Cambridge University Press 1996.

Berling, P; Martínez-de Albéniz, V.: Optimal inventory policies when purchase price and demand are stochastic. Operations Research 59 (2011) 1), pp.109-124.

Chan, L. M.; Shen, Z. M.; Simchi-Levi, D.; Swann. J. L.: Coordination of pricing and inventory decisions: a survey and classification. In: Simchi-Levi, D.; Wu, S.D.; Shen, Z.-J. (eds): Handbook of Quantitative Supply Chain Analysis. New York: Springer Science+Business Media 2004, pp. 335–392.

Cox, D. R.: Some statistical methods connected with series of events. Journal of the Royal Statistical Society. Series B (Methodological) 17 (1955) 2, pp. 129-164.

Erel, E.: The effect of continuous price change in the EOQ. Omega, 20 (1992) 4, pp. 523-527.

Goel, A.; Gutierrez, G.: Integrating commodity markets in the optimal procurement policies of a stochastic inventory system. Available at SSRN 930486, 2006.

Grubbström, R.: The newsboy problem when customer demand is a compound renewal process. European Journal of Operational Research 203 (2010) 1, pp. 134-142.

Haksöz, Ç; Seshadri, S.: Supply chain operations in the presence of a spot market: A review with discussion. Journal of the Operational Research Society 58 (2007) 11, pp. 1412-1429.

Hariga, M.: Effects of inflation and time-value of money on an inventory model with time-dependent demand rate and shortages. European Journal of Operational Research 81 (1995) 3, pp. 512-520.

Kalymon, B.: Stochastic prices in a single item inventory purchasing model. Operations Research 19 (1971) 6, pp. 1434-1458.

Khouja, M; Park, S.: Optimal lot sizing under continuous price decrease. Omega 31(2003) 6, pp. 539-545.

Lewis, P.A. W.; Shedler, G.A.: Simulation of Nonhomogeneous Poisson Processes by Thinning. Naval Research Logistics Quarterly 26 (1979) 3, pp. 403-414.

Secomandi, N.: Optimal commodity trading with a capacitated storage asset. Management Science 56(2010) 3, pp. 449-467.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Cost and Environmental Trade-offs in Supply Chain Network Design and Planning: The Merit of a Simulation-based Approach

## *Kosten und Austauschbeziehungen mit der Umwelt bei Entwurf und Planung von Supply Chains: Die Vorteile eines simulationsbasierten Ansatzes*

Christos Keramydas, Mallidis Ioannis, Vlachos Dimitrios, Iakovou Eleftherios,
Aristotle University of Thessaloniki, Thessaloniki (Greece), chkeramy@auth.gr,
imallidi@auth.gr, vlachos1@auth.gr, eiakovou@auth.gr

Abstract: In this paper, a simulation-based optimisation methodology is presented for the joint design and planning of globalised supply chains under cost and CO2 emissions minimisation objectives. The aim of this work is to investigate whether or not, and how, the critical decisions and insights – that derive from the respective analytical optimisation process – change, as their critical assumptions e.g. deterministic lead times, no-backorders constraints, power-of-two assumptions, etc., are relaxed. To that end, the proposed methodology is applied to a real-world supply chain, and the outcomes are directly compared to those derived through the related analytical process. The results reveal that the cost and emissions benefits, as well as the relevant managerial insights, gained through the simulation-based approach are meaningful and could provide new feasible, cost- and environmental-efficient options to the design and planning decisions of such networks.

## 1 Introduction

As consumers are nowadays interested in the environmental footprint of the products they consume, the greening of supply chain operations has turned into a top priority in the corporate agenda. In this context, a plethora of analytical methodologies have been developed to incorporate environmental considerations into the supply chain networks research. Given the complexity of such networks, several assumptions are required to be made in order to obtain meaningful analytical solutions. The trade-off is that the corresponding results probably differ from those expected when addressing the problem in its original form, i.e. with no assumptions made.

On the other hand, simulation-based optimisation methodologies have been traditionally employed in the cost optimisation of complex multi-echelon supply chain networks. A main advantage of simulation consists in its flexibility to relax critical

assumptions that are necessary within the relevant analytical optimisation processes, thus providing more realistic solutions. Indicatively, some of the latest simulation-based cost optimisation models regarding the efficient management of supply chains are presented in the works of Cigolini et al. (2014), He et al. (2014), and Henniesa et al. (2014). Additionally, more recently, a number of green simulation-based optimisation methodologies have been developed in order to further incorporate emissions minimisation routines in their optimisation processes (Mangla et al. 2014; Abduaziza et al. 2015; Garcia and You 2015).

However, there is a lack of papers dedicated to the quantitative investigation of the merit of a simulation-based approach – in comparison to an analytical one – when dealing with the design and planning optimisation of supply chain networks under cost and environmental trade-offs. In other words, there is a gap in identifying whether, or not, and how the critical decisions and the relevant insights change, as several critical assumptions of the associated analytical solution process are relaxed.

The contribution of the paper follows this direction. Specifically, a representative realistic example of a complex global multi-echelon supply chain is selected from the most recent literature (Mallidis et al., 2014). The problem deals with the joint design and planning of a supply chain under cost and $CO_2$-emissions minimisation objectives. To that end, a discrete event simulation (DES) optimisation methodology is developed allowing for the gradual relaxation of several critical constraints, such as the deterministic lead times, no-backorders constraints, power-of-two assumption, etc. Finally, the network design and planning decisions, as well as the results of the simulation approach, regarding the total costs and $CO_2$ emissions, are compared to the results of the analytical solution. The selection of the simulation technique was based on the fact that DES proved to be a reliable and effective approach to real-world multi-echelon supply chain management problems (Schmitt and Singh 2012). Alternative simulation techniques, e.g. Monte-Carlo simulation, could also be valuable. The aim of this paper is to examine whether it is meaningful, or not, to employ simulation-based approaches in the relevant decision-making processes, thus providing more realistic tools and alternatives to the involved companies.

The rest of the paper is organised as follows. In Section 2, we provide the description of the supply chain system under study, while in Section 3 we describe the simulation optimisation model. Section 4 demonstrates the applicability of the proposed methodology through a realistic case study. The relevant numerical results are presented in Section 5. Finally, Section 6 wraps up with the managerial insights.

## 2 System Description

This paper examines three of the most prevalent logistics network design options employed by multinational companies in the EU, as these options are presented in the work of Mallidis et al. (2014). In the first option (Option DC), the company considers the market as a single region and serves it through a number of retail stores (RSs). The RSs receive their orders from one large central distribution centre (DC) through truck transportation. To this end, the central DC places its orders to a faraway distant manufacturing site (a loading point, LP). These orders (containers) are transported to an entry port (EP) through deep-sea shipping and then by heavy-duty trucks to the central DC.

In the second option (Option DC/SDCs), the market is partitioned into smaller regions. In each region, the RSs are exclusively served through a peripheral DC (SDC). The peripheral DC is replenished by the central DC via intermodal connections, thus acting as a "satellite".

In the third option (Option M/DCs), there is no central DC. The market is split into the same number of regions, as in the second option, with one regional DC allocated to each region. Each regional DC is replenished independently from the LP via deep-sea shipping, through the respective EP. A simplified realisation of the problem is illustrated in Figure 1, including the three alternative options.

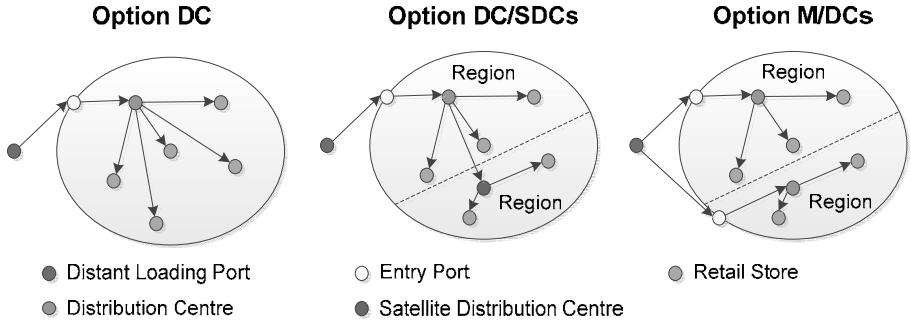



*Figure 1: Logistics network design options (Mallidis et al. 2014)*

In all options, we assume that transportation from the DCs (central, regional and satellite) to the RSs occurs by truck. More specifically, delivery trucks are employed for short transportation distances, within city limits, and heavy-duty trucks for long-haul regional transportation distances. Additionally, the company is fully accountable for its DC facilities and truck emissions, as trucks and DC facilities are assumed to be leased through long-term contracts. Specifically, with respect to truck transportation, this implies that the company will also be accountable for the return truck trip $CO_2$ emissions and costs. This is not the case however, for the barge, rail, and short sea shipping transportation modes employed in the intermodal connections of Option DC/SDCs and for the deep-sea ships employed in all options.

Moreover, the respective trips from the LP to the DC(s), in all options, are associated with lengthy, albeit deterministic, lead times. These lead times result in increased and variable waiting times for end-customers when a stock out occurs at the DC. Thus, in the corresponding analytical solution, the inventory level at the DCs (excluding SDCs) is considered to be large enough to practically satisfy all orders placed by the RSs they serve. This necessitates the operation of quite large DCs that can provide high service levels (negligible probability of a stock out).

On the other hand, the satellite DCs act as intermediate echelons, as they are replenished by the central DC. This further implies that: (i) their capacity could be smaller and thus the probability of a stock out occurrence could be significant; (ii) a simultaneous stock out occurrence at the satellite DC and the RS would increase the necessary time for a backorder (at the RS) to be fulfilled.

The critical decisions for managing the above logistics network design options involve the determination of: (i) the stock levels reserved at the RSs and the satellite DCs; (ii) the order delivery frequencies to the RSs; and (iii) the optimal logistics network design option in terms of cost and $CO_2$ emissions. The optimisation criteria employed are two. The first one concerns the total logistics costs, including the holding cost of a product per time unit at the RSs and the DCs (central, satellite and regional), as well as the backorder cost of a product per time unit at the RSs, and the backorder cost of a product at the satellite DC. Moreover, the total logistics costs include the transportation cost per trip/round trip (depending on the transportation mode employed) between the nodes of the alternative networks, and the (central, regional and satellite) DCs' operating costs. The second one is the logistics $CO_2$ emissions, including the transportation $CO_2$ emissions per trip/roundtrip between the nodes of the network, and the DCs' operating $CO_2$ emissions.

The review periods of the DCs (central, regional and satellite) are considered fixed as they are determined by the delivery frequency of the transportation mode employed for their replenishment. Specifically, the review period of the central DC (Option DC) and the regional DCs (Option M/DCs) are determined by the liner deep sea shipping schedule between the LP and their EP. In a similar manner, the review period of each satellite DC (of Option DC/SDCs) is dictated by the transportation schedules of the intermodal connections between the central and satellite DCs.

## 3 Model Development

A Discrete Event Simulation (DES) methodology is proposed to emulate the system's SC-related operational procedures, and the relevant cost- and emissions-related metrics. Indicatively, Figure 2 exhibits the DES algorithm logic diagram for the first network design option, i.e. Option DC. The model can be readily fine-tuned to also represent the other two alternative options under study.

From a technical point of view, the model includes eight simulation loops (SLs). Specifically, SL1 generates customer arrivals to the RS, determines the relevant demand size, and updates the on-hand inventory (OH), backorders (BO), and inventory position (IP) metrics at the RS. SL2 represents the RS inventory review process, including the periodical reviewer's "walk-in", the size determination of each order placed to the DC (order up-to-S policy), and the determination of the order quantity that the DC is able to deliver. Moreover, SL2 updates the OH, BO, IP, and on-order inventory (OO) metrics both at the RS and the DC. SL3 corresponds to the order delivery process from the DC to the RS. Specifically, SL3 carries out the calculation of the total truck load and the required number of trucks, and also the calculation of the corresponding transportation emissions. Moreover, it determines the relevant truck trip lead time, and updates BO, OH, and OO metrics at the RS.

SL4 presents the DC review process, including the periodic reviewer's "walk-in", the size determination of the order placed to the LP (order up-to-S policy), and the update of the OO and IP metrics at the DC. SL5 describes the first step of the overall LP to DC delivery process, i.e. the LP to EP shipment. Specifically, SL5 calculates the total ship load and the number of the required containers, the duration of the relevant ship trip, as well as the corresponding transportation emissions. SL6 describes the second step of the delivery process, i.e. from the EP to the DC. SL6

determines the total truck load and the number of trucks required, as well as the associated truck trip duration. Moreover it calculates the relevant transportation emissions, and updates the BO, OH, and OO indicators. Finally, SL7 and SL8 are responsible for calculating and updating the system's key performance metrics, i.e. holding (HC), backorders (BC), and transportation costs (TC) at the RS and the DC, and additionally the DC operational costs.

The input data (logistics and transportation cost data, transportation emission data, and demand parameters), which were used to carry out the simulation optimisation process is presented, in detail, in the work of Mallidis et al. (2014). The data were employed in their original form; no transformations or changes were applied.

The models were verified mainly through the checking of the model elements one-by-one, the tracking of the possible events along the simulation logic diagram compared to the events of the associated simulation run, and the available software error-checking procedures. The validation process was based on the comparison of the simulation results, for various scenarios and problem settings, to the real-system's performance, as well as on the experience of the involved managers.

The Arena® Simulation Software was used in modelling the alternative scenarios. The Optquest® for Arena® tool was employed to perform the necessary simulation optimisation procedures. A set of 100 replications was executed for each problem configuration corresponding to each one of the simulation cases. The simulation time-length of a replication was one year (time horizon of the analysis). The warm-up period was one month. The average time of an optimisation run was approximately 4 hours (Intel® CoreTM i7-3770 CPU@3.40GHz).

## 4 Realistic Case Study

In this Section, we illustrate the applicability of the proposed methodology in the case of a white goods retailer who serves the Greek market, as this case is described in the associated manuscript of Mallidis et al. (2014). The three alternative logistics networks follow the relevant configuration options presented in Section 2.

The first logistics network serves one large region and consists of one EP ($P_0$) (Piraeus) and one central DC that is located in the wider area of Aspropirgos and is served through the Piraeus EP. The region includes two privately owned RSs in the wider metropolitan Athens area (Marousi-$RM_1$, Argiroupoli-$RM_2$), two in the Peloponnese area (Tripoli-$RM_3$, Patra-$RM_4$), and one in each of the metropolitan areas of Ioannina-$RM_5$, Larissa-$RM_6$, Thessaloniki-$RM_7$, and Ksanthi-$RM_8$.

In the second option, the market is split into two regions. The first region is served through RSs (1-4), and the second one, through RSs (5-8). Both regions are indirectly served through the central DC via the respective peripheral SDCs. Specifically, in this case study, the central DC is located within the first region and involves short transportation distances to its serving RSs (up to 211 km). Therefore, the operation of a SDC for serving the RSs within the first region does not seem to be a realistic option. As a result, we only consider one SDC, which is employed for serving the second region (within a transportation range of 711 kms) by trucks, while the RSs in the first region are served directly by the central DC, also by trucks.

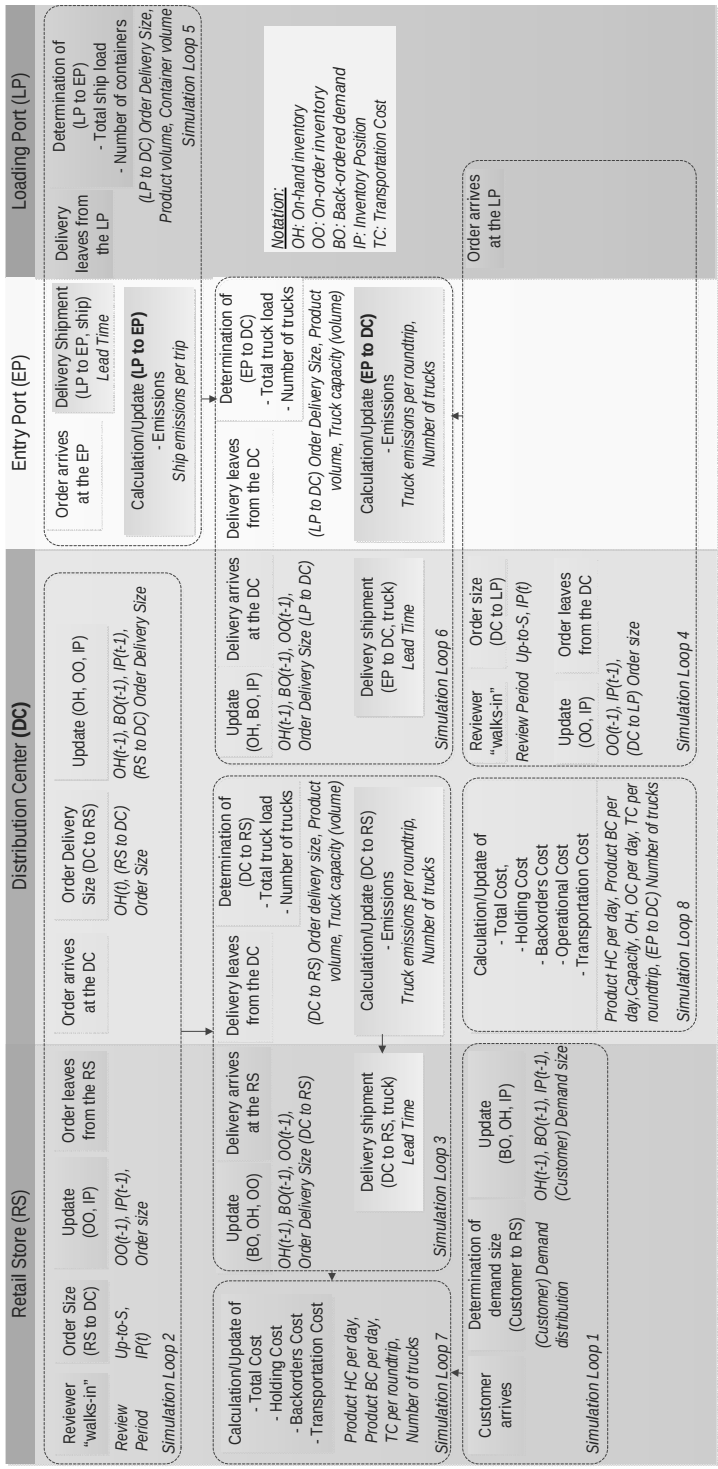



*Figure 2: Simulation algorithm logic diagram*

In the third option, there are two independent regional DCs ($DC_1$, $DC_2$) that order directly from the manufacturing site in Shanghai. $DC_1$ operates in place of the central DC of Option DC, while $DC_2$ operates in place of the SDC of Option DC/SDCs. Moreover, the additional EP of Thessaloniki ($P_2$) is employed for serving $DC_2$ in Kalochori, through heavy-duty truck transportation. The first entry port of Piraeus ($P_1$), still serves $DC_1$ in Aspropirgos. Both EPs receive container flows from Shanghai, China. To this end, the RSs of each region are served by the same transportation modes as those employed in the second option (DC/SDCs).

The review period of the central DC in the (DC) Option, and of the regional DCs in the (M/DCs) Option, is determined by the Shangai deep-sea shipping schedule and is equal to 30 days. Though, given that the RSs review periods are power-of-two time units, in order to coordinate orders arriving to the DCs, the DC review period is actually considered equal to 32 days. The review period of the SDC is determined by the rail transportation schedule between Aspropirgos and Thessaloniki and is equal to 7 days. Similarly, we consider a review period for the SDC equal to 8 days.

In this context, the proposed simulation-based optimisation approach includes the gradual relaxation of the most critical assumptions of the analytical optimisation process. Specifically, the first one relaxed is the “Powers of Two” (PoT) assumption regarding the coordination of the review periods at the RSs. The second one is the “no backorders” assumption at the DC, which is relaxed by treating the associated safety factor (SF) (order-up-to-S level) as a decision variable. The third one is the deterministic lead times (LT), and is relaxed by considering stochastic ship trip lead times, i.e. the duration of the LP to EP trip. These steps are separately performed for both the cost and $CO_2$ emissions minimisation criteria.

## 5 Numerical Results

In this section, we indicatively provide the numerical results regarding the first option, i.e. Option DC. The simulation optimisation results are compared to those of the analytical solution, which are presented and extensively discussed in the work of Mallidis et al. (2014). The latter are employed as the reference base scenario in order to evaluate the merit of the simulation-based approach. The results for the cost and the $CO_2$ emissions criteria are presented in Table 1 and Table 2, respectively, and the corresponding inventory control policies are presented in Table 3.

The relaxation of the PoT assumption, under the cost minimisation criterion, leads to a decrease of the total cost by 1.61%, mainly due to the decrease of the holding costs. Moreover, it seems that the RSs are classified into two groups, i.e. a group with a 3-day, and a group with a 5-day review period, respectively. This solution is associated with a $CO_2$ emissions decrease of 1.31%. Furthermore, relaxing the SF assumption raises up the overall cost benefit to 3.21%. Now, most of the RSs are coordinated at a 4-day review period. The results indicate that the operation of the DC at the lower (type I) service level of 22%, versus an always full DC, could prove to be a practically desirable option, as it is associated to a critically lower average inventory on-hand. Moreover, the total $CO_2$ emissions gains are estimated to be 1.70%. On the other hand, the relaxation of the LT assumption increase the stochasticity of the overall problem and incurs additional costs to the network. As the network tries to hedge against random lead times, there is a significant increase

of the holding costs. The overall cost at the new optimal solution is increased by 1.39% compared to the previous one. Moreover, the RSs are now classified, mainly, into three coordination groups, i.e. with review periods equal to 2, 5, and 8 days respectively. Consequently, the simulation-based optimisation leads to an average cumulative benefit of approximately 4.60%, in terms of the network's total costs, and a benefit of 1.76% in terms of the network's $CO_2$ emissions.

As concerns the $CO_2$ emissions minimisation criterion, the relaxation of the PoT assumption is associated to a decrease of the total cost by 1.10%. Moreover, it seems that the RSs are totally coordinated at the same review period, i.e. 17 days. The corresponding $CO_2$ emissions are reduced by 1.15%. Furthermore, relaxing the SF assumption provides an extra gain to the network, as the total cost benefit ramps up to 2.03%, given that holding costs are significantly decreased. The RSs retain their identical ordering frequencies, while the DC continues its operation at the high service level of 99%. Moreover, the total $CO_2$ emissions gains are estimated to 1.38%. However, when relaxing the LT assumption, all the nodes of the network exhibit some critically higher on-hand inventories. Thus, the new optimal solution increases the overall cost by 2.94%, compared to the last one. Most of the RSs retain their common ordering frequency. Finally, the simulation-based optimisation leads to an average cumulative benefit of approximately 4.97%, in terms of the network's total costs, and a benefit of 2.76% regarding the $CO_2$ emissions decrease.

*Table 1: Cost breakdown and $CO_2$ emissions of the cost optimal options*

| *Scenario* | *PoT* | *%* | *SF* | *%* | *LT* | *%* |
|---|---|---|---|---|---|---|
| Holding Cost [€/day] | 7108 | -3.24 | 6424 | -12.55 | 6970 | 8.49 |
| Backorder Cost [€/day] | 65 | -34.26 | 175 | 76.58 | 172 | -1.53 |
| Operating Cost [€/day] | 1491 | -1.06 | 1456 | -3.39 | 1479 | 1.60 |
| Transportation Cost [€/day] | 36679 | -1.22 | 36552 | -1.57 | 36606 | 0.15 |
| *Total Cost [€/day]* | *45343* | *-1.61* | *44607* | *-3.21* | *45227* | *1.39* |
| *Total Emissions [tons/day]* | *16.98* | *-1.31* | *16.91* | *-1.70* | *16.92* | *0.06* |

*Table 2: Cost breakdown and emissions of the $CO_2$ emissions optimal options*

| *Scenario* | *PoT* | *%* | *SF* | *%* | *LT* | *%* |
|---|---|---|---|---|---|---|
| Holding Cost [€/day] | 11279 | -1.83 | 10894 | -5.18 | 11653 | 6.61 |
| Backorder Cost [€/day] | 387 | 9.94 | 410 | 16.48 | 845 | 123.58 |
| Operating Cost [€/day] | 1500 | -0.46 | 1485 | -1.46 | 1504 | 1.26 |
| Transportation Cost [€/day] | 36102 | -1.00 | 36016 | -1.24 | 36269 | 0.69 |
| *Total Cost [€/day]* | *49268* | *-1.10* | *48805* | *-2.03* | *50271* | *2.94* |
| *Total Emissions [tons/day]* | *16.71* | *-1.15* | *16.67* | *-1.38* | *17.29* | *1.38* |

*Table 3: Optimal service levels (type I) and review periods of all problem instances*

| Criterion | Cost | | | | | | CO2 emissions | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scenario | PoT | | SF | | LT | | PoT | | SF | | LT | |
| Value | a[%] | R[days] | a | R | a | R | a | R | A | R | a | R |
| RS1 | 90 | 3 | 100 | 2 | 96 | 2 | 37 | 17 | 44 | 17 | 64 | 17 |
| RS2 | 93 | 3 | 90 | 4 | 100 | 2 | 37 | 17 | 41 | 17 | 76 | 18 |
| RS3 | 84 | 5 | 88 | 4 | 93 | 5 | 68 | 17 | 76 | 17 | 56 | 17 |
| RS4 | 84 | 5 | 84 | 4 | 84 | 5 | 34 | 17 | 31 | 17 | 56 | 17 |
| RS5 | 69 | 5 | 76 | 4 | 66 | 7 | 24 | 17 | 26 | 17 | 48 | 17 |
| RS6 | 79 | 5 | 82 | 4 | 76 | 8 | 24 | 17 | 24 | 17 | 48 | 18 |
| RS7 | 76 | 3 | 76 | 4 | 76 | 3 | 31 | 17 | 31 | 17 | 52 | 17 |
| RS8 | 62 | 5 | 69 | 5 | 58 | 8 | 24 | 17 | 24 | 17 | 64 | 19 |
| DC1 | 100 | 32 | 22 | 32 | 96 | 32 | 100 | 32 | 99 | 32 | 100 | 32 |

## 6 Summary and Conclusions

This paper presents a simulation-based optimisation methodology for the joint design and planning of globalised multi-echelon supply chains, under both cost and $CO_2$-emissions minimisation objectives. The methodology aims to investigate whether the decisions and the relevant insights change, or not, and how do they change, as the critical assumptions of the associated analytical optimisation processes are relaxed. An indicative example from the most recent literature is employed in order to access the merit of the simulation-based approach.

As regards the supply network planning decisions, the results indicate that the total cost benefit due to the simulation optimisation process, is approximately 4.5-5.0%. The respective $CO_2$ emissions benefit is lower, approximately 1.8-2.8%. More specifically, when relaxing the PoT assumption, the results reveal an average cost decrease of 1.1%-1.6%. Although, in all cases, the RSs ordering policies differ from those derived through the analytical solution process, the change of the total network costs are in line with those presented in the literature. The cost under the PoT assumption is not higher than the optimal cost by more than 6% (Muckstadt and Sapra 2010). Moreover, the paper documents the validity of this conclusion when also including the transportation cost in the total cost function. Additionally, the RSs seems to coordinate (in groups) regarding their ordering frequencies. Furthermore, the results reveal some new feasible network planning solutions that allows the central DC to efficiently operate at lower service levels, instead of operating under the “no backorders” constraint of the analytical solution. Finally, the extra “hidden” costs that arise due to the randomness of the lead times are estimated to 1.4%-3.0%.

Since this research is still in progress, more reliable insights would be available, in detail, at the time of the conference. However, the first preliminary results for the other two options reveal a similar cost and $CO_2$ emissions behaviour in terms of the network’s planning decisions. Moreover, the network design decisions do not seem to be critically affected by the difference in the optimisation approach.

The most critical limitation of the proposed optimisation technique is its model-specific character. Therefore, several methodological and modelling amendments are required in order to implement the proposed process to other supply chain management problems. Moreover, the relevant managerial insights should be critically generalised and interpreted. In the near future, the study of similar systems could enhance our knowledge and enlighten our view regarding the merit of simulation in the relevant real-world decision-making processes. These studies could further increase the reliability of the current insights and allow for their generalisation, or, alternatively, designate their case-specific characteristics. At present, the results, albeit being model-specific, renders simulation as a meaningful alternative approach when taking into account environmental considerations in the design and planning of multi-echelon supply networks.

## Acknowledgements

This research paper has been conducted in the context of the GREEN-AgriChains project that is funded from the European Community's Seventh Framework Programme (FP7-REGPOT-2012-2013-1) under Grant Agreement No. 316167. All the above reflect only the authors' views; The European Union is not liable for any use that may be made of the information contained herein.

## References

Abduaziza, O.; Cheng, J.; Tahara, R.; Varma, R.: A hybrid simulation model for green logistics assessment in automotive industry. Procedia Engineering 100 (2015), pp. 960-969.

Cigolini, R.; Pero, M.; Rossi, T.; Sianesi, A.: Linking supply chain configuration to supply chain performance: A discrete event simulation model. Simulation Modelling Practice and Theory 40 (2014), pp. 1-11.

Garcia, D.; You, F.: Supply chain design and optimization: Challenges and opportunities. Computers and Chemical Engineering, In Press (2015), http://www.sciencedirect.com/science/article/pii/S0098135415000861.

He, J.; Huang, Y.; Chang, D.: Simulation-based heuristic method for container supply chain network optimization. Advanced Engineering Informatics, In Press (2014), http://www.sciencedirect.com/science/article/pii/S1474034614000846.

Henniesa, T.; Reggelina, T.; Tolujewa, J.; Otto, P.: Mesoscopic supply chain simulation. Journal of Computational Science 5 (2014) 3, pp. 463-470.

Mallidis, I.; Vlachos, D.; Iakovou, E.; Dekker, R.: Design and planning for green global supply chains under periodic review replenishment policies. Transportation Research Part E 72 (2014), pp. 210-235.

Manglaa, S.; Kumarb, P.; Baruac, M.: Monte Carlo simulation based approach to manage risks in operational networks in green supply chain. Procedia Engineering 97 (2014), pp. 2186-2194.

Muckstadt, J.; Sapra, A.: Power-of-Two policies. In: Muckstadt, J.; Sapra, A. (eds.): Principles of Inventory Management: When you are Down to Four Order More. New York: Springer 2010, pp. 47-84.

Schmitt, A.; Singh, M.: A quantitative analysis of disruption risk in a multi-echelon supply chain. Int. Journal of Production Economics 139 (2012) 1, pp. 22-32.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Comparing Different Simulation-based Optimisation Approaches for Simultaneous Optimisation of Production Planning Parameters

## *Vergleich verschiedener simulationsbasierter Optimierungsmethoden für die simultane Optimierung der Produktionsplanungsparameter*

Thomas Felberbauer, René Schnirzer, Klaus Altendorfer, University of Applied Sciences Upper Austria, Steyr (Austria), thomas.felberbauer@fh-steyr.at, renekurt.schnirzer@fh-steyr.at, klaus.altendorfer@fh-steyr.at

Abstract: This paper investigates optimal MRP-parameter settings i.e. optimal lot size, planned lead time and safety stock for a production system with multi items, multi stages and stochastic production system parameters. The production system faces uncertainty in demand, setup times, processing times, and machine failures. For the optimisation of the MRP parameters a simulation-based optimisation approach is used, connecting a generic simulation model to a framework for heuristics and evolutionary algorithms. The objective is to minimise tardiness and inventory cost. The performance of a genetic and an evolutionary algorithm are compared to a grid search procedure. The observations deliver further insights into the optimal MRP parameter setting, the traceability of the optimal parameters and their practical applicability.

## 1 Introduction

The target of manufacturing companies is to deliver on time according to the customer requirements, but also to consider internal costs. To enable a good performance of manufacturing systems, the parameterisation of production planning methods has therefore to include the internal as well as the external perspective. For a manufacturing system applying a material-requirements-planning (MRP) approach, the aim of this article is to find near-optimal MRP parameters values for safety stock, planned lead time and lot size. The internal performance is measured as inventory costs for work-in-process (WIP) and finished-goods-inventory (FGI) while the external performance is measured as backorder costs. A multi stage, multi item, stochastic production planning problem is investigated using a simulation-based optimisation approach. The production system faces uncertainties like stochastic demand, set-up times, processing times and machines failures. The performance of the two metaheuristics genetic algorithm (GA) and evolution strategy (ES)

are compared to a grid search procedure. Furthermore, the optimal parameter values found with the metaheuristics are discussed for their practical applicability. The results of the grid search, using some available problem knowledge, are compared to the metaheuristics with respect to the overall costs and the practical applicability of the derived parameter values.

Both tested metaheuristics show a good performance with respect to solution quality, about 20% lower costs than the grid search, and converge in an acceptable amount of time. However, none of the metaheuristics finds a good solution within the grid search iteration number. Furthermore, the metaheuristics results show that they have disadvantages in practical applicability and traceability of the optimal planning parameters found.

## 2 Literature Review

According to Slack et al. (2004), MRP is a preferable method for production planning in systems with complex production system structures where extensive computations are needed to create an appropriate production schedule. Therefore, MRP is a core element of most recent ERP systems, and its position against other planning methods has strengthened since the 1990's (Jonsson and Mattsson 2006). The calculation of this algorithm is based on four steps: *Netting* (determine net requirements by subtracting on-hand inventory and scheduled receipts from the gross requirements), *Lot sizing* (combine the netted demand to appropriate lot sizes), *Backward scheduling* (calculating the planned start dates), *BOM explosion* (use start time, lot size and BOM to generate gross requirements for all required components) and *Iterate* (repeat step 1 to 4 until all components at all low level codes are processed). The MRPII concept (manufacturing resource planning) combines MRP with a hierarchical production planning concept, which is based on three planning levels: *Long-term planning* (strategy), *intermediate planning* (tactical) and *short-term control* (operational). The functions of MRPII are forecasting and aggregated production planning for long-term, MPS (master production scheduling) and MRP for the intermediate planning level and dispatching for short-term control. For a more detailed description of MRPII see Hopp and Spearman (2008).

The optimal setting of MRP planning parameters (i.e. lot size, safety stock and planned lead time) is a critical issue due to their influence on the performance of a production system. This influence is for example investigated in a recent work of Gansterer et al. (2014). Researchers are dealing with the problem of MRP parameter optimisation applying two kinds of methodologies. Firstly analytical models are used to do the parameter optimisation for streamlined production systems and secondly simulation-based optimisation, using metaheuristics, is applied to identify optimal parameter settings for more complex production systems. Many researchers (e.g. Yano 1987; Karmarkar 1987; Weng 1996 or Axsäter 2010) focus on an analytical solution for either one or two of the MRP parameters, while only few articles deal with the complexity of simultaneously optimising all three MRP parameters. However, the generalisation of the results derived from the analytical models to practical application is limited due to the streamlined settings in these models. Most of them have assumptions like single stage production systems, single material, or other production system parameters being assumed to be deterministic. This leads to the need of further ways to find a good parameter-setting for MRP

systems. Another way to find (near)-optimal MRP parameters in high complex stochastic production systems is simulation-based optimisation (e.g. Fu 2002). Molinder (1997), for example, has shown how to combine simulation and a metaheuristic to iteratively optimise the three kinds of MRP parameters. He proposed simultaneous optimisation of all three MRP parameters as a way to further decrease total costs, but did not show appropriate results. Gansterer et al. (2014) partially resolved this lack of exploration by the use of a combined simulation-optimisation approach using different kinds of metaheuristics. A disadvantage of simulation-based optimisation with metaheuristics is that the solution found might not be the global optimum of the defined problem.

# 3 Model Description

For the simulation study, a generic scalable simulation model (SimGen) designed for production planning problems was used. This simulation model is parameterised by a database, which makes it possible to define and run different simulation scenarios without any adaption of the simulation model itself. The simulation model was programmed in AnyLogic 6.9, which is a multi-method simulation software. More details of SimGen are outlined in Felberbauer et al. (2012) or Felberbauer and Altendorfer (2014).

For the simulation-based optimisation, SimGen is connected to HeuristicLab, which is a graphical user interface (GUI) based framework for heuristic and evolutionary algorithms (Elyasaf and Sipper 2014). A so-called *External Evaluation Problem* facilitates the connection between AnyLogic 6.9 and HeuristicLab. The chosen algorithm in HeuristicLab (GA and ES) creates a set of solutions (i.e. MRP planning parameters) and separately sends them to the simulation model. The simulation model then evaluates solutions qualities and forwards it to HeuristicLab. After the evaluation, the algorithm creates an improved set of solutions and the cycle starts again. HeuristicLab stops the optimisation after a predefined number of runs.

## 3.1 Search Procedure: Grid Search

For the grid search procedure, which is an enumeration of the predefined restricted solution space, the connection to HeuristicLab is not used, but a parameter variation experiment is defined within the simulation software. The grid search was carried out to find appropriate results as a base for the comparison to the metaheuristics. One decision variable per MRP-parameter, i.e. lot size, planned lead time and safety stock, is introduced. These decision variables adapt the initial value for each material, e.g. for lot size, by the same factor. Assume that the optimal factor for lot size is 0.5, then the optimal value for material 10, with initial lot size 110 pcs., is 55 pcs., whereas the optimal lot size for material 2, with initial lot size 330 pcs. is 165 pcs.. The decision variables are varied between a minimum and maximum with a fixed step size. The minimum lot size factor was determined in some manual simulation runs, where overall utilisation was checked to be lower than 100%. The minimum for the planned lead time and safety stock factor is on its natural minimum, which is zero. The maximum is determined as the fourfold of the initial value. The grid search is limited due to the fact that the ratio between optimal parameters is predetermined. Therefore, the optimal parameter of one material cannot change without modifying all other material parameter values. Nevertheless,

the latter explained limitation gives the possibility to enumerate in reasonable time through the whole solution space.

### 3.2 Search Procedure: Evolutionary Algorithms

For the evolutionary algorithms, all relevant MRP parameters are optimised for each single material, which lead to a number of 40 optimisation parameters. The parameterisation of the metaheuristics itself has been conducted in pre-tests.

Two algorithms, namely *evolution strategy* (ES) and *genetic algorithm* (GA) are tested for their performance. For both algorithms, the initial population was generated by a uniform random number generator. GA was run with a population size of 100, 1 elite, tournament selection with group size=2, blend-alpha-beta-crossover (alpha=0.75; beta=0.25) and a fixed-normal-all positions-manipulator with sigma=1. ES was parameterised with a population size of 20, 100 children, 2 parents per child, non-elitist population, self-adaptive-normal-all-positions-mutator and a uniform-all-positions-arithmetic-crossover with alpha=0.33.

Additionally, an adapted version of the ES starting with an initial population where all individuals are equal to the initial setting is tested. This *memetic evolution strategy* (mES), therefore uses problem specific knowledge, and applies ES as a local improvement search.

## 4 Production System

The modelled production system follows a flow shop structure which is inspired by a production company operating in the automotive sector. The flow shop consists of three machines (M1-M3) which correspond to turning, milling and grinding.

Figure 1 shows the bill of materials (BOM), the low level code (LLC), the work schedule and its associated machines. Each arrow represents one piece of material, which is part of another. The raw materials in LLC 3 are purchased parts which are assumed to be always available and no MRP parameters are optimised for these materials. All sales products and materials on LLC 1 and LLC 2 are MRP-planned.

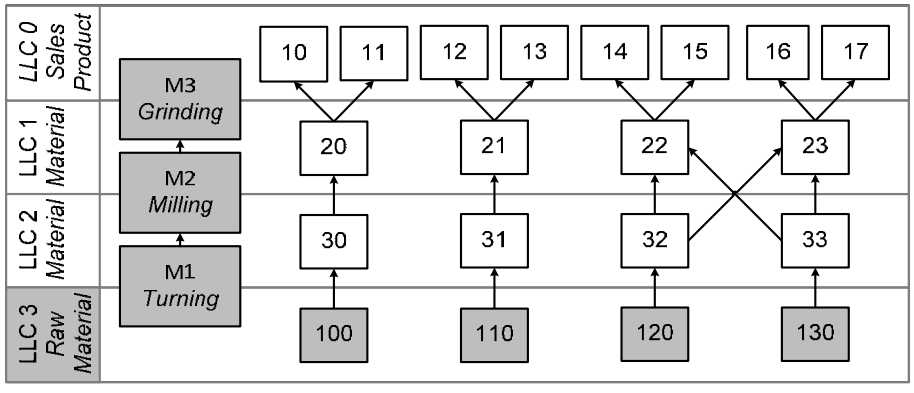



*Figure 1: Bill of Materials*

Items 10 to 17 are produced on machine M3 (grinding), items 20 to 23 are produced on machine M2 (milling) and items 30 to 33 are produced on machine M1 (turning). The mean setup time per order is for all machines about fifty times the mean processing time for a single item. Both parameters are randomly distributed. The specified values are equal for materials on the same LLC.

All even- and odd-numbered sales products have the same demand and mean amount of orders, which leads to the fact that at least two materials on all LLCs are similar (see Tab. 1). The demand per month for even-numbered sales products is 1,000 pcs. per month and the mean order amount is 10 pcs. (≈100 orders per month), for odd-numbered sales products this is 3,000 pcs. per month with a mean order amount of 15 pcs.. The demand per month is a fixed value with zero variance for all materials. The amount of orders is randomly distributed with a coefficient of variation (COV) of 0.25. Mean customer required lead time (CRLT) is 5 days for all orders and randomly distributed with a COV of 0.25.

Several random effects are implemented in the simulation model. Stochastic machine breakdowns and repair-times reduce the machine availability in this study to 95%.

## 4.1 Initial Lot Sizes

The initial values for lot sizes of all materials were calculated based on a planned utilisation factor of 80% for processing time and 85% with the consideration of machine breakdowns. 100% utilisation minus 85% planned utilisation results in a maximum of 15% time for setups. The number of possible setups per month is calculated by 15% (available time factor for setups) * 24 hours * 30 days / 1.5 hours (mean time per setup) = 72 setups. It has been assumed that every finished good needs the same number of setups per month resulting in 9 possible setups per month and material. Based on this information and the machine availability it was possible to calculate the initial lot size for all sales products, which is mean demand per month divided by 9. Table 1 outlines the appropriate initial lot sizes.

The initial lot sizes on lower stages of LLC were defined as the maximum lot size of the upstream materials.

## 4.2 Initial Planned Lead Times

Initial values for the planned lead times of the materials were set according to the necessary production lead time. The necessary production lead time was evaluated in a preliminary simulation study using 40 replications, no safety stocks and the initial values for the materials' lot sizes (see Section 4.1).

## 4.3 Initial Safety Stock

Initial safety stock was set to zero for all materials. For the following experiments safety stock is measured as a multiple of mean daily demand instead of pieces, which leads to improved comparability between materials. Furthermore, safety stock was only allowed for finished products, which leads to a reduced search space and is in line with the findings of Heisig (2002) and Whybark and Williams (1976).

### 4.4 Results for Initial Setting

The following Table 1 presents the initially determined production planning parameters for the considered system. The materials were grouped into 5 clusters by similar demand patterns and LLC. These clusters are used to compare the consistency of production planning parameters after the optimisation. A simulation run using the initial parameter setting of Table 1, leads to mean overall costs of 2,911.

*Table 1: Initial Parameter Setting*

| Cluster | Materials | Demand per Month | Lot Size | Planned Lead Time |
|---|---|---|---|---|
| 1 | {10,12,14,16} | 1,000 | 111 | 3 |
| 2 | {11,13,15,17} | 3,000 | 333 | 3 |
| 3 | {20,21,22,23} | 4,000 | 333 | 2 |
| 4 | {30,31} | 4,000 | 333 | 1 |
| 5 | {32,33} | 8,000 | 333 | 1 |

## 5 Numerical Study

The objective function value for all experiments is the mean overall cost, which is the sum of inventory and tardiness costs. To avoid situations with low costs due to excessively delaying customer orders, all customer orders that cannot be satisfied at the end of the simulation time are penalised with their tardiness. The ratio of the inventory cost rate to the tardiness cost rate is chosen to be 1:19 (see also Axsäter 2000 for details on this ratio). The holding costs are 1 for all sales products and 0.5 per day and piece for all other materials. In the experiments we simulated five whole years with the above described demand structure, whereby the first year is the warm up time of the simulation model and therefore excluded from the analysis. Each run is replicated 40 times to account for stochastic variance.

Each algorithm was run for 9,000 solution evaluations, which led to a total of 3.5 days computational time on a commercially available personal computer.

## 6 Results

In this section, on the one hand, the solution quality of the tested metaheuristics to optimise MRP parameters is evaluated. On the other hand, a measure for MRP parameter consistency is introduced to investigate their traceability and practical applicability.

*Solution quality - Convergence of the presented solution methods*

Figure 2 shows the solution value of the initial parameter setting and the performance of the four different solution procedures with respect to number of evaluated solutions. Please note that 2,500 evaluations are approximately 1 day parallel evaluation time. Therefore, the x-axis can easily be converted from evaluated solutions to computation time.

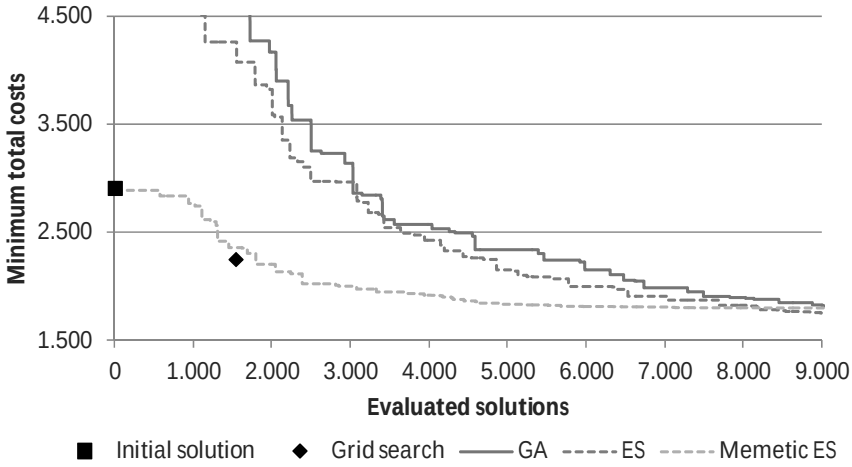



*Figure 2: Comparison of Solution Quality*

The solution value of the initial setting is 2,911. The best solution value derived from the grid search procedure is 2,252 using approximately 1,440 solution evaluations or 0.55 days computation time. The optimal parameter setting of the grid search procedure was 0.75*initial lot size, 0.50*initial planned lead time and 1.75*mean daily demand for safety stock.

*Solution quality - Cost improvement with metaheuristics*

The following Table 2 shows the best solution values of all solution procedures. All metaheuristic optimisation procedures found on average a 21% cost improvement in comparison to the grid search. However, the number of evaluations (computation time) needed is six times higher for the metaheuristics.

*Table 2: Comparison of Algorithm Qualities*

| | *Best Quality* | *better than* | | *% better than* | |
|---|---|---|---|---|---|
| | | Initial Setting | Grid Search | Initial Setting | Grid Search |
| Initial Setting | 2,911 | | | | |
| Grid Search | 2,252 | 659 | | 23% | |
| GA | 1,820 | 1,091 | 432 | 37% | 19% |
| ES | 1,739 | 1,171 | 512 | 40% | 23% |
| mES | 1,783 | 1,128 | 469 | 39% | 21% |

Note that none of the metaheuristics found a better solution than the grid search within the same amount of evaluations. We assume that reasons for the higher number of evaluations needed are that the applied metaheuristics do not use a predefined step-size and exploit the potential through single material parameter optimisation. ES delivered the best solution value with 1,739 mean overall costs,

which is a reduction of 40% compared to the initial setting and a reduction of 23% to the best grid search solution quality. A further interesting result concerning the metaheuristics is that the mES, which includes a problem specific start population, on the one hand provides a better solution value convergence, but leads on the other hand to slightly higher costs than the general ES.

*Optimal MRP parameters - Consistency and traceability*

Table 3 presents some details on the best MRP parameter values from the evaluated solution procedures. It shows the lot sizes for material 10, 12, 14 and 16 (Cluster 1 in Table 1) which have the same demand and production pattern and should therefore have similar planning parameters. From an analytical point of view, it is rather clear that the optimal solution for materials with equal parameters should be nearly the same for all of them. In the last row of Table 3, the coefficient of variance (COV) between the optimal parameter values for each solution procedure are calculated. The COV serves as measure for consistency between produced materials' parameter values. As production planning personnel will not apply MRP parameter settings which cannot be argued, this is a further argument why this COV value has to be low, to ensure the practical applicability of metaheuristic solution results.

*Table 3: Comparison of Optimal Lot Sizes*

| Material | Initial Solution | Grid Search | GA | ES | mES |
|---|---|---|---|---|---|
| 10 | 111 | 83 | 148 | 143 | 126 |
| 12 | | | 114 | 128 | 145 |
| 14 | | | 121 | 142 | 100 |
| 16 | | | 117 | 100 | 114 |
| *COV* | *0,0%* | *0,0%* | *12,4%* | *15,6%* | *15,7%* |
| *mean* | *111* | *83* | *125* | *128* | *121* |

From the grid search definition above, it is clear that this procedure leads to COV=0 and therefore a perfect consistent and traceable solution. All three metaheuristics tested in this study lead to a high COV, i.e. inconsistent MRP parameters for materials with equal demand and production patterns. An obvious reason for this observation is the optimisation of single materials' parameters instead of one parameter for all materials. This result points out a major shortcoming in the application of metaheuristic MRP parameter optimisation which is yet not discussed in literature but might be the key limitation for its practical applicability.

To improve solution consistency and traceability of results, a *mean parameter solution* is created within the material clusters for the related parameters (see last row *mean value* of Table 3 for the applied lot size of one cluster). This *mean parameter solution*, which is consistent and applicable, leads to a cost increase of about 10% for GA and ES. As this is still a cost advantage of 13% in comparison to the grid search algorithm, this might be a promising field of further research. Note

that for the best mES solution this *mean parameter solution* leads to a too high utilisation, and therefore a set of the best solutions has to be tested. For real size problems, the material clusters are not ex-ante to be identified which means that a consistent and traceable solution heuristic should probably have a pre-processing step to identify these clusters.

*Optimal MRP parameters - Lot size of components*

Table 4 presents the best parameter setting for lot sizes, planned lead times and safety stocks, of the ES algorithm. We observed *very small lot sizes* compared to their related finished goods for all components. A closer inspection of this observation shows that the large FOQ-lot sizes of finished goods override the small setting of subordinated materials. Based on this finding, a re-evaluation of the best parameter settings with absolute lot sizes for all components of 1 is performed showing a further cost reduction of 1% for GA and ES. However, the costs for the best mES solution doubled, which again shows that a set of the best solutions needs to be tested. This shows that the optimal lot size for components is 1 in a production system, where setup times of superior materials are greater or equal than the setup times of the subordinate materials. For such a setting we conjecture that the MRP calculation scheme, overruling a large part of the components lot sizing policy, reduces the solution space for MRP parameter optimisation.

*Table 4: Best Parameter Setting - ES*

| *Cluster* | *1* | | | | *2* | | | | *3* | | | | *4* | | *5* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Material | 10 | 12 | 14 | 16 | 11 | 13 | 15 | 17 | 20 | 21 | 22 | 23 | 30 | 31 | 32 | 33 |
| Lot Size | 143 | 128 | 142 | 100 | 233 | 241 | 176 | 194 | 13 | 22 | 1 | 7 | 8 | 2 | 1 | 1 |
| Planned Lead Time | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| Safety Stock | 61 | 46 | 53 | 41 | 178 | 221 | 147 | 228 | | | | | | | | |

## 7 Conclusion

In this study the performance and applicability of metaheuristics for simulation-based MRP parameter optimisation is investigated. To identify the applicability of the optimised MRP parameters a measure for consistency and traceability is developed. The results of the study show that a simple grid search leads to perfectly consistent results but to rather high costs in comparison to the metaheuristics. On the contrary, the metaheuristics lead to a good cost performance, even though we still expect a remarkable solution gap to the optimal solution, but rather inconsistent solutions. A first attempt to increase consistency of optimal MRP parameters from metaheuristics shows a moderate cost increase. In this attempt, the mean value of parameters for predefined material clusters with similar demand and production patterns is applied. The results of this study show that further research is needed to develop problem specific search heuristics, probably divided into clustering and optimisation, to improve the performance and applicability of simulation-based MRP parameter optimisation.

## References

Axsäter, S.: Inventory Control. Boston, Dordrecht, London: Kluwer Academic Publishers 2000.

Axsäter, S.: A capacity constrained production-inventory system with stochastic demand and production times. International Journal of Production Research 48 (2010) 20, pp. 6203–6209.

Elyasaf, A.; Sipper, M.: Software review: The HeuristicLab framework. Genetic Programming and Evolvable Machines 15 (2014) 2, pp. 215–218.

Felberbauer, T.; Altendorfer, K.: Comparing the performance of two different customer order behaviors within the hierarchical production planning (accepted). In: Proceedings of the Winter Simulation Conference, 2014, pp. 2227-2238.

Felberbauer, T.; Altendorfer, K.; Huebl, A.: Using a scalable simulation model to evaluate the performance of production system segmentation in a combined MRP and Kanban system. In: Proceedings of the Winter Simulation Conference, Berlin (Germany) 2012, pp. 3022-3033.

Fu, M.C.: Feature Article: Optimization for simulation: Theory vs. practice. INFORMS Journal on Computing 14 (2002) 3, pp. 192–215.

Gansterer, M.; Almeder, C.; Hartl, R.F.: Simulation-based optimization methods for setting production planning parameters. International Journal of Production Economics 151 (2014), pp. 206-213

Heisig, G.: Planning stability in material requirements planning systems. Berlin, Heidelberg: Springer 2002.

Hopp, W.J.; Spearman, M.L.: Factory Physics: Boston: Mc Graw Hill / Irwin 2008.

Jonsson, P.; Mattsson, S.: A longitudinal study of material planning applications in manufacturing companies. International Journal of Operations & Production Management 26 (2006) 9-10, pp. 971–995.

Karmarkar, U.S.: Lot sizes, lead times and in-pocess inventories. Management Science 33 (1987) 3, pp. 409–418.

Molinder, A.: Joint optimization of lot-sizes, safety stocks and safety lead times in an MRP system. International Journal of Production Research 35 (1997) 4, pp. 983–994.

Slack, N.; Chambers, S.; Johnston, R.; Chambers, S.: Operations management. New York: Prentice Hall/Financial Times 2004.

Weng, K.Z.: Manufacturing lead times, system utilization rates and lead-time-related demand. European Journal of Operational Research 89 (1996) 2, pp. 259–268.

Whybark, D.C.; Williams, J.G.: Material requirements planning under uncertainty. Decision Sciences 7 (1976) 4, pp. 595–606.

Yano, C.A.: Setting planned leadtimes in serial production systems with tardiness costs. Management Science 33 (1987) 1, pp. 95–106.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Influence of Customer Information Uncertainties on Production Order Variance Comparing two Different Kinds of Customer Order Behaviours

## *Einfluss von Kundeninformationsunsicherheit auf die Varianz der Fertigungsaufträge zweier unterschiedlicher Arten von Kundenbestellverhalten*

Thomas Felberbauer, Andreas Josef Peirleitner, Klaus Altendorfer, University of Applied Sciences Upper Austria, Steyr (Austria), thomas.felberbauer@fh-steyr.at, andreasjosef.peirleitner@fh-steyr.at, klaus.altendorfer@fh-steyr.at

Abstract: In this paper, two different order behaviours, differentiating how customers provide information about their required demand are discussed. On the one hand this is customer-required lead time order behaviour where customers order a stochastic order amount with a customer required lead time. On the other hand this is forecast evolution order behaviour where the customers provide a forecast quantity for a specific due date for a long horizon in advance and update their forecast quantities periodically. The performance is evaluated by measuring variance of customer demand, gross requirements and production orders and the information quality progress by varying order behaviour parameters as well as forecast error levels. Analysing this performance measures this study investigates how much master production schedule and material requirement planning mitigate customer information uncertainty.

## 1 Introduction

Two different kinds of how customers provide information about the required demand to their suppliers can be observed and are as well discussed in literature. Firstly, this is customer required lead time (CRL) order behaviour where customers state certain orders with specific CRL and stochastic order amounts. In this case, the manufacturing company generates an aggregated forecast for a long time in advance. The second customer behaviour is forecast evolution (FEV) order behaviour where customers provide a forecast quantity for a specific due date for a long horizon in advance and update their forecast quantities periodically. In both order behaviours, manufacturing companies are facing the problem of information dynamics and uncertainty, i.e. stochastic behaviour of arrivals, due dates and order amounts over time. In this study the stochastic nature of both customer order behaviours FEV and

CRL are compared by analysing the output of the intermediate planning level. Additionally, different levels of forecast errors, which is the deviation of the planned monthly demand to its realisation, are analysed. For the CRL order behaviour we investigate different expected values and variances of customer required lead time. The second source of randomness for CRL is the variance of the order amount. For the FEV order behaviour, the length of frozen zone, meaning the time window where a customer does not change its order anymore is investigated. Another interesting parameter studied for FEV is the change frequency of forecast and the variance of the order amount per change, indicating how volatile the forecast is. To evaluate the negative influence of these different kinds of uncertainty on the production system behaviour, this paper does not measure costs, but the input into the medium term planning, being the mean and standard deviation of gross requirements and its output, consisting of the mean and standard deviation of production order lot sizes. Another indicator for the negative influence of customer demand uncertainty is the variance of capacity needed. We monitor this indicator, to investigate how much the hierarchical planning structure already mitigates these uncertainty effects. As the interaction of the different planning levels cannot be treated analytically a simulation study is conducted for this investigation.

## 2 Literature Review

The Manufacturing resource planning concept (MRPII), is a structured approach which is widely used. MRPII consists of three planning levels: Long-term planning, intermediate planning and short term control. The long-term planning involves the functions forecasting and aggregate production planning (APP). The main functions of mid-term planning are master production scheduling (MPS), which compares the production plan with the actual customer orders, and material requirements planning (MRP), which identifies a list of production orders (see Orlicky 1975). The short-term planning conducts scheduling and dispatching tasks (see Panwalkar and Iskander 1977) for the released production orders and available resources.

Based on the decision hierarchy in manufacturing and the information uncertainties, there is still a research gap concerning the influence of information dynamics and uncertainties on the decisions taken on the upper hierarchical level. Ignoring this information uncertainty influence leads to a lack of coordination and integration also reported in surveys by Fleischmann and Meyr (2003), Kok and Fransoo (2003) and Missbauer and Uzsoy (2011).

Fildes and Kingsman (2010) develop a framework for examining the effect of demand uncertainty and forecast error on unit costs and customer service levels in supply chain planning, including MRP type manufacturing systems.

## 3 Problem Description

The hierarchical production planning approach is modelled in a simulation model. For a detailed description see Huebl et al. (2011) and for a further application see Felberbauer et al. (2012) or Felberbauer and Altendorfer (2014). The simulation model is running a rolling horizon planning on the long and intermediate range planning level. On the long range planning level, the functionality of APP is conducted. In this study we do not calculate an optimal production program but use the forecast

as production program. The intermediate-range planning uses the information of the production program. In the MPS, the production program is disaggregated and actual customer orders are used for the calculation of the gross requirements. The gross requirements are calculated taking the maximum of disaggregated production program and customer demand. MRP runs daily and planned order releases are calculated with the main MRP-functions netting, lot sizing, backward scheduling and bill of material (BOM) explosion (see also Hopp and Spearman (2008) for details on the MRP run). Planned orders from the MRP run are released to real production orders if all required sub materials are available. Production orders are produced on the shop floor according to the routing information and the dispatching rule.

## 3.1 Forecast Error

For both order behaviours the forecast error includes the difference between the monthly constant demand forecast $F_{p,t}$ and the stochastic realized monthly demand $D_{p,t}$ of sales product $p$ in time period $t$. The forecast error $\varepsilon_{p,t}$ is an identically independent truncated-normal-distributed random variable with an expected value $E[\varepsilon_{p,t}] = 0$ and variance $Var[\varepsilon_{p,t}] = (\alpha F_{p,t})^2$. The forecast error parameter $\alpha$ defines the quality of the forecast and is independent of time period $t$ and item $p$. The random monthly demand per sales product $p$ and time period $t$ is defined as $D_{p,t} = F_{p,t} + \varepsilon_{p,t}$.

## 3.2 CRL Order Behaviour

The order amount $O_p$ for item $p$ is log-normally distributed and $Var[O_p]$ is calculated based on the coefficient of variation $CV[O_p]$. The order arrival rate is $\lambda_{p,t} = \frac{D_{p,t}}{E[O_p]} = \frac{F_{p,t}+\varepsilon_{p,t}}{E[O_p]}$. Note that in the simulation study, the order rate $\lambda_{p,t}$ is adjusted to account for forecast error. Each customer order requests a stochastic customer required lead time $L$ based on the coefficient of variation $CV[L]$.

## 3.3 FEV Order Behaviour

For the evolution of forecasts in FEV behaviour we are using an extension of the Martingale Model for Forecast Evolution (MMFE), presented in Heath and Jackson (1994). This way of modelling forecasts is originally based on the model of Hausman (1969). In this FEV order behaviour, the customer provides a planned order amount per due date in advance of the real customer order for a certain forecast horizon. This amount stays unchanged until the forecast evolution horizon is reached. Between the forecast evolution horizon and the due date, the customer changes the order amount periodically. For each change, the identically independent truncated-normal-distributed evolution error $\omega_{p,t}$ is added to the actual order amount. To account for the forecast error $\varepsilon_{p,t}$, the evolution error is defined by $E[\omega_{p,t}] = \frac{\varepsilon_{p,t}}{n_{p,t}\lceil (h-fz)/cf \rceil}$ and $Var[\omega_{p,t}] = (\delta \bar{d}_p)^2$. Where $n_{p,t}$ is the number of orders per time period $t$ and sales product $p$, $h$ the forecast evolution horizon, $fz$ the frozen zone and $cf$ the change frequency. $(h - fz)/cf$ is the number of changes per customer order and therefore $n_{p,t}\lceil (h - fz)/cf \rceil$ is the total number of changes of all customer orders per time period $t$ and sales product $p$. The parameter $\delta$ defines

the variance of the evolution error $\omega_{p,t}$ and is independent of time period $t$ and sales product $p$. $\bar{d}_p$ is the constant demand forecast of sales product $p$ and due date $\tau$. Assuming that the error terms are independent of each other leads to the following Equation for the customer order of item $p$ in subperiod $\tau$. Note that Equation 1 defines the realized order amount.

$$O_{p,\tau} = \bar{d}_p + \sum_{k=1}^{\lceil (h-fz)/cf \rceil} \omega_{p,t,k} \tag{1}$$

## 3.4 Performance Indicators

In this study we analyse the effects of information uncertainty and forecast error of both customer order behaviours on customer demand, gross requirements and production orders. Therefore, the smoothing effects of the MPS and the MRP algorithm are investigated. Figure 1 shows the evaluated performance measures.

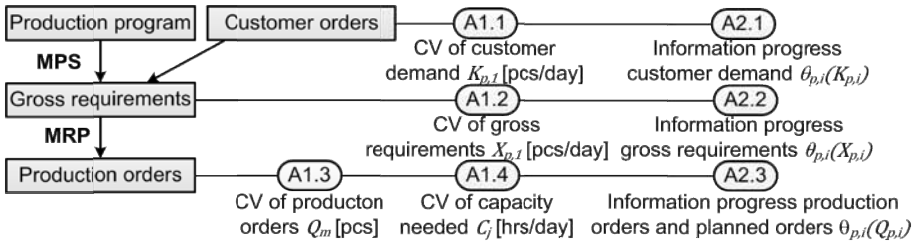



*Figure 1: Performance indicators*

Analyses A1.1 measures the coefficient of realized customer demand $CV[K_{p,1}]$ (note that index 1 again refers to the realized demand) for sales product $p$ which identifies the information uncertainty from the customer order process without any planning interaction. Analysis A1.2 measures coefficient of variation of realized gross requirements $CV[X_{p,1}]$ for sales product $p$ which is already smoothed by the MPS function. With A1.3, being the coefficient of variation of production order amounts $CV[Q_m]$ for material $m$, and A1.4, being the coefficient of variation of capacity requirements $E[C_j]$ per resource $j$, the influence of MRP is identified. A2.1, A2.2 and A2.3 show the evolution of the respective values. Equation (2) shows the calculation of the customer demand evolution quality measure. It evaluates the deviation between the value known $i$ periods in advance in comparison to the value realized ($i$=1) in percent. Note that the equations for $\theta_{p,i}(X_{p,i})$ (i.e. A2.2) and $\theta_{p,i}(Q_{p,i})$ (i.e. A2.3) are similar to $\theta_{p,i}(K_{p,i})$ and therefore omitted.

$$\theta_{p,i}(K_{p,i}) = \sqrt{E\left[(K_{p,1} - K_{p,i})^2\right]} \Big/ E[K_{p,1}] \tag{2}$$

To discuss the smoothing effect of the MPS and MRP on the gross requirements and production orders, respectively, in comparison to $\theta_{p,i}(K_{p,i})$, two normalized values are introduced in Equation (3).

$$P1_{p,i} = \theta_{p,i}\left(X_{p,i}\right)/\theta_{p,i}\left(K_{p,i}\right);\ P2_{p,i} = \theta_{p,i}\left(Q_{p,i}\right)/\theta_{p,i}\left(K_{p,i}\right) \tag{3}$$

# 4 Numerical Study Design

## 4.1 Production System

The modelled production system follows a flow shop structure and consists of 6 resources. The production system is inspired by different production facilities operating in the automotive sector. However, it is a streamlined version of such systems which clearly restricts the findings in this study. The arcs and numbers in Figure 2 represent the BOM structure with LLC being the low level code for the MRP run. The materials in LLC 3 are purchased parts which are always available and not further considered. Information about lot sizing policy, safety stock, planned lead time, and the monthly deterministic forecast are presented next to the sales products/materials $m$. In figure 2 the routing information is included showing at which resource $j$ the sales product or material is manufactured.

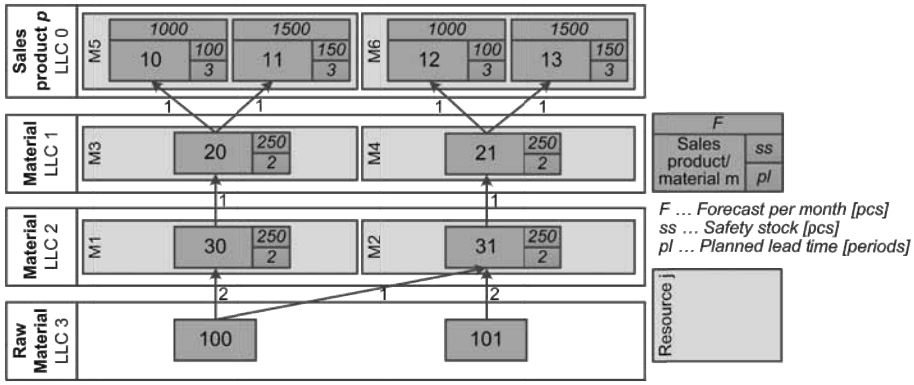



*Figure 2: Bill of material and routing information*

The applied lot sizing policy for all materials is lot-for-lot and no setup times are modelled. Each processing step requires a deterministic processing time of approximately 10 minutes and the internal capacity per resource is 480 hours per month. In the simulation study, five years are simulated with one year is the warm up time and 20 replications.

## 4.2 Parameter Settings

The forecast error parameter $\alpha$ is varied between 0 and 0.5 with a step size of 0.1. The parameters investigated for the CRL order behaviour are order amount: $E[O_p] = 30\ \forall\ p \in \{10,12\}$, $E[O_p] = 50\ \forall\ p \in \{11,13\}$, and $CV[O] \in \{0,0.15,0.5\}$; and customer required lead time: $E[L] \in \{3,6,12\}$ and $CV[L] \in \{0,0.5,1\}$. A basic scenario for detailed analysis is defined with $CV[O_p] = 0.15$, $E[L] = 6$ and $CV[L] = 0.5$. The parameters investigated for the FEV order behaviour are forecast

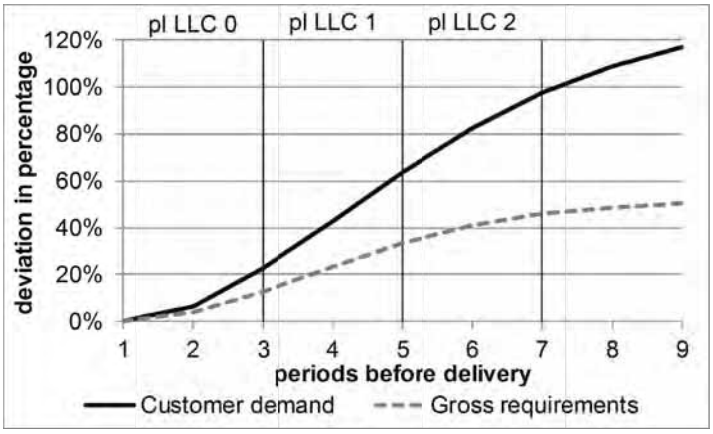



*Figure 3: Information progress CRL order behaviour with MPS functionality*

evolution horizon $h = 8$, frozen zone $fz \in \{0,2,4\}$, change frequency $cf \in \{1,2,4\}$, and variance of evolution error $\delta \in \{0.1,0.3,0.6\}$. Here the basic scenario is defined with $fz = 2$, $cf = 1$, and $\delta = 0.3$.

All results are shown for sales product 10, nevertheless, all other sales products and resources show a similar behaviour. The capacity requirements are shown for all resources (i.e. M5, M3 and M1 respectively) where sales product 10 is manufactured. Due to clarity the index of sales product 10 is not shown in the results.

# 5 Results

## 5.1 MPS Smoothing Effect for CRL Behaviour

The influence of the MPS functionality to smooth the demand uncertainty is studied here in detail for the CRL order behaviour without forecast error for the basic scenario. This order behaviour for the pure MTO system shows the following results:

$CV[K_1] = 0.91$, $CV[X_1] = 0.91$, $CV[C_{M5}] = 0.69$, $CV[C_{M3}] = 0.76$ and $CV[C_{M1}] = 0.74$.

Applying the MPS functionality this leads to:

$CV[K_1] = 0.91$, $CV[X_1] = 0.45$, $CV[C_{M5}] = 0.51$, $CV[C_{M3}] = 0.42$ and $CV[C_{M1}] = 0.38$.

Note that without MPS functionality $CV[X_1]$ is equal to $CV[K_1]$.

*A first observation from this setting is that the MPS functionality within the hierarchical planning structure mitigates the negative effect of demand uncertainty which is given by variance in order amount and CRL. The results show a lower CV of capacity requirements for the scenario with the MPS. The CV of capacity requirements decreases in the upstream manufacturing levels with MPS functionality. Such behaviour is not identified for the pure MTO system.*

Figure 3 shows the information progress of customer demand and gross requirements orders by applying the presented information evolution quality measures $\theta_{p,i}(K_{p,i})$ and $\theta_{p,i}(X_{p,i})$ (see 3.4) for the CRL order behaviour with MPS functionality.

*An observation from this setting is that the information quality of the customer demand continuously improves until the final delivery date. The information quality of the gross requirements is lower than the information quality of customer demand showing that the MPS functionality has a smoothing effect on the gross requirements.*

## 5.2 Effect of Frozen Zone for FEV Behaviour

The FEV order behaviour is discussed for the basic scenario with a frozen zone of two days and without a frozen zone. Without a frozen zone the following results can be observed:

$CV[K_1] = 0.70$, $CV[X_1] = 0.34$, $CV[C_{M5}] = 0.40$, $CV[C_{M3}] = 0.40$ and $CV[C_{M1}] = 0.38$.

With frozen zone the results are:

$CV[K_1] = 0.67$, $CV[X_1] = 0.33$, $CV[C_{M5}] = 0.38$, $CV[C_{M3}] = 0.37$ and $CV[C_{M1}] = 0.35$.

*An intuitive observation for the FEV order behaviour is that within the hierarchical production planning a frozen zone leads to a lower CV in customer demand and gross requirements as well as lower CV in capacity requirements.*

## 5.3 Effect of Forecast Error

Figure 4 shows the performance measures P1 (gross requirments, abbr. *gr*) and P2 (production orders, abbr. *po*) for different levels of the forecast error values $\alpha$ for both order behaviours. Note that P1 can be interpreted as gross requirements uncertainty in comparison to customer demand uncertainty and P2 can be interpreted as production order uncertainty in comparison to customer demand uncertainty.

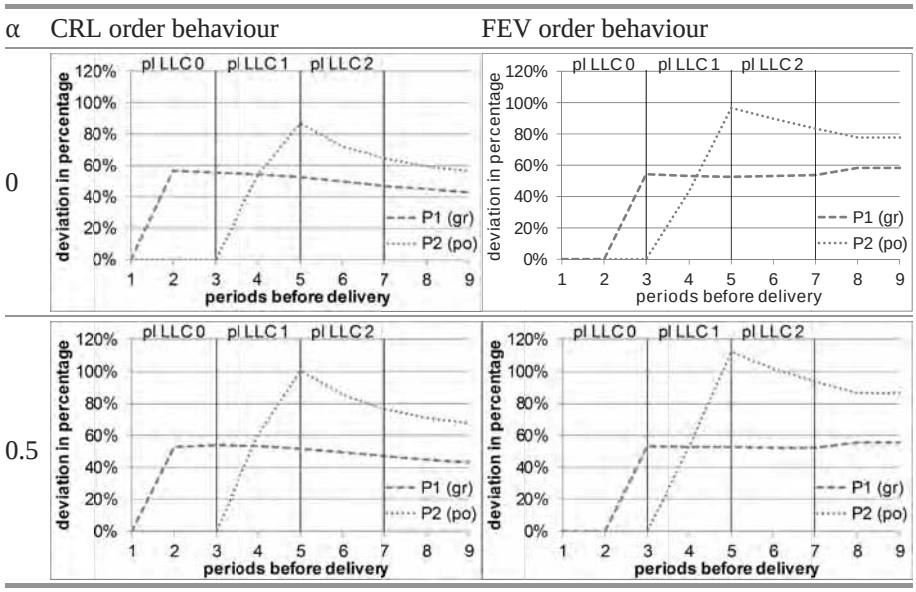



*Figure 4: Information progress with forecast error for both order behaviours*

*One unexpected observation from this setting is that the forecast error, which increases uncertainty of the customer order amounts and also leads to higher gross requirements variance (see fig. 5 and fig. 6), decreases the MPS improvement percentage only slightly.*

| | | a | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0 | | 0.1 | | 0.2 | | 0.3 | | 0.4 | | 0.5 | | Ø | |
| | | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po |
| $CV[O_p]$ | 0 | 0.33 | 0.43 | 0.33 | 0.45 | 0.34 | 0.50 | 0.35 | 0.55 | 0.36 | 0.60 | 0.37 | 0.68 | ***0.35*** | **0.54** |
| | 0.15 | 0.33 | 0.45 | 0.33 | 0.46 | 0.34 | 0.51 | 0.35 | 0.55 | 0.37 | 0.60 | 0.37 | 0.68 | ***0.35*** | **0.54** |
| | 0.5 | 0.38 | 0.54 | 0.38 | 0.55 | 0.39 | 0.58 | 0.40 | 0.62 | 0.41 | 0.66 | 0.41 | 0.73 | ***0.40*** | **0.61** |
| | Ø | **0.35** | **0.47** | **0.35** | **0.49** | **0.36** | **0.53** | **0.37** | **0.57** | **0.38** | **0.62** | **0.38** | **0.70** | 0.36 | 0.56 |
| | | a | | | | | | | | | | | | | |
| | | 0 | | 0.1 | | 0.2 | | 0.3 | | 0.4 | | 0.5 | | Ø | |
| | | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po |
| $E[L]$ | 3 | 0.43 | 0.44 | 0.43 | 0.45 | 0.45 | 0.50 | 0.47 | 0.56 | 0.48 | 0.62 | 0.49 | 0.70 | ***0.46*** | **0.55** |
| | 6 | 0.41 | 0.52 | 0.42 | 0.53 | 0.43 | 0.58 | 0.44 | 0.62 | 0.46 | 0.67 | 0.46 | 0.75 | ***0.44*** | **0.61** |
| | 12 | 0.19 | 0.46 | 0.19 | 0.47 | 0.19 | 0.50 | 0.19 | 0.54 | 0.20 | 0.57 | 0.20 | 0.64 | ***0.19*** | **0.53** |
| | Ø | **0.35** | **0.47** | **0.35** | **0.49** | **0.36** | **0.53** | **0.37** | **0.57** | **0.38** | **0.62** | **0.38** | **0.70** | 0.36 | 0.56 |
| | | a | | | | | | | | | | | | | |
| | | 0 | | 0.1 | | 0.2 | | 0.3 | | 0.4 | | 0.5 | | Ø | |
| | | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po |
| $CV[L]$ | 0 | 0.19 | 0.26 | 0.20 | 0.30 | 0.22 | 0.37 | 0.23 | 0.45 | 0.25 | 0.51 | 0.25 | 0.60 | ***0.22*** | **0.41** |
| | 0.5 | 0.40 | 0.58 | 0.40 | 0.59 | 0.41 | 0.61 | 0.41 | 0.64 | 0.42 | 0.68 | 0.43 | 0.75 | ***0.41*** | **0.64** |
| | 1 | 0.45 | 0.57 | 0.45 | 0.58 | 0.45 | 0.60 | 0.46 | 0.63 | 0.47 | 0.67 | 0.47 | 0.74 | ***0.46*** | **0.63** |
| | Ø | **0.35** | **0.47** | **0.35** | **0.49** | **0.36** | **0.53** | **0.37** | **0.57** | **0.38** | **0.62** | **0.38** | **0.70** | 0.36 | 0.56 |

*Figure 5: Parameter variation CRL order behaviour*

| | | a | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0 | | 0.1 | | 0.2 | | 0.3 | | 0.4 | | 0.5 | | Ø | |
| | | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po |
| $fz$ | 0 | 0.25 | 0.39 | 0.25 | 0.40 | 0.26 | 0.43 | 0.27 | 0.47 | 0.28 | 0.51 | 0.29 | 0.54 | ***0.27*** | **0.46** |
| | 2 | 0.23 | 0.37 | 0.23 | 0.38 | 0.24 | 0.42 | 0.25 | 0.46 | 0.27 | 0.50 | 0.27 | 0.53 | ***0.25*** | **0.44** |
| | 4 | 0.14 | 0.31 | 0.14 | 0.33 | 0.15 | 0.36 | 0.15 | 0.40 | 0.16 | 0.44 | 0.17 | 0.47 | ***0.15*** | **0.38** |
| | Ø | **0.20** | **0.36** | **0.21** | **0.37** | **0.22** | **0.40** | **0.23** | **0.44** | **0.24** | **0.48** | **0.24** | **0.51** | 0.22 | 0.43 |
| | | a | | | | | | | | | | | | | |
| | | 0 | | 0.1 | | 0.2 | | 0.3 | | 0.4 | | 0.5 | | Ø | |
| | | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po |
| $cf$ | 1 | 0.30 | 0.46 | 0.30 | 0.48 | 0.31 | 0.50 | 0.32 | 0.54 | 0.34 | 0.57 | 0.35 | 0.61 | ***0.32*** | **0.53** |
| | 2 | 0.21 | 0.36 | 0.21 | 0.37 | 0.22 | 0.40 | 0.23 | 0.44 | 0.25 | 0.48 | 0.26 | 0.51 | ***0.23*** | **0.43** |
| | 4 | 0.11 | 0.25 | 0.11 | 0.27 | 0.11 | 0.30 | 0.12 | 0.34 | 0.13 | 0.38 | 0.13 | 0.42 | ***0.12*** | **0.33** |
| | Ø | **0.20** | **0.36** | **0.21** | **0.37** | **0.22** | **0.40** | **0.23** | **0.44** | **0.24** | **0.48** | **0.24** | **0.51** | 0.22 | 0.43 |
| | | a | | | | | | | | | | | | | |
| | | 0 | | 0.1 | | 0.2 | | 0.3 | | 0.4 | | 0.5 | | Ø | |
| | | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po | gr | po |
| $\delta$ | 0.1 | 0.09 | 0.14 | 0.09 | 0.17 | 0.11 | 0.22 | 0.14 | 0.29 | 0.16 | 0.36 | 0.17 | 0.41 | ***0.13*** | **0.26** |
| | 0.3 | 0.22 | 0.38 | 0.22 | 0.39 | 0.23 | 0.42 | 0.23 | 0.46 | 0.24 | 0.50 | 0.25 | 0.53 | ***0.23*** | **0.45** |
| | 0.6 | 0.31 | 0.55 | 0.31 | 0.56 | 0.31 | 0.56 | 0.31 | 0.57 | 0.31 | 0.58 | 0.31 | 0.60 | ***0.31*** | **0.57** |
| | Ø | **0.20** | **0.36** | **0.21** | **0.37** | **0.22** | **0.40** | **0.23** | **0.44** | **0.24** | **0.48** | **0.24** | **0.51** | 0.22 | 0.43 |

*Figure 6: Parameter variation FEV order behaviour*

*Another counterintuitive effect is that the production order amount has over a long range a higher uncertainty than the gross requirements (only below 4 periods before delivery its uncertainty is lower) and the forecast error significantly decreases this performance.*

### 5.4 CRL and FEV parameter variation

The influence of different parameters on the information evolution quality measure for gross requirements $\theta_{p,i}(X_{p,i})$ (abbr. *gr*) and production orders $\theta_{p,i}(Q_{p,i})$ (abbr. *po*) is studied for sales product 10, 7 periods (sum of planned lead times) before the delivery date for the CRL, in figure 5, and FEV, in figure 6, order behaviours. Different levels of forecast error parameter $\alpha$ are discussed. Each figure consists of three sub tables for the respective parameters of CRL and FEV. The study shows for both order behaviours that with increasing forecast error parameter $\alpha$ the information evolution quality measure $gr$ and $po$ increase for all parameter combinations, which means more information progress deviation.

*An observation valid for both order behaviours is that the forecast error value $\alpha$ has a negative correlation with the information quality.*

*An observation from the CRL parameter variation in Figure* 4 *is that the longer the customer required lead time and the lower its variance the better is the information quality of gross requirements.*

The positive effect of a frozen zone on information quality is discussed in 5.2.

*An observation from the FEV parameter variation is that a higher number of changes per customer order and big changes are leading to a decrease of information quality.*

## 6 Conclusions

In this paper the effect of two different order behaviours, which differ how customers provide information about the demand, namely CRL and FEV is investigated. To identify the influence of information uncertainty within the hierarchical production planning a measure for information quality and information evolution is developed. The results of the simulation study show that for the CRL order behaviour the MPS mitigates the information uncertainty. For the FEV order behaviour the information quality improves with the use of a frozen zone. If the production system is facing forecast errors the system has to deal with higher information uncertainties. In further research we would like to discuss what information quality is needed to justify investments in information sharing between supply chain members.

## References

Felberbauer, T.; Altendorfer, K.: Comparing the performance of two different customer order behaviors within the hierarchical production planning. In: Proceedings of the Winter Simulation Conference, Savannah 2014, p. 2227-2238.

Felberbauer, T.; Altendorfer, K.; Huebl, A.: Using a scalable simulation model to evaluate the performance of production system segmentation in a combined MRP and Kanban system. In: Proceedings of the Winter Simulation Conference, Berlin (Germany) 2012, pp. 3022-3033.

Fildes, R.; Kingsman, B.: Incorporating demand uncertainty and forecast error in supply chain planning models. Journal of the Operational Research Society 62 (2010) 3, pp. 483–500.

Fleischmann, B.; Meyr, H.: Planning hierarchy, modeling and advanced planning systems. In: Graves, S.C.; de Kok, A.G. (eds.): Supply chain management: design, coordination and operation: Amsterdam: Elsevier 2003, pp. 455–523.

Hausman, W.H.: Sequential decision problems: A model to exploit existing forecasters. Management Science 16 (1969) 2, pp. B-93-111.

Heath, D.C.; Jackson, P.: Modeling the evolution of demand forecasts with application of safety stock analysis in production/distribution systems. IIE Transactions 26 (1994) 3, pp. 17–30.

Hopp, W.J.; Spearman, M.L.: Factory physics: Boston: Mc Graw Hill / Irwin 2008.

Huebl, A.; Altendorfer, K.; Jodlbauer, H.; Gansterer, M.; Hartl, R.: Flexible model for analyzing production systems with discrete event simulation. In: Winter Simulation Conference, Phoenix (AZ) 2011, pp. 1554–1565.

Kok, T.G. de; Fransoo, J.C.: Planning supply chain operations: Definition and comparison of planning concepts. In: Graves, S.C.; de Kok, A.G. (eds.): Supply chain management: Design, coordination and operation: Amsterdam: Elsevier 2003, pp. 597–675.

Missbauer, H.; Uzsoy, R.: Optimization models of production planning problems. In: Kempf, K.G.; Keskinocak, P.; Uzsoy, R. (eds.): Planning production and inventories in the extended enterprise. New York: Springer 2011, pp. 437–507.

Orlicky, J.: Material requirements planning;: The new way of life in production and inventory management. New York: McGraw-Hill 1975.

Panwalkar, S.S.; Iskander, W.: A survey of scheduling rules. Operations Research 25 (1977) 1, pp. 45–61.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Discrete Event Simulation of Modular Production System Models using Petri Nets

## *Ereignisdiskrete Simulation modularer Produktionssystemmodelle mittels Petri-Netzen*

Markus Rabe, Maik Deininger, TU Dortmund, Dortmund (Germany),
markus.rabe@tu-dortmund.de, maik.deininger@tu-dortmund.de

Abstract: Medium-term production planning is a complex task for manufacturers. The question whether accepting or declining a new order often cannot be answered well-grounded. Additionally, it cannot be predicted, which consequences an acceptance will have on the running production. This paper aims to provide a basis for medium-term production planning using Petri nets and discrete event simulation. Petri nets are used for creating Modular Production System Models (MPSM). MPSM are utilised for representing any process that exists within the considered production system as also processes that may be implemented to adapt to new requests. As consequence, the existing production system is modelled by interconnecting MPSMs and is varied by adding or removing modules. Each model is analysed using discrete event simulation based on Petri nets. Thus it can be verified whether all orders can be fulfilled or not.

## 1 Introduction

Production planning is a challenging task, especially for manufactures with custom-made products. The effects of accepting or declining an order without knowing how the production system is affected often cannot be forecasted adequately. In order to support planners with this decision, a method needs to be developed that is able to change and evaluate production systems automatically and provide the planner with the information needed. The changes made have to be target-oriented to avoid evaluating systems that do not reduce the target parameters as make span or costs. Based on such a method, a tool can be implemented to support the decision making within production planning.

Simulation is a wide-spread technique for evaluating existing or planned production systems as well as for analysing the impact of alterations to such systems (cp. Lefrançois et al. 1996; Kara et al. 2007; Frank et al. 2013). Since state changes in these systems in general occur at discrete points in time, Discrete Event Simulation (DES) is used in most cases. Different tools exist on the market that enable planners

to model and simulate production systems. The simulation models are usually created by connecting various building blocks manually. Some of the tools support automatic changes, but these changes are not operated through a superior mechanism as for example a heuristic optimisation that evaluates and selects the changes made. Such an algorithm needs a basis to work with. Therefore, a modular approach is investigated where a production system is interpreted as a compilation of modules. A first attempt was made by Rabe and Deininger (2013). This paper builds up on this work and uses Petri nets (PNs) for creating Modular Production System Models (MPSMs), where each module handles a single order. By connecting single modules, the path a customer order takes through the production system can be represented. Furthermore, by adding, exchanging, or removing modules, the system can be changed easily.

This paper is structured as follows: the next section gives a brief introduction to modelling and discrete event simulation. Afterwards the PN type used in this paper is introduced, before creating and simulating MPSMs in section 4 with PNs. The paper is closed with a conclusion that summarises the results and gives a short outlook.

## 2 Modelling and Simulation of Production Systems

In production and logistics simulation is one of the most important techniques. It is used for analysing and optimising existing as well as planned production systems. Due to the nature of these systems, where changes of interest only occur at discrete points in time, the most applied simulation method is DES as can be seen in many publications (cp. Lefrançois et al. 1996; März and Krug 2011). Tools like Plant Simulation (cp. Bangsow 2010) provide building blocks that are used to create complex simulation models. In general, the material flow is modelled and each building block represents an individual segment. These segments are means of production, means of transport, or resources. Additionally, if the provided building block does not meet the requirements, it is possible to modify existing or create new elements using a domain-specific language as for example SimTalk in Plant Simulation. Once the model is created, a DES can be used for analyses. In this paper, Petri Nets (PNs) are utilised to create modules that can be connected to model larger systems. These modules represent the flow of customer orders and internal orders instead of the material flow.

PNs are used to model various systems (cp. Viswanadham and Raghavan 2000; Dahms and Schmidt 2005; Koch 2014). A PN is a directed graph $PN = (T, P, E)$ where $T$ is a disjoint finite set of transitions, $P$ is a disjoint finite set of places, and $E \subseteq (P \times T) \cup (T \times P)$ is a finite set of directed arcs where each connects a transition with a place (Priese and Wimmel 2008; Reisig 2010). The concept of object orientation was introduced to PNs based on the concepts used in programming (cp. Desel, J. et al. 2004). There are different approaches for using object orientation in PNs (cp. Ceska et al. 1998; Dahms and Schmidt 2005; Koci et al. 2008). The most common is to treat tokens as instances of classes having member variables and functions. Using this approach, tokens contain the data that are manipulated through transitions they passed. Another approach is introduced with Object Oriented Petri Nets (OOPN), where PNs are organized in classes consisting of an object net and a set of method nets (Koci et al. 2008). OOPNs support

inheritance. Subclasses may redefine or add transitions and places as well as new methods. A token can either be a trivial object, e.g. a number, or an instance of an OOPN. A more complex type with higher flexibility is defined by Timed Hierarchical Object-Related Nets (THORNs) (Schöf et al. 1997) that are introduced within section 3 in detail.

By using THORNs the modeller is able to adapt predefined modules at two abstraction levels. It is possible to change the behaviour of a module by adding or removing places and transitions. Additionally, since THORNs use C++ for defining the behaviour of each element, individual elements can be implemented and used.

The simulation of PN models can be conducted using DES (cp. Zhou and Jeng 1998). Basically, there are two events: the transition fire event that consumes and produces tokens at places and the token move event that activates or deactivates transitions. How this concept is applied to THORNs is shown in section 4.

## 3 A Special Petri Net Type: Timed Hierarchical Object-Related Net

The PN type used in this paper is a THORN (cp. Wieting 1996; Schöf et al. 1997). THORNs extend PN by adding special types of places, transitions, and arcs, whereas places also offer a structure, as depicted in Figure 1. Additionally, object orientation is introduced not only to tokens but also for places, transitions, and arcs. Thus, a THORN can be used within many fields of application. THORNs implement two time concepts: duration of firing and a variation of window of invalidity (Starke 1995). Both times are determined using functions and thus can be computed dynamically.

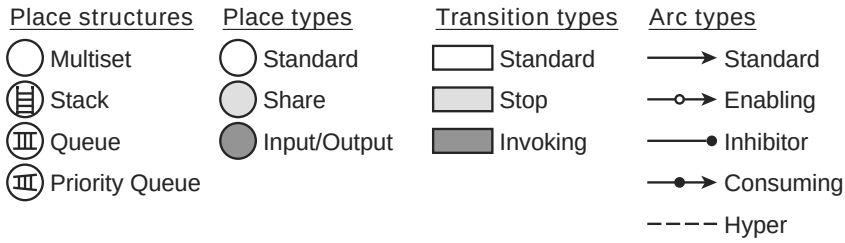



*Figure 1: THORN elements*

A transition contains multiple mandatory functions. Most important for the behaveiour of the THORN is the condition function. In order to activate the transition, the pre and post area conditions, as in standard PNs, need to be fulfilled and the condition function needs to be evaluated as true. This function enables the modeller to create transitions that are only activated if the tokens in the pre area meet specific conditions. For example, a token representing a worker can be tested for its qualification and working time spend within the current shift.

The three transition types indicate which kind of action is applied. Standard transitions manipulate or create new tokens, invoking transitions create new subnets that

exist and run parallel to the main THORN, and stop transitions destroy the subnet they are part of. The main THORN does not contain stop transitions.

Each transition always matches one of four states (inactive, delayed, completely delayed or active) and is only allowed to fire while being active. A transition is inactive if the pre area condition is not fulfilled. It is delayed if the pre area condition is fulfilled and completely delayed if it was delayed for a specified time (window of invalidity). If a transition is completely delayed and its post area condition is fulfilled, it switches to the state active where it is allowed to fire.

Next to the condition function of a transition, the place structure has a huge impact on the behaviour of the net. While multiset places allow taking any suitable token, stack, queue, and priority queue places only allow using the first token in line. A stack place implements the last in first out concept and may be used for modelling loading and unloading of means of transport. Queue and priority queue places apply first in first out, whereas priority queue places additionally allow sorting the tokens through a priority function. This can be used for defining in which sequence orders shall be processed.

The type of a place indicates where the tokens are used. Tokens from share places are equally used in multiple, parallel existing THORNs. Input and output places are used to transfer tokens between THORNs when creating or destroying a subnet. Tokens on standard places are only used within a single THORN.

Arcs influence the fulfilment of the pre and post area conditions. Five types are available. A standard arc can be used as input arc (place to transition) and as output arc (transition to place). It also can be weighted and thus determine how many tokens have to be moved. The enabling arc is always an input arc and can have a weight. In contrast to the standard arc, it only is tested whether the right amount of tokens exist on the place or not. They are not moved with the firing of the transition. Inhibitor and consuming arcs have no weight attached and only can be used as input arcs. The inhibitor arc inhibits the activation of a transition in case the connected place contains any token. The consuming arc activates a transition if the place contains at least one token. All tokens are removed from the place, the moment the connected transition fires. A hyper arc is used to connect share places with invoking transitions and thus indicating that the place is used within the subnets being created. Hyper arcs are undirected and in general are interpreted as a combined input and output arc. Within the subnet the share place is connected to transitions using arcs of the other four types.

## 4 Simulating Petri Nets without Conditional Events

The simulation of THORNs is based on a DES where the events are connected to places and transitions. The events are:

- a token arrives at a place (EAdd),
- a token is removed from a place (ERemove),
- a transition changes its state (EStateChange),
- a transition is completely delayed (EDelay), and
- a transition fires (EFire)

where EAdd, ERemove, EDelay, and EFire are booked events and EStateChange is a conditional event (for booked and conditional events see e.g., Robinson 2004). For example, a transition fires at $t = t_0$. Thus, the following events are created: at time $t_0$ ERemove for each place within the pre area and at time $t = t_0 + t_{Fire}$ EAdd for each place within the post area. Since only places connected to the transition are affected, it is not necessary to validate each EStateChange but only those that are connected to the places affected. If a transition's state changes to delayed, then an EDelay event is created with execution time $t = t_0 + t_{Delay}$. As a consequence, EStateChange can be integrated with EAdd and ERemove and is not needed as a separate event. EAdd now additionally validates the states of transitions in the post area and ERemove validates the states of transitions in the pre area of the place.

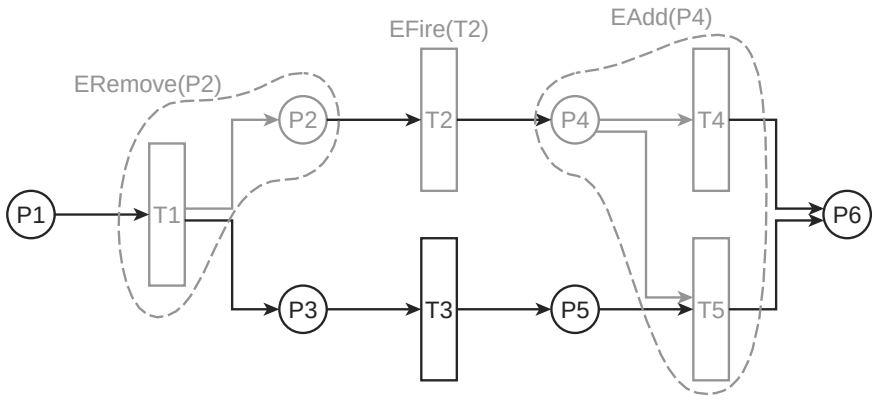



*Figure 2: Transitions tested within ERemove and EAdd event*

Figure 2 illustrates which transitions and places are handled by the events ERemove and EAdd when transition T2 fires. The remove event of place P2 includes a status update of transition T1 and the arrive event of place P4 includes the status updates of transitions T4 and T5. Thus, THORNs can be simulated by using four types of events. Through creating the initial mark-up using EAdd events, instead of just "placing" tokens, and ERemove events, for places connected by inhibitor arcs to transitions, at $t = t_0$, the simulation can be started without using conditional events.

*Table 1: Transitions to be checked if changes apply to its pre area places*

| | Token added to place | | | | Token removed from place | | | |
|---|---|---|---|---|---|---|---|---|
| | Multi-set | Stack | Queue | Priority queue | Multi-set | Stack | Queue | Priority queue |
| Inactive | + | + | + | + | – | + | + | + |
| Delayed | – | + | – | + | + | + | + | + |
| Completely delayed | – | + | – | + | + | + | + | + |
| Active | – | + | – | + | + | + | + | + |

Furthermore, based on the type of a place and the state of the transition, it can be decided, if the transition is affected and may be excluded from state validation. Table 1 and Table 2 show whether a transition needs to be checked (+) or not (–). Considering that in THORNs each transition has its own additional condition function, which can be very complex when checking and comparing specific token attributes, avoiding this test can improve the simulation speed, significantly.

For example, adding a token to a stack place of the pre area of a transition, the transition has always to be evaluated. Since the new token is available for use instead of the old one, the conditional function may lead to a different result. Adding a token to a multiset place in the pre area, inactive transitions have to be evaluated only. All tokens are still available for use and e.g. delayed transitions cannot become inactive.

*Table 2: Transitions to be checked if changes apply to its post area places*

| | Token added to place | | | | Token removed from place | | | |
|---|---|---|---|---|---|---|---|---|
| | Multi-set | Stack | Queue | Priority queue | Multi-set | Stack | Queue | Priority queue |
| Inactive | – | – | – | – | – | – | – | – |
| Delayed | – | – | – | – | – | – | – | – |
| Completely delayed | – | – | – | – | + | + | + | + |
| Active | + | + | + | + | – | – | – | – |

As Table 2 indicates, only few changes within the post area demand an evaluation of the transition's state. This is due to the fact that only the capacity and number of tokens at the place influence the transitions state. Adding a token may lead insufficient free capacity. Thus, only active transitions may change back to the state of being completely delayed. As the consequence, a completely delayed transition may be activated by removing a token. The states inactive and delayed do not take the post area into account and, therefore, are not affected by changes.

## 5 Modular Production System Models with Petri Nets

For modelling a production system, the system is considered as an aggregation of individual modules. Each module performs a single task and competes with other modules for resources such as workers, means of production, and materials. Figure 3a depicts the simplest form of a module. It takes an order (O1) and uses workers (W), means of production (P), and material (M). As output, a new order (O2) is created. Additionally, used resources are put back to their places and produced resources are placed at the appropriate places. For a more detailed specification of the performed actions, an invoking transition can be used that allows for creating a new subnet and now uses the places W, P, and M as share places instead of standard places. The subnet has its own places and transitions that specify how the order is handled. The changed module is shown in Figure 3b. Also a refinement of the transition could be used, but the invoking transition provides some advantages:

- multiple instances of the subnet are allowed (capacity > 1),
- with a capacity of one, the transition cannot be activated before the order is completely handled, and
- the subnet can be changed without touching the main net.

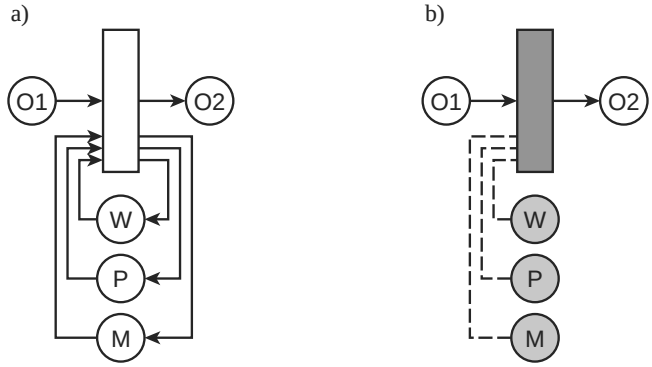



*Figure 3: Simple module (a) and simple module using an invoking transition (b)*

The modeller has to create a model, called MPSM, for each order that can be handled by the production system. These models are combined to a model of the complete production system by connecting places that store the same type of tokens.

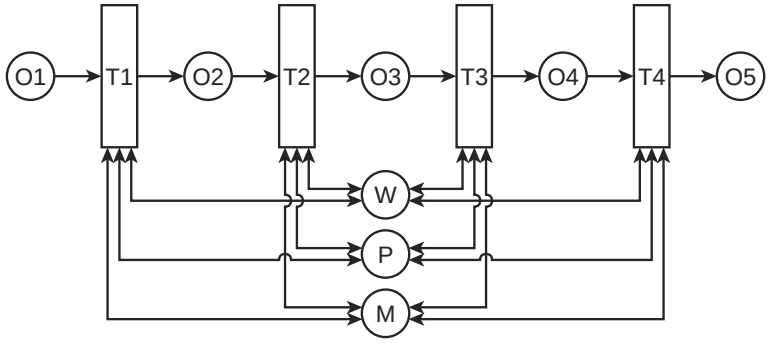



*Figure 4: Combination of modules*

For illustration, a production line that consists of four machines where each represents a single step of the production is used. Here, each machine can be modelled as a module as depicted in Figure 3a. The first module receives an order for processing step one and creates an order that activates module two which processes it. Each of the modules uses the same places for workers, means of production, and material and, therefore, they are competing for the resources. Figure 4 depicts this production line, where the double-headed arrow indicates the combination of a standard input and output arc. Invoking transitions can be used for refine-

ment and thus to improve the accuracy and to create different views: e.g. one for the manager and another one for the worker.

If analyses show that customer orders cannot be handled properly, the modular approach allows for easily adding or exchanging modules. Assuming the bottleneck is in handling the orders at place O2, a new transition T5 can be added, as shown in Figure 5. Here, transition T5 represents a production step that uses a new machine that is capable of performing the same steps as the two machines used in transitions T2 and T3. The places for resources (W, P, and M) are omitted for clarity.

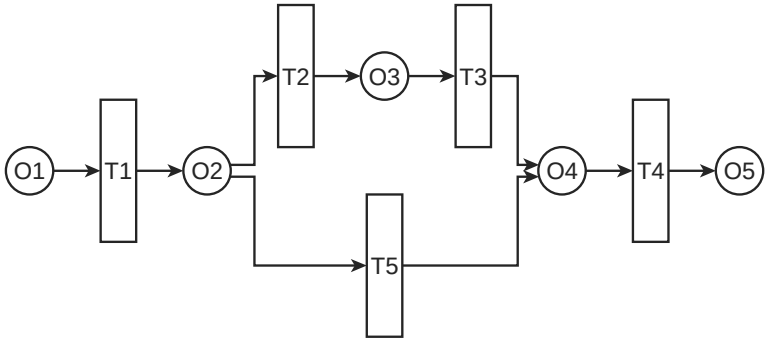



*Figure 5: Modified production line*

In reality, the acquisition of a machine is not only adding it to the production system. Among other things, the new machine, as well as the workers using it, needs to be installed and, if necessary, qualified. Therefore, the token representing the machine is not available for production since it has not the required qualifications. This has to be modelled by an individual module which exists in parallel to the model of the production system. This qualification module will use resources that initially had been scheduled for the production and are now no longer available. This conflict has to be solved. One opportunity is creating an additional production order for materials produced by machine 1 (represented by transition T1). As consequence, the work load for machine 1 is enlarged and the production has to be started earlier. Another option is adding a new module that is capable of performing the same steps as machine 1.

For evaluating different scenarios, a heuristic optimisation shall be developed that varies the model and analyses each version through simulation automatically. If shortages occur, those will be solved by adding MPSMs. Therefore, the simulation will be stopped. By predefined runtimes for each of the modules, it will be gone back in simulation, using backups that were made during the simulation run. Thus, a suitable point in time where the module has to be added can be selected. Now, the simulation can be restarted and the new module will support the production and the conflict will not arise.

## 6 Conclusion

This paper introduces an approach for applying discrete event simulation of THORNs without using conditional events. Therefore, the events that depend on conditions, as state changes of transitions, are integrated with booked events. The events of adding or removing tokens trigger the evaluation of a transition's state directly. Thus, it can be avoided to test each transition in every single time step and improve the performance of the simulation.

Furthermore, this simulation approach is used to evaluate models created with MPSMs. MPSMs are small models described through THORNs representing the handling of orders. By connecting MPSMs, the flow of orders within a production system can be modelled and afterwards analysed using DES. Single modules can easily be added, removed, or exchanged, which allows for evaluating multiple alternative systems.

Further research has to be conducted for developing a heuristic optimisation that enables an automated variation and evaluation of production systems. For this purpose, the MPSMs have to be stored in an appropriate manner: MPSM based models have to be adaptable within the simulation run and the simulation needs to be able to be stopped and restarted at suitable points in time. If these problems can be solved, a production planner will be able to decide if and how all customer orders can be handled.

## References

Bangsow, S.: Manufacturing simulation with Plant Simulation and Simtalk. Berlin, London: Springer 2010.

Ceska, M.; Janousek, V.; Vojnar, T.: Object-oriented Petri nets, their simulation, and analysis. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Piscataway, NJ, USA: IEEE 1998, pp. 262–267.

Dahms, M.; Schmidt, M.: Modeling of dispatching-rules for Job Shop Scheduling in manufacturing systems – A Petri net approach. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics. Waikoloa, Hawaii, USA, October 10-12 2005. Piscataway, NJ, USA: IEEE 2005, pp. 2025–2030.

Desel, J.; Reisig, W.; Rozenberg, G.: Lectures on concurrency and Petri nets: Advances in Petri nets. Berlin, New York (NY): Springer 2004.

Frank, M.; Laroque, C.; Uhlig, T.: Reducing computation time in simulations-based optimization of manufacturing systems. In: Pasupathy, R.; Kim, S.-H.; Tolk, A.; Hill, R.; Kuhl, M.E. (Eds.): Proceedings of the 2013 Winter Simulation Conference. Piscataway: IEEE 2013, pp. 2710–2721.

Kara, S.; Rugrungruang, F.; Kaebernick, H.: Simulation modelling of reverse logistics networks. International Journal of Production Economics 106 (2007) 1, pp. 61–69.

Koch, I.: Petrinetze in der Systembiologie. Informatik-Spektrum 37 (2014) 3, pp. 211–219.

Koci, R.; Janousek, V.; Zboril, F., Jr.: Object oriented Petri nets modelling techniques case study. In: Al-Dabass, D.; Nagar, A.; Tawfik, H.; Abraham, A.; Zobel, R. (Eds.): Proceedings of the Second UKSim European Symposium on

Computer Modelling and Simulation. Los Alamitos (CA): IEEE 2008, pp. 165–170.

Lefrançois, P.; Harvey, S.; Montreuil, B.; Moussa, B.: Modelling and simulation of fabrication and assembly plants: An object-driven approach. Journal of Intelligent Manufacturing 7 (1996) 6, pp. 467–478.

März, L.; Krug, W.: Kopplung von Simulation und Optimierung. In: März, L.; Krug, W.; Rose, O.; Weigert, G. (Eds.): Simulation und Optimierung in Produktion und Logistik. Berlin, Heidelberg: Springer 2011, pp. 41–45.

Priese, L.; Wimmel, H.: Petri-Netze. Berlin, Heidelberg: Springer 2008.

Rabe, M.; Deininger, M.: Fabrikmodelle für Job-Shop-Scheduling-Algorithmen in Changing-Steady-State-Systemen. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (eds.): Proceedings of 15. ASIM Fachtagung Simulation in Produktion und Logistik. Paderborn: Heinz-Nixdorf-Inst. 2013, pp. 579–589.

Reisig, W.: Petrinetze: Modellierungstechnik, Analysemethoden, Fallstudien. Wiesbaden: Vieweg + Teubner 2010.

Robinson, S.: Simulation: The practice of model development and use. Chichester, England: Wiley 2004.

Schöf, S.; Wieting, R.; Sonnenschein, M.: Distributed Net Simulation: DNS; Abschlußbericht: OFFIS, Oldenburger Forschungs- und Entwicklungsinstitut für Informatik-Werkzeuge und -Systeme, Forschungsbereich 4, Systemmodellierung 1997.

Starke, P.H.: A memo on time constraints in Petri nets. Berlin, HU Berlin, Institut für Informatik 1995.

Viswanadham, N.; Raghavan, S.N.: Performance analysis and design of supply chains: A Petri net approach. Journal of the Operational Research Society 51 (2000) 10, pp. 1158–1169.

Wieting, R.: Handbuch zur THORN Entwicklungsumgebung. Oldenburg: OFFIS 1996.

Zhou, M.; Jeng, M.D.: Modeling, analysis, simulation, scheduling, and control of semiconductor manufacturing systems: A Petri net approach. IEEE Transactions on Semiconductor Manufacturing 11 (1998) 3, pp. 333–357.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Scheduling under Consideration of the Machine Capability

## *Ablaufplanung unter Berücksichtigung der Maschinenfähigkeit*

Dirk Doleschal, Gerald Weigert, TU Dresden, Dresden (Germany),
Dirk.Doleschal@tu-dresden.de, Gerald.Weigert@tu-dresden.de

Abstract: A high yield is extremely important for the costs of production and thus also the price and competitiveness. In this paper a scheduling method is presented, which takes machine parameters into account. This is done exemplary for a SMD manufacturing. Here, not all machines within a parallel work centre are assumed to be equal, even if they are capable to process the same products. Typically, some machines perform better than other machines. This machine performance could be described by a so called machine health parameter. Also, some products could be more important than other products. This may be defined i.e. by the number of PCB layers. The goal is to schedule more important products to those machines with the best health value. Doing so will increase yield for these products. For this, different dispatching rules and a mixed integer programming approach are compared within a simulation model for practical test data.

## 1 Introduction

Yield is important in every area of current productions. This statement is equally valid, whether a semiconductor factory or a SMD (surface mounted device) factory is considered for example. In literature many different approaches exist to maximize the yield. For example Wein (1992) investigated the correlation between yield and cycle time. Other approaches try to optimize time dependencies between product steps (Klemmt and Mönch 2012). Also for SMD manufactory many investigations regarding yield optimization are done (Wohlrabe 2008). An important lesson here, among other things, is the process capability and, consequently, the machine capability. This parameter indicates how well a machine is able to maintain certain accuracies in production (Fig. 1). The process capability then depends, among other things, on the amount of equipment involved in the process and their machine capabilities. Due to this fact, a preliminary study will be presented for yield optimized scheduling. Using the example of the aforementioned SMD manufacturing, SMD components are assembled on printed circuit boards (PCB). For this, typically multiple placement machines are available.

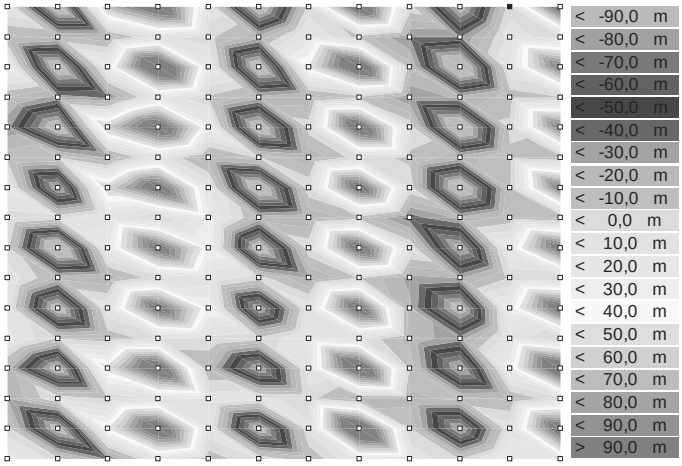



*Figure 1: Atlas accuracy of x-offset of a placement equipment (Wohlrabe 2008).*

Thereby not necessarily all machines are allowed for the placement of all components. Due to this there exists a dedication matrix, which defines the allowed machines for the different SMD devices. Furthermore there may be differences between the circuit boards, which should be equipped with the SMDs. These PCBs, for example, can differ between the numbers of layers (1-36) as well as the size and the complexity. Because of these differences, and thus also the cost of printed circuit boards, the products can be divided into high-priority and less prioritized products. Also the placement machines can have different machine capability values. These values may also depend on the product/machine combination. Now the task is to schedule the different products (PCB/SMD combinations) to the machines, whereby important products should be scheduled to machines with a high machine capability, respectively health value.

This problem could also be adapted to other production areas. For example, the complex production in the semiconductor manufacturing may take machine values into account to plan important or high cost products onto the best possible equipment. Since the production of today's chips in the semiconductor manufacturing partially takes up to three months and several hundred process steps (Potoradi et al. 2002; Yurtsever et al. 2009) until the completion of each layer, there is a large potential for optimization here.

In this paper the possible effect on using the provided machine health values should be estimated. Thereby the calculation of these health values is not taken into account here. The paper is structured as follows. In section 2 the problem is described. Section 3 is used to present the dispatching rules and define the mixed integer programming model. In section 4 the experimental setup is presented and afterwards in section 5 the results are shown. An outlook and conclusion is done in section 6.

## 2 Problem Description

The underlying work centre problem was derived from a practical point of view. A work centre with unrelated parallel machines is used as reference. Also an amount of jobs was created, whereby each job is assigned exactly to one product (PCB/SMD combination). Also each job has a release date and an operational due date (ODD), which defines the due date for the current operation. The defined products have a dedication matrix. This matrix contains the allowed machines for each product. Furthermore the products are divided into 2 groups – important products (IP) and normal products (NP). This classification is used for the yield integrated scheduling and the resulting objective function. The processing times differ for each product, whereby the processing time is equal for one product and different allowed machines. Each machine has a health value, which defines the actual health state of this machine. In this investigation this value is defined between zero and one. The health value is also assumed to be constant during the whole time horizon. This is done because of the simplicity of the model.

The schedules which are used to calculate the objectives are generated with a discrete event simulation model (Fig. 2). All properties, which are described within section 2, have been implemented in the simulation model.

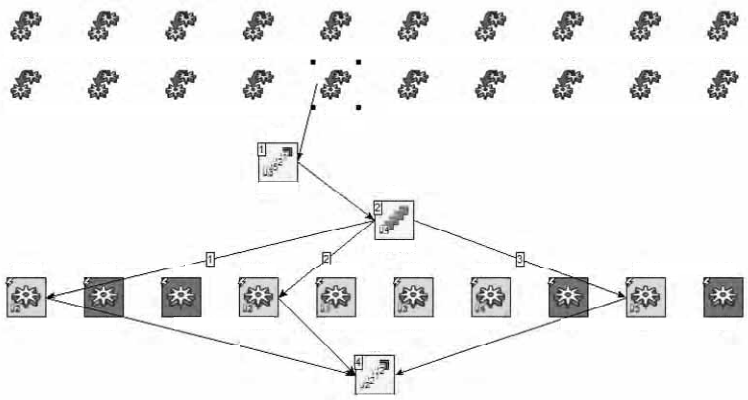


*Figure 2: Exemplary simulation model (simcron MODELLER)*

Also several dispatching rules are included within this simulation model. These different dispatching rules are compared with a mixed integer programming (MIP) approach. Furthermore, an interface was implemented where the results from the MIP model could be used.

The observed objectives in this investigation are:

- Flow factor – This is the ratio of the cycle time to the raw processing time
- Tardiness – The sum of delay for late jobs. Early jobs get a tardiness of zero
- Quality – This is defined as the sum of the machine health value for each important job processed on the corresponding machine

# 3 Methods

## 3.1 Dispatching Rules

As a reference, three different dispatching rules with different complexity are implemented:

- Operational due date - ODD
- Best Machine - BestM
- Practical rule - DePrio

The ODD rule is the simplest one. All jobs are ordered using their local operational due date. If a machine gets idle, the next allowed job regarding the due date is scheduled on this machine. Here no difference between important and non-important jobs is done. Also the health value of the machine is not observed. This rule is investigated by Rose (2003) and is usually used to minimize tardiness.

The next dispatching rule is called “Best Machine”. Here for all important products only the machine with the highest health value is allowed. All other machines which are released for this product in the dedication matrix get locked. So these products could only be processed on the “best” machine out of the set of allowed machines. Furthermore, the sorting of the jobs is also done by the ODD rule. Additionally, all jobs of the important products get a priority state. This means, these jobs have a higher priority compared to the normal jobs. This is done because of the hardly reduced dedication matrix for the important products. Due to the nature of this dispatching rule, the result concerning the quality of the jobs is an upper bound for this objective.

The last used dispatching rule is a practical orientated rule called DePrio. Due to the confidentiality, the rule is not described in detail here. It works with priority based dispatching lists for every machine. Here also the ODD rule is used as basis. Furthermore, the jobs get an additional priority in the case of important products. But the priority is reduced for important jobs depending on the actual load and machine health value. So in contrast to the basic ODD rule, where one global dispatching list exists, each machine has an own dispatching list now, where the jobs are sorted by their priority value.

Because mathematical methods for scheduling get more and more practicable, also a mixed integer programming approach was investigated.

## 3.2 Mixed Integer Programming

A mathematical capacity planning model was implemented using mixed integer programming methods. In this subsection this mathematical model is described in detail and afterwards the coupling with the simulation model is presented.

### *3.2.1 Input Parameter*

For describing the implementation of the mathematical model the input parameter should be explained first. These input parameters are retrieved from the simulation model and could be divided into dynamic and static parameters. This means, a static parameter does not change within the time horizon – in contrast to a dynamic parameter.

Static parameters:

- $n$ different products $P_i$ ($i$=1,…,$n$)
- $m$ unrelated parallel machines $M_k$ ($k$=1,…,$m$)
- Dedication matrix $D \in \{0,1\}^{n \times m}$. Also $D_k := \{ i \mid D_{i,k} = 1 \}$ is the set of products permitted for processing on machine $M_k$. In the same manner, $D_i := \{k \mid D_{i,k} = 1\}$ is the set of machines permitted for processing products $P_i$
- $pt_{i,k} > 0$ is the processing time for a job of product $P_i$ on machine $M_k$ if $D_{i,k} = 1$,
- Machine health value $mh_k \in (0,1]$ $(k = 1,..., m)$ for each machine
- Set of important products $IP \subset \{1,.., n\}$

Dynamic parameters:

- Remaining machine processing time $rpt_k \in \mathbb{N}$($k$ =1,…,$m$)
- Job volume $v_i \in \mathbb{N}$($i$ =1,…,$n$)

Additionally weighting parameters $\omega_1$ and $\omega_2$ exist, which are used to weight the MIP objective function.

### 3.2.2 Variables

The used decision variables in the mixed integer model are the following:

| | |
|---|---|
| $X_{i,k} \in \mathbb{N}$ | amount of jobs from product $P_i$ assigned to machine $M_k$ ($k$=1,…$m$; $i \in D_k$) |
| $C_{\max} \in \mathbb{N}$ | maximum makespan for all machines |

With the defined input parameters and variables the created mixed integer model could be described.

### 3.2.3 Mixed Integer Programming Model

The result from the mathematical model is an assignment of a number of jobs for each product to the machines. Due to the fact that the mathematical model is still a capacity planning model, this is done without knowledge of due dates or release dates. The objective for this MIP model is multi-criteria. On the one hand the maximal makespan $C_{\max}$ has to be minimized and on the other hand the quality $Q$ should be maximized. The quality within the MIP model is equally defined as explained in section 2:

$$\text{Quality: } Q = \sum_{\substack{i \in IP \\ k \in D_i}} X_{i,k} \cdot mh_k \tag{1}$$

The optimization model can be formulated as following:

$$\omega_1 \cdot C_{\max} - \omega_2 \cdot Q \rightarrow \min \qquad \text{subject to} \tag{2}$$

$$\sum_{k \in D_i} X_{i,k} = v_i \qquad i \in \{1, ..., n\} \tag{3}$$

$$rpt_k + \sum_{i \in D_k} X_{i,k} \cdot pt_{i,k} \leq C_{\max} \qquad k \in \{1, ..., m\} \tag{4}$$

(2) is the objective function. The parameters $\omega_1$ and $\omega_2$ are used to weight the two goals. Typically only one parameter is necessary, but for simplicity ($\omega_1$, $\omega_2 \in \mathbb{N}$) both parameters are used. (3) is used to ensure that all jobs are assigned to machines. With (4) the makespan $C_{\max}$ is calculated for the whole machine pool. For this the assignment matrix $X$ and the remaining processing times $rpt$ are used.

### 3.2.4 Result from Mathematical Model and Implementation

The result from this mathematical model is the assignment $X_{i,k}$ of a number of jobs to the machines. This result is directly used within the defined simulation model. For this a user specific script code has been implemented within the simulation model, which is needed to trigger the mathematical optimization run and to transpose the result to the simulation model as shown in Figure 3.

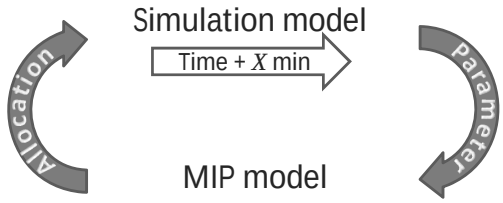



*Figure 3: Dynamic coupling between MIP model and simulation model*

The MIP model is calculated every 10 (simulation) minutes. The MIP model also gets a forecast of 10 minutes for incoming jobs. This means, all jobs with a release date within the next 10 minutes are considered. The implementation of the results within the simulation model is done via product volume lists for each machine. So each machine gets the information, how many jobs of a product are allowed to be processed. The sequencing of the jobs is done by the simulation model using the described ODD rule.

## 4 Experimental Setup

To test the presented methods, a set of test instances is generated. Table 1 gives an overview of the used parameters.

The unrelated parallel machine work centre consists of 10 machines, processing 20 different products. The number of important products differs between 1 and 4, where the corresponding products are chosen randomly. The maximum machine health is always 1. The minimum health varies between 0.1 and 0.7. The processing times for each product are chosen randomly between one and three hours. The release dates for all jobs are distributed between 0 and 1.03 * minimum makespan. The value “1.03” is calculated experimentally to gain an average flow factor of about three

using the ODD rule. The operational due dates are distributed using the average processing times.

*Table 1: Experimental setup (UD - uniform distribution)*

| Factor | Values used | Total values |
|---|---|---|
| Number of products $n$ | 20 | 1 |
| Number of jobs per product $n_i$ | UD ~ [100; 500] | 1 |
| Number of machines $m$ | 10 | 1 |
| Minimum machine health | 0.1; 0.4; 0.7 | 3 |
| Number of important products | 1; 2; 4 | 3 |
| Product dependent processing times $rpt_i$ | UD ~ [1h; 3h] | 1 |
| Release dates $rdd$ | UD ~ $[0, \frac{1.03}{m} \cdot \sum_{i=1,\ldots,n} rpt_i \cdot n_i\ ]$ | 1 |
| Operational due date $odd$ | $odd = rdd$ + UD ~$[\ 2 \cdot \overline{rpt}; 10 \cdot \overline{rpt}\ ]$ | 1 |
| | Number of independent instances | 50 |
| | Total problems | 450 |

On the basis of these test instances, the presented methods are tested and results are generated.

## 5 Results

In this section the results for the used methods and the observed objectives are presented. Thereby the parameter $\omega_2$ for the MIP model is chosen out of the set {1; 5; 25; 50}. The parameter $\omega_1$ is 1.

In Figure 4 the results for all test instances and the quality objective are shown.

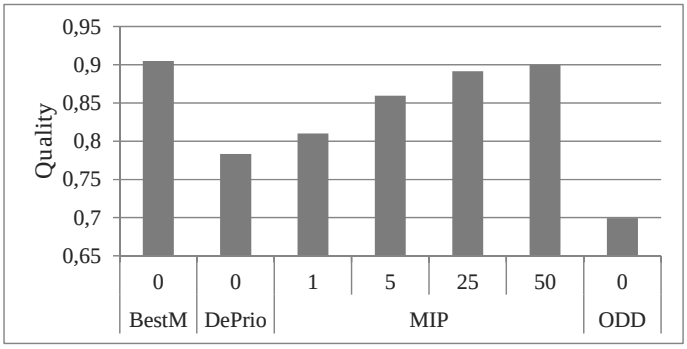



*Figure 4: Result for all test instances regarding quality objective*

Thereby, the x-axis has two rows. The first row describes the parameter $\omega_2$ and the second row the used method. As expected the Best Machine dispatching rule generates the highest quality value. The result of the MIP method depends on the used parameter $\omega_2$. Also, the ODD rule has the lowest quality result.

The results for the tardiness and the flow factor are divided into three bars for each method. The first bar shows the result for all products. In the second bar only the important products are considered and vice versa in the last bar only the result for the normal products are presented.

The tardiness results for all test instances are shown in Figure 5. Thereby the y-axis is cut at 10, because partly results are much higher.

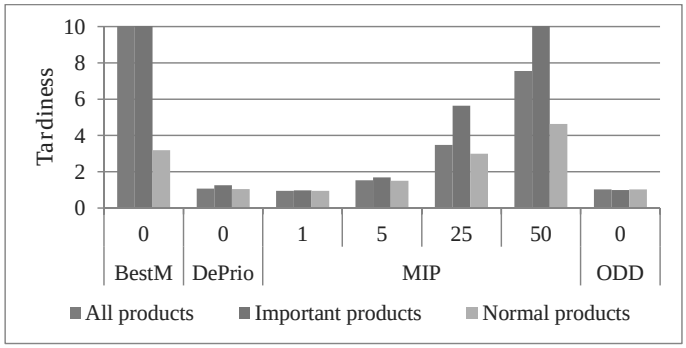



*Figure 5: Result for all test instances regarding tardiness (y-axis cut off)*

As seen in this result, the Best Machine rule and the mixed integer approach with a high parameter $\omega_2$ generates poor results regarding tardiness. This is due to the reduced dedication matrix especially for these test instances with a high number of important products. The DePrio rule performs a little bit poorer than the ODD rule for important products. This rule could also be parameterized. However, this has not been investigated until now. To show the influence of the number of important products, in Figure 6 and Figure 7 the results for one and four important products are presented. Thereby, this time the y-axis is not cut off. This result shows that the choice of parameters is important.

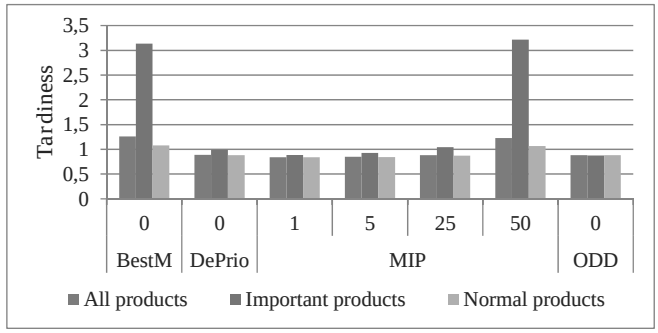



*Figure 6: Results for tardiness for one important product*

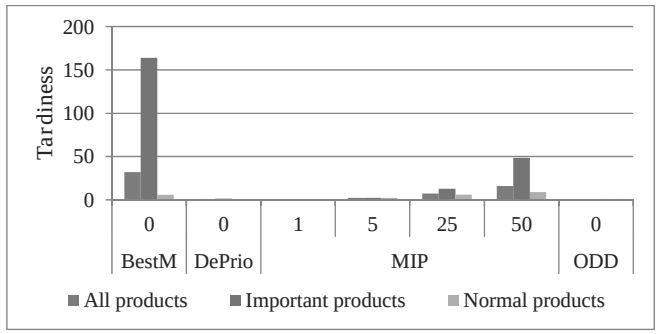



*Figure 7: Results for tardiness for four important products*

The last investigated objective is the flow factor. The results are shown in Figure 8.

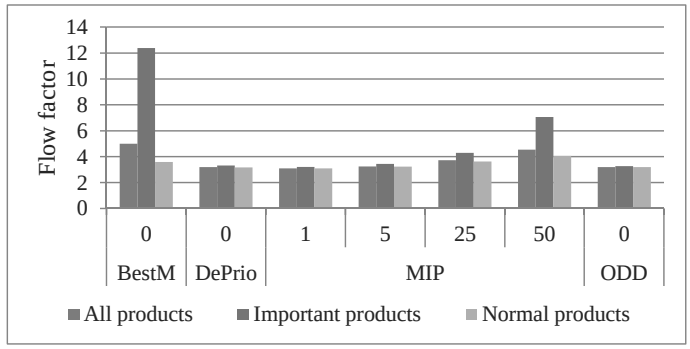



*Figure 8: Result regarding the flow factor for all test instances*

Here a similar result as seen for tardiness could be estimated. Also the minimum health values for the machines have an influence on the results. So in Figure 9 the result for a minimum health value of 0.7 is shown.

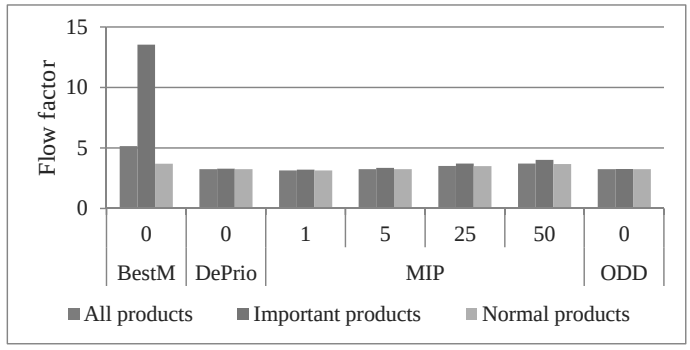



*Figure 9: Result for flow factor and a minimum machine health of 0.7*

Here the deviation of flow factor is much smaller than in Figure 8. This can be explained by the objective function of the MIP model. In the case where the deviation of the minimum health is smaller, also the possible optimization for quality objective is smaller and so the influence of parameter $\omega_2$ drops.

## 6 Conclusion and Outlook

In this paper a yield integrated scheduling method, which uses machine health parameters, is implemented. For this, different dispatching rules and a mixed integer programming model are compared with the well-known ODD dispatching rule. The results show that the mathematical approach gains slightly better results compared to the practical dispatching rule. Further, an improvement in the quality often concludes to a worsening of other objectives. The results in our study show an improvement in quality by an average of up to 20% compared to the ODD rule. Here, the formulation used for the quality objective is just an abstract improvement of the yield. The result cannot be directly converted into yield. This has to be done for each problem area specifically. Overall, the presented investigation is a proof of concept for implementing machine health parameter to scheduling methods.

Further research regarding the parameterization of the practical dispatching rule and the mathematical approach should be performed. Also more complex test instances as well as real data should be used to get a better overview of possible improvement.

## Acknowledgement

The work has been performed in the project eRamp (Grant Agreement N°621270), co-funded by grants from Austria, Germany, Slovakia and the ENIAC Joint Undertaking.

## References

Klemmt, A.; Mönch, L.: Scheduling jobs with time constraints between consecutive process steps in semiconductor manufacturing. In: Proceedings of the 2012 Winter Simulation Conference. Piscataway (NJ): IEEE 2012, S. 2173-2182.

Potoradi, J.; Boon, O.S.; Mason, S.J.; Fowler, J.W.; Pfund, M.E.: Using simulation-based scheduling to maximize demand fulfilment in a semiconductor assembly facility. In: Proceedings of the 2002 Winter Simulation Conference. Piscataway (NJ): IEEE 2002, S. 1857–1861.

Rose, O.: Accelerating products under due-date oriented dispatching rules in semiconductor manufacturing. In: Proceedings of the 2003 Winter Simulation Conference. Piscataway (NJ): IEEE 2003, S. 1346-1350.

Wein, L.M.: On the relationship between yield and cycle time in semiconductor wafer fabrication. IEEE Transactions on Semiconductor Manufacturing 5 (1992) 2, S. 156-158.

Wohlrabe, H.: Qualitätsoptimierung bei der Fertigung elektronischer Baugruppen mittels statistischer Analysemethoden. Templin: Detert 2008.

Yurtsever, T.; Kutanoglu, E.; Johns, J.: Heuristic based scheduling system for diffusion in semiconductor manufacturing. In: Proceedings of the 2009 Winter Simulation Conference. Picataway (NJ): IEEE 2009, S. 1677–1685.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Konsistente Simulationsmodelle für die virtuelle Inbetriebnahme fertigungstechnischer Anlagen mit Hilfe regelbasierter Modellverbinder

## *Consistent Simulation Models for the Virtual Commissioning of Manufacturing Plants Using Rule-based Model Connectors*

Philipp Puntel-Schmidt, Alexander Fay, Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg, Hamburg (Germany), philipp.puntelschmidt@hsu-hh.de, alexander.fay@hsu-hh.de

Abstract: Virtual commissioning requires a simulation model of the plant. The simulation model usually consists of many partial models of different levels of detail. In this contribution, modelling regulations and model connectors are introduced to ensure the interoperability of these partial simulation models. Model connectors allow thereby the representation of information flow between different parts of a plant that are modelled in an asset structure. This asset structure is modelled in a suitable database like AutomationML. The defined modelling regulations, presented by a set of rules, allow for the systematic identification of situations, where additional simulation models like coupling or termination elements are necessary to ensure the interoperability of the partial simulation models. The approach is exemplary shown in a short application example that covers a typical industrial application (windshield glazing in car production).

## 1 Einleitung und Klassifizierung fertigungstechnischer Anlagen

Simulationsmodelle einer Anlage bilden den Kern einer jeden virtuellen Inbetriebnahme (VIBN). Simulationsmodelle müssen die Anlage dabei in einem Maß abbilden, so dass die korrekte Ausführung eines Prozesses unter den bei einem Kunden vorherrschenden Bedingungen getestet werden kann (Haberfellner et. al 2012). Fertigungstechnische Anlagen bestehen aus einer Vielzahl unterschiedlichster Komponenten wie Sensoren, Aktoren und mechanischer/hydraulischer/pneumatischer Grundsysteme (z. B. Förderbänder, Drehtische, Eckumsetzer, Pumpen, Dosiergeräten und weiteren), die miteinander in Verbindung stehen. Stark zusammenhängende Komponenten bilden voneinander abgrenzbare Module, wobei jedes Modul für sich einen bestimmten Zweck bzw. eine bestimmte Funktion innerhalb einer Anlage erfüllt. Jedes Modul innerhalb einer Anlage kann dabei als „System im System“

klassifiziert werden, eine fertigungstechnische Anlage insgesamt unter der „System Of Systems“ (SoS) Sichtweise eingeordnet werden (Shenhar 1994). Um eine fertigungstechnische Anlage gemäß der Systemsichtweise des Systems Engineering zu typisieren, ist es, neben Art und Anzahl der vorhandenen Systemelemente, zusätzlich notwendig, die Interaktionen zwischen dieses Systemelementen zu betrachten.

## 1.1 Interaktionen zwischen typischen Elementen fertigungstechnischer Anlagen

Die Art und Struktur der Interaktion zwischen den modulbildenden Komponenten innerhalb einer fertigungstechnischen Anlage verändert sich in der Regel nicht. Beispielsweise ist die Art und Struktur der Verbindung von Aktoren zum jeweiligen mechanischen/hydraulischen/pneumatischen Grundsystemen als statisch zu bezeichnen (ein Motor ist immer fest mit dem Förderband oder einer Pumpe verbunden). Die Stärke der Verbindungen untereinander kann jedoch durchaus unterschiedlich sein: Versteht man Stärke als den Betrag einer zu übermittelnden (physikalischen) Größe, unterliegt diese mitunter starken Schwankungen und ist durchaus als dynamisch zu bezeichnen (beispielsweise ein aufgenommener Messwert, der mal kleiner, mal größer sein kann oder ein Drehmoment, das von einem Motor erzeugt wird und je nach Last mal kleiner, mal größer ist). Der Materialfluss als maßgebliche Verbindungsgröße einzelner Module untereinander kann in der Fertigungstechnik meist nicht aus einer ununterbrochenen Kette von Komponenten nachvollzogen werden, wie es wie es z. B. in der Verfahrenstechnik möglich ist (Güttel 2008).

## 1.2 Herausforderungen eines „Systems of Systems“-Ansatzes unter dem Aspekt der Simulation

Das System Engineering kennt eine Vielzahl von Herausforderungen, die bei der Betrachtung eines SoS-Ansatzes beachtet werden müssen – eine davon ist die Komplexität des Systems, der generell eine hohe Bedeutung beigemessen wird (INCOSE 2011). Simulationsmodelle von fertigungstechnischen Anlagen, die für eine VIBN genutzt werden und das komplexe System einer fertigungstechnischen Anlage, sind meist nicht monolithisch aufgebaut. Eine jeweils adäquate Modellierungstiefe, die auf standardisierten Modellarten basiert, kann insbesondere für die automatische Generierung von Simulationsmodellen gemäß der in Puntel Schmidt et al. (2015a) vorgestellten Ansätze von Vorteil sein. In der VDI 3693 werden für die Verwendung von Modellarten bei der VIBN in der Fertigungstechnik u. a. Zustandsautomaten, Kinematik- und Dynamikmodelle vorgeschlagen. Für die automatische Modellgenerierung im speziellen wird ein Lösungsraum für Modellarten in Puntel Schmidt und Fay (2015) vorgestellt, der eng mit der VDI 3693 verwandt ist. Verschiedene Modellarten können jedoch unterschiedliche Schnittstellen einzelner Teilsimulationsmodelle beinhalten, die nicht zwangsläufig miteinander kompatibel sind. Dies kann durch die verwendeten, je nach Modellierungstiefe und Modellart potentiell unterschiedlichen Modellierungsansätze (signalflussbasierte oder objektorientierte Modellierung, siehe auch Breitenecker und Popper 2009) begründet sein. Gemäß des Systems Engineering Ansatzes (INCOSE 2011) können „[…] im Widerspruch stehende oder fehlende Schnittstellenstandards den Datenaustausch über Systemgrenzen hinweg [hier: Systemgrenzen der Module] erschweren“. Dieses Problem wird durch einzelne Teilsimulationsmodelle, die in verschiedener Modellierungstiefe abgebildet werden, verschärft. Es ergibt sich daher die Notwendigkeit

einer systematisierten Betrachtung, Klassifizierung und Beschreibung von den Modellbildungsprozess unterstützenden Schnittstellenfunktionen, die die Erstellung eines konsistenten Simulationsmodells im Falle eines heterogenen Detaillierungsgrads erlauben.

## 2 Analyse des Ist-Stands

In Abhängigkeit der verwendeten Modellierungstiefe und Modellierungsart unterscheidet sich das von außen sichtbare Verhalten eines Simulationsmodells (Booch et al. 2005). Grundsätzlich kann dieses beobachtbare Verhalten als kausal (gerichtet, die Charakterisierung als Ein- bzw. Ausgangsgröße ist erforderlich) oder als a-kausal (nicht gerichtet, kann sowohl Eingangs- als Ausgangsgröße oder beides zusammen sein) gesehen werden (Breitenecker und Popper 2009). Während die Definition von kausalen Schnittstellen in der Regel problemlos möglich ist, sind a-kausale Schnittstellen vorzugsweise gemäß den aus der Elektrotechnik bekannten Kirchhoffschen Regeln zu betrachten. Sattler (2007) verallgemeinert diese wie folgt: „[…] [Die] Potentialdifferenzen zwischen zwei Klemmen treiben Flüsse [durch diese] hindurch“. Eine Liste typischer Potential- und Flussvariablen, die bei der Betrachtung fertigungstechnischer Anlagen beachtet werden sollten, kann aus Abbildung 1 entnommen werden.

| | Potential | Flow |
|---|---|---|
| Electrical | **U** [V] | **I** [A] |
| Translation | **v** [m/s] | **F** [N] |
| Rotational | $\mathbf{w_p}$ [rad/s] | **M** [Nm] |
| Fluidical | **p** [N/m²] | **Q** [m³/s] |
| Thermical | **T** [K] | $\mathbf{Q_{th}}$ [W] |

*Abbildung 1: Typische, korrespondierende Potential- und Flussgrößen in a-kausalen Modellierungsansätzen*

Ein typisches Beispiel, bei dem die Simulation von Potential- und Flussgrößen in der Fertigungstechnik eine Rolle spielen kann, sind Klebeapplikationen (z. B. Scheibenkleben in der Automobilmontage oder auch Klebeprozesse im Fahrzeug-Leichtbau/Rohbau). Schnittstellen in einem kausalen, signalflussbasierten Modellierungsansatz können im Gegensatz dazu entweder eine Fluss- oder eine Potentialgröße beschreiben, ihr Verhalten kann nicht nach den verallgemeinerten Kirchhoffschen Regeln beschrieben werden. Schnittstellen von signalflussbasierten Modellen ermöglichen daher nur, das explizite Signalverhalten zu betrachten. Ein typisches Beispiel für einen a-kausalen Modellierungsansatz liefert die Modellierungssprache Modelica, während Matlab-SimuLink ein typischer Vertreter kausaler Modellierungsansätze darstellt. Durch die Verwendung der StateGraph-Bibliothek (Otter et al. 2005) können auch diskrete (kausale) Sachverhalte in Modelica dargestellt werden. Probleme entstehen nun immer dann, wenn Simulationsmodelle, die in unterschiedlichen Modellierungsansätzen (auch innerhalb eines Simulationswerkzeugs) modelliert wurden, miteinander gekoppelt werden sollen. Dies ist nicht ohne weiteres möglich. Werden heterogen aufgebaute Simulationen betrachtet,

erfolgt die Umsetzung der Schnittstellen zwischen Modellen in verschiedenen Modellierungsansätzen meist mit Hilfe von Middleware, wie z. B. dem FMI-Standard (Modelisar 2010), oder auch gemäß des bekannten High-Level-Architecture (HLA) Ansatzes (Straßburger 2003). Jedoch steigt der Aufwand bei der Verwendung derartiger Middleware unter Umständen erheblich an, da diese Systeme derzeit manuell konfiguriert und eingerichtet werden müssen. Soll der Aufwand für den Modellbildungsprozess durch eine automatische Modellgenerierung minimiert werden, ist die vollständige, Darstellung der Schnittstellen als in der für die Modellgenerierung genutzten Datenbasis unerlässlich (Puntel Schmidt et al. 2015a). Ein Ansatz für die konsistente Darstellung von Schnittstellen wird dabei von Christiansen et al. (2014) vorgeschlagen. Die Autoren beschreiben darin das Konzept sogenannter „Modellverbinder“. Modellverbinder erlauben eine Kommunikation zwischen verschiedenen Teilen einer Anlage, die als Modelle in einem geeigneten Datenformat abgelegt sind und „[…] eine einheitliche und durchgängige Informationsverwaltung ermöglichen […]“. Verschiedene Typen von Modellverbindern können identifiziert werden, je nachdem, welche Situation welchen Modellverbinder-Typ erforderlich macht. Diese Typen können wie folgt beschrieben werden:

- *Senken und Quellen*: Terminierung von Schnittstellen wann immer notwendig
- *Modul*: Modellverbinder dieses Typs stellen die Konsistenz der zu übermittelnden Informationen über Modellgrenzen hinweg sicher
- *Jump Points*: Modellverbinder diesen Typs realisieren Beziehungen zwischen verschiedenen Modellen, z. B. zu einem Modell, dass erst während der Laufzeit aktiviert wird

Modellverbinder stellen eine vielversprechende Methode dar, den Informationsfluss zwischen verschiedenen Modellen sicher zu stellen. Jedoch sind im speziellen die Verbindungen zwischen (Teil-)Simulationsmodellen, die in unterschiedlicher Modellierungstiefe und damit potentiell in unterschiedlichen Modellierungsansätzen abgebildet werden, nicht Teil der bisherigen Betrachtung. Insbesondere für die virtuelle Inbetriebnahme und im speziellen eine automatische Modellgenerierung ist es notwendig, weitere Regeln und auch Vorschriften einzuführen, die die automatische Generierung eines konsistenten Anlagensimulationsmodells aus einer geeigneten Datenbasis ermöglichen.

## 3 Grundlegende Überlegungen zu Schnittstellen und mögliche Modellrestriktionen

Eine grundlegende Klassifizierung von Schnittstellen in einem VIBN Ansatz kann zunächst unter Zuhilfenahme der bekannten Automatisierungspyramide vorgenommen werden. Wechselwirkungen zwischen verschiedenen Ebenen der Automatisierungspyramide (z. B. von Steuerungsebene zur Feldebene) können dabei als vertikale Schnittstellen charakterisiert werden. Wechselwirkungen zwischen Elementen innerhalb einer Ebene (in diesem Beitrag: auf Feldebene) können als horizontale Schnittstellen beschrieben werden (Abb. 2).

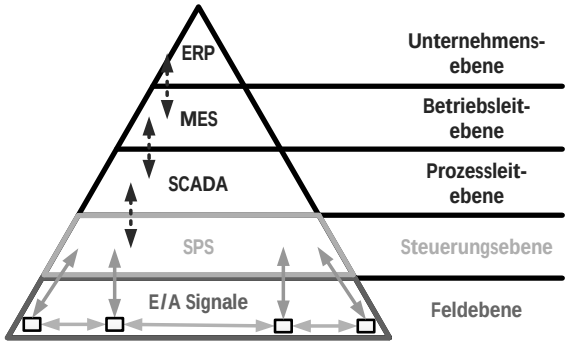



*Abbildung 2: Vertikale und horizontale Schnittstellen aus Sicht der Automatisierungspyramide: Die gestrichelten Pfeile zeigen weitere Schnittstellen zu übergeordneten Systemen an, die nicht Teil dieser Betrachtung sind*

Schnittstellen zwischen dem Anlagensimulationsmodell und der Steuerung müssen, unabhängig von der verwendeten Modellierungstiefe und der Granularität einzelner Teilsimulationsmodelle, stets so modelliert werden, dass Informationsverlust aufgrund nicht kompatibler Informationsstrukturen ausgeschlossen wird (Kufner et al. 2010). Schnittstellen der Steuerung (I/O Signale, Feldbus oder auch Ethernet) sind dabei unabhängig vom Detaillierungsgrad der Prozesssimulation. Horizontale Schnittstellen in Simulationsansätzen unterschieden sich stark von vertikalen Schnittstellen. Schnittstellen können nun an zwei Stellen auftreten: Schnittstellen zwischen Komponenten innerhalb eines Moduls (interne Schnittstellen) und Schnittstellen über Modulgrenzen hinweg (externe Schnittstellen). Abbildung 3 zeigt als Beispiel eine Klebstoffapplikationsanlage, bei der Klebstoff zum Kleben von Bauteilen verwendet wird.

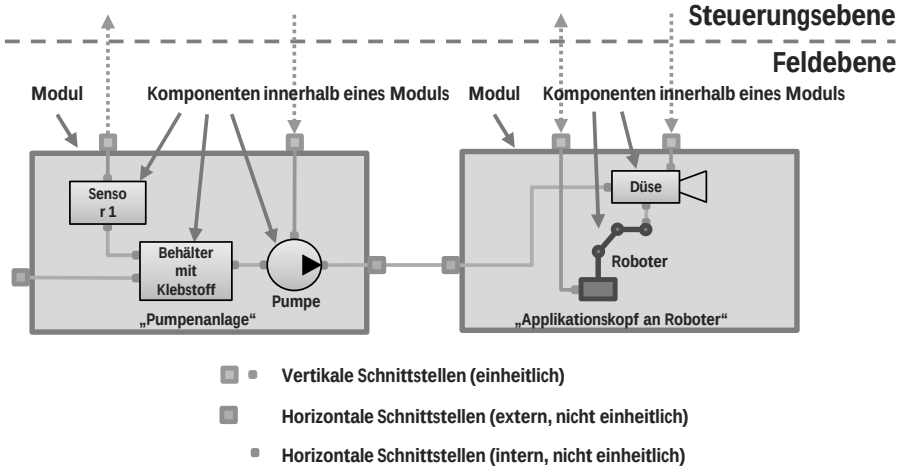



*Abbildung 3: Übersicht über vertikale und horizontale Schnittstellen am Beispiel zweier miteinander verbundener Module (hier: Pumpenanlage und Applikationskopf an Roboter einer Klebeanlage, stark vereinfachte Darstellung)*

### 3.1 Typisierung horizontaler Schnittstellen

Module innerhalb einer fertigungstechnischen Anlage führen meist abgekapselte Funktionen (z. B. Transport, Bearbeitung, Lagerung, Klebstoffapplikationen und weitere) durch und besitzen daher meist klare Systemgrenzen, die sie von anderen Modulen innerhalb einer Anlage abgrenzen. Abhängig von der modulspezifischen Aufgabe (also dem Prozess, der ausgeführt werden soll,) kann eine große Anzahl an Komponenten notwendig sein, um das korrekte Ausführen der jeweiligen Aufgabe zu ermöglichen. Aufgabenabhängig können diese Komponenten jeden denkbaren Typs sein (elektrisch, hydraulisch, pneumatisch, mechanisch, etc.). Es gibt keine Beschränkung hinsichtlich der Implementierungsmöglichkeiten und Kombinierbarkeit dieser Komponenten. Die Möglichkeit unterschiedlicher Schnittstellen innerhalb eines Moduls kann daher (im Kontrast zu Schnittstellen zwischen Modulen) deutlich höher sein und sich in auch in Ihrer Art und Struktur unterscheiden. Jedes Simulationsmodell einer einzelnen Komponente (insbesondere wenn diese in unterschiedlicher Modellierungstiefe dargestellt werden) besitzt daher *potentiell unterschiedliche* Schnittstellen. Um das Handling und damit auch den Aufwand bei der Betrachtung von Anlagensimulationsmodellen zu vereinfachen, sollten Simulationsmodelle von Komponentenmodellen innerhalb eines Moduls daher stets in derselben Modellierungstiefe abgebildet werden. Eine zusätzliche Komplexitätssteigerung des Anlagensimulationsmodells durch potentiell nicht kompatible Schnittstellen kann dadurch vermieden und Situationen ausgeschlossen werden, bei denen „Komplexität zu unerwartetem und unvorhersehbarem Verhalten […] führen kann […]“ (INCOSE 2011).

## 4 Definition von Schnittstellen zwischen Modulen für konsistente Simulationsmodelle

Das bereits in Kapitel 2 beschriebene Modellverbinder-Konzept kann eine Basis für die Definition von Beschreibungen von Schnittstellen sein, um ein konsistentes Simulationsmodell zu gewährleisten. Modellverbinder werden durch virtuelle Objekte innerhalb des Anlagenstrukturmodells (als Teil des Datenmodells der Digitalen Fabrik) definiert, das wiederum die Basis für eine automatische Modellgenerierung bildet (Puntel Schmidt et al. 2015a). Innerhalb der Anlagenstruktur sind Modellverbinder entsprechenden Modulen als hierarchisch tiefer liegende Objekte zugeordnet. Modellverbinder werden durch Attribute ergänzt, die Informationen über die Art und Struktur der über diese Schnittstelle zu übermittelnden Signale/Werte beinhalten. Typischerweise repräsentieren diese Attribute daher die über die Schnittstelle auszutauschenden Potential- und/oder Flussgrößen. Diese Informationen müssen, gemäß den von Bracht et al. (2010) definierten Anforderungen an das Datenmodell der Digitalen Fabrik, stets vollständig und konsistent abgebildet sein, um eine geeignete Basis für die automatische Modellgenerierung zu bilden (Puntel Schmidt et al. 2015a). Modellverbinder sollten granular so fein aufgelöst sein, dass immer nur eine Potential- und/oder (korrespondierende) Flussgröße beschrieben wird. Ein Modellverbinder, der alle Schnittstellen eines Teil-Simulationsmodells zu einer zusammenfasst, wäre zwar möglich und würde die Betrachtung auf Anlagenstrukturebene vereinfachen, jedoch den Aufwand zur Kopplung in Situationen, wo unterschiedliche Schnittstellen existieren, stark erhöhen, wie sich im nachfolgenden Kapitel zeigen wird.

## 4.1 Definition von angebotenen und benötigten Schnittstellen

Um die Anforderungen an eine automatische Modellgenerierung zu erfüllen und das Erstellen von konsistenten Simulationsmodellen zu ermöglichen, muss das Modellverbinder-Konzept zunächst erweitert werden. Das Ziel ist, Modellverbinder inhärent so zu gestalten, dass ein Modellgenerator Inkonsistenzen zunächst identifizieren kann und diese in einem zweiten Schritt beheben kann. Eine grundlegende Klassifizierung von (horizontalen) Schnittstellen, die sich durch einen Modellgenerator interpretieren lassen, ist daher notwendig. Dies kann mit Hilfe der UML-Klassifikation von Schnittstellen geschehen: Anhand der Einteilung in angebotene Schnittstellen und benötigte Schnittstellen. Diese Terminologie soll genutzt werden, um das Modellverbinder-Konzept mit grundlegenden Informationen anzureichern, die von einem Modellgenerator genutzt werden können, um ein konsistentes Simulationsmodell zu erzeugen. Benötigt meint in diesem Fall, dass die „Erfüllung“ der Schnittstelle obligatorisch ist: die Schnittstelle muss immer angeschlossen sein. Bei der Betrachtung horizontaler Schnittstellen sind benötigte Schnittstellen z. B. die Eingangsschnittstellen von kausalen, signalflussbasierten Simulationsmodellen, wobei angenommen wird, dass die Aktionen innerhalb eines signalflussbasierten Simulationsmodells immer durch einen äußeren Einfluss (z. B. ankommendes Material) angestoßen werden. Eine andere Art von benötigter Schnittstelle sind die Schnittstellen von a-kausalen Simulationsmodellen, die gemäß den verallgemeinerten Kirchhoffschen Regeln (siehe Kapitel 2) beschrieben werden können. Angebotene Schnittstellen sind die Ausgangsschnittstellen von kausalen, (signalflussbasierten) Simulationsmodellen. Diese werden bereitgestellt und können von anderen Teilsimulationsmodellen genutzt werden, haben jedoch für das eigentliche Modell keine Funktionsrelevanz. Diskrepanzen zwischen Schnittstellen von Teilsimulationsmodellen können nun immer dort auftreten, wo Attribute der Schnittstellen, definiert innerhalb der jeweiligen Modellverbinder, nicht zueinander passend sind bzw. angebotene auf benötigte Schnittstellen treffen.

## 4.2 Regelbasierte Definition von Schnittstellenfunktionen

Ein regelbasierter Ansatz soll helfen, diese und weitere denkbare Situationen, bei denen es zu Inkonsistenzen kommen kann, zu lösen. Ziel ist es, die Schnittstelle resp. deren Darstellung im Datenmodell der Digitalen Fabrik so zu modifizieren, dass ein Datenaustausch zwischen Teilsimulationsmodellen ermöglicht wird und damit die Konsistenz des Gesamt-Simulationsmodells an sich gewährleistet wird. Hierfür sind, je nach Situation und unter Betrachtung des entstehenden Gesamtsimulationsmodells, zusätzliche Simulationsmodelle, die z. B. die a-kausale Schnittstelle abschließen, notwendig. Diese Simulationsmodelle müssen für jede mögliche a-kausale Schnittstelle vorgesehen werden und den Übergang zwischen detaillierten und weniger detaillierten Anlagenteilen und umgekehrt ermöglichen. Dadurch, dass die Granularität der Schnittstellen und der sie repräsentierenden Modellverbinder fein-granular gewählt wurde (vgl. Kapitel 4.1), ist nur die Definition von wenigen Kopplungs- resp. Terminierungssimulationsmodellen notwendig, die die in Kapitel 2 bereits identifizierten Fluss- und Potentialgrößen umfassen (siehe Abb. 1). Diese Simulationsmodelle müssen in einer Modellbibliothek entsprechend vorgehalten werden. Die folgenden Regeln zeigen eine allgemeine Umsetzung im Rahmen einer automatischen Modellgenerierung für diejenigen Fälle, in denen zwei Schnittstellen aufgrund unterschiedlicher Attribute/Typen nicht miteinander kompatibel sind.

- Regel 1 (Fall 1 Abb. 4):

WENN *eine benötigte Schnittstelle mit einer angebotenen Schnittstelle gekoppelt werden soll*

DANN *füge ein Kopplungselement als ergänzendes Simulationsmodell ein*

UND *leite die dafür notwendigen Daten aus den Attributen des Modellverbinders der angebotenen und benötigten Schnittstelle sowie der Prozessbeschreibung ab*

- Regel 2 (Fall 2 Abb. 4):

WENN *beide Schnittstellen als benötigte Schnittstellen klassifiziert sind*

UND *die Schnittstellenattribute nicht zueinander passen*

DANN *füge ein zusätzliches Simulationsmodell als Abschlussmodell (Terminierungs-Modell) mit einer Schnittstelle ein, dass dieselben Schnittstellenattribute wie die korrespondierende, benötigte Schnittstelle aufweist*

UND *greife den für das nachfolgende Simulationsmodell notwendigen Wert aus der Schnittstellenverbindung ab (Analog dem Spannungs-Messprinzip)*

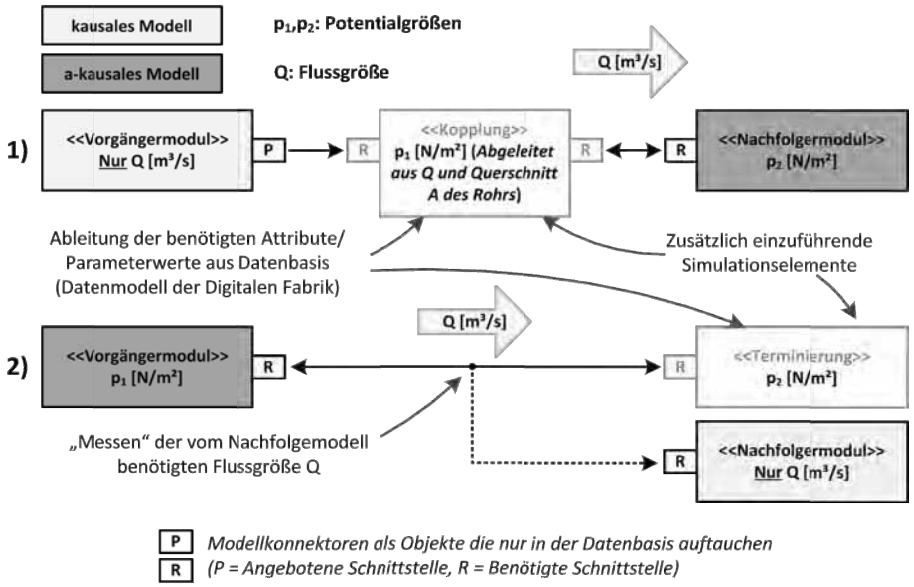



*Abbildung 4: Anwendung des Ansatzes innerhalb eines einfachen Anlagensimulationsmodells (Klebstoffapplikation, siehe Kapitel 3). Betrachtet wird die Flussgröße Q, die zwischen zwei Teilsimulationsmodellen ausgetauscht werden soll*

Abbildung 5 zeigt im speziellen die Anwendung von Regel 1 (Fall 1) im Rahmen des in Kapitel 3 eingeführten Beispiels und unter Verwendung eines Kopplungselements, das direkt aus der Modelica Standard-Simulationsbibliothek abgeleitet wurde (für das Kopplungselement konnte das Modell „MassFlowSource_T“ genutzt werden). Die für die Parametrierung des Kopplungselements und seiner Schnittstellen notwendigen Daten können direkt aus dem Datenmodell der betrachteten Anlage abgeleitet werden. Die genaue softwaretechnische Umsetzung des in Abbildung 5

dargestellten Kopplungselements kann der Modelica-Standardbibliothek entnommen werden.

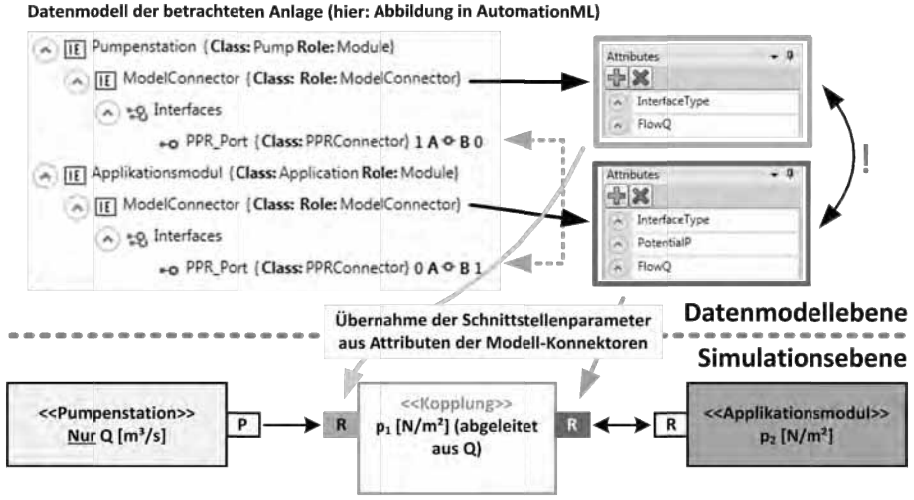



*Abbildung 5: Exemplarische Anwendung von Regel 1, aufbauend auf dem vorgestellten Beispiel aus Kapitel 3 und unter Verwendung der Modelica-Standardbibliothek*

## 5 Zusammenfassung

Die Sicherstellung der Konsistenz eines Anlagensimulationsmodells, insbesondere bei der Verwendung einer heterogenen Modellierungstiefe, stellt eine Herausforderung bei der Modellerstellung dar und kann zu erhöhtem Aufwand führen. Die im Beitrag vorgestellte Anwendung und Erweiterung des Modellverbinder-Konzepts erlaubt es, Teilsimulationsmodelle von fertigungstechnischen Anlagen, die in unterschiedlicher Modellierungstiefe abgebildet sind, konsistent miteinander zu koppeln. Die vorgestellten Methoden und Modellierungsvorschriften, die in einem einfachen Regelwerk hinterlegt sind, tragen dazu bei, die Modellbildungsphase von Anlagensimulationsmodellen, die für die virtuelle Inbetriebnahme genutzt werden, effizienter zu gestalten. Der Ansatz ist dabei unabhängig von der genutzten Integrationsplattform, so dass z. B. eine reine Modelica-Umgebung genutzt werden kann, aber auch z. B. eine Kombination aus Simulink und Modelica oder anderen, für die VIBN genutzten Modellierungsansätzen. Zum Einsatz kommen könnte der vorgestellte Ansatz z. B. auch in einem High-Level-Architecture-Ansatz, und die Verbindung einzelner Teilsimulationen ermöglicht. Durch die hier beschriebene Ergänzungen des Modellverbinder-Konzepts und die Umsetzung von ergänzenden Terminierungs- wie Kopplungsmodellen konnte der in Puntel Schmidt et al. (2015a) beschriebene Modellgenerierungsalgorithmus so erweitert werden, dass die Erstellung von Anlagensimulationsmodellen, bei der einzelne Teilsimulationsmodelle in unterschiedlichen Modellierungstiefen abgebildet werden, automatisiert ermöglicht wird. Zukünftige Betrachtungen konzentrieren sich auf Anwendungsfälle, bei denen sich die Modellierungstiefe einzelner Teilsimulationsmodelle dynamisch während des

Simulationslaufs ändern kann – bisher ist das Anlagensimulationsmodell statisch und kann nicht während der Laufzeit verändert werden.

## Literatur

Booch, G.; Rumbaugh, J.; Jacobson, I.: The unified modeling language user guide, 2. Aufl. Upper Saddle River (NJ): Addison-Wesley 2005.

Bracht, U.; Wenzel, S.; Geckler, D.: Digitale Fabrik. Methoden und Praxisbeispiele, 1. Aufl. Berlin: Springer 2010.

Breitenecker, F.; Popper, N.: Classification and evaluation of features in advanced simulators. In: Tagungsband "MATHMOD 2009 - 6th Vienna International Conference on Mathematical Modeling", Vienna 2009.

Christiansen, L.; Hoernicke, M.; Fay, A.: Entwicklung eines Modellverbinders zur Vermeidung von Inkonsistenzen zwischen Modellen im Kontext des modellbasierten Engineerings. In Tagungsband: "Automation 2014", Baden-Baden, 1.-2. Juli 2014.

Güttel, K.; Fay, A.: Beschreibung von fertigungstechnischen Anlagen mittels CAEX. In: atp edition 50 (2008) 5, S. 34–39.

Haberfellner, R.; Weck, O. de; Fricke, E.; Vössner, S.: Systems Engineering. Grundlagen und Anwendungen. Zürich: Orell Füssli 2012.

INCOSE 2011: Systems Engineering Handbook A Guide for System Life Cycle Processes and Activities, Version 3.2.2. International Council on Systems Engineering, 2011.

Kufner, A., Haug, K., Klemm, P.: Modellierung von Montagemaschinen für die Hardware-in-the-Loop-Simulation. 7. Paderborner Workshop Entwurf Mechatronischer Systeme (EMS), Paderborn 2010.

Modelisar: FMI for Model Exchange 1.0 Specification: https://svn.modelica.org /fmi/branches/public/specifications/FMI_for_ModelExchange_v1.0.pdf, 2010.

Otter, M.; Årzén, K.-E.; Dressler, I.: StateGraph - A Modelica library for Hierarchical State Machines. In: Proceedings of the 4th International Modelica Conference, Hamburg, 7.-8. März 2005.

Puntel Schmidt, P.; Fay, A.; Riediger, W.; Schulte, T.; Köslin, F.; Diehl, S. (2015a): Validierung von Steuerungscode mit Hilfe automatisch generierter Simulationsmodelle. In: at – Automatisierungstechnik 63 (2015) 2, S. 111-120.

Puntel Schmidt, P.; Fay, A. (2015): Levels of detail and appropriate model types for virtual commissioning in manufacturing engineering. In: Proceedings MATHMOD 15 – Vienna Conference on Mathematical Modelling, Vienna, 17.-20. Februar 2015.

Sattler, R.: Physikalisch basierte Mixed-Level Modellierung von gedämpften elektromechanischen Mikrosystemen. TUM München, Fakultät für Elektrotechnik und Informationstechnik, 2007.

Shenhar, A.: A new systems engineering taxonomy. In: Proceedings of the 4th International Symposium of the National Council on System Engineering, Vol. 2, 1994.

Straßburger, S.: HLA als Basis eines Simulationsbackbones für die Digitale Fabrik. In: Proceedings of the 17th Symposium Simulationstechnik (ASIM 2003), Magdeburg 2003, S. 355-360.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Ein Beitrag zum Lebenszyklusmanagement von Simulationsmodellen in der Digitalen Fabrik

## *A Contribution to Simulation Model Lifecycle Management in the Context of Digital Manufacturing*

Sven Völker, Hochschule Ulm, Ulm (Germany), voelker@hs-ulm.de
Jörg W. Fischer, Hochschule Karlsruhe, Karlsruhe (Germany), joerg.fischer@hs-karlsruhe.de

Abstract: A variety of simulation models is created in the context of Digital Manufacturing. Each of these models is used for specific purposes and has an individual lifecycle. Identifying and managing interdependencies between the models along the product and factory lifecycle is an essential part of Digital Manufacturing. The paper proposes a classification of the different types of relations between simulation models. Furthermore, it explains how Lifecycle Mapping can be used for documenting model relations along the product and factory lifecycle. Lifecycle Mapping provides a semiformal, easy-to-understand notation that proved to be useful in PLM projects already. The paper shows that Lifecycle Mapping is applicable in the context of Digital Manufacturing as well.

## 1 Motivation

Der Einsatz von Simulation zur Planung, Absicherung, Inbetriebnahme und Optimierung produktionslogistischer Systeme ist im Kontext der Digitalen Fabrik heute Stand der Technik. Dabei wird Simulation nicht nur einmal im Fabrikplanungsprozess genutzt, sondern kommt in verschiedenen Planungsphasen zur Anwendung (Grundig 2014, S. 241 f.). Im Rahmen der Digitalen Fabrik entstehen dabei mehrere Modelle, die teilweise aufeinander aufbauen, sich gegenseitig ergänzen und die parallel weiterentwickelt werden. Auch kommen je nach untersuchter Fragestellung unterschiedliche Modelltypen zum Einsatz. In der Praxis der Digitalen Fabrik gibt es Defizite beim Management der Interdependenzen zwischen Modellen, und die mangelnde Interoperabilität der Modelle wird beklagt (Wiendahl et al. 2014, S. 518).

Bereits die Definition des Begriffs „Digitale Fabrik“ hebt darauf ab, dass es sich um ein umfassendes Netzwerk von digitalen Modellen und Methoden handelt (VDI 2008, S. 3). Die vernetzten Modelle müssen in ihrer Gesamtheit jedoch nicht unbedingt ein redundanzfreies und global konsistentes Gesamtmodell der Fabrik ergeben. Redundanzen und Widersprüche zwischen den Teilmodellen der Digitalen Fabrik

sind nicht nur unvermeidlich, sondern gelegentlich auch gewollt. Trotz beachtlicher Fortschritte auf dem Gebiet der Digitalen Fabrik (siehe z. B. Zülch et al. 2002) wird das Management der wechselseitigen Modellabhängigkeiten in der Praxis nicht vollständig beherrscht. Zwar umfassen Implementierungsprojekte der Digitalen Fabrik meist Analysen darüber, welche Planungs- und Simulationsmodelle benötigt werden. Welche Beziehungen zwischen den Modellen bestehen sollen, wird aber in der Regel nicht geklärt. Das führt in der Praxis dazu, dass das Modellmanagement nicht organisatorisch verankert wird und die technischen Schnittstellen nicht bedarfsgerecht oder gar nicht implementiert werden.

Daher bedarf es einer Methode zum Erfassen, Dokumentieren und Visualisieren des Zusammenwirkens von Simulationsmodellen. Die erfassten Zusammenhänge bilden dann die Grundlage für die Soll-Konzeption des Lebenszyklusmanagements von Simulationsmodellen. Der vorliegende Beitrag stellt zunächst eine Klassifikation der möglichen Modellbeziehungen vor, geht dann kurz auf gängige Modellierungsmethoden ein und schlägt schließlich Lifecycle Mapping als Methode zur Dokumentation von Modelllebenszyklen in der Digitalen Fabrik vor.

## 2 Beziehungen zwischen Simulationsmodellen in der Digitalen Fabrik

Ebenso wie das Produkt hat auch die Fabrik einen Lebenszyklus (Bauernhansl 2014, S. 25). In unterschiedlichen Lebenszyklusphasen der Fabrik werden Simulationsmodelle unterschiedlicher Granularität für unterschiedliche Zwecke eingesetzt. Am Beginn jeder Simulationsuntersuchung steht die Festlegung des Untersuchungszwecks, idealerweise in Form einer quantitativen Fragestellung (Williams und Ülgen 2012, S. 3668). Die Fragestellung determiniert den zu modellierenden Gegenstand, die Systemgrenzen und den erforderlichen Detaillierungsgrad des Modells. Daraus wiederrum ergeben sich der geeignete Modelltyp sowie die einzusetzenden Softwarewerkzeuge. Klar voneinander abgegrenzte Fragestellungen führen in der Regel auch zu separaten Modellen.

Die verschiedenen Modelle der Digitalen Fabrik stehen untereinander in Beziehung, wobei nur einige dieser Beziehungen explizit, also mit dem Austausch von Daten oder Modellbestandteilen verbunden sind. Viele Beziehungen sind implizit, sollten aber dennoch gemanagt werden. Wie Abbildung 1 zeigt, können vier grundlegende Beziehungstypen unterschieden werden, die zum Teil weiter zu differenzieren sind:

*Typ A*: Zwei Modelle sind alternativ, wenn sie konkurrierende Konzepte für ein zu realisierendes Produktionssystem abbilden (Planungsalternativen). Als Ergebnis der Simulation erweist sich eine der Planungsalternativen überlegen und wird weiter verfolgt, die andere wird verworfen.

*Typ B*: Zwei Modelle sind komplementär, wenn sie verschiedene Subsysteme des zu realisierenden Produktionssystems abbilden, die sich gegenseitig ergänzen.

*Typ C*: Ein Modell baut auf einem anderen Modell auf, wenn es dieses verfeinert oder komplett ersetzt. Dieser Fall ist zu unterscheiden vom normalen Reifungsprozess eines Modells innerhalb einer Planungsphase. Aufeinander aufbauende Modelle unterscheiden sich darin, dass sie der Beantwortung verschiedener Fragestellungen oder derselben Fragestellung auf unterschiedlichem Detaillierungsgrad

dienen. Ein Spezialfall ist die Ableitung eines Grobmodells für die Konzeptplanung aus einem Feinmodell, das den aktuellen Ist-Zustand repräsentiert.

*Typ D*: Zwei Modelle können den gleichen Ausschnitt des Realsystems abbilden, aber unterschiedliche zeitliche Gültigkeit besitzen. Beispielsweise könnte ein Modell den aktuellen Zustand des Produktionssystems repräsentieren, während ein weiteres Modell dem künftigen Zustand der Fabrik nach Abschluss geplanter Umbaumaßnahmen entspricht.

*Typ E*: Für das Verständnis des Zusammenwirkens der Modelle einer Digitalen Fabrik ist es auch wichtig zu erkennen, welche Simulationsmodelle unabhängig voneinander sind, obwohl sie verschiedene Aspekte derselben Fabrik abbilden.

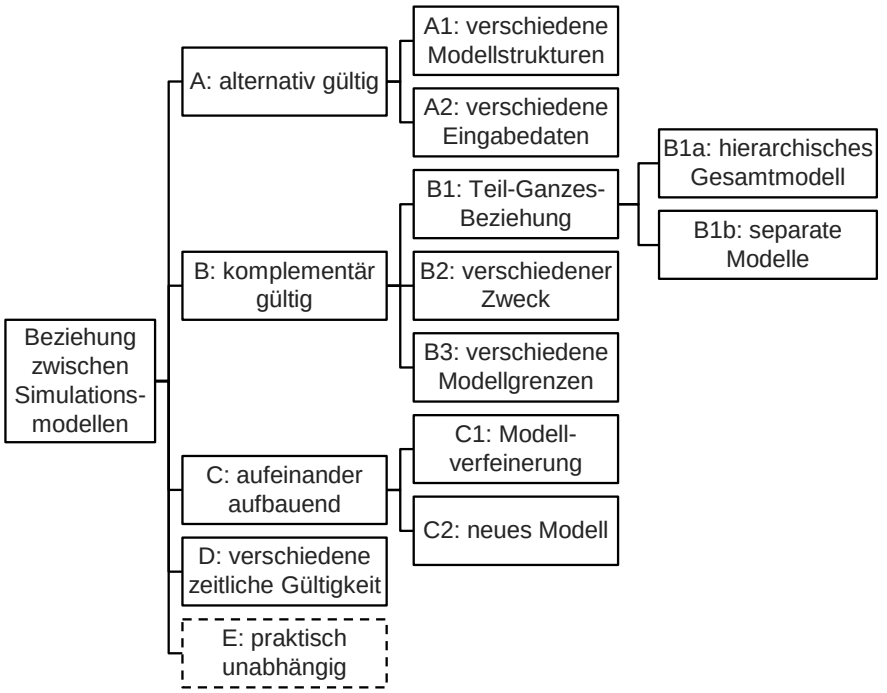



*Abbildung 1: Mögliche Beziehungen zwischen Simulationsmodellen*

Im Zuge der Implementierung der Digitalen Fabrik müssen die für ein Unternehmen relevanten Simulationsmodelle und deren Beziehungen über den Lebenszyklus hinweg geklärt werden. Die dafür infrage kommenden Methoden werden im Folgenden diskutiert.

## 3 Bekannte Methoden zur Erfassung und Beschreibung von Modellbeziehungen

In Implementierungsprojekten der Digitalen Fabrik werden heute meist bewährte Analyse- und Modellierungsmethoden aus dem Umfeld der Geschäftsprozessanalyse (Weske 2012) und der objektorientierten Analyse (O'Docherty 2005) eingesetzt. Die

Analyseergebnisse werden oft in den standardisierten Notationen der BPMN (OMG 2013) bzw. der UML (OMG 2015) dokumentiert. Wird bei der Analyse eine serviceorientierte Sichtweise betont, können auch Notationen aus dem Umfeld Serviceorientierter Architekturen eingesetzt werden, wie z. B. SRML (Fidaeiro et al. 2006).

Trotz der Mächtigkeit der genannten Notationen sind sie zur Darstellung von Modelllebenszyklen nur bedingt geeignet, weil sie nicht modell-, sondern aktivitäts- (BPMN), objekt- (UML) bzw. serviceorientiert (SRML) sind. Aus diesem Grund wurden von Fischer (2014) Lifecycle Maps als einfache, leicht verständliche und semiformale Notation vorgeschlagen. Lifecycle Maps können im Kontext von Digitale-Fabrik-Projekten genutzt werden, um gemeinsam mit den Fachexperten aus verschiedenen Planungsabteilungen den Lebenszyklus der benötigten Planungs- und Simulationsmodelle zu diskutieren und zu dokumentieren.

## 4 Lifecycle Mapping zur Modellierung der Lebenszyklen von Simulationsmodellen

Die als Lifecycle Maps bezeichnete Notation wurde ursprünglich im Kontext des Product Lifecycle Management (PLM) zur Erfassung und Visualisierung von Modellbeziehungen im Produkt- und Fabriklebenszyklus vorgeschlagen (Fischer und Stowasser 2013). Da Modelle der Digitalen Fabrik eine Teilmenge der für PLM relevanten Modelle sind (Eigner und Stelzer 2009, S. 2 ff.), eignen sich Lifecycle Maps auch zur Darstellung der Beziehungen zwischen Simulationsmodellen.

In der praktischen Anwendung werden Lifecycle Maps sowohl vom Ist-Prozess als auch vom zu entwickelnden Soll-Prozess angefertigt: Die Lifecycle Map des Ist-Prozesses hilft beim Identifizieren von Bereichen mit Optimierungspotenzial, für die dann ein Soll-Prozess entworfen wird. Im Folgenden wird die Notation der Lifecycle Maps an einem einfachen Beispiel erläutert (Abb. 2).

Basis der Lifecycle Map ist die Zeitachse. Auf ihr werden die für die Betrachtung wesentlichen Meilensteine des Produktentstehungsprozesses (PEP) notiert. Im nächsten Schritt werden die lebenden Teilmodelle als Rechtecke eingezeichnet. Dabei wird sich auf die Teilmodelle eines repräsentativen Produktes oder Produktionsablaufs konzentriert. Die Rechtecke beginnen zu dem Zeitpunkt, zu dem die Modelle im PEP erstmalig initiiert werden. Ihre Länge über dem Zeitstrahl symbolisiert ihre zeitliche Entwicklung. Streng genommen endet ein Teilmodell nicht, da die Daten über den PEP hinaus vorgehalten und auch "fertige" Teilmodelle oft zu einem späteren Zeitpunkt noch geändert werden. Das Rechteck endet in der Lifecycle Map daher zu dem Zeitpunkt, an dem die Arbeit am Teilmodell im Wesentlichen abgeschlossen ist. Falls es notwendig ist den Aspekt des "Weiterlebens" darzustellen, wird ein verkleinertes, gefülltes Rechteck und/oder eine über den Zeitstrahl weiterführende Linie verwendet. Die geringere Breite symbolisiert die geringere Arbeitsintensität am Teilmodell.

Bei der Analyse von Lifecycle Maps ist es wichtig, vorab einen Fokus zu definieren und diesen einzuhalten. Um zu verhindern, dass die Fachdiskussionen über diesen Fokus hinausgehen, werden "Out-of-Scope-Wolken" verwendet: Sie bezeichnen Aspekte, die außerhalb des Fokus der aktuellen Lifecycle Map liegen. Damit ist in der Diskussion jederzeit klar, dass es diesen Bereich gibt, er gegenwärtig aber nicht betrachtet wird.

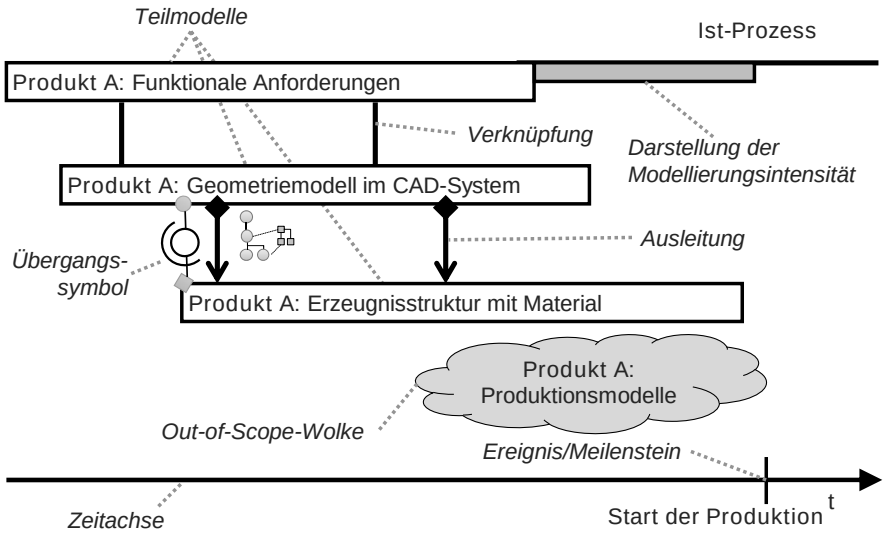



*Abbildung 2: Lifecycle Map eines einfachen Ist-Prozesses*

Die Analyse der Modelle im PEP umfasst zwei wesentliche Aspekte: die Entwicklung der Teilmodelle an sich (d. h. was innerhalb der gezeichneten Rechtecke geschieht) sowie die Interaktion zwischen den Teilmodellen im Lebenszyklus. Lifecycle Maps bilden den zweiten Aspekt, die Interaktion zwischen Teilmodellen, ab. Der erste Aspekt wird mit Hilfe anderer Diagrammtypen untersucht und ist nicht Gegenstand des vorliegenden Beitrags. Zur Darstellung der Interaktion von PLM-Modellen werden in Lifecycle Maps unterschiedliche Verbindungslinien (Linien, Pfeile, gestrichelte Pfeile etc.) sowie Übergangssymbole genutzt. Die Verbindungslinien beschreiben, wie die Modelle zueinander in Beziehung stehen bzw. gesetzt werden. Die Übergangssymbole zeigen, wie die jeweilige Verknüpfung organisiert wird. Die in Abbildung 2 dargestellten Übergangssymbole bedeuten beispielweise, dass die Verbindung zwischen den Modellen von einer verantwortlichen Person aufrechterhalten wird und dass an der Stelle eine semantische Integration der Daten durchgeführt wird. Wie dies genau geschieht, kann mit Hilfe des Diagrammtyps "Semantic Map" des Lifecycle Mappings im Detail analysiert werden. Für weitere Symbole sei auf das in Abschnitt 5 beschriebene Beispiel verwiesen.

# 5 Lifecycle Mapping für die Digitale Fabrik eines Serienfertigers

## 5.1 Betriebliches Szenario

Nach der Klassifikation von Modellbeziehungen und der Vorstellung der Notation sollen die Beziehungstypen und ihre Darstellung in Lifecycle Maps an einem Beispiel erläutert werden. Das betriebliche Szenario liefert ein Hersteller von Elektrogeräten. Zur Produktpalette gehören Rasenmäher-Roboter, die im Folgenden Gegenstand der Betrachtung sind. Das Produkt wird in mehreren Varianten hergestellt. Die Produktion der Rasenmäher-Roboter erfolgt auf einer eigenen Montage-

linie, die sich in derselben Werkhalle befindet wie Montagelinien für andere Produkte des Unternehmens.

Da der Arbeitsaufwand pro Station von der Variante des jeweils bearbeiteten Produkts abhängt, müssen die Montageaufträge sequenziert werden. Diesem Zweck dient ein simulationsbasierter Optimierungsalgorithmus (siehe z. B. Iltzsche et al. 2010), der in das Fertigungssteuerungssystem integriert ist. Die werksinterne Materialversorgung erfolgt über Zugmaschinen und Stapler, die von einem eigenen Leitsystem gesteuert werden, das ein simulationsbasiertes Verfahren zur Allokation der Transportaufträge zu den Fahrzeugen nutzt.

Im Rahmen des Kontinuierlichen Verbesserungsprozesses (KVP, siehe z. B. Kostka und Kostka 2013) werden alle Montagelinien des Werkes regelmäßig hinsichtlich Abtaktung, Arbeitsplatzgestaltung und Materialbereitstellung überprüft. Dabei kommen 3D-Planungs- und -Simulationswerkzeuge zum Einsatz.

Die Unternehmensplanung sieht vor, eine neue Generation von Rasenmäher-Robotern auf den Markt zu bringen. Dabei soll auch die Produktionskapazität verdoppelt werden. Im Rahmen des Projektes werden verschiedene alternative Linienkonzepte diskutiert, die simulativ zu bewerten sind, u. a. eine Duplikation der vorhandenen Linie und eine Verlängerung der Linie bei Halbierung der Taktzeit.

Außerdem erteilt der Produktionsleiter den Auftrag, eine Teilautomatisierung zu evaluieren: Konkret soll untersucht werden, ob das Palettieren der fertigen Rasenmäher künftig manuell oder durch einen Industrieroboter erfolgen soll.

## 5.2 Lifecycle Map der Simulationsmodelle

In dem beschriebenen Szenario kommen mehrere Simulationsmodelle zum Einsatz, die in bestimmten Phasen des Produkt- und Fabriklebenszyklus entwickelt, genutzt und miteinander synchronisiert werden. Die in Abbildung 3 gezeigte Lifecycle Map illustriert die Beziehungen zwischen den Modellen im Lebenszyklus.

Um simulationsbasierte Verfahren zur Auftragssequenzierung und Transportplanung einsetzen zu können, müssen zunächst geeignete Simulationsmodelle entwickelt und validiert werden. Nach Freigabe der Modelle werden sie im laufenden Fabrikbetrieb nicht mehr verändert, sondern nur noch genutzt, wobei sie über automatisierte Schnittstellen mit Auftragsdaten aus den jeweiligen Leitsystemen versorgt werden. Die beiden Modelle bilden verschiedene Bereiche des Produktionssystems ab, sind also komplementär (Beziehungstyp B3).

Die Simulationsmodelle sind ereignisdiskret, der modellierte Ausschnitt der Realität ist aus Performancegründen auf ein Minimum beschränkt: Feingranular sind die Modelle lediglich in Bezug auf die Montagezeiten (Sequenzierungsmodell) sowie die Lade- und Transportzeiten (Transportplanungsmodell). Das Sequenzierungsmodell ist nicht animiert, das Transportplanungsmodell stellt nur eine schematische 2D-Animation bereit. Obwohl beide Modelle parallel genutzt werden, besteht keine Notwendigkeit einer Modellkopplung, das heißt die Modelle sind praktisch voneinander unabhängig (Beziehungstyp E).

Parallel zur laufenden Produktion finden regelmäßig KVP-Workshops statt. Im Rahmen dieser Workshops wird ein 3D-Planungs- und Simulationsmodell genutzt, das insbesondere die Gestalt der Arbeitsplätze, die Abtaktung und das Bereitstell-Layout feingranular abbildet (Abb. 3: KVP-Modell). Das KVP-Modell überlappt

sich inhaltlich mit dem Sequenzierungs- und Transportplanungsmodell, dient jedoch einem anderen Zweck (Beziehungstyp B2) und hat einen anderen Typ.

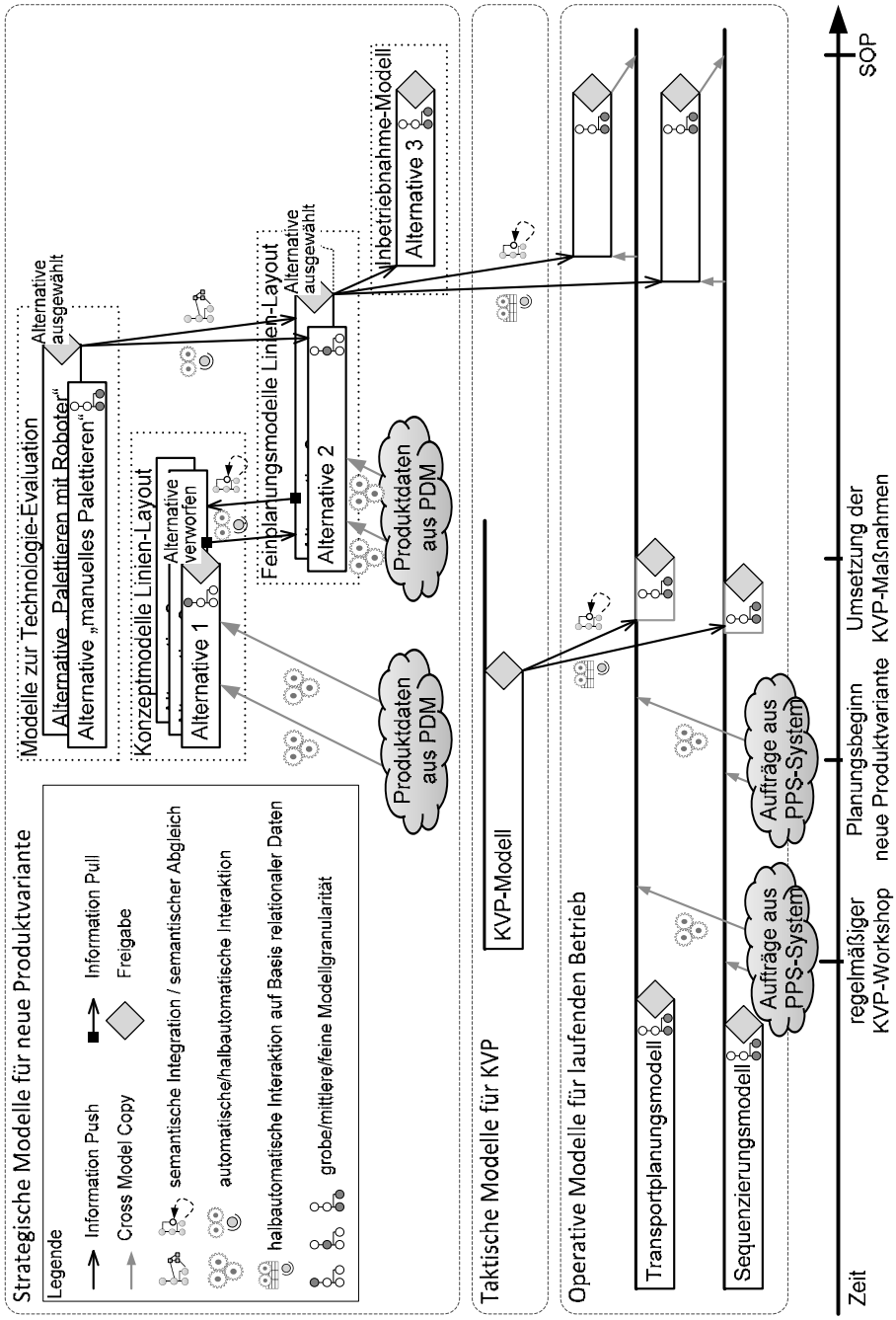



*Abbildung 3: Lifecycle Map der Simulationsmodelle eines Serienfertigers*

Nach Festlegung sinnvoller KVP-Maßnahmen müssen das Sequenzierungs- und das Transportplanungsmodell angepasst werden. Da die originalen Modelle während dieser Modellierungsphase weiterhin benötigt werden, müssen sie kopiert werden, so dass neue Modelle entstehen, die zwar denselben Ausschnitt der Realität abbilden, aber eine andere zeitliche Gültigkeit besitzen (Beziehungstyp D). Die Übertragung der KVP-Maßnahmen aus dem KVP-Modell in das Sequenzierungs- und das Transportplanungsmodell kann nur manuell oder halbautomatisch erfolgen. Erst wenn die modifizierten Modelle validiert und freigegeben wurden, können die KVP-Maßnahmen in der realen Fabrik umgesetzt werden.

Parallel zu Betrieb und fortlaufender Optimierung des Produktionssystems erfolgt die Fabrikplanung für die Einführung der neuen Produktgeneration. Dabei werden zunächst parallel und unabhängig voneinander zwei Planungsaufgaben bearbeitet: die Evaluation der Technologie-Alternativen für das Palettieren und der Entwurf eines Konzeptlayouts für die Montagelinie.

Zur Evaluation der Technologie-Alternativen werden zwei alternative, feingranulare Modelle des Palettier-Arbeitsplatzes benötigt (Beziehungstyp A1). Diese Modelle müssen trotz gleicher Aufgabenstellung nicht zwangsläufig vom gleichen Typ sein: Während die Analyse des automatischen Palettierens eine 3D-Roboter-Simulation erfordert, kann das manuelle Palettieren auch mittels Cardboard Engineering (Gorecki und Pautsch 2014, S. 170 ff) untersucht werden.

In der Konzeptplanung zur Neugestaltung der Montagelinie werden verschiedene Linienkonzepte in ereignisdiskreten Simulationsmodellen mit 2D-Visualisierung abgebildet. Die Modelle können mit Hilfe von Simulationsbausteinen zur Repräsentation einzelner Arbeitsplätze hierarchisch strukturiert werden (Beziehungstyp B1a). Einige Alternativen können bereits in dieser Planungsphase verworfen werden, andere müssen in der Feinplanung ausdetailliert werden. Dazu werden neue Simulationsmodelle gebildet (Beziehungstyp C2): Die Feinplanungsmodelle bilden die 3D-Geometrie der Arbeitsplätze sowie den Verbauzustand des Produktes an den einzelnen Stationen ab. So ermöglichen sie Einbausimulationen und Werkzeugabsicherungen. Im Gegensatz zu den ereignisdiskreten Simulationsmodellen der Konzeptplanungsphase sind diese Modelle jedoch nicht geeignet, um Fragen aus dem Gebiet der Ablaufsimulation zu beantworten.

Aufgrund des Simultaneous Engineering überlappen sich die Modellierungsphasen von Konzept- und Feinplanungsmodellen. Bei der Arbeit mit diesen Modellen werden unter Umständen Erkenntnisse gewonnen, die eine Aktualisierung der jeweils anderen Modellgruppe erforderlich machen. Auch diese Aktualisierung kann nicht voll automatisiert werden, da sie ein ingenieurtechnisches Verständnis der zu übertragenden Sachverhalte erfordert. Gegenstand einer automatischen Modellaktualisierung sind jedoch Updates der Produktdaten, die über Standardschnittstellen aus dem PDM-System bezogen werden. Nach der Entscheidung für eine Technologie zum Palettieren muss das entsprechende Arbeitsplatzmodell in das Feinplanungsmodell integriert werden (Beziehungstyp B1b). Der Aufwand für diesen Schritt hängt von den verwendeten Simulationswerkzeugen ab. Die Integration erfordert aber auch im günstigsten Fall den manuellen Eingriff eines Simulationsanwenders.

Nach Abschluss der Feinplanung beginnt die Vorbereitung der Inbetriebnahme. Zur Absicherung automatisierter Komponenten der neuen Linie mittels virtueller Inbetriebnahme (Bracht et al. 2011, S. 215 ff.) wird ein Modell benötigt, dessen Gra-

nularität gegenüber den Feinplanungsmodellen noch einmal erhöht ist (Beziehungstyp C1): Das Inbetriebnahme-Modell muss das Verhalten der Anlage auf der Ebene der Aktor- und Sensor-Signale abbilden. Erst wenn die Programme zur Anlagensteuerung in der virtuellen Umgebung abgesichert wurden, werden sie auf der realen Anlage (nach deren Aufbau) implementiert. Parallel dazu werden auch neue Versionen des Sequenzierungs- und Transportplanungsmodells entwickelt, die ebenfalls bei Produktionsstart zur Verfügung stehen müssen.

## 6 Zusammenfassung und Ausblick

Es gibt in der Digitalen Fabrik eine Vielzahl schwer zu überblickender Modellbeziehungen, die oft nur in bestimmten Lebenszyklusphasen relevant sind oder ihren Charakter im Zeitverlauf ändern. Die Vielzahl möglicher Modellbeziehungen kann jedoch auf relativ wenige Grundtypen zurückgeführt werden.

Erfahrene Simulationsexperten und Fabrikplaner besitzen umfangreiches – wenn auch meist nur implizites – Wissen über das Zusammenwirken von Simulationsmodellen im Fabriklebenszyklus. Dieses Wissen kann mit Hilfe von Lifecycle Maps explizit formuliert werden. Auf diese Weise schaffen Lifecycle Maps Transparenz über Modellbeziehungen und bilden so die Grundlage für die Soll-Konzeption der Digitalen Fabrik.

Bei der Soll-Konzeption werden erfahrungsgemäß immer wieder ähnliche Probleme diskutiert und üblicherweise ähnliche Lösungen gefunden. Das legt die Entwicklung und Anwendung von Entwurfsmustern nahe, wie z. B. der Lifecycle Sandbox (Fischer 2015). Entwurfsmuster spezifizieren die für einen bestimmten Anwendungsfall benötigten Modelltypen und –granularitäten sowie die Beziehungen und Informationsflüsse zwischen den Modellen im Lebenszyklus. Die Anwendung von Entwurfsmustern verkürzt die Konzeptionsphase bei der Implementierung der Digitalen Fabrik und reduziert das Risiko von Fehlern bei Modellpartitionierung und Schnittstellendefinition.

## Literatur

Bauernhansl, T: Die Vierte Industrielle Revolution – Der Weg in ein wertschaffendes Produktionsparadigma. In: Bauernhansl, T.; ten Hompel, M.; Vogel-Heuser, B. (Hrsg.): Industrie 4.0 in Produktion, Automatisierung und Logistik. Wiesbaden: Springer 2014.

Bracht, U.; Geckler, D.; Wenzel, S.: Digitale Fabrik. Heidelberg: Springer 2011.

Eigner, M.; Stelzer, R.: Product Lifecycle Management. Heidelberg: Springer 2009.

Fiadeiro, J. L.; Lopes, A.; Bocchi, L.: A formal approach to service-oriented architecture. In: Bravetti, M.; Nunez, M.; Zavattaro, G. (Hrsg.): Web Services and Formal Methods. Berlin, Heidelberg: Springer 2006, S. 193–213.

Fischer, J. W.: Lifecycle Mapping – PLM verstehen und gestalten. Gezeigt an einem Praxisbeispiel des Anforderungsmanagements. Zeitschrift für wirtschaftlichen Fabrikbetrieb 109 (2014) 3, S.138-141.

Fischer, J. W.: Licht ins Dunkle – PLM verstehen heißt Lebenszykluseffekte (er)kennen. Zeitschrift für wirtschaftlichen Fabrikbetrieb 110 (2015) 1-2, S. 36-39.

Fischer, J. W.; Stowasser, S.: Industrial Engineering und Lean Product Development. Industrial Engineering 66 (2013) 2, S. 20–27.

Gorecki, P.; Pautsch, P.: Praxishandbuch Lean Management, 2. Aufl. München: Hanser 2014.

Grundig, C.-G.: Fabrikplanung. Planungssystematik – Methoden – Anwendungen, 5. Aufl. München: Hanser 2014.

Iltzsche, L.; Schmidt, P.-M.; Völker, S.: Simulationsbasierte Optimierung der Einsteuerungsreihenfolge für die Automobil-Endmontage. In: März, L.; Krug, W.; Rose, O.; Weigert, G.: Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen. Heidelberg: Springer 2010, S. 117-132.

Kostka, C.; Kostka, S.: Der Kontinuierliche Verbesserungsprozess, 6. Aufl. München: Hanser 2013.

Object Management Group: OMG Unified Modeling Language (OMG UML), Superstructure, Version 2.5. 6/2015. Verfügbar unter: www.omg.org/spec/UML.

Object Management Group: Business Process Model and Notation (BPMN), Version 2.0.2. 12/2013. Verfügbar unter: www.omg.org/spec/BPMN.

O'Docherty, M.: Object-oriented analysis and design. Chichester: Wiley 2005.

VDI: Richtlinie VDI 4499, Blatt 1. Digitale Fabrik – Grundlagen. Berlin: Beuth 2008.

Weske, M.: Business Process Management, 2. Aufl. Heidelberg: Springer 2012.

Wiendahl, H.-P.; Reichardt, J.; Nyhuis, P.: Handbuch Fabrikplanung, 2. Aufl. München: Hanser 2014.

Williams, E. J.; Ülgen, O. M.: Pitfalls in managing a simulation project. In: Laroque, C.; Himmelspach, J.; Pasupathy, R.; Rose, O.; Uhrmacher, A. M. (Hrsg.): Proceedings of the 2012 Winter Simulation Conference. Berlin 2012, S. 3667-3674.

Zülch, G.; Jonsson, U.; Fischer, J.: Hierarchical simulation of complex production systems by coupling of models. International Journal of Production Economics 77 (2002) 1, S. 39–51.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# A Case of Demonstrating Digital Factory Potentials through a Modular Approach

## *Anwendungs-Demonstration von Potenzialen der Digitalen Fabrik mit einem modularen Ansatz*

Giacomo Liotta, Aalborg University, Copenhagen (Denmark), gl@business.aau.dk
Bent Aksel Jørgensen, Xcelgo A/S, Ry (Denmark), baj@xcelgo.com
Charles Møller, Aalborg University, Aalborg (Denmark), charles@business.aau.dk

Abstract: The digital factory concept still needs practical application examples to demonstrate and scale-up its potential benefits. This work presents a pilot project of factory digitalisation that aims to demonstrate the simulation usage in specific projects and daily management practices. A modular approach to discrete-event simulation modelling, project execution and architectural IT framework has been implemented. The architecture consists of a detailed discrete-event simulation model that can be used for virtual commissioning and connected to a performance dashboard. A shared database enables the link with a strategic simulation model and manufacturing system data. The overall approach to the project and outcomes of demonstration activities preliminarily support the following digitalisation hypotheses: a) simulation can be successfully utilised by non-expert end-users; b) multiple performance dimensions can be more easily and quickly analysed and related; c) simulation can be used by industry domain experts and managers to support specific projects and day-to-day management decisions.

## 1 Introduction

The overarching concept of the digital factory or virtual enterprise (Møller et al. 2008) has been around since the early days of Computer Integrated Manufacturing (CIM). However, the reality has shown than many of the proposed benefits have been difficult to realise. Lessons learned from the CIM crisis include, e.g., the need for complexity reduction by using modularisation and lean technologies, decentralised hierarchies, standardisation, human-centred solutions (Zuehlke 2010). Additionally, problem fields for the digital factory implementation (identified a decade ago) included technical and organisational issues such as networking, maintenance and management of versions, data and knowledge, organisation of companies for efficient design processes across multiple departments as well as commitment

towards the involvement of and cooperation with suppliers including small and medium-sized partners (Bracht and Masurat 2005).

Simulation is an integral part of the digital factory concept (IMAGINE 2013; Azevedo and Almeida 2011; Zuehlke 2010; Kühn 2006; Bracht and Masurat 2005).

Two sets of simulation challenges are addressed in this paper: (i) simulation within the broad scope of digital factory applications, and (ii) development of simulation projects and day-to-day simulation usage. The digital factory concept and the "Industrie 4.0" paradigm are considered.

Concerning the first set of simulation challenges, the concept of digital factory includes the simulation of all the activities during the lifecycle of a factory by means of networked digital models, methodologies and applications to integrate design and planning of manufacturing facilities and processes (Azevedo and Almeida 2011). Additionally, it supports the implementation, control and continuous improvement of plant processes and resources (IMAGINE 2013). Recommendations at EU level are to improve accuracy, reliability, performance, scalability and functionalities of simulation and optimisation tools within the context of digital factories which should foster, for example, the development of modular, flexible and highly adaptable solutions, more efficient/reliable design and re-design, shorter time to production (IMAGINE 2013). Even though significant advances have been recently realised in terms of simulation user-friendliness and data management, further R&D and innovations are needed, e. g., for concretely exploiting the connection with physical systems (re-)design, planning and commissioning in industrial applications. This work is based on a discrete-event simulation environment which strongly relies on modularity of simulation objects, control logics and data management in order to (i) facilitate the building, (re-)usability and maintenance of simulation modules in systems (re-)planning, and (ii) enable the connection with Programmable Logic Controllers (PLCs) for the implementation of virtual commissioning in which virtual prototypes simultaneously support the commissioning of control software, fabrication and assembly of a production system (Reinhart and Wunsch 2007). This approach is in line with four of the envisioned design principles of the Industrie 4.0 paradigm in terms of modularity, virtualisation, real-time capability and interoperability (Hermann et al. 2015). Moreover, simulation tools should access data from actual systems, feed and possibly interoperate with advanced analytics tools for strategic alignment and concurrent measurement and prediction of multiple metrics. The latter aspect is addressed in this work by using a modular architecture to link actual data from the manufacturing system, simulation models, as well as advanced analytics and customised visualisations tools.

Concerning the second set of challenges, future models, methods and tools for the digital factory should address not only technological challenges and technical features but also organisational and engineering knowledge aspects (Azevedo and Almeida 2011; Bracht and Masurat 2005). This embraces the approaches by which simulation projects are developed. The theoretical background here includes conventional step-based approaches for discrete-event simulation studies/projects (Banks et al. 2010; Law 2009). Moreover, Johansson (2006) presented a novel methodology for modular discrete-event simulation of manufacturing systems. We build upon the Johansson's approach to contribute to simulation project developments compliant with the future needs of the digital factory concept and Industrie 4.0 paradigm by

using a five-stage approach presented in Section 4. In the Johansson's methodology, the simulation models consist of standard or special purpose modules (e.g., machines, material handling equipment) that are pre-constructed, verified, validated and maintained by simulation experts while embodying their engineering knowledge in terms of logics, data, interface and graphics (knowledge externalisation). Simulation users (e.g., domain experts such as system technicians or production engineers) then (re-)utilise the same modules with the possibility to change and set only a limited number of parameters in order to (re-)build models and add new perspectives, thus realising knowledge internalisation (Johansson 2006). So doing, the project development lead-times can be significantly reduced (Johansson 2006). Such approaches open space for wider studies of existing or not yet existing manufacturing systems potentially with a stronger engagement of stakeholders (managers without specific knowledge and skills), less misunderstanding risks or data analysis/collection issues with respect to conventional simulation projects carried out end-to-end by simulation experts (Johansson 2006).

We try to extend the Johansson's approach by exploring: a) the simulation usage also by domain experts not necessarily skilled on simulation, b) a wide scope of decisions that can be addressed; c) usage contexts of simulation in industrial practices. Our study demonstrates the benefits of large-scale digitalisation and modelling of an entire factory. The proposed approach aims to streamline and facilitate data gathering, modelling processes, simulation model (re-)use and maintenance. It also contributes to overcome potential difficulties related to stakeholder engagement and data management as well as level of detail required by discrete event simulation (Jahangirian et al. 2010).

The hypotheses and the detailed digital model are presented in Section 2 and 3, respectively. Section 4 includes the application, architectural aspects and related discussion. The work contribution is summarised in Section 5. Conclusions and outlook follow in the final section.

## 2 Digitalisation Hypotheses

We formulate the following hypotheses that appear to be plausible according to the first interviews and discussions with the stakeholders as well as preliminary indications:

1. *Simulation models can be built and used by professionals not necessarily skilled on simulation modelling*. In our approach there is a division of labour and roles between simulation experts and two types of end-users: domain experts (e.g., industrial engineers working in the manufacturing system) and managers (higher-level decision makers for investment or re-engineering decisions). Domain experts combine pre-built simulation modules, partially set a limited number of parameters, and evaluate the results on multiple Key Performance Indicators (KPIs) dashboards with managers.
2. *Modular simulation models developed with the proposed approach can simultaneously answer questions based on standard and highly customised KPIs at multiple decision levels*. The approach followed focuses on KPIs and charts interrelating multiple metrics in articulated dashboards with focus, e.g., on sizing and performance of material handling systems, capacity utilisation of manufacturing equipment, production and service rates of machines and

handling systems, as well as demand scenarios. Nevertheless, higher level decision-making is supported as explained in Section 4.

3. *Modular simulation models developed with the proposed approach can benefit domain experts and managers for daily management practices as well as strategic decisions*. Simulation models can be traditionally used within the context of an investment or re-engineering project in order to assess and justify the related cost. Even though this conventional usage is also supported in this application, the proposed overall approach aims to build libraries of modules that can be (re-)combined in living models for the management of operations in daily practices.

These hypotheses have been guiding the engagement throughout the project and have been the assumptions on which the digital model has been made.

## 3 The Digital Model

The digital model reproduces the following system. The system transports empty and full totes in four sizes between a production area and a tote handling areas with buffers, robot handling, tote washing and conveyor transport to a mini-load storage system and dispatch area.

The production area consists of 12 moulding modules each of up to 64 moulding machines, served by 3-5 Automated Guided Vehicles (AVGs) for tote transport between each moulding machine and tote loading/unloading stations, connected to a decentralised tote buffer.

The backbone is a shuttle transport system of almost 1000 meters with approximately 70 shuttle vehicles, transporting totes between the moulding modules and the tote handling area.

Discrete event simulations are normally performed by highly trained simulation experts, and the simulation models are built in specialised simulation software requiring property programming skills.

By structuring the tasks and the model in a new way, the workload can be divided across multiple professionals (Johannsson 2006), and the models can be prepared for future usage such as virtual commissioning.

The simulation model and modelling process are divided into the following main work packages:

1. Material handling objects in a 3D model for discrete event simulation;
2. Flow controllers for applying the overall material flow logics;
3. Input files for generating events such as tote flow, flow rates, order release, breakdown, etc.;
4. Statistics module for monitoring the system characteristics such as performance, utilisation of equipment, buffering, queuing – All relevant statistics are logged in database and files for analysis.

Applying this alternative structure opens for involving the domain experts (system technicians) and managers (system owners) more actively in the modelling process as well as the use and maintenance of the model – an important driver for ensuring the modelling process to reach into the future and towards virtual manufacturing.

The model elements are structured for fast model building and modification by non-simulation experts, building the model by dedicated base objects and in accordance to layout drawing and system documentation by means of pick'n'place. No specific simulation skills are required for model building, but the ability to understand layout drawing and use a 3D modelling environment are the main skills required. Knowledge and insight in the systems functions, operation, controls logics, etc. are important for firstly model verification and adjustment, and secondly model modification, maintenance, and specification of future developments.

The base objects, a total of 14 objects, are developed by skilled object oriented software developers in accordance to specifications and few defined parameters important for preparing the objects for pick'n'place model building and simulation. This includes embedding basic controls logics and functions as well as opening for few relevant parameters settings. A base structure for, e.g., conveyors, AGVs, shuttles, robots and a strong API for simulation features and 3D modelling reduces the development process of the objects. The system flow controller is programmed in C# and successively enhanced for accommodating base simulation runs for this particular system and its characteristics. This also includes logging statistics and event into the repository. Input files for generating events, flow traffic are retrieved from production statistics by the IT department and modified for applying future production scenarios. A screenshot of this detailed discrete-event simulation model created in Experior (Xcelgo A/S) is presented in Figure 1.

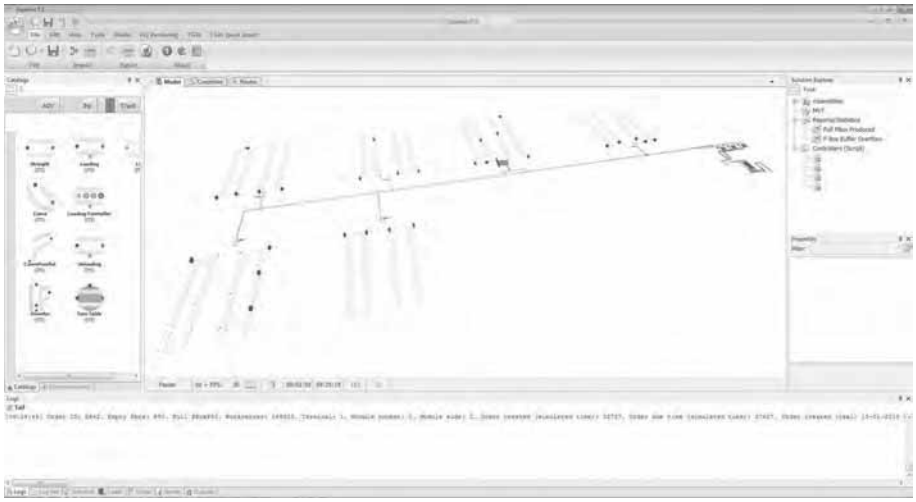


*Figure 1: Overview of a simulation model in Experior (Xcelgo A/S).*

# 4 Project Development, Applications and Evaluation

## 4.1 Project Development

The project development follows an approach which aims to streamline model building, (re-)use and maintenance.

By dividing the tasks into work packages and assign technicians with the right competences, the workload is divided across multiple professionals. The necessary simulation competences are applied to the project by assigning experienced simulation experts to the project group. Specific tasks related to simulation are handled by the simulation experts such as validating the input data, preparing scenario input, defining statistics and especially defining setting for simulation runs. This includes arranging simulation workshop and crash-classes in simulation theory for training the users (i.e., the domain experts and managers) in simulation basics.

The entire digitalisation demonstrator project has been carried out as a pilot over the course of less than one year, and encompassing the following stages.

- *Stage 1: Management buy-in and scoping*. The initial interviews and discussions between simulation experts and users take place while defining goals and scoping the system modelling.
- *Stage 2: Establishing the foundation and building the model*. The core aspects of the model (system components, input data, KPIs, etc.) and of the model usage are established. The simulation experts build and validate the modules of the model.
- *Stage 3: Creating applications of the model*. The domain experts, initially supported by the simulation experts, finalise data gathering and create the applications.
- *State 4: Formulating and quantifying the benefits*. Results and insights are analysed and discussed at two levels: (i) domain experts and simulation experts for the sake of validation and further experiments of the users, and (ii) domain experts and managers.
- *Stage 5: Communicating and extending use*. The results and indications for future usage in daily practices are shared among the stakeholders (domain experts and managers).

The aim of the project is to take this first detailed model as a platform for extending the ideas into the entire factory. So doing, the knowledge of simulation experts can be scaled-up and internalised at end-user level for the finalisation of this first digital factory simulation modelling, its usage in daily practices as well as in new simulation projects for manufacturing system modifications.

## 4.2 Architecture, Preliminary Results and Evaluation

The detailed digital model presented above has been framed in a wider modular architecture consisting of a KPI dashboard for advanced analytics and a strategic discrete-event simulation model built in Simio (Simio LLC). The dashboard is an analytics module which has a business intelligence function to connect the manufacturing system performance with the supply chain and financial requirements. The function of the strategic discrete-event simulation model is twofold: first, it can be used for setting-up high level requirements of the detailed discrete-event simulation model built in Experior from Xcelgo A/S; second, it can be used to make strategic analyses by using the detailed model as a virtual copy of the physical system to compare with for overall cross-validation and alignment purposes. A shared database connected to the actual system data enables the data exchange among the architecture components. The overall architecture is presented in Figure 2 while the strategic simulation model in Figure 3.

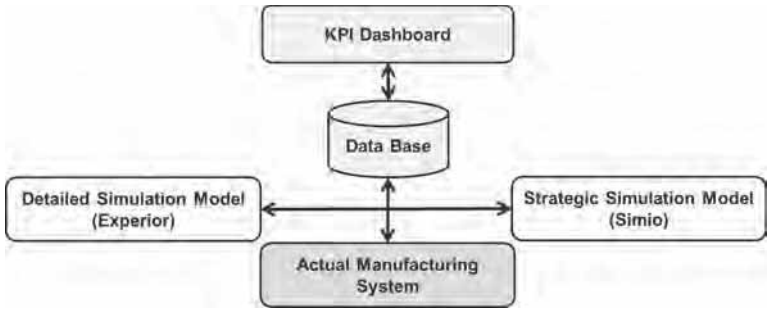



*Figure 2: Overview of the architecture for discrete event simulations and analytics*

The modular structure of the architecture provides the simulation experts and the end-users with the possibility to identify roles and tasks of all the involved professionals, usage profiles and goals, data flow requirements and formats, required output visualisations depending on the usage needs, easier and faster model (re-)building, maintenance, experimentation and analysis in complex decision making situations.

Recalling our digitalisation hypotheses presented in Section 2, we can preliminary argue that:

1. *Simulation models can be built and used by professionals not necessarily skilled on simulation modelling*. It appears that through the proposed application framework young professionals (domain experts and managers) are more inclined and naturally skilled in approaching simulation models and articulated performance dashboards most likely because they are already aware and familiar with visual technologies and usage of objects in dynamic tools such as, e.g., games or other professional IT tools. Detailed simulation models can be built within the context of a specific project or daily management tasks. According to our first findings, observations and user reactions, we therefore believe that simulation modelling skills are not strictly necessary in applications developed in the proposed way.
2. *Modular simulation models developed with the proposed approach can simultaneously answer questions based on standard and highly customised KPIs at multiple decision levels*. Standard KPIs are included in the modules. However, KPIs can also be added on-demand in specific module instances by simulation experts and, more importantly, combined for business intelligence purposes. So doing, high added-value predictive modelling for advanced analytics of systems performance can be carried out. More importantly, the dashboard of the modular architecture can relate highly detailed results of the simulations and the actual system with higher-level supply-chain and financial performance of primary importance of managers.
3. *Modular simulation models developed with the proposed approach can benefit domain experts and managers for daily management practices as well as strategic decisions*. We believe that both domain experts and managers can benefit from modular models and architectures because of the possibility to more easily and quickly combine and test different perspectives in the medium-short term or in the long term planning. Additionally, a validated simulation module or an entire digital factory model can be connected downstream to PLCs

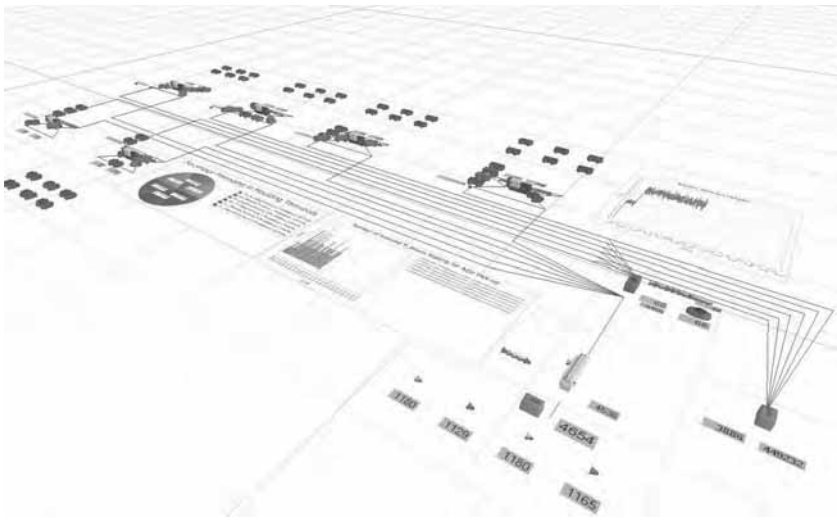


*Figure 3: Overview of the strategic discrete-event simulation model (Aalborg University uses Simio simulation software under a grant from Simio LLC).*

to optimise the equipment commissioning, thus realising virtual commissioning. A detailed simulation model can also share the output upstream with other tools for different analyses such as performance dashboards and higher level simulations to bridge potential alignment gaps with strategic production, supply-chain and financial goals.

## 5 Work Contribution

This work contributes to enhanced simulation usage and exploitation within the context of the digital factory by implementing modularity at two levels: model and architecture.

At the model level, modularity of simulation objects, control logics and data management allows the users to create digital models of physical systems that can be easily built, maintained, (re-)used and connected to PLCs to realise virtual commissioning. At the architecture level, modularity is implemented in terms of: potentially different decision levels supported by different simulation models; advanced analytics and customised visualisation; and actual manufacturing data management.

The proposed approach also contributes to the ways by which simulation models are developed and daily utilised, thus addressing the knowledge and organisational dimensions. A five-stage approach for the digitalisation demonstrator project has been implemented (Section 4.1). More importantly, we build upon the modularity-based simulation by Johansson (2006), and we try to extend it in terms of domain experts not necessarily skilled on simulation, decisions that can be addressed, and multiple usages in industrial practices, from manufacturing system modifications, to production, supply chain and financial management. This also supports stakeholder engagement, a better data management and a suitable level of detail for simulations which may represent issues for discrete event simulation (Jahangirian et al. 2010).

To the best of our knowledge, an overall approach concurrently addressing all the mentioned challenges in a single application case has not been presented so far.

The digitalisation hypotheses presented in this paper, whilst preliminarily discussed, suggest interesting insights and implications. First, the number of users and applications in manufacturing industry can significantly grow both within the context of a single company and across multiple (new user) companies because of the non-compulsory expert requirement as well as the economies of scale and scope in the development of simulation models. Second, in manufacturing industry, this may also open spaces for a concretely beneficial usage of simulation and analytics by small and medium sized players. Furthermore, such simulation-based digitalisation approaches and related potentials may represent an opportunity to spread a beneficial usage of simulation and analytics also in service industry and public administrations.

## 6 Conclusion and Outlook

The overall approach allowed conceiving a modular architecture based on modular detailed simulation modelling and advanced analytics. The work has been developed by using division of labour according to the roles of the involved professionals. A synergetic way of communication among them has been used to pinpoint practical goals and required results. The outcomes of the demonstrator developed for this application appear promising. On the basis of our assumptions, first indications and results discussed with the stakeholders, we believe that it is plausible to spread a quicker and user-friendlier usage of simulation in practical applications even if domain experts and managers are not skilled on simulation modelling. Multiple KPI analyses and decision levels as well as different planning purposes can be potentially supported.

Examples of definitions of these approaches to digitalisation of manufacturing systems and supply chains seem not well defined or aligned yet and need to be clearly linked with available definitions of digital, smart and virtual factory that can be found, for example, in (European Commission 2015; Constantinescu et al. 2006).

With this application, we aim to contribute with a normative approach to define innovative ways of developing simulation projects, using simulations in daily practices, and obtaining practical benefits. We believe that this is highly important in relation to the practical, beneficial usage of digital factory models and the preliminary principles of the Industrie 4.0 paradigm. Our future works will address and evaluate the execution of further experiments, the validation of the overall architecture, the efficiency of simulation usage and related analytics potentials in combination with other techniques and tools.

## Acknowledgement

This work has been developed within the context of MADE – Manufacturing Academy of Denmark, a Danish Strategic Platform for Innovation and Research (SPIR) supported by the Danish Ministry of Education and Science. The authors wish to thank the MADE consortium.

## References

Azevedo, A.; Almeida A.: Factory templates for digital factories frameworks. Robotics and Computer-Integrated Manufacturing 27 (2011) 4, pp. 755-771.

Banks, J.; Carson, J.S.; Nelson, B.L.; Nicol, D.M.: Discrete-event system simulation. Harlow: Pearson, 2010.

Bracht, U., Masurat, T.: The Digital Factory between vision and reality. Computers in Industry 56 (2005) 4, pp. 325-333.

Constantinescu, C.; Hummel, V.; Westkämper, E.: The migration of life cycle paradigm into the manufacturing engineering. In: Duflou, J.R.; Dewulf, W.; Willems, B.; Devoldere, T. (eds.): Proceedings of the 13th CIRP International Conference on Life Cycle Engineering, Leuven (Belgium), 31 May-2 June 2006, pp.705-710.

European Commission: Digital Agenda for Europe - The Factories of the Future, Completed projects. https://ec.europa.eu/digital-agenda/en/completed-projects. Accessed on May 11, 2015.

Hermann, M.; Pentek, T., Otto, B.: Design principles for Industrie 4.0 scenarios: A literature review. Technische Universität Dortmund, Fakultät Maschinenbau, Audi Stiftungslehrstuhl Supply Net Order Management, Working paper no. 01/2015. http://www.snom.mb.tu-dortmund.de/cms/de/forschung/Arbeitsberichte/Design-Principles-for-Industrie-4_0-Scenarios.pdf. Accessed on August 24, 2015.

IMAGINE-FOF-H2020-Event: Session A: PLM / Digital Factory / Simulation/ Optimization for the Smart Factory, Geneva (Switzerland), June 13, 2013. http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=3228. Accessed on May 7, 2015.

Jahangirian, M.; Eldabi, E.; Naseer, A.; Stergioulas, L.K.; Young, T.: Simulation in manufacturing and business: A review. European Journal of Operational Research 203 (2010) 1, pp. 1-13.

Johansson, B.: On virtual development of manufacturing systems - Proposal for a modular discrete event simulation methodology. Ph.D. Thesis, Chalmers University of Technology, Department of Materials and Manufacturing Engineering, Göteborg, 2006.

Kühn, W.: Digital factory - Simulation enhancing the product and production engineering process. In: Perrone, L.F.; Wieland, F.P.; Liu, J.; Lawson, B.G.; Nicol, D.M.; Fujimoto, R.M. (Eds.): Proceedings of the Winter Simulation Conference, Monterey (USA), 03.-06. December 2006, pp. 1899-1906.

Law, A.M.: How to build valid and credible simulation models. In: Rossetti, M.D.; Hill, R.R.; Johansson, B.; Dunkin, A.; Ingalls, R.G. (Eds.): Proceedings of the Winter Simulation Conference, Austin (USA), 13.-16. December 2009, pp. 25-33.

Møller, C.; Jørgensen, B.A.; Chaudhry, S.: Complex service design: A virtual enterprise architecture for logistics service. Information Systems Frontiers 10 (2008) 5, pp. 503-518.

Reinhart, G., Wunsch, G.: Economic application of virtual commissioning to mechatronic production systems. Production Engineering – Research and Development 1 (2007) 4, pp. 371-379.

Zuehlke, D.: SmartFactory - Towards a factory-of-things. Annual Reviews in Control 34 (2010) 1, pp. 129-138.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Optimal Maintenance Resources Allocation Using Automated Simulation-based Optimisation and Data Management

## *Optimale Allokation von Wartungsressourcen durch den Einsatz von automatisierter simulationsbasierter Optimierung und Datenmanagement*

Amos H.C. Ng, University of Skövde, Skövde (Sweden),
amos.ng@his.se

Anders Skoogh, Chalmers University of Technology, Gothenburg (Sweden),
anders.skoogh@chalmers.se

Dan Lämkull, Volvo Car Corporation, Gothenburg (Sweden)
dan.lamkull@volvocars.com

Abstract: This paper introduces a Streamlined Modelling and Decision Support (StreaMod) approach in which input data management, simulation model generation/update and simulation-based optimisation are synergistically integrated into a largely automated process. The aim of this automated process is to support decision making related to the optimal maintenance resources allocation that could improve the performance of production/logistic systems. The basic novelty of the StreaMod optimisation methodology lies on the formulation of an optimal maintenance allocation problem of a production/logistic system into a bi-objective optimisation problem, so that optimal resources/changes to improve the throughput of the system can be sought in a single optimisation run. The successful application of this methodology in a real-world automotive factory will also be addressed in this paper.

## 1 Introduction

Global industry, especially the automotive sector, is constantly challenged towards reduced lead times in product realisation projects and increased efficiency in existing operations. Virtual production development tools are of great importance to address these challenges. The use of simulation with the purpose to improve maintenance operations and, thus, increase equipment availability and Overall Equipment Effectiveness (OEE), is a common industrial practice nowadays. Nevertheless, although simulation is commonly used as a performance evaluation tool in industry, to identify the bottleneck of a production system is still not a trivial task. For example, in a recent study at an automotive manufacturer, it was revealed that no

employee (operators, maintenance personal, system developers, and production managers included) was able to point out the current constraint (bottleneck) of their system (Andersson and Danielsson 2013). In the literature, there are various methods for detecting both momentary and steady-state bottlenecks. These include utilisation of machines (Hopp and Spearman 2000), blocking and starving patterns (Kuo et al. 1996), data-driven approach (Li 2009), shifting bottleneck detection (Roser et al. 2001), multiple bottlenecks (Aneja and Punnen 1999) as well as a method based on inter-departure time and failure cycle data (Sengupta et al. 2008). Nevertheless, all of these existing methods suffer from the same deficiency: even if the overall constraint of the system can be identified down to a specific workstation or machine in the system, there is not enough information provided for determining what improvement action(s) has to be taken. These improvement actions can be in form of reducing processing time, e.g., through new tooling. But in many cases, they are related to improving the availability of the workstation. As maintenance is the process that keeps a workstation operational and reduces the variability and failure in its operational availability (Malakooti 2014), improvement actions can therefore be related to how maintenance resources are allocated. For instance, to increase the availability of the system through preventive maintenance to ensure a process to stay operational (i.e. increase Mean-Time-To-Failure/MTTF) and/or effective breakdown maintenance to reduce its down-time. While these improvement actions can be purely technological, e.g. through investing condition-based maintenance, they can also be related to elevating the capability of the maintanence personell through training. Therefore, to identify and improve the bottleneck of a system can be formulated into an optimal maintenance resources allocation problem, which in turn can be solved as an optimisation problem.

The aim of this paper is to introduce a simulation-based optimisation methodology for solving the above-said optimal maintenance resources allocation problem. By using simulation to model the behaviour of the production/logistic system, optimisation can be carried out to predict where the right improvement actions can be put in order to optimally improve the performance of the entire system. It has been observed that the development of a simulation model as well as the input data analysis and update could be too time-consuming, prohibiting the interest of using simulation and hence the optimisation method for decision support in industry. Therefore, the StreaMod (Strealined Modelling) application framework is introduced so that automated data management, automated model generation and the above-said simulation-based optimisation are integrated to provide a ‘streamlined’ decision support system for production and maintenance engineers. The rest of the paper is organised as follows: Section 2 is about the theory of applying for multi-objective optimisation for optimal maintenance resources allocation, illustrated by a real-world, motivating example in Section 3. Section 4 provides details of the StreaMod framework and why it is so important for decision-making support in industry. Conclusions and current work are given in Section 5.

## 2 Theory

In principle, improving any system is a kind of multi-objective optimisation (MOO) problems because there are almost always some constraints that restrain the total investment/budget and resources that can be put to the improvement. Interestingly,

to our best knowledge, bottleneck improvement has not been studied adequately within the context of MOO. Consider an MOO problem in its general mathematical form:

$$\text{Minimise/Maximise } f_m(\boldsymbol{x}), \qquad m = 1,2,\dots,M \tag{1}$$

Subject to $g_j(\boldsymbol{x}) \geq 0, h_k(\boldsymbol{x}) = 0, \qquad j = 1,2,\dots,J; k = 1,2,\dots,K$

With respect to $\boldsymbol{x} = (x_1, x_2, \dots, x_N)^T$, where $x_i^L \leq x_i \leq x_i^U$ and $i = 1,2,\dots N$.

Here, $f_m(x)$ represents the $M \geq 2$ objectives, which can be minimised and maximised with $x \in$ R N as a solution vector (or simply solution), consisted of $N$ decision variables within their respective lower bounds ($x_i^L$) and upper bounds ($x_i^U$), which maps a solution in the decision space to the objective space as fi : R N → R M. These solutions also have to satisfy the inequality constraints, $g_j(x)$, and equality constraints $h_k(x)$. In many MOO applications, where the objectives $f_m(x)$ are in conflict with each other, finding a single best optimal solution is impossible because improving one objective would deteriorate the others. This gives rise to the concept of Pareto-optimality, describing the set of optimal solutions which are the best trade-offs with respect to $f_m(x)$. In order to determine such a set of optimal solutions, popularly known as Pareto-optimal solutions, the concept of dominance is commonly used by many MOO algorithms:

*Definition 1:* A solution $x_1$ is said to dominate the other solution $x_2$, if both of the following two conditions hold true:

1. The solution $x_1$ is no worse than $x_2$ in all $M$ objectives. So without loss of generality, if we consider a problem of minimising all $f_m(x)$ objectives, then $f_m(x_1) \leq f_m(x_2)$, $\forall\ m = 1,2, \dots, M$.
2. The solution $x_1$ is strictly better than $x_2$ in at least one objective, i.e. $\exists\ j \in (1,2,\dots,M)$ such that $f_j(x_1) < f_j(x_2)$.

The basic idea proposed in (Pehrsson et al. 2011) was based on an observation that many decision-making situations in production system improvement projects can be effectively formulated into an MOO problem. While the primary objective is usually related to a key performance measure, such as system throughput (TH) or cycle time, the novelty of our approach is on formulating the investments needed to improve various attributes of the system as a summation function to represent the second objective of the MOO problem. For example, if the production system throughput, TH is the primary objective for the improvement so that $f_1(x)$=maximise(TH), then we can define the total number of changes, i.e. improvement actions, as minimise($f_2(x)$) in the MOO problem. In this paper, we consider three types of discrete, two-level parameters that can either be set to the system's original value or to a value representing an improvement action: workstation cycle times ($C$), workstation availability ($A$), and workstation mean down time ($D$). In other words, mathematically, the second objective in the MOO problem can be represented by the summation function if there are totally $N$ workstations:

$$f_2(x) = \min\left(\sum_{i=1}^{N} \hat{A}_i + \sum_{i=1}^{N} \hat{C}_i + \sum_{i=1}^{N} \hat{D}_i\right) \tag{2}$$

where:

$\hat{A}_i = 0$, if availability of workstation *i* is not improved and remains to be $\beta_i$, or;

$\hat{A}_i = 1$, if availability of workstation *i* is improved (increased) from $\beta_i$ to $\hat{\beta}_i$.

$\hat{C}_i = 0$, if cycle time of workstation *i* is not improved and remains to be $\alpha_i$, or;

$\hat{C}_i = 1$, if cycle time of workstation *i* is reduced from $\alpha_i$ to $\hat{\alpha}_i$.

Similarly, $\hat{D}_i = 0$, if the mean down time of workstation *i* is not improved and remains to be $\gamma_i$, or; $\hat{D}_i = 1$, if mean down time of workstation *i* is improved (reduced) from $\gamma_i$ to $\hat{\gamma}_i$.

In this way, the second objective function is a discrete function that varies in the range [0, 3*N*], because the maximum number of changes can only be 3*N*. Figure 1 illustrates this bi-objective problem in a graphical way, showing how the Pareto-optimal solutions in the objective space can be used for supporting the decision making in choosing the optimal (minimal) changes to improve the system to achieve the target throughput (e.g., the throughput $TH_t$ in fig. 1).

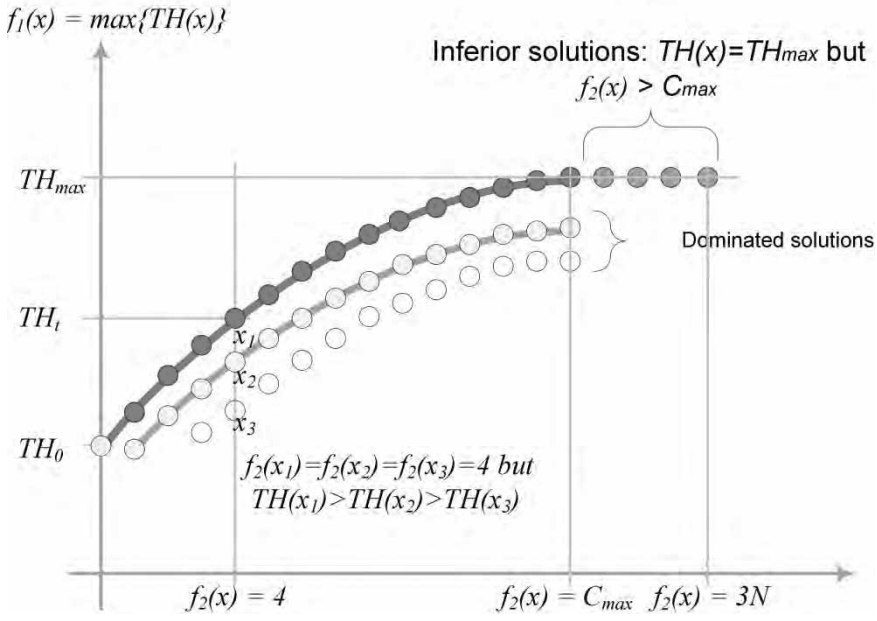



*Figure 1: MOO results for decision making of system improvement.*

## 3 A Motivating Industrial Application

This section addresses how a real-world improvement project was solved by MOO which effectively identified the exact areas of optimal improvements in production

and maintenance parameters to reach the desired target condition. The complex production line is an automotive component machining line that includes multiple parallel sections, portal cranes, machining centres and assembly stations which conduct multiple operations. Apart from some maintenance issues that affect the machine availabilities, tens of variants have to be processed in the line and variations in the weekly volume contribute to the high variability of the system. The crucial issue of the company was the poor production capacity of this line, far below the required capacity to keep up production, which made it the "bottleneck" of the entire plant. An improvement project was in order but given the size, complexity and variability of the line, it was believed to be extremely hard to locate where and what to improve, let alone the effect of performing the improvements, if only traditional simulation tool had been used. And it was obvious that not a single, but multiple improvement actions had to be made in order to achieve the targeted throughput level demanded by the management of the company. The engineers in charge decided to build a simulation model. Several versions had been developed but the latest complete model as shown in figure 2 was built in FACTS Analyzer (Ng et al. 2007). The rapid modelling, easy to learn and use features of FACTS Analyzer has endowed the possibility for the engineers to build/update simulation models for their own production lines.

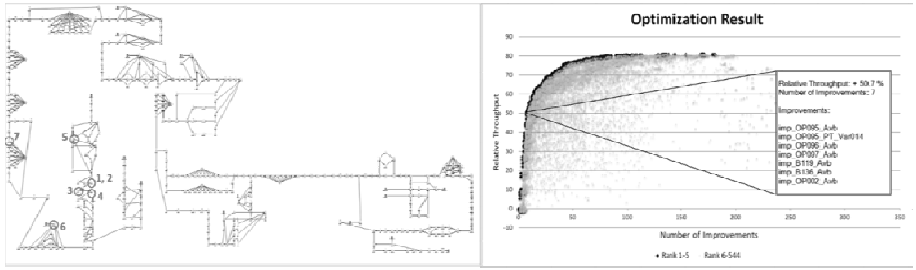



*Figure 2: The simulation model and MOO results from a real-world production system improvement project.*

When the model was completed and validated, the project manager called upon all production and maintenance engineers in charge of the production line together in some workshops in order to propose all possible improvement options, including reduced processing times (per variant where applicable), increased availabilities, and reduced mean times to repair (MTTR). The levels of the improvement (from the original value) are varied between different work-stations. Table 1 lists the number of improvement variables of each type and also the range of the improvements of that type, e.g. the processing time reduction ranges from only 0.2 % for one station to 41.8 % for another station. That summed up to be 464 improvement alternatives represented in the optimisation problem as binary Multiple Choice Set (MCS) variables, for more detail see (Bernedixen and Ng 2014).

*Table 1: Number of decision variables and the range for their improvements*

| *Number of variables of this type* | *Number of variables* | *Range (min-max)* |
|---|---|---|
| Processing time | 317 | -0.2% to -41.8% |
| Availability | 82 | +0.1% to +23.8% |
| MTTR | 65 | -5.5% to -92.1% |

The optimisation results, in form of Pareto front, showing the optimal trade-offs between minimising the number of improvement actions and maximising throughput generated in the project using the well-known NSGA-II algorithm (Deb et al. 2002) are also shown in the data plot in figure 2. In order to conceal the real figure of the company, here only relative throughput, expressed as relative change in percent from the original state (i.e. no improvements), is displayed. It can be seen that the optimal improvement actions found in the optimisation could improve the system substantially up to 80 % and a significant improvement (about 50 %) can already be reached with only seven improvements, as indicated in the data plot in figure 2. These seven areas of improvement are labelled in the snapshot of the model and note that four of them are centred around one single production area. These optimisation results actually surprised the production engineers as they had believed that the major problematic areas that should be improved and prioritised were somewhere else. With further engineering/maintenance investigations on how to achieve the optimal improvement actions, subsequent decision making was made largely based on this set of optimisation results plus more simulation runs to verify the final decisions before implementations. The engineering team also concluded that the optimisation results also greatly facilitated their discussions and cooperation as well as supported their decision making when compared with only traditional lean methods and tools had been used. Nevertheless, the most daunting task identified in the project which yet to be solved if the same method has to be used for other lines, was data collection for updating the input data of the model. As a matter of fact, data collection alone is a critical issue in any simulation modelling project and an automated process to collect and update the model data is badly needed if such an optimisation method has to be used in a regular (weekly) basis to continuously improve the performance of complex production system.

## 4 StreaMod: System Framework

As hinted above, a major challenge for increasing the use of virtual tools for the purpose of increasing the efficiency of existing production system is to reduce the often extensive lead times of simulation studies. There is also a need to increase the usability of these tools especially among production and maintenance engineers. A current Swedish research project called StreaMod is addressing these challenges by building a framework for automated connection of input data collection and analysis, simulation model generation, and optimisation.

Most previous research in the area has been directed towards increasing efficiency of the separate activities of the simulation modelling procedure, e.g. input data

management, model generation, and output analysis. These efforts have, for example, resulted in software solutions such as the GDM-Tool (for automated input data management) and FACTS Analyzer (for efficient model building and optimisation). Nevertheless, to reach the desired lead-time in simulation studies, these steps need to be fully integrated.

## 4.1 StreaMod Purpose and Aim

Based on the challenges outlined above, the purpose of the StreaMod framework is to enable more facts-based decisions in production development related to lean operations and plant design. This will be reached through dramatic reductions of the times used in both data collection, modelling, and generation of decision support. Such a new type of decision support system is also necessary to turn today's substantial system losses into a potential for improved economic and ecologic sustainability.

Advanced analysis of production data, in combination with simulation model building and optimisation, can provide both detailed and visual input to continuous improvements as well as conceptual design. A key component for such decision support system is dramatically reduced time from data collection through model building and optimisation. Improved interfaces and user focus is also necessary to facilitate dissemination among production and maintenance engineers. A major challenge is to provide the necessary technical solutions and to combine them with manual work procedures making the results generic in terms of different organisational structures.

In order to fulfil the purpose, the specific aims of the StreaMod framework are:

- Reduce lead-time for simulation of production flows to <24h response time, including data collection, model building, optimisation, and improvement suggestions.
- Proven identification of improvement potential never detected by analytical analysis, resulting in 10% increased throughput at our demo-lines.
- Extended user base among production and maintenance engineers and more frequent use of modelling techniques in production systems analysis.
- Increase the amount of quality assured data compared to current assumptions and guesses.

## 4.2 Framework Components

The StreaMod framework is divided into several integrated components. The decision support system is the "user interface" providing a set of automated analyses for the targeted production and maintenance engineers. Further, the components for efficient input data management, model generation and optimisation are integrated using a solution based on Intelligent Data Points (IDP). There is also a component providing manual work procedures, mainly for assuring high quality of input data and for complementing missing parts of raw data.

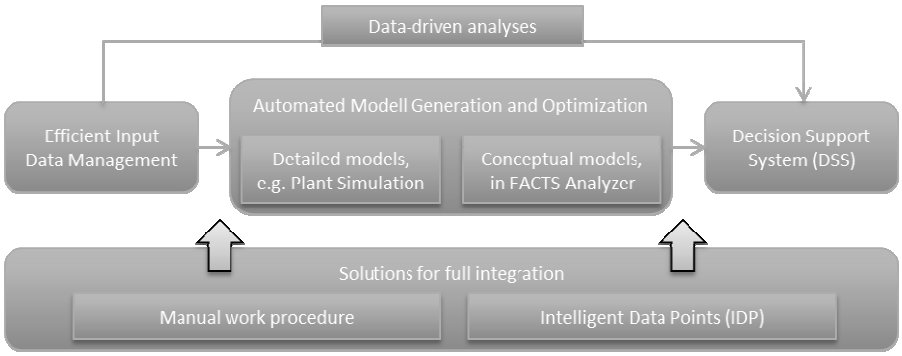



*Figure 3: StreaMod framework*

### 4.2.1 Decision Support System

This component provides a decision support system (DSS) that can be used to support high-quality decision making in production systems design and improvement in industry. Apart from the common features in an ordinary DSS, like visualisation, scenario development, user interaction, database and knowledge-base management, the uniqueness of the new DSS is built on the concept of integrating data-driven analyses, simulation, optimisation and data mining technologies within a unified framework. An aim is to show how production constraints and corresponding improvement proposals as shown in Section 2-3 can be generated as daily reports automatically using the simulation model generated for real-world projects. In order to generate valid improvement proposals automatically, the simulation models must also be generated automatically.

### 4.2.2 Efficient Input Data Management

This component increases the applicability of simulation and optimisation as well as purely data-driven approaches by reducing the time-consumption for collecting and preparing the necessary input data. Correct input data is crucial for facts-based decision support using any detailed analysis method, specifically when considering dynamic aspects of production systems. The procedure of managing this data is currently a major contributor for extensive lead-times of simulation studies due to the vast amount of samples required. Hence, this is a common reason for companies to renounce these powerful approaches in favour of less detailed and precise methods. This is specifically problematic in operational phases when answers are expected in just days or hours.

A crucial capability is to automatically combine and integrate data from major PLM systems (e.g. SIEMENS' Teamcenter) with local data sources (e.g. systems for real-time collection of downtimes and cycle times). This is necessary for both simulation and data-driven analyses. The development of automated input data management is based on prior research on the GDM-Tool (Skoogh et al. 2012).

### 4.2.3 Automated Model Generation and Optimisation

This component provides the required automated model generation and optimisation algorithms for the DSS as exemplified earlier in this paper. The concept of auto-

mated model generation was first explored and successfully demonstrated in FACTS Analyzer, but unlike the technology introduced in FACTS that a detailed model (e.g. Siemens Plant Simulation) can be generated and subsequently extended from a simple line specification, the StreaMod framework requires exploring the technology to generate simulation models from the automated data management system in an on-line basis. Another major challenge is on how the automated model generation can seamlessly support the automated optimisation required by the DSS.

### 4.2.4 Integrated Data Management and Model Generation

The integration of previously explained components is enabled by the automated, seamless model generation based on IDP. IDP is used for automating the connection between model generation and data collection/processing using data intelligence, i.e. data will be automatically requested by recognising model entities and finding related data and information from global and local sources online. Automated import of model entities and logics from PLM system is also provided through IDP – i.e. customised IDPs that connect to PLM entities. The integration of FACTS and the GDM-Tool is then realised by a standard format for sharing information between these data and modelling applications.

### 4.2.5 Integrating Work Procedure

This component describes a general work procedure for integrating data sources and simulation model with a focus on identifying and quality assuring available data and completing unavailable data with manual gathering. One important example is instructions specifying responsibility and frequency for updating sources production data. Despite the focus on data sources and gathering, the work procedure aims to cover all parts of the information chain of the StreaMod framework, e.g. including instructions on the use of the DSS. The major challenges are to turn the existing, general approach, into practical work procedures similar to project management models. Furthermore, the work procedure has to be adaptable to the different organisational structures exemplified by major automotive companies as well for Small and Medium-sized Enterprises (SMEs) and clients to consultancy firms.

## 5 Conclusions and Current Work

This paper has introduced how optimal maintenance allocation can be formulated into a bi-objective optimisation problem in order to improve bottleneck of production systems. As a matter of fact, by formulating a bottleneck improvement into an MOO and then generate an optimal efficient solution set is only part of a complete solution because various issues, including computationally expensiveness; incorporating preference of decision maker, if any; the analysis of the factors which contribute to the best trade-offs, as well as automated input data collection and update for the simulation model, are equally important. The last issue is addressed in this paper as a key objective of the StreaMod project in Sweden. The StreaMod framework is now being implemented in this industrial-based project, aiming at a complete system prototype for the try-out in Swedish automotive industry.

## References

Andersson, J.,; Danielsson. M.: Analysis of system losses: Capacity study at Volvo Cars Torslanda. MSc Thesis, Dept. of Product and Production Development, Chalmers University of Technology, Sweden, 2013.

Aneja, Y.P.; Punnen, A.P.: Multiple bottleneck assignment problems. European Journal of Operational Research 112 (1999) 1, pp. 167-173.

Bernedixen, J.; Ng, A.H.C.: Practical production systems optimization using multiple-choice sets and Manhattan Distance based constraints handling, Proceedings of the 12th Industrial Simulation Conference, 2014, pp. 97-103.

Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T.: A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Transaction on Evolutionary Computation 6 (2002) 2, pp. 181-197.

Hopp, W.J.; Spearman, M.L.: Factory physics: foundations of manufacturing management, 2nd ed. Burr Ridge, IL: Irwin McGraw-Hill Higher Education 2000.

Kuo, C.-T.; Lim, J.-T.; Meerkov, S.M.: Bottlenecks in serial production lines: A systematic approach. Mathematical Problems in Engineering 2 (1996) 3, pp. 233-276.

Li, L.: Bottleneck detection of complex manufacturing systems using a data-driven method, International Journal of Production Research 47 (2009) 24, pp. 6929-6940.

Malakooti, B.: *Operations and Production Systems with Multiple Objectives*, Wiley Series in System Engineering and Management. New Jersey: Wiley & Sons 2014.

Ng, A.H.C.; Bernedixen, J.; Urenda, M.; Jägstam, M.: Factory flow design and analysis using internet-enabled simulation-based optimization and automatic model generation. In: Jain, S.; Creasey, R.R.; Himmelspach, J.; White, K.P.; Fu, M (eds.): Proceedings of the 2011 Winter Simulation Conference, 2011, pp. 2181-2193.

Pehrsson, L.; Ng, A.H.C.; Bernedixen, J.: Multi-objective production system optimisation with investment and running costs. In: Wang, L.; Ng, A.; Deb, K. (eds.): Evolutionary Multi-objective Optimization in Product Design and Manufacturing. London: Springer 2011, pp. 431-453.

Roser, C.; Nakano, M.; Tanaka M.: A practical bottleneck detection method, Proceedings of the 2001 Winter Simulation Conference, 2001, pp. 949-953.

Sengupta, S.; Das, K.; VanTil, R.P.: A new method for bottleneck detection. In: Mason, S. J.; Hill, R.R.; Mönch, L.; Rose, O.; Jefferson, T.; Fowler, J.W. (eds): Proceedings of the 2008 Winter Simulation Conference, 2008, pp. 1741-1745.

Skoogh, A., B. Johansson, and J. Stahre. 2012. Automated input data management: Evaluation of a concept for reduced time-consumption in discrete event simulation. Simulation: Transactions of the Society for Modeling and Simulation International 88(2012) 11, pp. 1279-1293.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# A Comparison of Discrete-Event Simulation Approaches for Complex Manufacturing Systems and Healthcare Systems

## *Ein Vergleich von Ansätzen der diskreten Simulation für komplexe Produktionssysteme und für Systeme im Gesundheitswesen*

John W. Fowler, Arizona State University, Tempe (USA), John.Fowler@asu.edu
Lars Mönch, University of Hagen, Hagen (Germany),
Lars.Moench@fernuni-hagen.de

Abstract: In this paper, we compare discrete-event simulation approaches for complex manufacturing systems and healthcare systems. The aim of this paper is to explore differences in simulation of complex manufacturing systems and healthcare systems as a first step to improving simulation of healthcare systems. We are interested in identifying obstacles and challenges for simulation application in healthcare systems. We focus on applications of discrete-event simulation to mainly address operational questions. We describe the manufacturing and healthcare domains, review the relevant literature, describe our research approach, point out the major structural differences in these domains, identify how these structural differences lead to modelling differences, and highlight some challenges in simulating healthcare systems.

## 1 Introduction

Discrete-Event Simulation (DES) has a long tradition in manufacturing. A large number of researchers have worked on this topic for decades. DES is used to design, understand, and improve manufacturing systems. Successful applications exist for different types of manufacturing systems (Negalban and Smith 2014). While there are still grand challenges (Fowler and Rose 2004), DES is in a relative mature state in the complex manufacturing domain.

Healthcare is now the world's largest industry. Compared to manufacturing, a smaller number of researchers have worked in simulation for healthcare for a long time (Günal and Pidd 2010 and Jacobson et al. 2013 and the references therein). On the one hand, there is an increasing interest in healthcare topics in the last decade. On the other hand, it seems that in contrast to the complex manufacturing domain, simulation techniques have not found their full way to hospitals very often (Kirchhof

and Meseth 2012 for the results of a survey dealing with applications of simulation in German healthcare institutions).

The main goal of the present paper is to explore differences in simulation of complex manufacturing systems and healthcare systems as a first step to improving simulation of healthcare systems. We are especially interested in identifying obstacles and challenges for simulation application in healthcare systems. Due to space limitations, we focus on applications of DES to address operational questions.

This paper is organised as follows. In Section 2, we briefly describe the two different domains. We then discuss related literature with respect to comparing DES techniques for these domains and describe our research approach. The comparison of the two domains is performed in Section 3, and major differences in modelling of the two domains are identified in Section 4. Specific challenges for simulating healthcare systems are discussed in Section 5. Finally, we provide some concluding remarks and mention a couple of areas for future research.

# 2 Problem Setting and Related Work

## 2.1 Domain Description

### *2.1.1 Complex Manufacturing Systems*

Manufacturing systems have the purpose to produce goods. A manufacturing system consists of a base system and an information system. The system components of the base system are used to transform raw materials and intermediate products into final products. The information system and the corresponding planning and control processes are responsible for making planning and control decisions related to the production of goods.

The base system of a manufacturing system can be further divided into a job processing system and a material flow system. The job processing system contains all the system components that allow for value-added processing of working objects. Jobs are the most important working objects (i.e. entities). They are formed based on customer orders or forecast information. They are released into the manufacturing system after formation. Capacity for processing is offered by the system components of the job processing system. The job processing system is formed by resources including machines, operators, and auxiliary resources. The facilities that are used to store, transport, and supply raw material, working objects, and auxiliary resources form the material flow system. The base process in a manufacturing system is responsible for the usage of system components of the base system by working objects. The base process is specified by process flows, also called routes (sequences of operations), and a given set of working objects. A set of possible machines is assigned to each operation within a route.

Complex manufacturing systems are characterised by a number of specific process restrictions including a large number of products and routes that involve a large number of process steps. Resource flexibility is also typical, i.e., machines that offer the same functionality (parallel machines), are organised in machine groups. In addition, sequence-dependent setup times and batch processing machines exist. A batch is a set of jobs that are processed together at the same time on a single

machine. Frequent disturbances like machine breakdowns are also common. Semiconductor wafer fabrication facilities can be considered as complex manufacturing systems (Mönch et al. 2013).

### 2.1.2 Healthcare Systems

Healthcare systems deal with health diagnosis, treatment, and prevention of disease, illness, and injury of human beings. Similar to a manufacturing system, a healthcare system can be divided into a base system and an information system. The system components of the base system are given by system components that are responsible to provide care to patients. Patients are the working objects in healthcare systems. In contrast to jobs in manufacturing, patients have preferences with respect to be served by specific individuals, to be seen at a specific place, and to be seen at a specific time (Hall 2012). The base system can be differentiated in a subsystem that tries to control and improve the health status of patients, while there is another subsystem that deals with the transport of patients, supplies, and implantable devices, and organs. The base process in a healthcare system is responsible for the usage of resources by (or for) patients. It is represented by a sequence of subtasks that specify how to use the resources in specific situations by a set of patients. Care is provided through the provision of resources that have to meet the needs of the patients. Important resources in healthcare include doctors and other health professionals, different type of rooms, equipment, instruments, medical supplies, and implantable devices and organs (Hall 2012).

The basic components of a very simple queuing system in manufacturing and healthcare are shown in figure 1. Examples of resources and working objects are given. The healthcare system-related entities are marked in bold. We see that the principle structure of complex manufacturing and healthcare systems are similar.

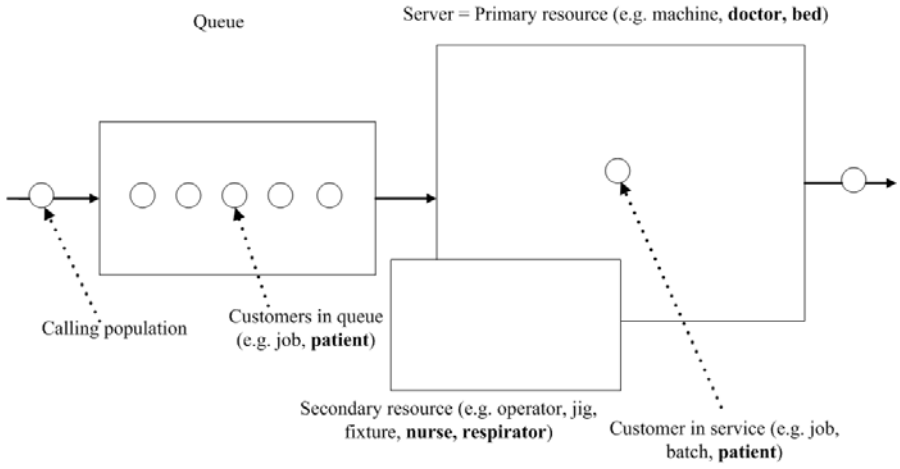



*Figure 1: Elements of a simple manufacturing or healthcare system*

## 2.2 Discussion of Related Literature

Next, we discuss related work with respect to DES approaches in healthcare and with respect to attempts to transfer knowledge from the manufacturing to the healthcare domain. There are many papers that deal with specific applications of DES in healthcare systems. Several recent surveys related to applications of discrete-event simulation in healthcare exist. Günal and Pidd (2010) discuss specific healthcare applications that are supported by simulation. Jacobson et al. (2012) provide a taxonomy of literature related to DES over the past 30 years.

Some researchers deal with transferring knowledge from other domains to healthcare. The possibility to transfer methods from job shop scheduling to appointment scheduling in healthcare is discussed by McKay (2011). Naseer et al. (2009) discuss the possibility to transfer knowledge from the defense domain to healthcare. The authors conclude that healthcare lags behind other domains, especially in terms of stakeholder engagement and in terms of implementation of simulation outcomes. Jahangirian et al. (2012) conduct a simulation-related literature analysis for the commerce, defense, and healthcare domains. Their findings are somewhat similar to those from Naseer et al. (2009). Verification and validation tasks are compared by means of case studies for the manufacturing and healthcare domains in Jenkins et al. (1998). The authors argue that healthcare is characterised by a level of complexity in handling resources in simulation models which does not often occur in the manufacturing domain. Kujlis et al. (2007) collect arguments for and against a possible transfer of methods from manufacturing to healthcare. The main findings of this paper are that methods from business and manufacturing are potentially applicable to healthcare, but their practical application is not straightforward because of the diversity and unique attitudes of the human actors involved in healthcare systems (Kujlis et al. 2007). Tako and Robinson (2012) conduct a survey to answer the question whether or not simulation in healthcare is different and more difficult than in other domains.

## 2.3 Research Approach

In the present paper, we are interested in comparing DES approaches for complex manufacturing and healthcare systems. The scope of the comparison is limited to discrete manufacturing systems because for this class of manufacturing systems we see some analogy to healthcare systems where the individual patient is modelled. We focus on applications of DES to mainly address operational questions. While we are aware of other simulation paradigms for healthcare systems including system dynamics, Monte Carlo simulation, agent-based simulation, or hybrid approaches (Brailsford et al. 2009), DES techniques are attractive for operational questions because this simulation paradigm allows for high-fidelity models. By analysing the comparison results, we strive for identifying obstacles and challenges for applying simulation in healthcare systems.

While there are several attempts made in the literature to discuss the differences of simulation approaches for the manufacturing and healthcare domain (see Subsection 2.2), to the best of our knowledge a clear comparison of the two domains based on analysing the differences in the underlying base systems and processes is missing.

# 3 Comparison of Complex Manufacturing and Healthcare Systems

## 3.1 Criteria for Comparison

We derive criteria for comparison from the domain description in Subsection 2.1. Therefore, we compare the base system and process and the related planning and control processes. This approach is reasonable since the abstract structure of the different subsystems and processes is somewhat similar for complex manufacturing systems and for healthcare systems. The specific criteria are shown in table 1. Note that we do not further differentiate between the information system and the corresponding information process. For comparison purposes, we are more interested in the related planning and control processes that describe under which circumstances specific control actions are carried out, while the information system that contains human decision-makers and application systems is responsible for providing planning and control instructions.

*Table 1: Criteria for comparing complex manufacturing and healthcare systems*

| Subsystem/Subprocess | Criterion |
|---|---|
| Base system | resource types |
| | offered capacity |
| | reliability and maintenance policies and procedures |
| | emergency events, e.g. resource breakdowns |
| Base process | working objects, i.e. moving entities |
| | product vs. service structure and the resulting routes vs. sequences of subtasks |
| | nature of demand |
| | consumed capacity |
| | importance and impact of cycle time |
| | emergency events, e.g. patient emergencies |
| Information system/process | decision-making procedures |

## 3.2 Results of the Comparison

### *3.2.1 Base System*

We start by describing the different resource types as the main ingredients of the base system. In complex manufacturing systems, we have consumable resources, i.e. raw material, and renewable resources including machines and operators. Resources are used for job processing and material handling. There are auxiliary resources like photolithography masks in semiconductor manufacturing or jigs in mixed-model assembly lines in the aerospace industry. There is often a high degree of automation in complex manufacturing systems. Therefore, operators are often less important.

In healthcare systems, resources are responsible for patient treatment and transportation tasks. Important resources are different types of rooms, for instance, for examination, procedures, surgeries, diagnostic tests and recuperation. Equipment, e.g., a CT scanner, is either permanently or temporarily assigned to an operating room or an examination room. Instruments form another important class of resources since they support medical procedures. Examples of instruments include thermometers and surgical instruments. Another class of resources is given by human resources such as doctors and different types of nurses. The human resources often have multi-tasking capabilities in time and space. For instance, a nurse might be responsible for several patients at the same time even though the patients stay in different rooms. Doctors and nurses perform direct patient care activities when they treat a patient. However, at the same time, indirect activities interrupt direct activities. Examples of indirect activities are given by writing or reading the file of a patient, discussing the status of a patient with a colleague or the family of the patient, and cleaning the operating room and supplying it with new medical supplies and instruments. In contrast to manufacturing, healthcare personnel have some independency that allows them to choose the next task to be performed on their own.

Supplies including medications and blood, are items that are consumed when care is delivered. Implantable devises and organs, for instance a liver or a kidney, are required for certain medical procedures. The implantation procedure cannot start until the device or organ is available.

The different resources in healthcare systems have to be often strongly synchronised in such a way that the requested resources are available when the patients need them. This includes the right amount and type of resources. For example, most surgeries cannot start if an anesthesia doctor or a surgical nurse is unavailable. The different types of resources found in both domains are summarised in table 2.

*Table 2: Resources in complex manufacturing and healthcare systems*

| Renewable resource | Manufacturing | Healthcare |
|---|---|---|
| Primary | Machines (stationary) | medical doctors (mobile), operating rooms (stationary), equipment and instruments (mobile or stationary) |
| Secondary | Operators (mobile), auxiliary resources (often mobile) | anesthesia doctors (mobile), various types of nurses (mobile), equipment (mobile or stationary) |
| Tertiary | maintenance staff (mobile) | N/A |
| Consumable resource | raw materials | medical supplies, implantable devices, organs |

Next we discuss the offered capacity in both domains. The capacity of resources in manufacturing systems is given in parts (jobs) per hour. Usually, capacity modelling is difficult in manufacturing (Elmaghraby 2011). The rate when the resources are available is variable since machine breakdowns occur quite often. In healthcare,

capacity is given in man hours available. The rate when resources are available is generally relatively constant.

The reliability of machines in complex manufacturing is often rather low. Machine-related emergency events, i.e. machine breakdowns, occur quite often. Therefore, various kinds of preventive maintenance policies exist in manufacturing plants. The maintenance activities consume scarce capacity. In healthcare systems, however, the reliability of resources is usually much higher. Equipment breakdowns are generally not a major source of difficulties. The preventive maintenance policies for equipment are different compared to manufacturing. The related activities usually do not reduce the capacity because they are performed outside regular working hours.

### 3.2.2 Base Process and Decision-making

Jobs, also called parts, are the working objects in complex manufacturing systems. Important attributes include the priority of the job, its size, and its due date. Quality-related attributes also occur quite often. The specification of how to manufacture a part, i.e. its route, is given by the corresponding product specification. The manufacturing of a specific product often requires assembly activities as described by the bill of materials (BOM). Assembly activities require a synchronisation of subparts.

In healthcare systems, however, a patient with a medical condition is the working object. The required speciality is a result of this medical condition. There are different types of due dates for a patient. We find patients with regular, urgent, and emergency due dates. Since the medical condition of a patient can change over time, certain attribute values like the temperature might also vary.

A product is well-defined through a set of BOMs and routes. It is relative static, i.e., the product specification does not change often. However, we find sequence, process, and resource (machine) flexibility to manufacture a product. The amount of operations to be performed is often large. For instance, up to 800 process steps are required in semiconductor manufacturing to produce a single integrated circuit. The manufacturing process has to follow the product specification closely.

In healthcare, we do not find a static product description. Each medical condition (there can be more than one medical condition for a single patient) requires a specific service (or set of services). Each service might require subservices. On the one hand, the sequence of these subservices is often unknown in advance. On the other hand, the number of subservices is rather small compared to complex manufacturing. The offered flexibility in healthcare is much smaller than in manufacturing. We find almost no process flexibility and only some sequence and resource flexibility.

The demand in manufacturing is often levelled even though it might be unpredictable. However, because of order release mechanisms, inventory, and backlog handling strategies, we know with more certainty which requests will arrive on the shop floor. In contrast to manufacturing, the demand is unlevelled in healthcare. In the long term it may follow a specific pattern, for example, it might depend on the day of the week or it might be based on the population size. However, in the short term demand in healthcare is highly unpredictable. Inventory strategies through preproduction are almost impossible. The demand situation in emergency rooms is quite different from the one in outpatient clinics. Therefore, the character of the medical facility or the hospital unit is important to describe the nature of demand.

From a consumed capacity point of view, in manufacturing systems we observe relatively low arrival variability (see the demand discussion above). The amount of consumption generally has a low or medium uncertainty. On the other hand, we find rather high arrival variability in healthcare systems. This arrival variability goes along with a highly uncertain amount of consumption (e.g. two patients with the same medical condition may require a significantly different length of stay).

In manufacturing, cycle time is defined as the amount of time elapsing between work being released into the manufacturing system and its emerging as finished product that can be used to meet demand. In healthcare, the counterpart of the cycle time is the time span between the arrival of a patient and his or her departure. The amount of cycle time ranges from days to several weeks in complex manufacturing, while in healthcare a cycle time of minutes, hours, or days is more typical.

Emergency events like emergency patients occur quite often in healthcare systems. High-priority jobs, so called hot or rocket jobs, are common in many manufacturing settings. But they are not really a counterpart to emergency patients since their appearance is less uncertain and are not a matter of life and death.

Operational decisions in complex manufacturing are often heavily influenced by previous decisions because of the complex product structure and the rather long cycle times. The impact of previous decisions is much smaller in healthcare systems. Subproblems based on time decomposition, for example appointment scheduling for a single day, are not heavily coupled.

## 4 Identification of Major Differences in Modelling

The following modelling differences are identified based on the results of the comparison performed in Section 3:

1. As Jenkins et al. (1998) point out “the health sector is characterized by a level of complexity in handling ‘resources’ (as understood in simulation modelling) which is not frequently found in the manufacturing sector”. While in manufacturing activities can be modelled as happening at discrete points of time and the machines are often the most important resources, in healthcare the modelling of secondary resources is generally more important. Multi-tasking abilities often need to be modelled, e.g. nurses may monitor many patients at the same time. While it is possible to crudely model these activities with a discrete modelling paradigm, it may not be adequate to capture their impact for a high fidelity model.
2. The modelling of human behavior is generally more important in healthcare compared to complex manufacturing. Three things make the modelling of human behavior in healthcare particularly challenging: 1) many activities involve multiple humans who act independently (cf. Lim et al., 2013); 2) healthcare providers are involved in many indirect activities that are typically not modelled resulting in underestimation of patient waiting times (cf. Rohleder et al., 2011); and 3) patients sometimes make irrational make choices that impact their future health (cf. Brailsford and Schmidt, 2003).
3. The dynamic nature and composition of services and subservices in healthcare is very different from the rather static product and route structure in most manufacturing systems. In the survey by Tako and Robinson (2012), 66% of the

respondents indicated that they felt there was “less evident structure” compared to other domains.
4. For the working objects in healthcare (i.e. the patients), the modelling of entity attributes is generally more important and more difficult because they are not known ahead of time and they might change over time (e.g. the patient temperature).
5. Unexpected or emergency events have a different sense of urgency in manufacturing and healthcare settings (life or death) and may require different treatments.

## 5 Challenges for Healthcare Simulation

The following challenges exist for modelling and simulation of healthcare systems:

1. The modelling of primary and secondary resources needs to be improved in simulation models for healthcare systems. We believe that hybrids of continuous, discrete-event, and agent-based simulations are a fruitful avenue for further exploration.
2. The modelling of human behavior needs to be better incorporated in simulation models. Some initial steps towards this goal are described by Brailsford and Schmidt (2003). One way to reach this goal might be the usage of agent-based simulation techniques.
3. Compared to complex manufacturing, there is a larger need to model the stochastic behavior of resources and working objects.
4. In contrast to manufacturing, healthcare simulation models are generally smaller in spatial and time scope. At the same time, the more complicated service structure of healthcare is a serious modelling issue.

## 6 Conclusions and Future Research

In this paper, we discussed the commonalities and differences in simulating complex manufacturing and healthcare systems. By differentiating between base system and process and information system, we were able to derive criteria to conduct such a comparison. Our main findings were that there are a couple of important differences. The modelling of resources is generally more sophisticated in healthcare simulations. The composition of services in healthcare is quite different from the more static product structure and route definition in complex manufacturing systems. The simulation models in healthcare are often smaller in size and need more details compared to the large-scale simulation models in complex manufacturing. This is not to say that healthcare systems cannot be adequately modelled using DES constructs very similar to those used for manufacturing systems for some applications, but refinement of these techniques or other approaches may be needed to achieve high fidelity models.

There are some directions for future work. First of all, we are interested in completing the discussion in this paper by considering simulation paradigms different from DES including system dynamics and agent-based simulation. We propose to look at simulation applications at a more strategic level rather than focusing only on the operational level. Secondly, while in the present paper simulation is primarily used to represent the base system and process of the underlying system, it seems

interesting to compare the usage of simulation for decision-making, i.e. simulation is part of the planning and control system.

## References

Brailsford, S.C.; Schmidt, B.: Towards incorporating human behaviour in models of health care systems: An approach using discrete event simulation. European Journal of Operational Research 150 (2003) 1, pp. 19–31.

Brailsford, S.C.; Harper, P.R.; Patel, B.; Pitt, M.: An analysis of the academic literature on simulation and modelling in health care. Journal of Simulation 3 (2009) 3, pp. 130–140.

Elmaghraby, S.E.: Production Capacity: Its bases, functions, and measurement. In Kempf, K., Keskinocak, P., Uzsoy, R. (eds). Planning Production and Inventories in the Extended Enterprise. A State of the Art Handbook, Volume 1. New York: Springer 2011, pp. 119-166.

Fowler, J.W.; Rose, O.: Grand challenges in modeling and simulation of complex manufacturing systems. Simulation – Transactions of the Society for Modeling and Simulation International 80 (2004) 9, pp. 469-476.

Günal, M.M.; Pidd, M.: Discrete event simulation for performance modelling in health care: a review of the literature. Journal of Simulation 4 (2010) 1, pp. 42-51.

Jacobson, S.H.; Hall, S.N.; Swisher, J.R.: Discrete-event simulation in health care systems. In Hall, R. (ed.): Patient Flow: Reducing Delay in Healthcare Delivery. New York: Springer 2013, pp. 211-252.

Jahangirian, M.; Naseer, A.; Stergioulas, L.; Young, T.; Eldabi, T.; Brailsford, S.; Patel, B.; Harper, P.: Simulation in health-care: Lessons from other sectors. International Journal of Operational Research 12 (2012) 1, pp. 45–55, (2012).

Jenkins, R.; Deshpande, Y; Davison, G.: Verification and validation and complex environments: A study in service sector. In: Medeiros, D.J.; Watson E.F.; Carson, J.S.; Manivannan, M.S. (Eds.): Proceedings of the 1998 Winter Simulation Conference, Washington (DC), 13.-16. December 1998, pp. 1433-1440.

Hall, R.: Matching healthcare resources and patient needs. In Hall, R. (ed.): Handbook of Healthcare System Scheduling. New York: Springer 2012, pp. 1-9.

Kirchhof, P.; Meseth, N.: A survey on the use of simulation in German healthcare. In: Laroque, C.; Himmelspach, J.; Pasupathy, R.; Rose, O.; Uhrmacher, A. M. (eds.): Proceedings of the 2012 Winter Simulation Conference, Berlin (Germany), 9.-12. December 2012, pp. 1082-1091.

Kujlis, J.; Paul, R. J.; Stergioulas, L. K.: Can health care benefit from modeling and simulation methods in the same way as business and manufacturing has? In: Henderson, S.G.; Biller, B.; Hsieh, M.-H.; Shortle, J.; Tew, J.D.; Barton, R.R. (eds.) Proceedings of the 2007 Winter Simulation Conference, Washington (DC), 9-12 December 2007, 1449-1453.

Lim, E.M.; Worster, A.; Goeree, R.; Tarride, J.E.: Simulating an emergency department: the importance of modeling the interactions between physicians and delegates in a discrete event simulation. BMC Medical Informatics and Decision Making 13 (2013), pp. 59-69.

McKay, K.N.: Inter-domain translational research on planning and scheduling: Operating room vs. job shops. International Journal on Planning and Scheduling, 1 (2011) 1/2, pp. 42-57.

Mönch, L.; Fowler, J.W.; Mason, S.J.: Production planning and control for semiconductor wafer fabrication facilities: Modeling, analysis, and systems. New York: Springer 2013.

Naseer, A.; Eldabi, T.; Jahangirian, M.: Cross-sector analysis of simulation methods: a survey of defense and healthcare. Transforming Government: People, Process and Policy 3 (2009) 2, pp. 181 – 189.

Negalban, A.; Smith, J.S.: Simulation for manufacturing system design and operation: Literature review and analysis. Journal of Manufacturing Systems 33 (2014) 2, pp. 241-261.

Rohleder, T.R.; Lewkonia, P.; Bischak, D.P.; Duffy, P.; Hendijani, R.: Using simulation modeling to improve patient flow at an outpatient orthopaedic clinic. Health Care Management Sciences 14 (2011) 2, pp. 135-145.

Tako, A. A.; Robinson, S. Is simulation in health different and is it more difficult? In: Tjahjono, B.; Heavey, C.; Onggo, S.; van der Zee, D.-J. (eds.) Proceedings of the Operational Research Society Simulation Workshop 2012, Worcestershire (UK), 27.-28. March 2012, pp. 93-101.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Einsatzmöglichkeiten der Simulation in einem cyber-physischen Produktionssystem

## *Application of Simulation in Cyber-physical Production Systems*

Christina Reuter, Felix Brambring, Christiane Müller,
Werkzeugmaschinenlabor (WZL), RWTH Aachen, Aachen (Germany),
c.reuter@wzl.rwth-aachen.de, f.brambring@wzl.rwth-aachen.de,
ch.müller@wzl.rwth-aachen.de

Abstract: In a dynamic and volatile market environment, excellent production planning and control (PPC) processes enable manufacturing companies to establish a competitive advantage since short lead times and a high adherence to promised delivery dates are valued highly by B2B and B2C customers alike. However, especially in small-scale series production with highly customised products, complex work plans and material flows, manufacturing companies struggle to excel in logistic performance. In the future, cyper-physical production systems (CPPS) will provide a higher transparency in manufacturing and production and facilitate PPC processes. The approach presented in this paper, addresses the application of simulation as the basis of a cyber-physical support system for PPC processes which is being developed as part of the research project ProSense.

## 1 Einführung

In volatilen Märkten, die durch zunehmende Konkurrenz aus aufstrebenden Schwellenländern geprägt sind, stellt eine hohe logistische Leistungsfähigkeit einen Wettbewerbsvorteil dar. Für 67% aller Unternehmen aus dem Maschinen- und Anlagenbau gilt die Erreichung einer hohen Liefertermintreue daher als die führende logistische Zielgröße. Eine hohe Liefertermintreue wird jedoch insbesondere in einer komplexen Einzel- und Kleinserienfertigung mit stark kundenindividuell geprägten Erzeugnissen regelmäßig nicht in zufriedenstellendem Maße erreicht (Schuh und Stich 2013). Offensichtlich sind die Prozesse der Produktionsplanung und -steuerung (PSS) der Komplexität und Volatilität der Produktionsbedingungen nicht gewachsen, da ihnen bei der Erreichung einer hohen Termintreue trotz kurzfristiger Änderungswünsche seitens des Kunden und anderer unvorhersehbarer Störungen eine herausragende Bedeutung zukommt (Schuh und Gierth 2006, S. 3). Je nach benötigtem Funktionsumfang und Detaillierungsgrad werden für die PPS heutzutage meist Enterprise Resource Planning (ERP) Systeme, Manufacturing Execution Systeme (MES) oder Advanced Planning and Scheduling (APS) Systeme eingesetzt

(Schuh und Stich 2013). Dass diese IT-Systeme eine hohe logistische Leistungsfähigkeit nicht garantieren können, hat im Wesentlichen zwei Gründe: Zum einen stehen für die Planung und Steuerung notwendigen Daten nicht in ausreichender Granularität und Qualität zur Verfügung (Schuh et al. 2013b, S. 347f.). Zum Anderen benötigt es zu viel Zeit, aus den vorliegenden Daten die entscheidenden Analysen und Rückschlüsse zu ziehen, um angemessen handeln zu können (Schuh et al. 2011). Aufgrund der damit entstehenden Intransparenz bleibt trotz umfangreichem Einsatz von IT ein Großteil der Entscheidungsvorbereitung dem Bauchgefühl und impliziten Wissen der verantwortlichen Mitarbeiter überlassen (Schuh 2013a, S. 651f.).

Ansätze, die im Zuge einer sich anbahnenden Industrie 4.0 entwickelt werden, versprechen eine bessere Beherrschung der Produktionssteuerung. In cyber-physischen Produktionssystemen (CPPS), die als Basis für die sog. vierte industrielle Revolution dienen, werden Maschinen, Produkte, intelligente Sensorik und Betriebsmittel in der Lage sein, eigenständig Informationen auszutauschen, Vorschläge für zu ergreifende Maßnahmen anzubieten sowie im begrenzten Rahmen Aktionen selbstständig auszulösen und sich somit teilautonom zu steuern (Kagerman et al. 2013, S. 5). Im Rahmen des Forschungsprojekts ProSense (Förderzeichen 02PJ2495) wird dafür u. a. intelligente Sensorik entwickelt, durch die bislang nicht verfügbare Daten über Materialflüsse generiert und in einem kybernetischen Unterstützungssystem mit bereits verfügbaren Daten zu hochwertigen Informationen verknüpft werden können. Das kybernetische Unterstützungssystem wird dabei als iPad-App realisiert, das den Fertigungssteuerer im Sinne einer Kybernetik zweiter Ordnung als Beobachter und Beeinflusser der Fertigungsprozesse in die Lage versetzt, Steuerungseingriffe in das Produktionssystem zu verstehen und zu hinterfragen. Die Kybernetik zweiter Ordnung geht davon aus, dass der Beobachter Teil des beobachtenden Systems ist und es schon allein dadurch beeinflusst und somit in der Modellierung berücksichtigt werden muss (Foerster 1995). Dazu werden mögliche Handlungsmaßnahmen in einem automatisch generierten Simulationsmodell analysiert, vorbewertet und ihre Auswirkungen für den Fertigungssteuerer intuitiv erfassbar visualisiert. Im Sinne einer hochauflösenden Produktionssteuerung werden feingranulare Daten für die Abbildung des derzeitigen Fertigungsstatus genutzt und die Ergebnisse über mehrere Aggregationsebenen hinweg transparent dargestellt.

Die Nutzung automatisch generierter Simulationsmodelle stellt damit eine zentrale Methode einer cyber-physischen Produktionssteuerung dar. Übergeordnetes Ziel ist es, dem Fertigungssteuerer eine höhere Transparenz über den aktuellen Fertigungsstatus und die Auswirkungen verschiedener Steuerungseingriffe zu bieten und somit die logistische Leistungsfähigkeit produzierender Unternehmen nachhaltig zu erhöhen.

Im Folgenden wird zunächst der Stand der Technik im Hinblick auf den Einsatz automatisch generierter Simulationsmodelle und die Einbindung von sensorgenerierten Rückmeldedaten dargestellt. Es folgt die Beschreibung eines Ansatzes und dessen prototypischer Umsetzung zur Nutzung von Simulation als Basis für eine cyber-physische Produktionssteuerung, die im Rahmen des Forschungsprojekts ProSense (Förderkennzeichen 02PJ2490) entwickelt und erprobt wird. Der Beitrag schließt mit einer kurzen Zusammenfassung und einem Ausblick auf zukünftige Forschungsschwerpunkte.

## 2 Stand der Technik

Für die Entwicklung eines kybernetischen Unterstützungssystems für eine hochauflösende, cyber-physische Produktionssteuerung auf Basis von Simulationen sind im Wesentlichen drei Forschungsfelder relevant, deren Stand der Technik im Folgenden dargestellt werden soll:

1. Erzeugung hochauflösender Bewegungsdaten in der Produktion mittels Sensorik sowie die Speicherung historischer Planungsstände
2. Automatisierte Generierung von Simulationsmodellen als Basis für die Simulation geeigneter Handlungsalternativen
3. Grafische Oberflächen zur Darstellung der aktuellen Planung und zur Auswahl möglicher Steuerungseingriffe

Produzierende Unternehmen mit Einzel- und Kleinserienfertigung setzen für ihre Produktionsplanung ERP- oder APS-Systeme ein, in denen der aktuelle Produktionsplan erstellt und gespeichert wird. Diese Systeme sind jedoch nicht dafür ausgelegt, historische Plandaten für einen späteren Abgleich zu speichern. Vielmehr wird mit jedem neuen Planungslauf, der meist einmal täglich über Nacht erfolgt, der aktuelle Produktionsplan überschrieben. Erfolgt keine manuelle Sicherung des täglich aktualisierten Produktionsplans fehlt die Möglichkeit zur Überprüfung der Planungsqualität (Hauptvogel 2015, S.7). Zur Ermittlung der logistischen Leistungsfähigkeit sammeln produzierende Unternehmen eine Vielzahl von Daten während der Produktionsprozesse. Für die Erzeugung dieser Rückmeldedaten werden verschiedene Technologien für die Erfassung von Bewegungsdaten in der Produktion eingesetzt. Dazu zählen neben klassischen Terminals zur Betriebsdatenerfassung (BDE) und die als Maschinendatenerfassung (MDE) bezeichnete Erzeugung von Daten direkt aus der Maschinensteuerung moderner CNC-Maschinen, vor allem nach wie vor die manuelle, schriftliche Dokumentation der Daten auf handschriftlich zu füllenden Vorlagen oder in Tabellenkalkulationen wie Microsoft Excel. Technologien wie Radio-frequency identification (RFID) oder auf Basis von optischer Sensorik haben bislang noch keine flächendeckende Verbreitung gefunden (Schuh und Stich 2013, S. 20). Die für die Materialflusssimulation einer werkstattorientierten Einzel- und Kleinserienfertigung relevantesten Daten betreffen mindestens Start- und Endzeitpunkte der einzelnen Arbeitsschritte von Fertigungsaufträgen, die jeweiligen Rüst- und Bearbeitungszeiten auf den eingesetzten Arbeitssystemen sowie die produzierten Mengen bzw. Losgrößen. Je nach Komplexität des Simulationsmodells können Daten zum Personaleinsatz relevant sein (Perzyk et al. 2008, S. 1503f.). Aufgrund der verschiedenen eingesetzten Technologien und dem Fehlen eines führenden IT-Systems sind diese Daten meist nicht in einer zentralen Datenbank gespeichert, sondern sind mit unterschiedlichen Aktualisierungsfrequenzen über mehrere IT-Systeme verteilt (Loos 1999, S 230). Es ist offensichtlich, dass eine in dieser Weise aufgebaute IT-Infrastruktur dem Paradigma einer „Single Source of Truth“ nicht entspricht und eine schnelle Datenaggregation im Sinne einer hochauflösenden Produktionssteuerung nur schwer zu realisieren ist (Kagermann et al. 2013, S. 5). Im Rahmen von ProSense werden daher sowohl neue Einsatzfelder von Sensorik zur Erhöhung der Datengranularität als auch eine Datenbank zur Speicherung aller notwendigen Rückmeldedaten sowie historischen Planungsständen entwickelt.

Für eine hochauflösende Produktionssteuerung ist eine schnelle Reaktionsfähigkeit auf Planabweichungen von entscheidender Bedeutung. Ansätze zur automatisierten Generierung von Simulationsmodellen ermöglichen eine stärkere Einbindung von Simulationsergebnissen in die Entscheidungsfindung der Produktionssteuerung, da sie den notwendigen Einsatz von Expertenwissen reduziert und damit eine Zeitersparnis in der Modellgenerierung ermöglicht (Schuh et al. 2013a, S. 2). Für eine automatisierte Generierung des Simulationsmodells sollen Daten über spezielle Schnittstellen aus den datenführenden IT-Systemen herangezogen werden, ohne dass manuelle Nacharbeiten notwendig sind (Straßburger et al. 2010, S.37 f.). Besonders die betriebsbegleitende Simulation, mittels derer die kurzfristige Steuerung eines zuvor erstellten Produktionsplans ermöglicht werden soll, kann von einer automatisierten Modellgenerierung profitieren, da die notwendige Aktualisierung des Simulationsmodells aufwendig und entsprechend kostenintensiv ist (Schuh et al. 2013a, S. 2). Eine zentrale Herausforderung in der automatisierten Modellgenerierung ist die Sicherstellung einer ausreichenden Datenqualität, d. h. Vollständigkeit, Korrektheit und Konsistenz der verwendeten Daten müssen anhand von Plausibilitätsprüfungen validiert werden (Rabe et al. 2008, S. 132). Die von verschiedenen Autoren aufgezeigten Ansätze zur automatisierten Modellgenerierung können u. a. aufgrund der Mängel der Datenqualität nicht vollkommen ohne manuelle Nacharbeiten angewendet werden (vgl. Bergmann 2011; Jensen 2007; Schuh et al. 2013a). Im Rahmen von ProSense wird das von Schuh (Schuh et al. 2013) entwickelte Vorgehen zur automatisierten Generierung von Simulationsmodellen weiterentwickelt, um eine automatisierte Überprüfung der Datenqualität gewährleisten zu können.

Aufgrund der komplexen Planungs- und Steuerungsentscheidungen in hochdynamischen Produktionsumgebungen ist der Mensch als Entscheider in der PPS unverzichtbar und soll diese Stellung als letzte Entscheidungsinstanz auch im Zeitalter von Industrie 4.0 weiter innehaben (Kagermann et al. 2013, S. 5). Da bereits in der heutigen PPS die Datenerfassung, -aufbereitung und -nutzung weitestgehend digital erfolgt, kommt den Benutzeroberflächen von ERP-, MES- und APS-Systemen als wichtigste Schnittstelle zum Menschen eine herausragende Bedeutung zu. Neben einfachen tabellarischen Darstellungen haben sich vor allem Gantt-Charts als Visualisierungsmethode für die Simulationsergebnisse von APS-Systemen durchgesetzt, die u. a. die Ermittlung des kritischen Pfads eines Auftrags erleichtert. Die Gestaltung der grafischen Oberflächen dieser Systeme ist jedoch meist mit einer Vielzahl von Informationen überladen, wenig intuitiv und benutzerfreundlich (Schuh und Stich 2013, S. 46). Im Rahmen von ProSense wird daher an innovativen Visualisierungsmöglichkeiten gearbeitet, welche die zielgerichtete Bereitstellung aller relevanten Informationen auf intuitive Weise ermöglichen.

## 3 Simulation in der Produktionssteuerung cyber-physischer Produktionssysteme

Das Modell einer hochauflösenden Produktionssteuerung in cyber-physischen Produktionssystemen, wie es im Rahmen des Forschungsprojekts ProSense entwickelt wurde, ist in Abbildung 1 dargestellt.

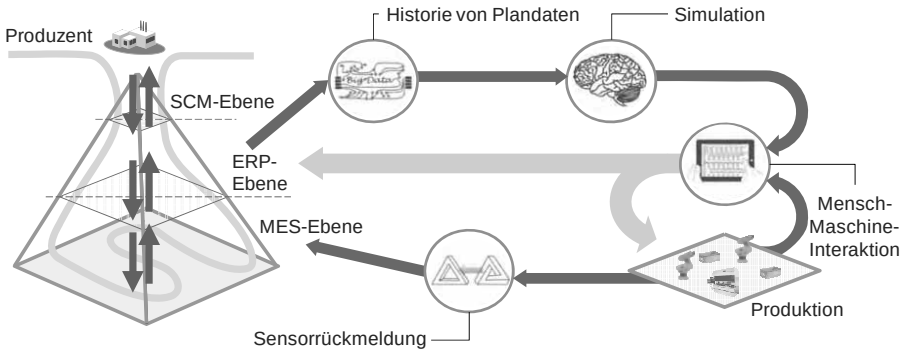



*Abbildung 1: Modell einer cyber-physischen Produktionssteuerung (Quelle: ProSense)*

Die Nutzung von automatisch generierten Simulationsmodellen stellt in diesem Modell eine zentrale Methode dar, die eine Verkürzung der Entscheidungsfindung in der Produktionssteuerung ermöglicht. Dazu werden die klassischen Quellen von Rückmeldedaten aus der Produktion wie BDE- und MDE-Systeme durch zusätzliche Sensorik, wie RFID-Sensoren, Laser- und Barcode- bzw. Data-Matrix-Scanner ergänzt. Diese zusätzlichen Datenquellen dienen u. a. zur feineren Auflösung der Übergangszeiten von Fertigungsaufträgen in Liegezeiten nach Bearbeitung, Transportzeiten und Liegezeiten vor der Bearbeitung, die erst dadurch in der Simulation berücksichtigt werden können. Alle generierten Rückmeldedaten werden in der Datenbank des datenführenden ERP-Systems zwischengespeichert. Eine speziell entwickelte Datenbankanwendung, der sog. Data-Collector, legt auf Basis dieser Daten eine Historie der Rückmeldedaten sowie der vergangenen und aktuellen Plandaten in standardisierter Form an und bildet damit die Schnittstelle zwischen ERP-System und der Simulationssoftware Plant Simulation der Siemens AG, in der das Simulationsmodell aufgebaut wird.

Zwei Tabellen des Data-Collectors sind beispielhaft in Tabelle 1 und Tabelle 2 angegeben.

*Tabelle 1: Fertigungsauftragsbezogene Plan- und Rückmeldedaten*

| Ferigungsauftrag | Artikelnummer | Beschreibung | Menge | Freigabetermin | Bedarfstermin | Ist-Start | Ist-Ende | Priorität |
|---|---|---|---|---|---|---|---|---|
| *String* | *String* | *String* | *Integer* | *Datetime* | *Datetime* | *Datetime* | *Datetime* | *Integer* |
| 1314770 | 4021 | Welle | 4 | 29.04.2014 | 07.02.2015 | 30.04.2014 | 30.09.2014 | 0 |
| 1315872 | 3033 | Schaufel | 65 | 05.05.2014 | 23.02.2015 | 08.05.2014 | 18.10.2014 | 1 |
| 1315915 | 6790 | Rad | 51 | 21.07.2014 | 08.02.2015 | 21.07.2014 | 30.11.2014 | 0 |

Die fertigungsauftragsbezogenen Daten in Tabelle 1 umfassen die eineindeutige Nummer des Fertigungsauftrags, die zugehörige Artikelnummer und -beschreibung, die Losgröße, den Freigabe- und Bedarfstermin sowie den tatsächlichen Start- und

Fertigstellungstermin nach erfolgter Bearbeitung des gesamten Auftrags. Über den Kenner „Priorität“ werden zudem Informationen über ggf. definierte Eilaufträge mitgeführt.

Tabelle 2 ergänzt die fertigungsauftragsbezogenen Daten um die in den Arbeitsplänen hinterlegten arbeitsschrittbezogenen Daten, womit für den Prozessschritt des Fertigungsauftrags alle Vorgänge mit Plan- und Istzeiten beschrieben sind.

*Tabelle 2: Arbeitspläne der Fertigungsaufträge*

| Auftrag | Vorgang | Rüstzeit | Hauptzeit | Arbeitsplatz | Geplanter Start | Geplantes Ende | Ist-Start | Ist-Ende |
|---|---|---|---|---|---|---|---|---|
| String | Integer | Float | Float | String | Datetime | Datetime | Datetime | Datetime |
| 1314770 | 10 | 1182 | 360 | 741 | 29.04.2014 | 30.04.2014 | 14.05.2014 | 16.05.2014 |
| 1314770 | 20 | 1200 | 883 | 72 | 04.05.2014 | 05.05.2014 | 18.05.2014 | 19.05.2014 |
| 1314770 | 30 | 0 | 364 | 677 | 15.06.2014 | 17.06.2014 | 20.07.2014 | 22.07.2014 |
| 1314770 | 40 | 1980 | 1080 | 669 | 08.09.2014 | 10.09.2014 | 29.08.2014 | 30.09.2014 |

Auf Basis dieser und weiterer produktionslogistischer Daten, wie das angewendete Schichtmodell, Maschinenäquivalenzen und Qualifikationsprofilen der Mitarbeiter kann das Simulationsmodell automatisch generiert werden. Dazu werden die Maschinen, das Produktionsprogramm inklusiver aller Arbeitspläne sowie alle benötigten Stammdaten aus der PostgreSQL Datenbank des Data-Collectors über eine definierte Schnittstelle in die Simulationsumgebung PlantSimulation übergeben. Mittels im Simulationsmodell hinterlegter, generischer Baukastenmodule wird ein Simulationsmodell automatisch erzeugt und die Parameter konfiguriert. Zur Validierung des Simulationsmodells werden im Anschluss automatisch Simulationsläufe durchgeführt, deren Ergebnisse anhand von Kennzahlen, wie Mittelwert und Standardabweichung des Umlaufbestands, Durchsatz pro Zeiteinheit und Maschinenauslastungen zu einem Vertrauenskoeffizienten aggregiert werden. Dieser Vertrauenskoeffizient erleichtert dem Benutzer den Vergleich zwischen Simulationsmodell und realem Basisdatensatz und dient damit als Ergänzung zum Bauchgefühl des Benutzers bei der Abschätzung der Validität der vorgeschlagenen Handlungsmaßnahmen, wodurch die Objektivität der Entscheidungsfindung erhöht wird. Bei einem zu niedrigen Vertrauenskoeffizienten sind nachträgliche manuelle Anpassungen im Simulationsmodell zu empfehlen, die aber unabhängig von dessen Wert immer möglich sind. Einige Anpassungen, wie die Schichtplanung, Freigabe- oder Reihenfolgeregeln können direkt über die Benutzeroberfläche der App vorgenommen werden.

Auf Basis des automatisch generierten Simulationsmodells werden anschließend mögliche Handlungsmaßnahmen simuliert, die zu einer Verbesserung in den produktionslogistischen Kennzahlen Termintreue, Durchlaufzeit, Auslastung und Umlaufbestand führen können. Zur Identifikation geeigneter Maßnahmen dient zum einen eine definierte Liste von zehn generischen Maßnahmen sowie zum Anderen die vier vordefinierten typischen Handlungsfelder der Produktionssteuerung zur Überwachung der maschinellen bzw. personellen Kapazität, der Eilauftragslage sowie der Auftragsfreigabe. Tabelle 3 gibt einen Überblick über die generischen Maßnahmen und die durch sie beeinflussbaren Handlungsfelder.

*Tabelle 3: Generische Maßnahmen mit ihrer Wirkung auf definierte Handlungsfelder*

| Maßnahme | Maßnahme in/ auf | Maschinelle Kapazität | Personelle Kapazität | Eilauftragslage | Auftragsfreigabe |
|---|---|---|---|---|---|
| Maschine substituieren | IT-System | X | | X | |
| Aufträge extern vergeben | IT-System | X | X | X | |
| Rüstzeiten minimieren (Reihenfolge anpassen) | Shopfloor | X | X | X | |
| Mehrmaschinenbedienung einführen | Shopfloor | | X | | |
| Mitarbeiter aus anderen Bereichen zuordnen | Shopfloor | | X | X | |
| Überstunden anordnen | Shopfloor | | X | X | |
| Zusatzschicht anordnen | Shopfloor | | X | X | |
| Prioritäten umverteilen | IT-System | | | X | |
| Mehr Aufträge freigeben | IT-System | | | | X |
| Weniger Aufträge freigeben | IT-System | | | | X |

Vor den jeweiligen Simulationsläufen muss eine Einschränkung der zu simulierenden Handlungsmaßnahmen durch ihre gezielte Parametrisierung erfolgen, um die Simulationszeit auf ein sinnvolles Maß zu reduzieren. So könnte z. B. bei einem drohenden Engpass durch eine kapazitive Überlastung die Auswirkungen von ein, zwei oder mehr Überstunden bzw. Zusatzschichten simuliert werden. Sollten sämtliche, parametrisierbaren Maßnahmen des Produktionssteuerers in den vier Handlungsfeldern in Kombination durchsimuliert werden, so kann bei einer angenommenen Simulationszeit von einer Minute pro Simulationslauf und einer zehnfachen Wiederholung eines jeden Simulationslaufs zur Glättung von statistischen Abwiechungen mit einer kumulierten Simulationszeit von mehreren Jahren gerechnet werden. Daher werden durch eine Vorfilterung Abschätzungen getroffen, welche Parametrisierung zu tatsächlichen Verbesserungen in den vier Handlungsfeldern führen können und nur diese Handlungsmaßnahmen für einen Simulationslauf freigegeben.

Die simulativ ermittelten Auswirkungen der so definierten Handlungsmaßnahmen auf die vier logistischen Zielgrößen werden mittels einer Zielfunktion erfasst. Die Aggregation der logistischen Ziele durch die Zielfunktion erlaubt dann eine auf die vier Handlungsfelder bezogene Priorisierung der Handlungsmaßnahmen. Wie in Abbildung 1 dargestellt, erhält der Fertigungssteuerer die priorisierten Maßnahmenvorschläge über eine innovative Benutzeroberfläche, die es ihm ermöglicht, die Auswirkungen der einzelnen Maßnahmen im Detail nachzuvollziehen. Er hat dann die Möglichkeit entweder neu parametrisierte Simulationsläufe zu starten oder die

entsprechenden Maßnahmen über Änderungen in den jeweiligen IT-Systemen bzw. durch Rücksprache mit den Mitarbeitern in der Produktion umzusetzen. Der Mensch bleibt auch in einer cyber-physischen Produktionssteuerung damit immer die letzte Entscheidungsinstanz. Durch das automatisch generierte Simulationsmodell, das auf einem zentralen Server läuft und die Priorisierung der Maßnahmen vorbereitet, kann sich der Fertigungssteuerer auf seine zentralen Aufgaben konzentrieren, ohne sich um die komplexe Datenaggregation und -auswertung kümmern zu müssen.

## 4 Prototypische Umsetzung und Anwendung

Das beschriebene Modell für eine cyber-physische Produktionssteuerung wurde im Rahmen des Forschungsprojekts ProSense in der Demonstrationsfabrik Aachen prototypisch umgesetzt. Die Demonstrationsfabrik Aachen GmbH ist Teil des Campus-Clusters Logistik der RWTH Aachen und stellt als Referenzfabrik für Industrie 4.0 marktfähige Produkte in Einzel- und Kleinserie her. Sie bietet zugleich die Möglichkeit, produktionssystematische und -logistische Fragestellungen in Experimenten in einer realen Produktionsumgebung zu erforschen (Schuh et al. 2014, S. 571).

In diesem Labormaßstab funktioniert die automatische Generierung des Simulationsmodells zuverlässig mit einem Validierungskoeffizienten von bis zu 90 %. Die von dem kybernetischen Unterstützungssystem vorgeschlagenen Handlungsmaßnahmen sind prinzipiell geeignet Verbesserungen in der logistischen Zielerreichung zu ermöglichen (Abb. 2).

*Abbildung 2: Prototypische Umsetzung des kybernetischen Unterstützungssystems mit angeschlossenem Simulationsmodell*

Die Produktionsstruktur der Demonstrationsfabrik mit sieben Maschinen, vier zu fertigenden Produktvarianten und einem relativ gerichteten Materialfluss unterscheidet sich jedoch noch deutlich von einer Werkstattfertigung mit komplexen Materialflüssen, die im Maschinen- und Anlagenbau häufig vorzufinden ist. Eine erste Implementierung des sich aus den Rückmeldedaten selbst generierenden Simulationsmodells bei einem Konsortialpartner des Forschungsprojekts erzielte dementsprechend lediglich einen Validierungskoeffizienten von 63 %. Diese signifikant niedrigere Validität ergibt sich aus Differenzen zwischen den Plan- und Ist-Zeiten der Arbeitsschritte, Blockrückmeldungen zum Schichtende und Abweichungen von vorgegebenen Arbeitsgangreihenfolge. Durch manuelle Bereinigung der Daten ist zwar ebenfalls eine Übereinstimmung in der Nähe von 90 % erzielbar gewesen, ohne Details an der Struktur des Simulationsmodells anpassen zu müssen. Aufgrund des dadurch jedoch weiterhin bestehenden hohen Aufwands kann dieses Vorgehen nur ein Zwischenschritt zu einer praxistauglichen Lösung einer cyberphysischen Produktionssteuerung im Sinne von Industrie 4.0 sein.

## 5 Zusammenfassung und Ausblick

Der in diesem Artikel beschriebene Ansatz einer cyber-physischen Produktionssteuerung beinhaltet als zentralen Baustein Simulationsmodelle, die auf Basis von Rückmeldedaten automatisch generiert werden können. Im Rahmen des Forschungsprojekts ProSense wurde dabei ein Modell einer cyberphysischen Produktionssteuerung in Zeiten von Industrie 4.0 entwickelt und in der Demonstrationsfabrik Aachen die notwendige Infrastruktur für eine prototypische Anwendung geschaffen, die im Labormaßstab funktioniert.

Einer praxistauglichen Anwendung der vorgeschlagenen Methode wird derzeit noch durch die noch nicht vollständig erreichbare Realitätstreue des automatisch generierten Simulationsmodells verhindert. Aktuell wird das Problem durch eine teilautomatische Generierung des Modells gelöst, indem das Grundmodell der Simulation an unternehmensspezifische Rahmenbedingungen angepasst wird. Zukünftige Forschungsschwerpunkte liegen daher insbesondere auf der Detaillierung des generischen Baukastens für die automatische Modellgenerierung, sodass das Simulationsmodell die komplexen und volatilen Produktionsstrukturen des Maschinen- und Anlagenbaus realitätstreuer abbilden kann.

## Danksagung

Die Entwicklung einer hochauflösenden Produktionssteuerung auf Basis kybernetischer Unterstützungssysteme und intelligenter Sensorik ist Gegenstand des Forschungs- und Entwicklungsprojekts „ProSense“, das mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) im Rahmenkonzept „Forschung für die Produktion von morgen“ sowie der Förderinitiative "Intelligente Vernetzung in der Produktion – Ein Beitrag zum Zukunftsprojekt "Industrie 4.0" (Förderkennzeichen 02PJ2490) gefördert und vom Projektträger Karlsruhe (PTKA) betreut wird (www.prosense.info).

## Literatur

Bergmann, S.: Automatische Generierung adaptiver und lernfähiger Modelle zur Simulation von Produktionssystemen. Dissertation, TU Ilmenau 2011.

Foerster H. v.: Cybernetics of Cybernetics: The control of control and the communication of communication, 2. Aufl. Minneapolis (MN): Future Systems 1995.

Hauptvogel, A.: Bewertung und Gestaltung von cyber-physischer Feinplanung. Aachen: Apprimus 2015.

Jensen, S.: Eine Methodik zur teilautomatisierten Generierung von Simulationsmodellen aus Produktionssystemen am Beispiel einer Job Shop Fertigung. Dissertation, Universität Kassel 2007.

Kagermann, H.; Wahlser, W.; Helbig, H. (Hrsg.): Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0 - Abschlussbericht des Arbeitskreises Industrie 4.0 2013.

Loos, P.: Grunddatenverwaltung und Betriebsdatenerfassung als Basis der Produktionsplanung und -steuerung. In: Corsten, H.; Friedl, B. (Hrsg.): Produktionscontrolling. München: Vahlen 1999, S. 227-252.

Perzyk, M.; Biernacki, R.; Kozlowski, J.: Data mining in manufacturing: Significance analysis of process parameters. Journal of Engineering Manufacture 222 (2008) 11, S. 1503-1516.

Schuh, G.; Gierth, A.: Einführung. In: Schuh, G. (Hrsg.): Produktionsplanung und -steuerung - Grundlagen, Gestaltung und Konzepte, 3. Aufl. Berlin: Springer 2006, S. 3-7.

Schuh, G.; Stich, V.: Produktion am Standort Deutschland - Ergebnisse der Untersuchung. Aachen, 2013, www.fir.rwth-aachen.de.

Schuh, G.; Lödding, H.; Stich, V.; Reuter, C.; Schmidt, O.; Potente, T. et al.: High Resolution Production Management. In: Brecher, C.; Klocke, F.; Schmitt, R.; Schuh, G. (Hrsg.): Wettbewerbsfaktor Produktionstechnik: Aachener Perspektiven. Aachener Werkzeugmaschinenkolloquium 2011. Herzogenrath: Shaker 2011, S. 61-80.

Schuh, G.; Potente, T.; Hauptvogel, A.; Thomas, C.; Müller, C.: Wertstromorientierte Konfiguration der Produktionssteuerung durch webbasierte Simulation. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik, Paderborn: HNI 2013a, S. 651 - 660.

Schuh, G.; Potente, T.; Thomas, C.; Brambring, F.: Approach for reducing data inconsistencies in production control. In: Zäh, M. (Hrsg.): Proceedings of the Fifth International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV), München 2013b, S. 347-351.

Schuh, G.; Gartzen, T.; Basse, F.: Scientific Management 2.0. In: WGP Congress 2014, Advanced Materials Research vol. 1018, S. 571 - 579.

Straßburger, S.; Bergmann, S.; Müller-Sommer, H.: Modellgenerierung im Kontext der Digitalen Fabrik - Stand der Technik und Herausforderungen. Karlsruhe: KIT Scientific Publishing 2010.

Rabe, M.; Spieckermann, S.; Wenzel, S.: Verifikation und Validierung für die Simulation in Produktion und Logistik, Vorgehensmodelle und Techniken. Berlin: Springer 2008.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulationstool zur alters- und belastungsorientierten Entwicklung und Bewertung von Arbeitssystemen in der industriellen Fertigung

## *Simulation Tool for the Age- and Stress-based Development and Assessment of Work Systems in Industrial Manufacturing*

Ulf Müller, Nico Feller, Michael Schiffmann, FH Köln, Köln (Germany),
ulf.mueller@fh-koeln.de, nico.feller@fh-koeln.de, michael.schiffmann@fh-koeln.de

Peter Gust, Bergische Universität Wuppertal, Wuppertal (Germany),
peter.gust@uni-wuppertal.de

Abstract: In the context of demographic changes in industrial nations and the common global megatrend internet of things, enterprises have to tackle technological, social and economic challenges. To maintain their market position, amongst others, enterprises have to change their manufacturing processes. The following article presents the design and implementation of a simulation software application for the age- and stress-based development and assessment of work systems in industrial manufacturing. The approach for the simulation of aging workforce and calculation of overall stress of work systems is introduced. Implementation and usage of the software tool will be explained by an example. Concluding, presented results and future development are discussed.

## 1 Einleitung

Sowohl aktuell als auch in den kommenden Jahren müssen sich Unternehmen mit wesentlichen gesellschaftlichen, technologischen und ökonomischen Wandlungsprozessen auseinandersetzen. Hierbei ist zum einen der demographische Wandel zu nennen, zum anderen muss die Digitalisierung von Produkten und Produktionsprozessen im Rahmen der Industrie 4.0 angeführt werden.

Bis 2060 ist bspw. in Deutschland ein Bevölkerungsrückgang der Menschen im Erwerbsalter von rund 50 auf 34,5 Mio. prognostiziert; über 30% der Erwerbstätigen werden älter sein als 50 Jahre (vgl. Statistisches Bundesamt 2015, S. 19-25). Aufgrund ihrer altersdifferenzierten physischen und psychischen Fähigkeiten, haben ältere Menschen gegenüber jüngeren sowohl im Alltag, als auch am Arbeitsplatz

unterschiedliche Anforderungen an ihre Umwelt und die Interaktion mit technischen Systemen (Boot et al. 2012, S. 1444).

Im Kontext des globalen Megatrends Industrie 4.0 sind Unternehmen gezwungen ihre jeweiligen Produkte, Geschäftsmodelle und vor allem Fertigungsprozesse zu modifizieren respektive neu zu entwickeln. D.h. Unternehmen müssen sich digital transformieren. Veränderungen hin zu einer intelligenten Fabrik führen u.a. zu einer Dezentralisierung der Produktion. Neben der Integration von neuen Technologien beinhaltet dies zudem die Anpassung sowohl von Tätigkeitsfeldern als auch von Verantwortungsbereichen, was letztlich zu neuen, individuellen Beanspruchungsprofilen der Mitarbeiterinnen und Mitarbeiter führt. (Gienke und Kämpf 2007, S. 338-339)

Zusammenfassend besteht ein wachsender Handlungsbedarf zur Erhaltung und Förderung von Wohlbefinden, Gesundheit und Leistungsfähigkeit von Belegschaften im demographischen und digitalen Wandel.

Aktuelle Softwarelösungen zur Bewertung und Gestaltung von Arbeitssystemen basieren zum einen auf digitalen Menschmodellen und fokussieren kurzfristige Anwendungszeiträume. Zum anderen basieren sie auf personalintegrierenden Simulationen, welche die Entwicklung von Arbeitszeit- und Organisationsmodellen für mittel- und langfristige Zeiträume ermöglichen (Zülch 2010; Leupold et al. 2010; Wischniewski und Adolph 2013). Ein ganzheitlicher mittel- und langfristiger Belastungseinfluss von Arbeitssystemen auf die zukünftige Gesundheit sowie Leistungsfähigkeit von Mitarbeitern – als Alterungsprozess – ist aktuell nicht hinreichend erforscht (Zülch 2013, S. 59). Dies muss bei der Entwicklung und Bewertung von Arbeitssystemen vor allem für ältere Erwerbstätige berücksichtigt werden.

Mit dem Simulationstool zur alters- und belastungsorientierten Entwicklung und Bewertung von Arbeitssystemen in der industriellen Fertigung wird nachfolgend ein Lösungsbeitrag zum dargestellten Handlungsbedarf vorgestellt. Durch eine ganzheitliche Untersuchung individueller und altersabhängiger Belastungen der Erwerbstätigen sowie Simulation einer alternden Belegschaft, ermöglicht das Simulationsmodell die Entwicklung und Bewertung von Arbeitssystemen im Sinne einer prospektiven Arbeitsgestaltung.

Im weiteren Beitrag wird zunächst der verwendete Systemansatz vorgestellt. Anschließend beschreibt Kapitel 3 zum einen die Implementierung des Simulationssystems, zum anderen wird ein Anwendungsbeispiel gegeben. Abschließend werden die präsentierten Arbeitsergebnisse in einer kurzen Diskussion reflektiert. Alle im Beitrag verwendeten Parameter sind nachfolgend in Tabelle 1 aufgelistet.

## 2 Systemansatz

Nachfolgend sind u.a. die wesentlichen Vorarbeiten von Feller und Müller (2014) sowie Müller et al. (2015) zusammengefasst. Fokussiert werden sowohl die Struktur des Simulationssystems und die Modellbildung als auch die Zustandsänderungen des zentralen Modellelementes „Mitarbeiter“.

*Tabelle 1: Nomenklatur*

| Parameter | Beschreibung | Einheit |
|---|---|---|
| PAL | Physische Leistungsfähigkeit | % |
| k | Konstante für individuelle Fitness: k < 1 niedrige Fitness; k < 2 hohe Fitness | |
| WAL | Arbeitsleistungsfähigkeit | % |
| s | Beanspruchung | |
| $s_i$ | Beanspruchung zu einem diskreten Zeitpunkt i | |
| $t_w$ | Dauer des Arbeitsprozesses | h |
| $t_b$ | Dauer der Pause | h |
| OS | Ganzheitliche Belastung | % |
| $S_i$ | Belastungsfaktor | |
| $S_i^*$ | Gewichteter Belastungsfaktor | |
| $k_i$ | Gewichtungsfaktor | |
| $k_s$ | Gewichtungsfaktor Geschlecht | |
| $k_o$ | Gewichtungsfaktor Sauerstoffaufnahme | |
| $k_v$ | Gewichtungsfaktor visuelle Fähigkeiten | |
| $k_a$ | Gewichtungsfaktor auditive Fähigkeiten | |

## 2.1 Alters- und belastungsorientiertes Simulationsmodell

Wie Abbildung 1 zeigt, untergliedert sich das von Feller und Müller (2014) konzipierte Simulationsmodell in drei übergeordnete Segmente: Generator, Simulator und Analysator.

Der Generator dient der Modellbildung. Hierbei werden die Elemente Mitarbeiter und Arbeitsplatz differenziert. Ein Mitarbeiter wird durch die Attribute Alter, Geschlecht und eine optionale Konstante für die individuelle Fitness beschrieben. Basierend auf diesen drei Attributen werden während einer Simulation die Parameter physische Leistungsfähigkeit und Arbeitsleistungsfähigkeit zeitdiskret berechnet. Die diesbezüglich interne Logik ist im Folgenden unter 2.1.1 näher beschrieben.

Das Modellelement Arbeitsplatz wird zunächst hinreichend genau durch die fünf quantitativ mess- respektive eindeutig bewertbaren Attribute Licht, Klima, Lärm, Arbeitsposition und Arbeitsintensität beschrieben. Nach Spath et al. (2012, S. 1655) sind die angeführten drei Umweltfaktoren als wesentliche Belastungsfaktoren in der industriellen Fertigung zu identifizieren respektive zu berücksichtigen. Ergänzend benennt Bullinger (1994, S. 198) die angeführten ergonomischen Parameter als die wesentlichen Faktoren in Bezug auf ergonomische Arbeitsgestaltung. Alle fünf Attribute bzw. Belastungsfaktoren werden tätigkeitsspezifisch zwischen 1 „hervorragend“ und 6 „defizitär“ bewertet. Die Parametrierung des Modellelements Arbeitsplatz wird ausführlich beschrieben von Feller und Müller (2014, S. 143 ff.), die interne Logik ist unter 2.1.2 beschrieben.

Der Simulator bildet den Kern des Systems. Entsprechend des gewählten Simulationszeitraums sowie der jeweils definierten Tätigkeitsabfolgen (inkl. Pausen) intera-

gieren die Mitarbeiter zeitdiskret mit den zugeordneten Arbeitsplätzen und ihre individuelle Beanspruchung wird berechnet. Wie in den folgenden Unterkapiteln näher beschrieben, wird hierbei durch eine mitarbeiterindividuelle Gewichtung der einzelnen Arbeitsplatz- bzw. Belastungsparameter ein Altern der Belegschaft während der Simulationsdauer berücksichtigt.

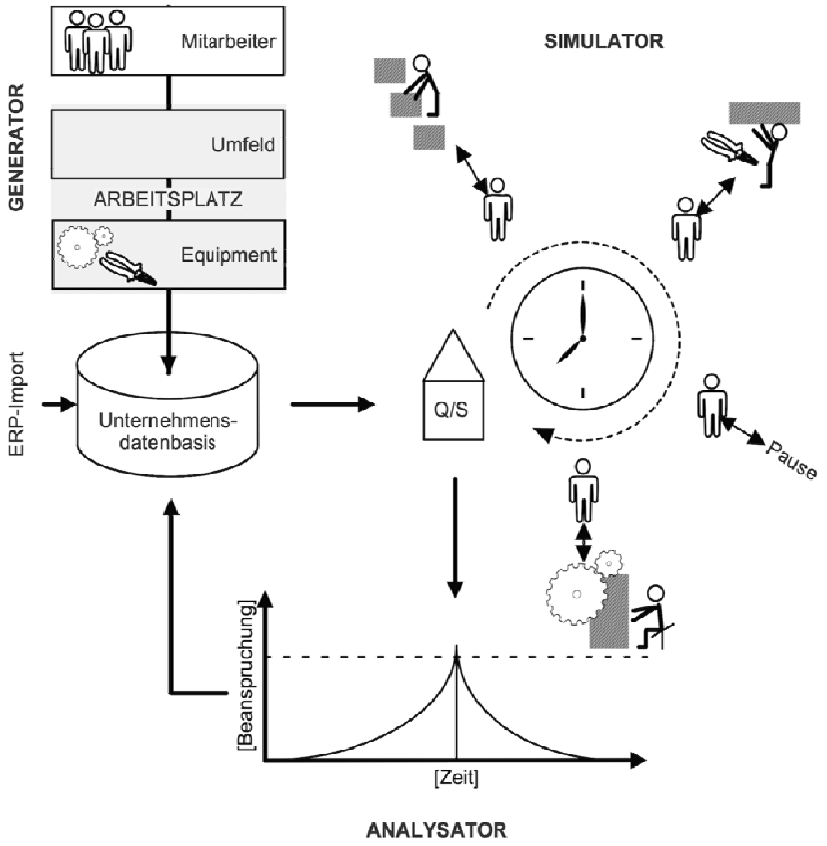



*Abbildung 1: Systemstruktur (Feller und Müller 2014, S. 142)*

Abbildung 2 visualisiert die logischen Verknüpfungen der Modellelemente Mitarbeiter und Arbeitsplatz. Die den jeweiligen Parametern vorangestellten Nummern beziffern die Reihenfolge der Abarbeitung. Zudem ist auf die Gleichungen der entsprechenden Übertragungsfunktionen verwiesen, welche in den folgenden Unterkapiteln näher beschrieben sind. Ist die individuelle Fitness bei der Modellierung des Mitarbeiters nicht explizit bewertet worden, wird k = 1,0 gesetzt.

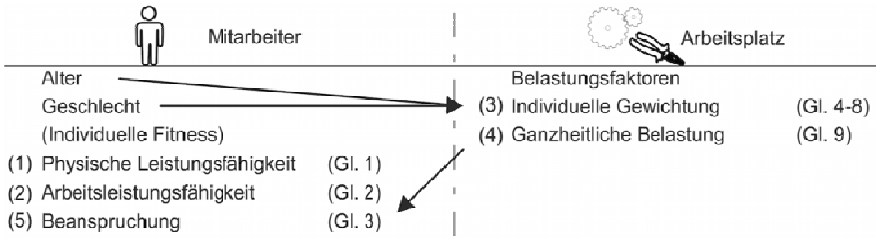



*Abbildung 2: Ablaufstruktur Mitarbeiter und Arbeitsplatz*

Der Analysator dient der Visualisierung und Auswertung der Simulationsergebnisse. Wie von Müller et al. (2015) beschrieben, wird hierbei die Beantwortung der folgenden Leitfragen fokussiert:

- welche Arbeitsprozesse und -abfolgen verursachen eine Überbeanspruchung bzw. eine Beanspruchung des jeweiligen Mitarbeiters oberhalb der individuellen Arbeitsleistungsfähigkeit? (vgl. Bullinger 1994, S. 45-46)
- welcher Arbeitsplatz hat diesbezüglich den wesentlichen Einfluss?
- zu welchem Zeitpunkt tritt eine Überbeanspruchung auf und für welche Dauer?
- welche spezifischen Arbeitsplatzparameter sind für die Überbeanspruchung verantwortlich?
- wie setzt sich die Überbeanspruchung zusammen: Parameter, Dauer, Abfolge?

### 2.1.1 Ablauflogik Mitarbeiter

Die nachfolgenden Gleichungen beschreiben die Übertragungsfunktionen für das Modellelement Mitarbeiter. Hierbei bauen die Gleichungen Gl. (1) und Gl. (3) auf Arbeiten zur Leistungsfähigkeit von Schlick et al. (2010, S. 121, 202) auf. Gl. (2) beruht auf einem Motivationsschema von Bullinger (1994, S. 46). (Feller und Müller 2014)

$$\mathrm{PAL} = \begin{cases} \left(5\left(1 - \mathrm{e}^{-\left(\frac{\mathrm{Alter}-16}{2}\right)}\right) + 95\right)\mathrm{k}, & 16 \ldots 30\ Jahre \\ (-0{,}6\mathrm{Alter} + 118)\mathrm{k}, & 31 \ldots 67\ Jahre \end{cases} \tag{1}$$

$$\mathrm{WAL} = 0{,}6\ \mathrm{PAL} \tag{2}$$

$$\mathrm{s} = \begin{cases} \mathrm{s_i} + \mathrm{e}^{\mathrm{t_w}\left(\frac{\mathrm{OS}}{\mathrm{WAL}}\right)}, & Ermüdung \\ \mathrm{s_i}\mathrm{e}^{-\mathrm{t_b}\mathrm{k}}, & Regeneration \end{cases} \tag{3}$$

Neben den individuellen leistungsbezogenen Parametern sind auch die alters- und geschlechtsspezifischen Gewichtungsfaktoren dem Modellelement Mitarbeiter zugeordnet. Gl. (4) und Gl. (5) drücken die geschlechtsspezifische und altersabhängig linear abnehmende Sauerstoffaufnahmefähigkeit aus. Gl. (6) beschreibt die altersbedingte Zustandsänderung der Sehfähigkeit unter der Annahme, dass das Auge als proportional geregeltes System mit Dämpfung Lichtimpulse in Informationen verarbeitet. Die alters- und geschlechtsspezifische lineare Abnahme des Hörvermögens drückt Gl. (7) aus. Die literaturbasierte Herleitung der Gleichungen Gl. (4) bis Gl. (7) ist ausführlich bei Feller und Müller (2014, S. 145 ff.) beschrieben.

$$k_s = \begin{cases} 1 & männlich \\ 1{,}4 & weiblich \end{cases} \tag{4}$$

$$k_o = \begin{cases} 2 - (-0{,}007\text{Alter} + 1{,}14), & Alter \geq 20\ Jahre \\ 1, & Alter < 20\ Jahre \end{cases} \tag{5}$$

$$k_v = \begin{cases} 2 - \left(-0{,}286e^{-\left(\frac{\text{Alter}-20}{8}\right)}\right) + \left(-1{,}286e^{-\left(\frac{\text{Alter}-20}{36}\right)}\right), & Alter \geq 20\ Jahre \\ 1, & Alter < 20\ Jahre \end{cases} \tag{6}$$

$$k_a = \begin{cases} 0{,}00625\text{Alter} + 0{,}8225, & (\text{männlich})\text{Alter} \geq 30\ \text{Jahre} \\ 0{,}00388\text{Alter} + 0{,}8886, & (\text{weiblich})\text{Alter} \geq 30\ \text{Jahre} \\ 1, & Alter < 30\ Jahre \end{cases} \tag{7}$$

### *2.1.2 Ablauflogik Arbeitsplatz*

Die Übertragungsfunktionen für das Modellelement Arbeitsplatz werden durch die Gleichungen GL. (9) und Gl. (10) beschrieben. (Feller und Müller 2014, S. 149)

$$S_i^* = \begin{cases} 6, & S_i \prod_{i=1}^{n} k_i(\text{Geschlecht}, \text{Alter}) > 6 \\ S_i \prod_{i=1}^{n} k_i(\text{Geschlecht}, \text{Alter}), & S_i \prod_{i=1}^{n} k_i(\text{Geschlecht}, \text{Alter}) = 1 \ldots 6 \\ 1, & S_i \prod_{i=1}^{n} k_i(\text{Geschlecht}, \text{Alter}) < 1 \end{cases} \tag{8}$$

$$OS = \left(S_{i,max}^* + \frac{\sum_{i=1}^{n} S_i^*}{n-1} \times \frac{6 - S_{i,max}^*}{6}\right) \times \frac{\ln 100}{48} \tag{9}$$

Die Berechnung von OS basiert hierbei auf zwei wesentlichen Annahmen. Zum einen ist definiert, dass nach acht Stunden Arbeit in einem maximal-schlecht bewerteten Arbeitssystem eine hundertprozentige Auslastung der Arbeitsleistungsfähigkeit (WAL) eines 30-jährigen, männlichen Mitarbeiters durchschnittlicher Fitness (k = 1,0, PAL = 100 %) erreicht wird. Tätigkeiten welche darüber hinaus gehen, rufen eine Überbeanspruchung hervor. Zum anderen ist festgelegt, dass der maximale gewichtete Belastungsfaktor den vorherrschenden Einfluss auf die Belastung des Mitarbeiters hat; die übrigen gewichteten Belastungsfaktoren werden als gewogener Mittelwert aufsummiert (Feller und Müller 2014, S. 148).

## 2.2 Mitarbeiter als zentrales Simulationselement

Wie zuvor beschrieben durchlaufen die Modellelemente „Mitarbeiter“ als zentrale Elemente zeitdiskret den Simulator (vgl. Abb. 1). Hierbei werden im Laufe jedes simulierten Tages fünf Zustände differenziert: Alter, Arbeiten, Pause, Warten und Freizeit. Abbildung 3 zeigt das Zustandsdiagramm Mitarbeiter, welches sowohl die Stimuli für Zustandsänderungen d.h. die Entscheidungsregeln für Handlungsalternativen als auch die zustandsspezifischen Aktionen in Pseudocode beschreibt. Hierdurch wird die Konsistenz des Simulationsmodells gewährleistet.

Die Definition von Kapazitäten für Arbeitsplätze reglementiert eine widerspruchsfreie Interaktion der Modellelemente Mitarbeiter und Arbeitsplatz. Handlungsalter-

nativen sind das Abarbeiten einer möglichen anderen Tätigkeit oder das Warten, jeweils gem. der dargestellten Entscheidungsregeln. Bspw. wird so der Gebrauch von ein und desselben Fertigungsmittels durch zwei Mitarbeiter zum gleichen Zeitpunkt unterbunden.

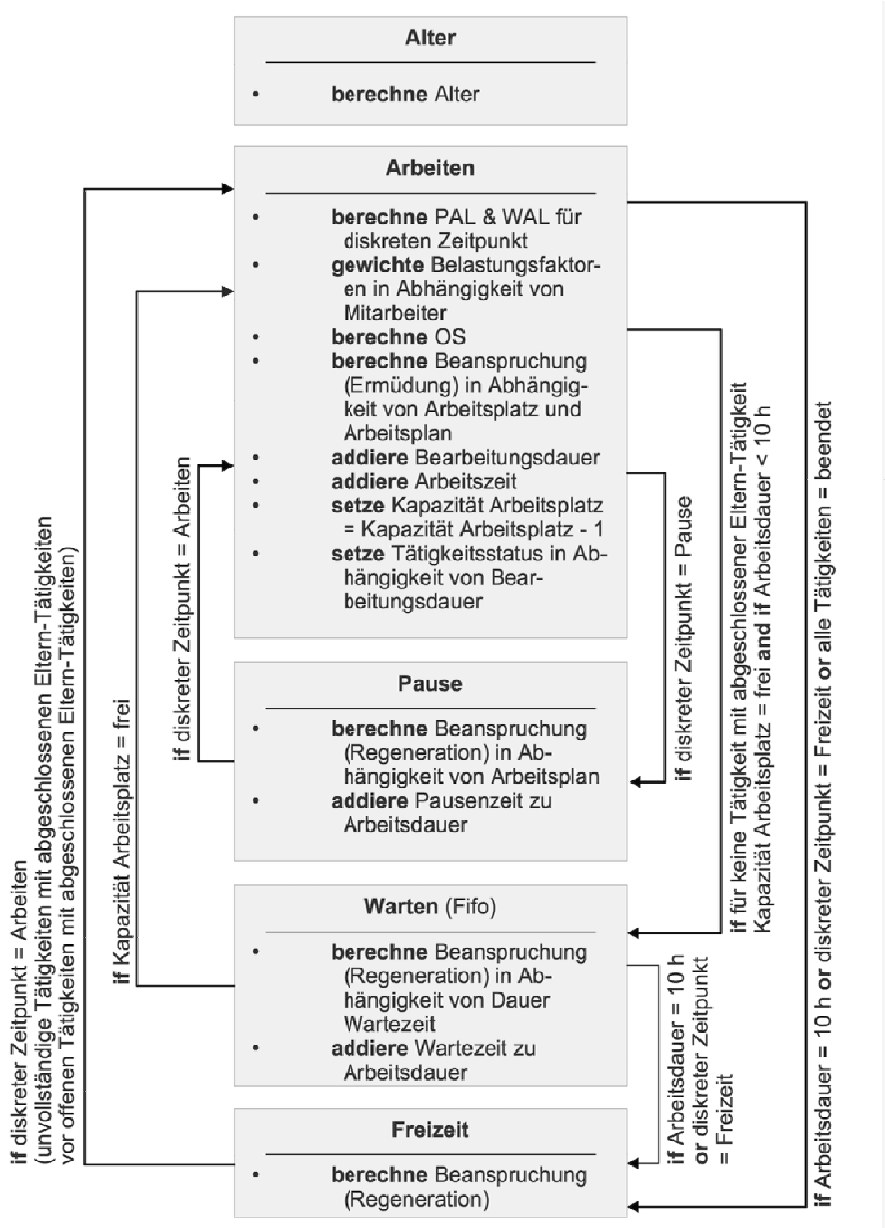



*Abbildung 3: Zustandsdiagramm Mitarbeiter (Müller et al. 2015)*

Der Status einer Tätigkeit (offen; geschlossen) richtet sich nach ihrer jeweiligen Bearbeitungsdauer (Soll-Ist-Vergleich). Nach Erreichen einer täglichen maximalen Arbeitszeit von zehn Stunden werden offene Tätigkeiten auf den folgenden Tag übertragen. Zudem werden den Tätigkeiten bei Bedarf Eltern-Tätigkeiten zugeordnet. Bspw. macht es nur Sinn ein Produkt nach der Fertigung aller Komponenten zu montieren, die Herstellungsreihenfolge kann jedoch variieren.

# 3 App-basierte Systemlösung zur Simulation und Analyse von Arbeitssystemen

Mit der Realisierung des Simulationstools als App wird besonders auf den Bedarf kleiner und mittelgroßer Unternehmen (KMU) nach schlanken, intuitiven und kostengünstigen Systemlösungen reagiert. Vor allem für KMU in der industriellen Fertigung stellen die einleitend skizzierten Wandlungsprozesse große Herausforderungen dar. Als Lohnfertigungsunternehmen sind KMU i.d.R. konfrontiert mit häufig wechselnden Produkten sowie hohem Wettbewerbsdruck und verfügen meist über stark limitierte Ressourcen. Gängige IT-Systemlösungen im Bereich der Digitalen Fabrik sind hinsichtlich ihrer Kosten sowie ihres Umfangs unzureichend auf die Anforderungen von KMU abgestimmt.

## 3.1 Implementierung

Algorithmus 1 beschreibt die Implementierung des Simulationskerns in SWIFT. In der Funktion „doSimulation“ ist die inkrementelle Simulation aller Modellelemente durch drei geschachtelte for-Schleifen realisiert. Der Funktion wird die Anzahl zu simulierender Tage, die zeitliche Auflösung in Sekunden und eine Liste mit Mitarbeiterobjekten übergeben. Die Mitarbeiterobjekte enthalten jeweils den Arbeitsplatz und den Arbeitsplan des jeweiligen Tages. Das Mitarbeiterobjekt wird innerhalb der Simulationsschleife über die Funktion „do_action“ angesprochen. In dieser Funktion wird anhand der Tageszeit, des Zeitinkrements und des Tages der Zustandsautomat des Mitarbeiters und somit die jeweiligen Mitarbeiterattribute aktualisiert (vgl. Abb. 3).

*Algorithmus 1: Implementierung Simulationskern*

```
func doSimulation(days:Int,time_step:Int,EMP_List:
EmployeeList){
      var Abs_Steps: Int = 86400 / time_step
      var DayTime: Int = 0
      for (var d: Int = 0; d < days; d++){
            DayTime = 0
            for (var step: Int = 0;step < Abs_Steps;step++){
                  DayTime = DayTime + time_step
                  for Emp in EMP_List{
                  Emp.do_action(DayTime,inkr:time_step,Day:d)
                  }
            }
      }
}
```

## 3.2 Anwendungsbeispiel

Für ein erstes Anwendungsbeispiel sind zwei männliche Erwerbstätige über 24 Stunden simuliert worden. Für beide Mitarbeiter ist der gleiche repräsentative Arbeitstag mit zwei unterschiedlichen Pausenmodellen durchlaufen worden. Tabelle 2 führt die jeweilige Tätigkeit und entsprechende Dauer chronologisch auf.

Herr Meier ist ein 45 Jahre alter Montagearbeiter. Er hat eine unterdurchschnittliche Fitness (k = 0,8), seine Arbeitsleistungsfähigkeit beträgt 43,68 Prozent. Herr Mustermann ist ein 62-jähriger Montagearbeiter. Er hat eine durchschnittliche Fitness (k = 1,0), seine Arbeitsleistungsfähigkeit beträgt 48,48 Prozent. Die jeweilige Arbeitsleistungsfähigkeit bezieht sich auf einen 30-jährigen, männlichen Referenzmitarbeiter von durchschnittlicher Fitness: k = 1,0, PAL = 100 %, WAL = 60 % (vgl. Gl. 1 und Gl. 2).

*Tabelle 2: Arbeitstag mit variiertem Pausenmodell*

| Tätigkeit | Dauer V1 [h] | Dauer V2 [h] |
|---|---|---|
| Arbeitsvorbereitung, Rüsten | 0,5 | 0,5 |
| Montage | 1,5 | 1,5 |
| Frühstückspause | 0,25 | 0 |
| Montage | 3 | 3 |
| Mittagspause | 0,5 | 0,75 |
| Montage | 2,5 | 2,5 |
| Nachbereitung | 0,5 | 0,5 |

Abbildung 4 zeigt exemplarisch für die Erstellung von Modellelementen mit WorkDesigner die Konfiguration des Arbeitsplatzes für die Tätigkeit Montage. Im Hauptbereich der graphischen Benutzeroberfläche werden die Attribute des jeweiligen Arbeitsplatzes editiert (vgl. Abschnitt 2.1). Das linke Tabellensteuerelement dient einer komfortablen Verwaltung der Modellelemente „Arbeitsplatz".

Die Programmnavigation erfolgt über die Steuerleiste im unteren Bereich der Benutzeroberfläche. Neben der dargestellten Ansicht „Manage Workplaces" (Abb. 4) stehen dem Anwender zudem vier weitere, analog aufgebaute Ansichten zur Verfügung. „Manage Work Processing Sheets" für die Verwaltung und Beschreibung von Arbeitsplänen und Schichtmodellen. „Manage Employees" zur Verwaltung und Beschreibung der Modellelemente „Mitarbeiter". „Setup a Simulation" für die Verwaltung und Definition von Simulationsstudien; d.h. Festlegung eines Simulationszeitraums und der beteiligten Modellelemente. „Analyzer" für die graphische Aufbereitung der Simulationsergebnisse (Abb. 5).

Abbildung 5 zeigt die Aufbereitung der Simulationsergebnisse des beschriebenen Anwendungsbeispiels. Das Diagramm im Hauptbereich der Benutzeroberfläche zeigt für die simulierten Mitarbeiter Herrn Meier und Herrn Mustermann jeweils für beide Arbeitstagvarianten den Verlauf der Auslastung ihrer individuellen Arbeitsleistungsfähigkeit in Prozent über der Tageszeit in Stunden. Die beiden oberen

Graphen beschreiben den Beanspruchungsverlauf von Herrn Meier, die unteren beiden stellen den Beanspruchungsverlauf von Herrn Mustermann dar.

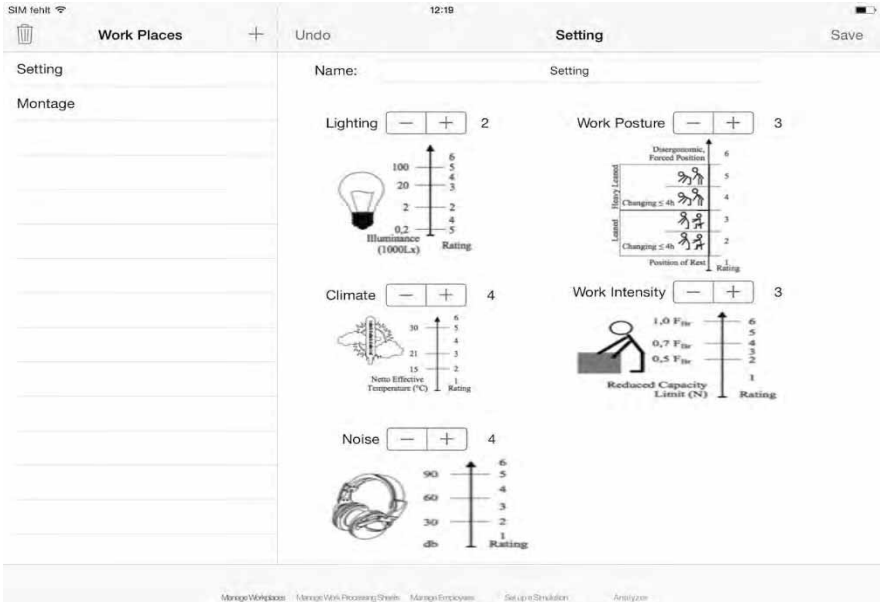



*Abbildung 4: iPad Screenshot Erstellung Arbeitsplatz mit WorkDesigner 1.0*

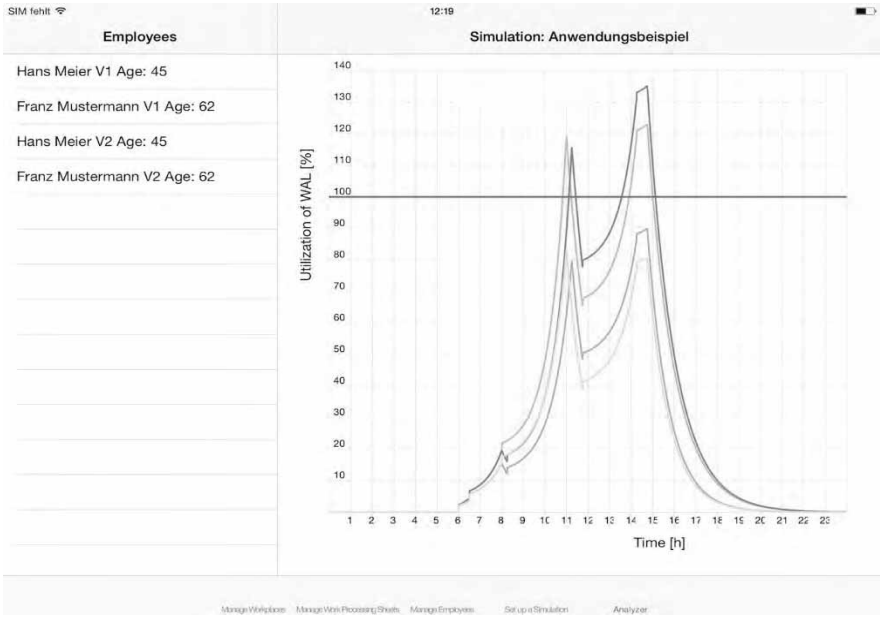



*Abbildung 5: iPad Screenshot Aufbereitung Simulationsergebnisse mit WorkDesigner 1.0*

Wesentliche Ergebnisse der Simulationsstudie sind:

- bedingt durch seine unterdurchschnittliche Fitness wird Herr Meier deutlich überbeansprucht.
- Herr Meier benötigt im Gegensatz zum älteren Herrn Mustermann mehr als eine Stunde länger für die vollständige Regeneration seiner Arbeitsleistungsfähigkeit, was neben der höheren Beanspruchung ebenfalls seiner unterdurchschnittlichen Fitness zuzuordnen ist.
- die geringfügige Modifikation des Pausenmodells (vgl. Tab. 2) führt zu einer Reduzierung der jeweils maximalen Beanspruchung von durchschnittlich zehn Prozent.

## 4 Diskussion

Mit dem app-basierten Simulationstool ist eine schlanke und intuitive Systemlösung zur alters- und belastungsorientierten Gestaltung und Bewertung von Arbeitssystemen vorgestellt worden. Die vorgefertigten Modellelemente Mitarbeiter und Arbeitsplatz ermöglichen eine schnelle und hinreichend genaue Abbildung von Arbeitssystemen. Die Gestaltung von Arbeitssystemen kann schnell hinsichtlich ihrer Auswirkungen auf die Gesundheit und Leistungsfähigkeit der Belegschaft bewertet werden.

Als Schwachstelle des Systems muss die Simulation von Freizeitperioden angeführt werden. Durch nicht bekannte Aktivitäten wie Sport oder physische sowie psychische Erkrankungen und Stress lässt sich die Beanspruchung in Freizeitphasen nicht eindeutig abbilden. Auch ermöglicht das Simulationstool aktuell lediglich eine rein physische Betrachtung der Beanspruchung. Jedoch erlaubt der verwendete Berechnungsansatz eine spätere Ergänzung von Attributen bspw. zur Beschreibung von Mitarbeiterqualifikation und kognitivem Anspruch einer Tätigkeit.

Die empirische Validierung des Simulationstools mit einer heterogenen Probandengruppe in der industriellen Fertigung ist als nächster Entwicklungsschritt eingeplant. Diesbezüglich werden Arbeitssysteme mit mehrjährig bekannter Historie rückwirkend simuliert.

## Literatur

Boot, W. R.; Nichols, T. A.; Rogers, W. A.; Fisk, A. D.: DESIGN FOR AGING. In: Salvendy, G. (Hrsg.): Handbook of human factors and ergonomics. Hoboken, NJ: Wiley 2012, S. 1442-1465.

Bullinger, H.-J.: Ergonomie - Produkt- und Arbeitsplatzgestaltung. Stuttgart: Teubner 1994.

Feller, N.; Müller, U.: Development and Assessment of Work Systems for Elder Employees in Industrial Manufacturing. In: Trzcielinski, S.; Karwowski, W. (Hrsg.): Advances in the Ergonomics in Manufacturing – Managing the Enterprise of the Future. USA: AHFE Conference © 2014, S. 140-151.

Gienke, H.; Kämpf, R.: Handbuch Produktion – Innovatives Produktionsmanagement: Organisation, Konzepte, Controlling. München: Hanser 2007.

Leupold, M.; Stock, P.; Schmidt, D.; Zülch, G.: Simulationsbasierte Bewertung von Arbeitszeitmodellen bezüglich der Work-Life-Balance. In: Zülch, G.; Stock, P.

(Hrsg.): Integrationsaspekte der Simulation - Technik, Organisation und Personal. Karlsruhe: KIT Scientific Publishing 2010, S. 365-372.

Müller, U; Gust, P.; Feller, N.; Schiffmann, M.: WorkDesigner: Consulting Application Software for the Strain-Based Staffing and Design of Work Processes. Angenommen zur Veröffentlichung von AHFE Conference © in Kooperation mit Elsevier Procedia; Elsevier Science Direct. 6th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conference (AHFE), Las Vegas (USA), 26.-30. Juli 2015.

Schlick, C.; Bruder, R.; Luczak, H.: Arbeitswissenschaft. Berlin: Springer 2010.

Spath, D.; Braun, M.; Meinken, K.: HUMAN FACTORS IN MANUFACTURING. In: Salvendy, G. (Hrsg.): Handbook of human factors and ergonomics. Hoboken, NJ: Wiley 2012, S. 1643-1665.

Statistisches Bundesamt: BEVÖLKERUNG DEUTSCHLANDS BIS 2060 - 13. koordinierte Bevölkerungsvorausberechnung. Wiesbaden: Statistisches Bundesamt 2015.

Statistisches Bundesamt: BEVÖLKERUNG DEUTSCHLANDS BIS 2060 - 12. koordinierte Bevölkerungsvorausberechnung. Wiesbaden: Statistisches Bundesamt 2009.

Wischniewski, S.; Adolph, L. (Vorsitz): Tagungsbericht des Symposiums der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin im November 2012. Digitale Ergonomie - Trend und Strategien für den Einsatz digitaler Menschmodelle. Dortmund 2013.

Zülch, G.: Stand und Entwicklungstendenzen der personalorientierten Simulation. In: Zülch, G.; Stock, P. (Hrsg.): Integrationsaspekte der Simulation - Technik, Organisation und Personal. Karlsruhe: KIT Scientific Publishing 2010, S. 1-19.

Zülch, G.: Ergonomic representation of humans in the digital factory - The new guideline VDI 4499-4. In: Dangelmaier, W; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik - Entscheidungsunterstützung von der Planung bis zur Steuerung. Paderborn: HNI-Verlagsschriftenreihe 2013, S. 53-60.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Abbildung von Steuerungslogiken durch maschinelles Lernen für die Simulation von Produktionssystemen

## *Emulation of Control Strategies through Machine Learning in Manufacturing Simulations*

Sören Bergmann, Niclas Feldkamp, Ulrich Hinze, Steffen Straßburger, TU Ilmenau, Ilmenau (Germany), soeren.bergmann@tu-ilmenau.de, niclas.feldkamp@tu-ilmenau.de, ulrich.hinze@tu-ilmenau.de, steffen.strassburger@tu-ilmenau.de.

Abstract: In the context of discrete-event simulation of production and logistics systems, modelling an exact representation of the real system is needed for obtaining sound and reliable results. The automatic generation of simulation models is an approach for saving time and effort for creating models and, therefore, it is a recurring issue in modelling methodology research. In automatic model generation, the approximation of dynamic behaviour is a challenging problem. This is for example the case when the dispatching and scheduling of production jobs needs to be adequately emulated, but the underlying rules are unknown. In previous work, we presented an approach for approximating dynamic behaviour through artificial neural networks. In this paper, we propose an improved approach and investigate its suitability again with artificial neuronal networks as well as with other data mining and supervised machine learning methods.

## 1 Einführung

Die automatische Generierung von Simulationsmodellen wird immer wieder als ein probates Mittel genannt, um zum einen Simulationsstudien zu beschleunigen, sowie zum anderen neue Nutzergruppen für die Simulation zu gewinnen (Selke 2005, S.23; Bergmann und Straßburger 2010, S.545f; Bergmann 2013, S.25ff). Eine spezielle Herausforderung bzw. Lücke innerhalb der automatischen Modellgenerierung (AMG) stellt die Abbildung von Steuerungsregeln dar (Selke 2005, S.30; Bracht et al. 2008, S.30; Bergmann und Straßburger 2010, S.545f; Bergmann 2013, S.71ff).

Steuerungsregeln können hierbei in vier verschiedene Klassen eingeordnet werden, konkret in Strategien zur Auftragsfreigabe, zur Losgrößenbestimmung, zur Ressourcenbelegung und Reihenfolgegestaltung (Selke 2005, S.45), wobei gerade Entscheidungen der letztgenannten Klasse in vielen Systemen von enormer Bedeutung sind und stellvertretend im Folgenden betrachtet werden. In der Praxis werden in allen

Klassen häufig Entscheidungsregeln, z. B. First In First Out (FIFO) oder Shortest Processing Time (SPT), genutzt (Pinedo 2012).

Zur Behandlung der skizzierten Lücke im Prozess der AMG sind drei Klassen von Ansätzen mit je eigenen Vor- und Nachteilen bzw. Restriktionen denkbar.

Die Triviallösung und erste Klasse besteht in dem Modellieren der Regeln durch einen Experten, entweder direkt in der Simulationsumgebung oder in einer systemneutralen Sprache, z. B. SysML (Schönherr und Rose 2009).

Die zweite Klasse versucht, bestehende Regeln bzw. eine Kombination von Basisregeln zu erkennen (Wiedemann 1999; Selke 2005; Reinhart und Gyger 2008).

Der in die dritte Klasse fallende im Folgenden aufgegriffene Ansatz erkennt keine vorgefertigten Steuerungsregeln, sondern approximiert das Verhalten der (Realwelt-)Regeln. Basis können hier beispielsweise Daten der Betriebsdatenerfassung (BDE-Daten) sein, welche direkt zum Erzeugen bzw. Trainieren der Regel genutzt werden. Dieser Ansatz wurde bereits vereinzelt in Publikationen aufgegriffen. So wurden bspw. künstliche neuronale Netze (KNN) zur Approximation von Reihenfolgeregeln prototypisch eingesetzt (Bergmann 2013; Bergmann et al. 2014). Das genannte approximierende Verfahren zeigte das Potential solcher Ansätze auf, unterlag aber verschiedensten Restriktionen bezüglich Datenqualität oder Skalierbarkeit; so wurde beispielsweise eine fixe Puffergröße unterstellt.

Dieser Beitrag beschreibt eine Erweiterung bzw. Generalisierung des bestehenden Ansatzes. Es wird ein an den KDD-Prozess (Knowledge Discovery in Databases) (Fayyad 1996) angelehntes Framework zur Approximation von Steuerungsregeln aus BDE-Daten mittels Methoden des maschinellen Lernens vorgestellt. Hierbei werden mittels Data-Mining-Methoden auf Basis von BDE-Daten Entscheidungsmodelle erzeugt, welche im Folgenden aus der Simulation heraus genutzt werden können.

Hierzu wird in Kapitel 2 das Framework zur Datenaufbereitung und der Abbildung von Steuerungsregeln durch maschinelles Lernen vorgestellt. Ebenfalls werden die aktuell genutzten Methoden des maschinellen Lernens / Data-Mining-Methoden und eine Auswahl ihrer wichtigsten Parameter kurz vorgestellt. Im 3. Kapitel wird zunächst ein kleines Testszenario definiert, bevor die verschiedenen vorgestellten Data-Mining-Methoden unter unterschiedlichen Randbedingungen und mit verschiedenen Parametrisierungen der Methoden sowie unter Nutzung verschiedener Datenaufbereitungsverfahren des Frameworks hinsichtlich ihrer Abbildungsgenauigkeit validiert werden. Die Ergebnisse werden anschließend miteinander verglichen und diskutiert. Abschließend wird ein kurzes Fazit gezogen und ein Ausblick auf mögliche Erweiterungen gegeben.

## 2 Konzept eines Frameworks zur Abbildung von Steuerungsregeln

Ziel des Frameworks ist es, ein direkt nutzbares Entscheidungsmodell auf Basis eines bestehenden Datenbestands zu erzeugen. Das finale Entscheidungsmodell kann nach Abschluss des Erzeugungsprozesses direkt genutzt werden, z.B. im Kontext der AMG direkt aus der Simulation heraus.

Der Basisdatenbestand umfasst hierbei zunächst je einen Datensatz pro Entscheidungsvorgang, welcher den aktuellen Systemzustand und die getroffene Entscheidung beinhaltet. Aus Gründen der Skalierbarkeit und Robustheit werden alle Entscheidungen auf binäre Entscheidungen zurückgeführt und entsprechende Datensätze erzeugt. Muss zum Beispiel eine Auswahl zwischen 3 Aufträgen (A, B, C) getroffen werden, wird zunächst ein Datensatz für den Vergleich zwischen A und B und anschließend zwischen dem „Gewinner“ und C durchgeführt.

Um gute Ergebnisse zu erreichen, sind zwei wesentliche Teilbereiche zu untersuchen (Abb. 1). Zum einen müssen geeignete Verfahren des maschinellen Lernens ausgewählt und deren Parameter geeignet gesetzt werden, zum anderen müssen die Daten für gute Ergebnisse meist vorab aufgearbeitet werden (Datentransformation).

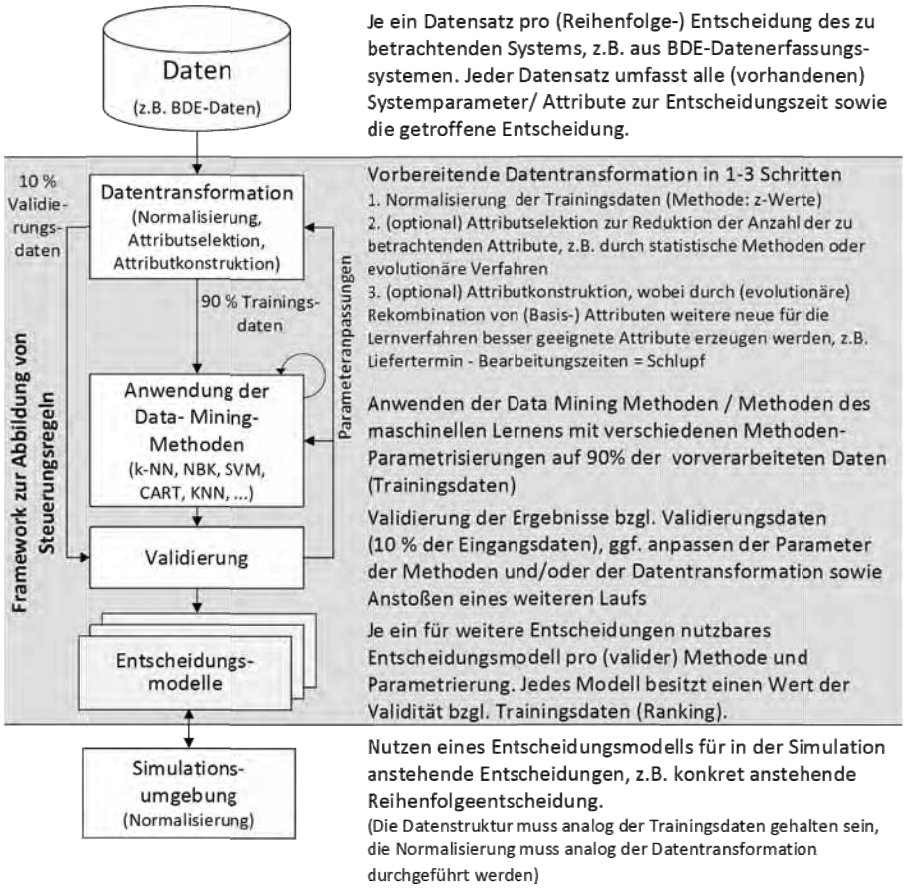



*Abbildung 1: Schematischer Ablauf bei der Abbildung von Steuerungsregeln*

Das Framework bietet für die Datentransformation verschiedene Methoden zur Datenaufbereitung, z. B. die konstruktive Induktion von Attributen, die Attributselektion und die Normalisierung von Attributwerten. Die Normalisierung von Attributwerten wird obligatorisch durchgeführt, das aktuell genutzte Verfahren

basiert auf dem statistischen Standardverfahren der z-Transformation (Bortz und Schuster 2001).

Die Attributselektion und -konstruktion sind optionale Schritte. Im aktuellen Entwicklungsstand werden u.a. evolutionär mögliche Varianten zur Auswahl von Attributen und zur Rekombination von Attributen durchlaufen. Jede Variante wird anschließend angewendet und validiert.

Allgemein sind beliebige Methoden des maschinellen Lernens im Framework einsetzbar und aktuell sind die fünf bekanntesten Vertreter implementiert (vgl. Kapitel 3). Jede Methode hat wiederum eine Vielzahl von Parametern, die das Ergebnis beeinflussen können. Viele dieser Parameter können im Framework gesetzt und entsprechend angepasst werden.

# 3 Methoden des maschinellen Lernens

Unter Methoden des maschinellen Lernens bzw. den im Folgenden synonym verwendeten Data-Mining-Methoden werden allgemein Methoden verstanden, welche aus heterogenen und großen Datenmengen Muster erkennen, um diese für zukünftige Entscheidungen nutzbar zu machen (Fayyad 1996). Für die Nachbildung der Steuerungsregeln wurden fünf verschiedene Methoden genutzt, und zwar k-nächste-Nachbarn (k-NN), Naiver-Bayes-Klassifikator (NBK), Support-Vektor-Maschine (SVM), künstliche neuronale Netze (KNN) und der CART-Algorithmus zur Erzeugung von Entscheidungsbäumen, jeweils mit verschiedenen Parametervariationen. Die prinzipielle Funktionsweise der jeweiligen Methode soll im Folgenden kurz besprochen werden. Um einen Vergleich zur vorherigen Arbeit und dem Ansatz mit neuronalen Netzen zu ziehen, wird auch die Funktionsweise der neuronalen Netze (Multilayer Perceptrons mit Backpropagation) kurz erläutert.

## 3.1 k-nächste-Nachbarn-Algorithmus

Dieser Algorithmus ordnet unklassifizierte Datensätze jener Klasse zu, die der Mehrheit seiner k-nächsten Nachbarn entspricht. Die Nachbarschaft von Datensätzen wird durch deren Ähnlichkeit bestimmt. Für k=1 sucht der Algorithmus beispielsweise den ähnlichsten Datensatz in der Trainingsdatenmenge und weist dem zu klassifizierenden Datensatz dieselbe Klasse entsprechend seines Nachbarn zu. Die Ähnlichkeit von zwei Datensätzen wird über ein Distanzmaß berechnet.

$$d(p,q) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2} \qquad (1)$$

Gleichung 1 zeigt, wie sich die Distanz d zwischen zwei Datensätzen p und q mit Hilfe des euklidischen Distanzmaßes berechnen würde. Die Wahl des k-Parameters hat großen Einfluss auf die Qualität der Ergebnisse und hängt von der Beschaffenheit der zu untersuchenden Datenmenge ab. Größere k-Werte machen den Algorithmus üblicherweise robuster gegen Rauschen, während kleinere k-Werte eine schärfere Abgrenzung zwischen den Klassen herstellen. Tatsächlich generiert der k-NN-Algorithmus kein konkretes Entscheidungsmodell oder -funktion, sondern vergleicht lediglich paarweise den zu klassifizierenden Datensatz mit den Trainingsdaten und berechnet jeweils die Distanz (Altman 1992).

## 3.2 Naiver-Bayes-Klassifikator

Dieser Algorithmus ist ein wahrscheinlichkeitsbasierter Klassifikator. Das Entscheidungsmodell weist demnach einem unklassifizierten Datensatz diejenige Klasse mit der höchsten Wahrscheinlichkeit zu.

Während der Trainingsphase, d.h. der Generierung des Entscheidungsmodells, werden Wahrscheinlichkeitsverteilungen für die Parametervektoren jeder Klasse gebildet. Soll nun ein neuer Datensatz klassifiziert werden, wird aus den Wahrscheinlichkeiten jedes Parameters eine Gesamtwahrscheinlichkeit für jede Klasse gebildet. Abschließend wird nun also jene Klasse zugewiesen, die die höchste Wahrscheinlichkeit aufweist, dass der zu klassifizierende Datensatz dieser Klasse angehört. Ein großer Nachteil hierbei ist, dass die Verteilungsfunktion für kontinuierliche Parameter im Vorhinein bekannt sein muss. Eine falsche Verteilungsfunktion führt eventuell zu falschen Wahrscheinlichkeiten und damit zu falschen Klassifizierungen. Durch die Anwendung von komplexeren Verteilungsfunktionen wie der multivariaten Normalverteilung kann dieses Problem abgeschwächt werden (Rennie et al. 2003).

## 3.3 Support-Vektor-Maschine

Im Gegensatz zu den anderen Methoden teilt dieser Algorithmus die Trainingsdaten immer in zwei Klassen auf. Da wir in unserem Ansatz hinsichtlich der Reihenfolgeplanung immer zwei Aufträge (und damit zwei Klassen) paarweise miteinander vergleichen, ist dies aber keine hinderliche Einschränkung. Ordnet man alle Trainingsdatensätze im Raum an, spannt der Algorithmus eine Ebene bzw. bei mehr als 3 Parametern eine Hyperebene auf, die als Grenze zwischen den zwei Klassen dient. Neue Datensätze können dann relativ einfach anhand ihrer Position im Raum klassifiziert werden. Wenn die Trainingsdaten nicht linear trennbar sind, lässt sich der sogenannte Kernel-Trick anwenden, welcher die Daten in einen Raum mit höherer Dimension übersetzt, in welcher diese dann separierbar sind. Die dafür benötigte Kernelfunktion ist ein zusätzlicher Stellparameter der SVM und kann deren Leistung beeinflussen (Han et al. 2012; Boser et al. 1992).

## 3.4 Entscheidungsbäume

Für die Generierung von Entscheidungsmodellen sind Entscheidungsbäume sicherlich die bekannteste Methode. Ein Entscheidungsbaum besteht aus Knoten und Blättern. Von oben nach unten werden dabei entlang der Baumstruktur an den Knoten jeweils Entscheidungen über bestimmte Parameterwerte getroffen, sodass am Ende eine an den Blättern hinterlegte Klasse zugewiesen werden kann. Für den Aufbau von Entscheidungsbäumen existiert eine Vielzahl an verschiedenen Algorithmen. In dieser Arbeit wurde für die Induktion von Bäumen der Algorithmus CART (Classification and Regression Trees) verwendet. CART verwendet binäre Bäume, was bedeutet, dass jeder Knoten im Baum jeweils exakt zwei Kinder hat. Welcher Parameter an welchem Knoten im Baum abgefragt wird, hängt von der Reinheit der zu erwartenden Trennung ab, also dem Gewinn an Informationsgehalt (Entropie). Dies wird durch das sog. Splitkriterium, wie beispielsweise der Gini-Index (Breiman 1984), berechnet.

## 3.5 Künstliche neuronale Netze

Künstliche neuronale Netzte (KNN) sind vielfältig einsetzbar und in verschiedenen Varianten bzw. Topologien zu finden. Allgemein kann gesagt werden, dass KNN als hoch flexible Funktionsapproximierer genutzt werden können (Zell 1997). Der Grundansatz ist inspiriert vom menschlichen Gehirn. Die Basiskomponenten sind einfache Entitäten, die Neuronen, welche untereinander durch gewichtete Verbindungen verbunden sind. Im einfachsten Fall besitzt ein Neuron ein Aktivierungslevel, welches sich aus der Aktivierung der verbunden Neuronen ergibt, d.h. die Summe der gewichteten Eingangssignale darstellt. Durch eine im Neuron definierte Aktivierungsfunktion wird dessen Ausgangssignal auf Basis des Aktivierungslevels berechnet. In den hier genutzten Multilayer Perceptrons sind die Neuronen in Schichten angeordnet, wobei nur Verbindungen von Neuronen einer Schicht zur nächst höheren erlaubt sind. Alle Eingaben an das Netz werden über die Aktivierung von Neuronen der ersten Schicht realisiert. Die Ergebnisse sind in der Aktivierung der Neuronen der letzten Schicht abzulesen. Das Trainieren der Netzte kann mit verschiedenen Verfahren erfolgen, wobei die Gewichte von Verbindungen und die Parameter der Aktivierungsfunktionen angepasst werden. Das verbreitetste Verfahren für überwachtes Lernen ist das Back-Propagation-Verfahren (Zell 1997).

# 4 Testszenario und Ergebnisse der Versuche

Der Validierung des Ansatzes bzw. des Frameworks erfolgt in drei Phasen. In Phase 1 werden die (BDE-) Daten erzeugt. Da die Untersuchungen zunächst darauf abzielen, Methoden, Methodenparametrierungen und Einstellungen der Datentransformation zu testen und zu bewerten, wurde bewusst ein überschaubares Simulationsmodell zu Datengenerierung herangezogen. Die Simulation ermöglicht hierbei, schnell verschiedene Daten mit unterschiedlichen Eigenschaften zu erzeugen; so können u. a. schnell Daten für verschiedene Reihenfolgeregeln, welche dann approximiert werden sollen, generiert werden. Als Simulationssystem wurde Tecnomatix Plant Simulation gewählt. Im Modell (Abb. 2) wurden fünf Auftragstypen mit unterschiedlichen Attributen (z.B. Bearbeitungszeiten, Lieferzeiten, Rüstzuständen) definiert.

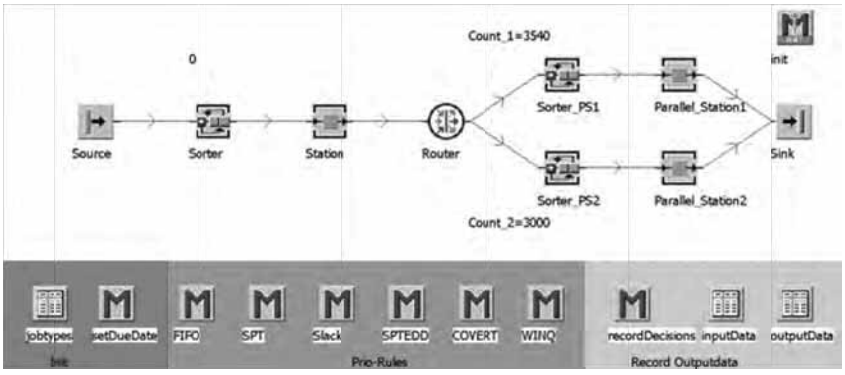



*Abbildung 2: Screenshot des Plant-Simulation-Modells zur Datengenerierung*

Die Reihenfolgeentscheidung erfolgt in einem Sortierpuffer vor der ersten Bearbeitungsstation. Im Sortierpuffer können verschiedene Reihenfolgeregeln (z.B. FIFO, SPT, EDD) hinterlegt werden, wobei je Datenerzeugungsphase nur eine sich nicht ändernde Regel zum Einsatz kommt. Des Weiteren wurde eine BDE-Daten-Erfassung emuliert.

Die 2. Phase umfasst den eigentlichen Prozess der Erzeugung der Entscheidungsmodelle wie in Kapitel 2 (vgl. Abb. 1) beschrieben. Hierbei kommen für jedes Trainingsdatenset aus Phase 1 alle vorgestellten Data-Mining-Methoden zum Einsatz, wobei je Methode verschiedene Parametrierungen getestet werden. Zusätzlich werden Versuche mit unterschiedlichen Datentransformationen durchgeführt. Der Großteil der prototypischen Implementierung des Frameworks erfolgte in MathWork MATLAB unter Nutzung der enthaltenen Toolboxen.

In der letzten Phase (Phase 3) werden die in Phase 2 erzeugten Modelle für die Validierungsdaten getestet, hierbei wird der Anteil der korrekten (binären) Entscheidungen an der Menge aller (binären) Entscheidungen gespeichert.

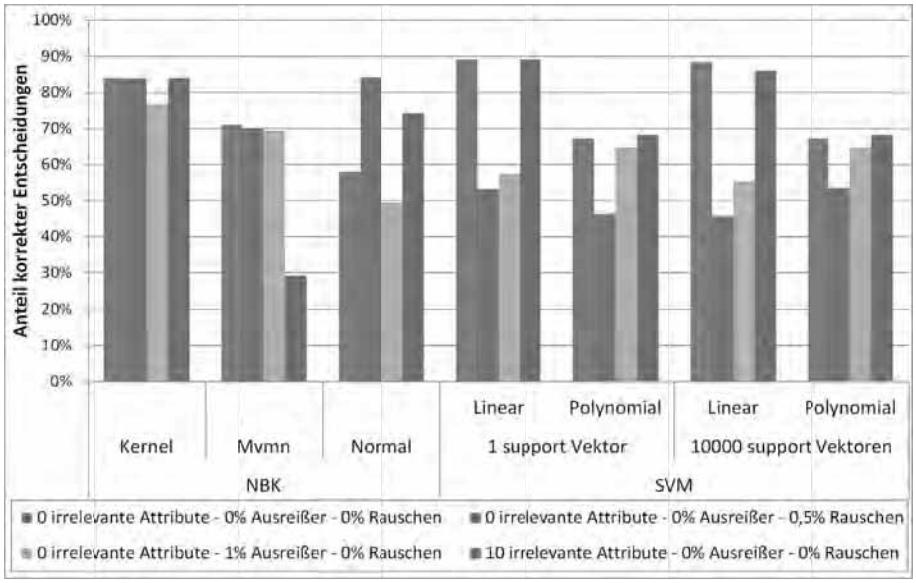



*Abbildung 3: Vergleich der Data-Mining-Methoden NBK und SVM für verschiedene Parametrierungen*

Erste Versuche und Auswertungen wurden für jede Data-Mining-Methode durchgeführt, um zum einen deren prinzipielle Eignung und zum anderen eine, für das Szenario, gute Parametrierung zu ermitteln. Hierbei wurde zugleich die Robustheit bzgl. der Datenqualität (Stufe A – blau; B – violett; C – grün; D – rot) untersucht. Beispielhaft sind die gemittelten Resultate für verschiedene Systemkonfigurationen (wie verschiedene Reihenfolgeregeln im Generatormodell) für die Methoden NBK und SVM in Abbildung 3 dargestellt.

Gut zu erkennen ist der hohe Einfluss der Parametrierung auf den Anteil der korrekten Entscheidungen. Des Weiteren zeigten die Versuche, dass einzelne Methoden

erhebliche Unterschiede bzgl. ihrer Robustheit gegenüber niedrigerer Datenqualität aufweisen. Auffällig ist, dass einzelne Methoden bei bestimmten Parametrierungen besonders bei bestimmten Datenverfälschungen deutlich unterdurchschnittlich abschneiden, z. B. reagiert der NBK mit multivariater Multinomialverteilung (Mvmn) besonders schlecht auf zusätzliche Attribute. Wird anstatt dessen eine Kernelfunktion genutzt, ist der SVM robust gegenüber sinkender Datenqualität. Der SVM hingegen schneidet bei Rauschen und Ausreißern prinzipiell unterdurchschnittlich ab. Die skizzierten Fakten sind auch bei allen weiteren Methoden entsprechend beobachtbar. Nebenergebnis dieser ersten Untersuchungen war die Erkenntnis, dass die Methoden sich deutlich in ihrer Performance des Lernalgorithmus unterscheiden. So wurden Zeiten gemessen, die je nach Datenqualität, -menge und Parametrierung von wenigen Sekunden für beispielswiese den k-NN bis hin zu mehreren Stunden für SVM und KNN lagen. Bei der späteren Nutzung der Entscheidungsmodelle hingegen waren kaum Performanceunterschiede erkennbar.

Neben der Ermittlung der besten Parametrierung für einzelne Methoden ist natürlich die Auswahl der besten Methode sowie Datentransformation Ziel der Versuche. Alle weiteren Versuche und Vergleiche nutzen die für jede Data-Mining-Methode ermittelte beste Parametrierung.

Abbildung 4 zeigt den Vergleich der fünf Data-Mining-Methoden. Gut ersichtlich ist, dass Entscheidungsbäume unter Anwendung des CART-Algorithmus und die Methode k-NN prinzipiell zu sehr guten Ergebnissen führen. Für den Anwendungsfall der AMG sind Werte von >95% korrekte Entscheidungen hervorragend.

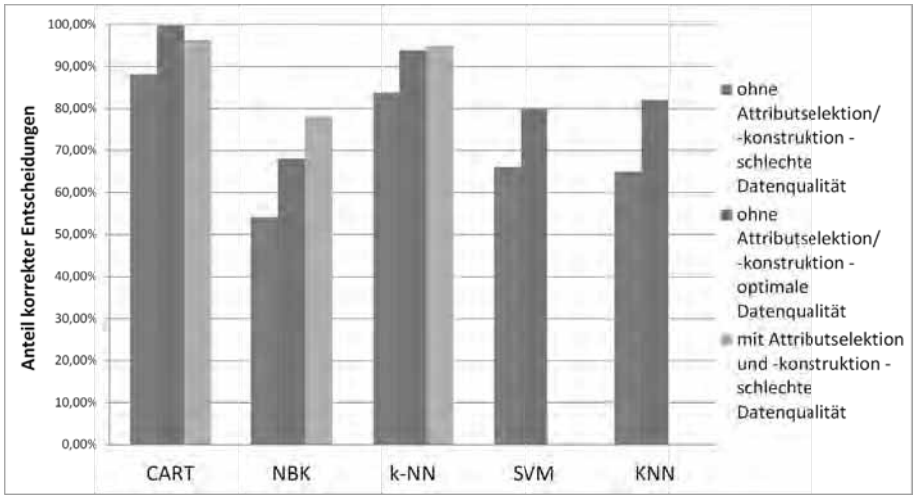



*Abbildung 4: Vergleich der verschiedenen Data-Mining-Methoden (je beste Parametrierung) für verschiedene Datenqualitäten und bei Einsatz der Attributselektion und -konstruktion*

Des Weiteren ist ersichtlich, dass Datentransformation, mit Attributselektion und -konstruktion, schlechte Datenqualität kompensieren kann. Für die Methoden SVM und KNN sind keine Daten bzgl. der Korrektheit bei aktivierter Attributselektion und -konstruktion sinnvoll zu erheben, da aufgrund der Laufzeit der Algorithmen

der im Framework vorgesehene evolutionärer Ansatz, welcher eine große Zahl von Lernvorgängen benötigt, versagt. Alle anderen Methoden profitieren aber deutlich von der Datentransformation.

Generell hat sich gezeigt, dass die Methode der Entscheidungsbäume unter Nutzung des CART-Algorithmus zu den besten Ergebnissen (bis zu 99% korrekte Entscheidungen) führt. Erstaunlich ist aber, dass der sehr simple k-NN, besonders bei Nutzung entsprechender Datentransformation, zu ähnlich guten Ergebnissen führt. Der NBK hingegen scheint eher ungeeignet für die Abbildung von Steuerungsregeln; in keinem Versuch, auch bei bester Datenqualität, konnten Ergebnisse mit 80% oder mehr korrekten Entscheidungen beobachtet werden.

## 5 Zusammenfassung und Ausblick

Die Ergebnisse der Untersuchung zeigen, dass ausgewählte Methoden des maschinellen Lernens selbst für komplexe Steuerungsregeln prinzipiell gute Approximationen bieten. Somit wurde veranschaulicht, dass die Anwendung von Data-Mining-Methoden eine effektive Lösung für das Abbilden von dynamischem Verhalten im Kontext der automatischen Modellgenerierung darstellt. In den bisherigen Untersuchungen wurden keinerlei Einschränkungen bzgl. möglicher abbildbarer Regeln sichtbar, weitere Untersuchungen in komplexeren Realwelt-Szenarien stehen aber noch aus. Auch lassen sich keine Aussagen darüber machen, wie gut und wie schnell sich die Algorithmen an ein sich änderndes Systemverhalten während der Trainingsphase anpassen können. Dies wäre in weiteren Forschungsarbeiten zu überprüfen.

Unser Ansatz, Reihenfolgeentscheidungen auf binäre Vergleiche zwischen zwei Aufträgen zurückzuführen, bringt dabei eine große Flexibilität mit sich. Auf der anderen Seite steigt damit jedoch der Rechenaufwand mit der Anzahl der zu vergleichenden Aufträge im System. Die Qualität der Entscheidungsmodelle schwankt stark in Abhängigkeit von der gewählten Methode und deren Konfiguration und Parametrisierung. Hier besteht im Rahmen weiterer Forschungsarbeit noch weiteres Optimierungspotenzial. Durch vorbereitende Datentransformationsschritte konnte zudem die Entscheidungsqualität weiter erhöht werden. Selbst bei schlechter Datenqualität der Trainingsdaten (d. h. Rauschen und absichtliche Fehlklassifizierungen) konnte hierbei mit Hilfe des CART-Algorithmus eine Korrektheit von bis zu 99% erreicht werden.

Allerdings richtet sich das von uns verwendete Kriterium zur Bewertung der Entscheidungsqualität rein nach der Anzahl der korrekt getroffenen Entscheidungen. Eine interessante Erweiterungsmöglichkeit zur Bewertung der Verfahren wäre hierbei, die generierten Entscheidungsmodelle wieder zurück in das Ausgangsmodell zu implementieren und somit zusätzlich zu der Korrektheit der Entscheidung zu messen, ob und wie stark sich eine falsch getroffene Entscheidung überhaupt auf das System auswirkt. Eine faktisch falsch getroffene Entscheidung muss sich nicht zwangsläufig in veränderten Leistungskennzahlen des Systems niederschlagen. Des Weiteren ist zu überprüfen, ob die Leistungsfähigkeit der untersuchten Algorithmen in komplexeren Realszenarios bestätigt werden kann.

## Literatur

Altman, N. S.: An introduction to Kernel and Nearest-Neighbor Nonparametric regression. The American Statistician (1992) 46, S. 175–185.

Bergmann, S.: Automatische Generierung adaptiver Modelle zur Simulation von Produktionssystemen. Ilmenau: Univ.-Verl. Ilmenau 2013.

Bergmann, S.; Strassburger, S.: Challenges for the automatic generation of simulation models for production systems. In: Proceedings of the International Summer Simulation Multiconference. Ottawa (Canada), 12.-14. Juli 2010, S. 545–549.

Bergmann, S.; Stelzer, S.; Strassburger, S.: On the use of artificial neural networks in simulation-based manufacturing control. Journal of Simulation 8 (2014) 1, S. 76–90.

Bortz, J.; Schuster, S.: Statistik für Human- und Sozialwissenschaftler, 7 .Auflage. Berlin: Springer 2001.

Boser, B. E.; Guyon, I. M.; Vapnik. V. N.: A training algorithm for optimal margin classifiers. In: Haussler, D. (Hrsg.): Proceedings of the fifth annual workshop on Computational learning theory, Pittsburgh (USA), 27 -29 July 1992, S. 144–152.

Bracht, U.; Rooks, T.; Adrian, R.: Virtuelle Logistikplanung für die Montage im Rahmen der Digitalen Fabrik. In: Rabe, M. (Hrsg.): Advances in Simulation for Production and Logistics Applications. Stuttgart: Fraunhofer IRB 2008.

Breiman, L.: Classification and regression trees. New York: Chapman & Hall 1984.

Fayyad, U. M.: Advances in knowledge discovery and data mining. Menlo Park, Calif.: American Association for Artificial Intelligence; AAAI 1996.

Han, J.; Kamber, M.; Pei J.: Data mining concepts and techniques, 3. Auflage. Amsterdam: Elsevier/Morgan Kaufmann 2012.

Pinedo, M.: Scheduling. Theory, algorithms, and systems, 4. Auflage. New York: Springer 2012.

Reinhart, G.; Gyger, T.: Identification of implicit strategies in production control. In: Proceedings of the Industrial Engineering and Engineering Management 2008 (IEEM 2008). Singapore, 8.-11. Dezember 2008, S. 302–306.

Rennie, J. D.; Shih, L.; Teevan, J.; Karger, D. R.: Tackling the poor assumptions of naive bayes text classifiers. In: Fawcett, T.; Mishra N. (Hrsg.): Proceedings of the Twentieth International Conference on Machine Learning, Washington, DC (USA), 21-24 August 2003, S. 616–623.

Schönherr, O.; Rose, O.: First steps towards a general SysML model for discrete processes in production systems. I: Rossetti, M. D.; Hill, R. R.; Johansson, B.; Dunkin A.; Ingalls, R. G. (Hrsg.): Proceedings of the 2009 Winter Simulation Conference (WSC2009). Austin (USA), 13.-16. Dezember 2009, S. 1711–1718.

Selke, C.: Entwicklung von Methoden zur automatischen Simulations-modellgenerierung. München: Utz 2005.

Wiedemann, T.: Database oriented modeling with simulation microfunctions. In: Farrington, P. A.; Nembhard, H. B.; Sturrock, D. T.; Evans, G. W. (Hrsg.): Proceedings of the 1999 Winter Simulation Conference (WSC1999). Phoenix (USA), 5.-8. Dezember 1999, S. 586–590.

Zell, A.: Simulation Neuronaler Netze. München: Oldenbourg 1997.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Integration of Ergonomic Analysis into Simulation Modelling of Manual Operations

## *Integration der ergonomischen Analyse in die Simulationsmodellierung von manuellen Operationen*

Alireza Golabchi, SangUk Han, Simaan AbouRizk, University of Alberta, Edmonton (Canada), alireza1@ualberta.ca, sanguk@ualberta.ca, abourizk@ualberta.ca

Abstract: Considering the prevalence of simulation modelling in production design, integrating ergonomic safety considerations into simulation models can help production planners identify and prevent ergonomic risks before workers encounter unsafe conditions, in addition to productivity improvements. This study proposes integrating ergonomic analysis into micro-motion-level simulation modelling by developing a simulation modelling template that enables simultaneous estimation of the standard duration of manual activities, as well as biomechanical risk factors inherent in the production design. The proposed approach enables modelling of manual activities at a human motion level in order to provide standard task durations in addition to initial insight into the level of ergonomic risks associated with a planned production scenario. As a case study, the proposed methodology has been implemented in a construction steel fabrication company and the results confirm the validity of the approach in identifying ergonomic risks associated with a production design.

## 1 Introduction

Simulation modelling is an effective tool for design and planning of construction operations. A digitalised model representing a physical system allows prediction of the system performance in a real world application, as well as understanding of the system behaviour under various conditions. In construction, advances in simulation modelling techniques have been made to estimate the duration of projects and the productivity of different operational systems. However, it still remains challenging to analyse manual operations due to the uncertainties in modelling human movements and estimating the cycle time of manual tasks. Furthermore, the impact of the designed workplace and workflow on human capacity (e.g., durations and frequencies of physical tasks) for ergonomic safety analysis has not been thoroughly considered. In this regard, integrating ergonomic considerations into simulation models can help production planners identify and prevent ergonomic risk factors

affecting human movements before the workers encounter the unsafe conditions, as well as understand the impact of production systems on human behaviour.

To address this issue, this study proposes integrating ergonomic analysis into micro-motion-level simulation modelling. A discrete-event simulation modelling template has been developed which enables simultaneous estimation of the standard duration of manual activities and biomechanical risk factors inherent in the production design. Specifically, the concept of a Predetermined Motion Time System (PMTS) is adopted to model human movements (e.g., retrieving, carrying, placing an object) and estimate cycle time of manual tasks based on the unit cycle time of each movement. The human postures and motions involved in the manual operation are then evaluated from an ergonomic perspective, using an ergonomic assessment tool. The proposed approach thus provides the user with initial insight into the level of ergonomic risks associated with a planned production scenario. By inputting basic design data available during workplace design, the simulation model will report on the ergonomic risks associated with the task based on postural strains, external force exertion, and frequency of the task. As a result, the simulation model not only provides information regarding the standard duration of the envisaged production, but also reports on the level of ergonomic risks associated with the design; which potentially allows for the in-depth understanding of the trade-off between productivity and safety. Consequently, the proposed simulation model enables modifying the workplace design in order to achieve both ergonomically safe and productive operation systems.

## 2 Method

This research uses PMTS and ergonomic assessment in order to develop an integrated micro-motion-level simulation modelling environment. The framework of the research is shown in Figure 1.

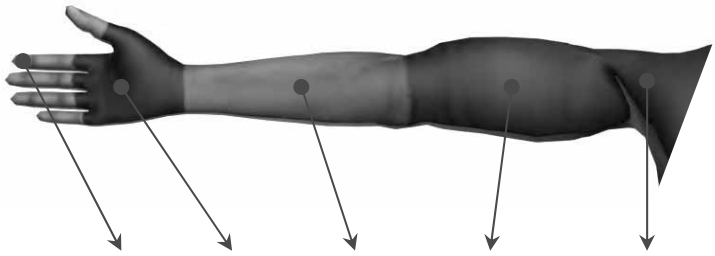
*Figure 1: Integrated micro-motion-level simulation framework*

### 2.1 Predetermined Motion Time System

PMTSs are used in various industries to determine the amount of time required to perform a manual process by breaking down a task into sequential motions carried out by the worker to complete the task (Farrell 1993). By using PMTSs, production planners can make more productive decisions on designing work processes, as they

are provided with a standard duration in which a worker is expected to complete an action, as well as a comparison of the efficiency of different alternatives of performing the operations (ILO 1986). Researchers have studied large samples of various manual activities in order to develop PMTS methods that can provide the standard time required to carry out a manual task. Currently, the most widely used PMTSs include Methods-Time Measurement (MTM) (Maynard et al. 1948), Modular Arrangement of Predetermined Time Standards (MODAPTS) (Heyde 1966), and Maynard Operation Sequence Technique (MOST) (Zandin 1980). This study uses the MODAPTS method for modelling worker motions due to its simplicity and quick application. However, the other methods can also be similarly applied using the same approach.

MODAPTS classifies the complexity of manual actions (e.g., get, move, put) by the amount of sensory feedback required to carry out the action rather than the geometrical properties of the material to be handled. It was developed in an attempt to reduce the time and cost required to apply prior PMTS methods and create a system that could be memorised, is simple to learn and apply and yields consistent results (Heyde 1989). MODAPTS assumes that the duration of any body motion can be expressed as a multiple of the time required to move a single finger, called a MOD. The MOD unit, equal to 0.129 second, is used assuming that the movements are carried out with minimal energy expenditure and that the movement duration is proportional to the fifth root of the moment of inertia of the body part moved (Minter 1983). Figure 2 shows the moving distance of the different body parts and the corresponding MOD values for MODAPTS.



| Body Part | Finger | Hand | Arm | Whole arm | Extended arm | Trunk |
|---|---|---|---|---|---|---|
| MOD | 1 | 2 | 3 | 4 | 5 | 7 |
| Distance Moved | 2.5 cm | 5.0 cm | 15 cm | 30 cm | 45 cm | 75 cm |

*Figure 2: Basic movements as defined by MODAPTS*

In order to analyse a manual task using MODAPTS, the task has to be broken down into basic motions, known as modules, that can be described using MODAPTS classes, and for each motion, a MOD value has to be assigned. By adding the MOD values and converting the sum of MODs to seconds, the total amount of time required to complete the task is obtained. MODAPTS codes are used to represent different classes of movements which describe the type of motion. Some basic

classes include: Get (G), Move (M), Put (P) and Use (U). For example, a worker moving his hand 15 centimetres to pick up a wrench is represented by the MODAPTS code M3G3. M3 represents the hand movement motion that takes 3 MODs, and G3 represents performing an impeded grasp which also takes 3 MODs. For the purpose of this study, the MODAPTS method is integrated into discrete-event simulation in order to enable the modelling of manual activities in simulation models of construction operations.

## 2.2 Ergonomic Assessment

Various ergonomic assessment methods have been developed in order to evaluate the risks associated with manual tasks and identify unsafe actions that can lead to Work-related Musculoskeletal Disorders (WMSDs). These methods consider human posture as well as external risk factors (e.g., frequency and duration) to provide a global risk assessment of a motion, which is used to take required corrective actions (Golabchi et al. 2015b). Different methods prioritise different factors and focus on various body parts in order to assess different cases of manual handling activities. The cube model (Kadefors 1994) is one of these ergonomic assessment tools which provides an ergonomic evaluation of a manual task by inputting basic design data. This study uses the cube model for ergonomic evaluation due to its simplicity and efficiency, which make it suitable for incorporating into simulation modelling.

The cube model was developed based on the premise that the risk of acquiring WMSDs is mainly dependent on three interrelated variables (Laring et al. 2002). These variables include: working posture, muscular force, and time. For each of these factors, three levels of demands are defined based on severity, including low demand, medium demand, and high demand. These demand levels can be assigned using scientific evidence (e.g., Snook and Ciriello's (1991) tables for maximum acceptable weight limits), or by using a consensus approach to identify the severity of each demand (Kadefors 1993). The three factors (i.e., posture, force and time) are considered as axes of a cube, yielding 27 sub cubes, which is the number of possible combinations. For each combination, a score is calculated by multiplying the scores of each of the three demands, assuming 1, 2, and 3 as low, medium, and high demand, respectively (Kadefors 1997). A manual task is considered ergonomically acceptable for a final score of less than 5, conditionally acceptable for a score between 5 and 10, and ergonomically unacceptable for a score of higher than 10.

## 2.3 Integrated Micro Motion Level Simulation

The authors use Simphony (Hajjar and AbouRizk 1999) as the platform for simulation modelling as it enables development of special purpose templates that can encompass elements that work in conjunction with the general template. For the purpose of this study, an integrated micro-motion-level simulation template is developed that consists of elements that enable applying the MODAPTS standard in the simulation environment, without requiring knowledge about the details of implementing MODAPTS, and also provide insight into the level of ergonomic risks associated with manual tasks of the model, based on the principles of the cube model. The modelling elements developed include one primary element, named MHE (Manual Handling and Ergonomics), and six secondary elements. The MHE element requires basic design data pertaining to a manual task as input and

calculates the corresponding MODAPTS duration for the task and uses it for simulation purposes, and also reports on the level of ergonomic risks for each manual task. Figure 3 shows the design inputs required for the MHE element. As shown in Figure 3, the required inputs include design information that a production planner has when designing or redesigning a process.

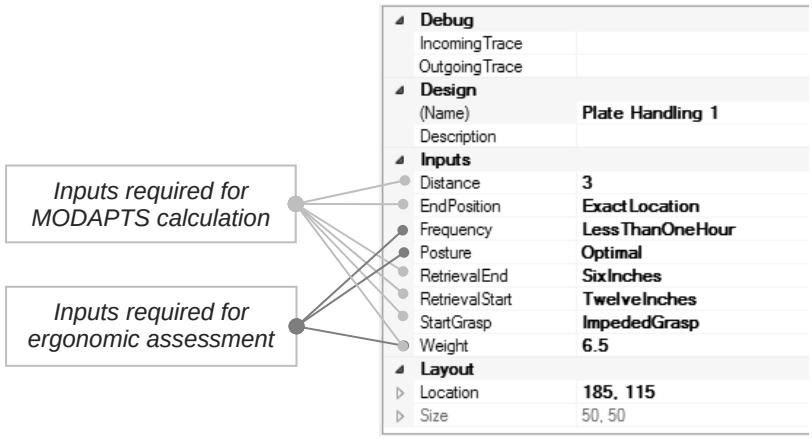



*Figure 3: Inputs of the MHE element*

The six secondary elements represent lower level manual activities that are designed based on the MODAPTS classes and include: Move, Get, Put, Walk, BendAndArise and SitAndStand. For example, the Move element can represent the motion of a worker moving his hand to grasp or put down an object, with the required input being the distance between the worker's hand and the object. These elements are designed and implemented to provide more flexibility in designing detail-level manual motions, in cases where a designer intends to investigate the effect of changing the attributes of one motion on the simulation results (Golabchi et al. 2015c). Furthermore, these models are used when the model requires placing one lower-level motion between non-manual elements. Figure 4 shows the Simphony environment and the elements of the special purpose template.

By using the developed elements, designers will have the opportunity to use reliable estimates of the duration required to perform a manual task within the simulation environment, without requiring prior knowledge about the details of predetermined motion time systems. Furthermore, designers are provided with an assessment of the level of ergonomic risks associated with each manual activity, which enables them to prevent WMSDs during the workplace design phase.

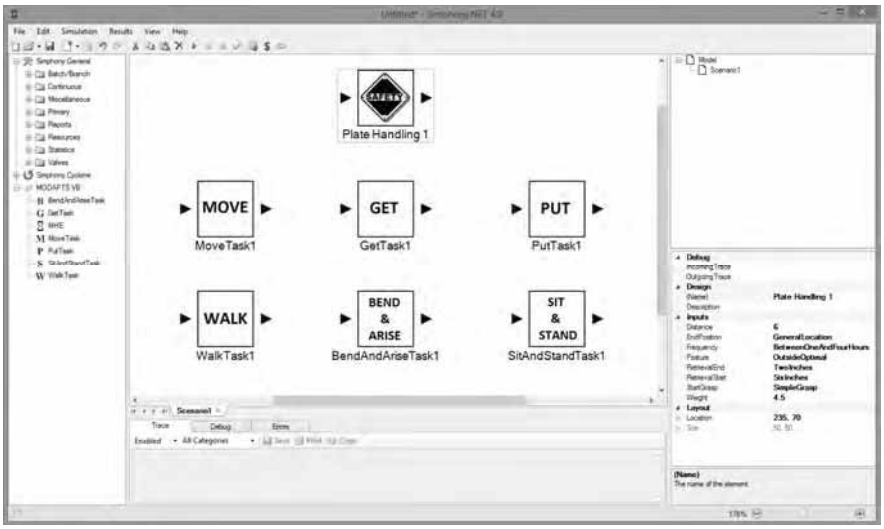



*Figure 4: Simphony modelling environment*

## 3 Implementation: Case Study

In order to implement and validate the proposed approach, the integrated micro motion level simulation has been used to model a steel plate handling task in a steel fabrication and construction service provider company in Canada (Golabchi et al. 2015c). Due to the nature of manual tasks in the construction industry, workers are frequently involved in physically challenging activities. This makes the integrated simulation approach suitable for modelling these manual tasks since both a standard duration for worker actions and a feedback on the level of ergonomic risks associated with the action is obtained. The steel plate handling task investigated as the case study includes a worker removing a steel plate from a drilling machine after the drilling is done, carrying the plate to designated bins, and placing it in the appropriate spot inside the bin. This cycle can be modelled for simulation purposes using the developed MHE element. The drilled plates have various dimensions and weights with an average size of 40 cm x 40 cm x 5 cm and an average weight of 20 kilogrammes. The task has been observed at the jobsite in order to collect the required inputs for the MHE element. Table 1 shows the inputs required and the description of each input for modelling the plate handling task. Although the proposed approach can be highly effective in designing new processes, an existing manual task has been selected for the case study in order to firstly enable studying the validity of the approach and secondly demonstrate its functionality in redesigning existing operations, besides designing new processes. It should be noted that the MHE element modelling the steel plate handling task will be used as part of a larger simulation model in conjunction with other simulation modelling elements, representing the whole steel fabrication operation. The steel plate handling activity can also be broken down into lower level motions and be modelled using the secondary elements.

*Table 1: Inputs required for modelling the steel plate handling task*

| *Input* | *Description for steel plate handling* | *Alternatives* |
|---|---|---|
| Distance | The distance that the worker carries the plate. | - |
| EndPosition | The precision required for placing the plate at the destination. | GeneralLocation, WithTidiness, ExactLocation |
| Frequency | The frequency of the task per day. | LessThanOneHour, BetweenOneAndFour, MoreThanFourHours |
| Posture | The optimality of the worker's posture while handling the plate. | Optimal, NearOptimal, OutsideOptimal |
| RetrievalEnd/ RetrievalStart | The distance between the plate and the worker's hands when retrieving/placing the plate. | OneInch, TwoInches, SixInches, TwelveInches, EighteenInches, ThirtyInches |
| StartGrasp | The ease of grasping the plate. | SimpleGrasp, ImpededGrasp |
| Weight | The weight of the plate. | - |

Validating the proposed approach using the case study involves two steps: (1) comparing the MODAPTS duration calculated by the model with actual time data collected from the jobsite, and (2) investigating the reliability of the ergonomic assessment provided by the simulation model. In order to perform the first part of the validation study, the actual duration for the worker to complete the task is also recorded for each instance, besides collecting the data related to the inputs shown in Table 1. Table 2 shows the results of running the simulation model using the MHE element for 10 instances of the plate handling task as an example. The MODAPTS code and duration are also calculated manually to ensure that the simulation model is calculating the durations accurately, and the results are consistent with the durations calculated by the simulation model, as shown in Table 2. A correlation analysis has been carried out and a Pearson correlation coefficient of 0.956 and a Spearman correlation coefficient of 0.913 is calculated between the actual time and the simulation MODAPTS time, which indicates very strong association between the two data sets.

In order to investigate the reliability of the ergonomic assessment results from the developed simulation template, for each instance of the steel plate handling task, the level of ergonomic risks associated with the instance is analysed using the Rapid Upper Limb Assessment (RULA) (McAtamney and Corlett 1993) method. RULA is a widely-used approach to ergonomic assessment which focuses on human posture and also considers the load involved as well as the frequency of the task (Golabchi et al. 2015a). The result of a RULA analysis is a total score between 1 and 7, with 1 and 2 indicating acceptable posture, 3 and 4 indicating that changes may be needed, 5 and 6 indicating that changes are required soon, and 7 indicating that immediate

investigation and modification is required. Since the result of a cube model analysis is represented in one of the three assigned categories (i.e., acceptable, conditionally acceptable, and unacceptable), the RULA scores are also categorised as 1 and 2 representing safe motion, 3 and 4 indicating further investigation required, and 5, 6 and 7 specifying unsafe motion, to enable comparison of the results. Table 3 shows the result of the ergonomic analysis for 10 instances of the steel plate handling task. The results indicate consistency between the output of the ergonomic assessment from the simulation model and the RULA analysis, which confirms the reliability of the approach in evaluating the ergonomic risks of manual activities.

*Table 2: Comparison between actual time and simulation time for case study*

| *Instance* | *Duration (Sec)* | | *MODAPTS Code (Manual)* | *MODAPTS Duration (Manual)* |
|---|---|---|---|---|
| | *Actual* | *Simulation* | | |
| 1 | 8.2 | 7.095 | M2G1L3W47M2P0 | 7.095 (=55*0.129) |
| 2 | 6.7 | 6.966 | M3G3L3W39M4P2 | 6.966 (=54*0.129) |
| 3 | 6.2 | 6.450 | M3G3L3W35M4P2 | 6.45 (=50*0.129) |
| 4 | 6.8 | 6.450 | M2G1L3W39M5P0 | 6.45 (=50*0.129) |
| 5 | 5.4 | 5.934 | M2G1L3W35M5P0 | 5.934 (=46*0.129) |
| 6 | 5.6 | 6.063 | M3G1L3W35M5P0 | 6.063 (=47*0.129) |
| 7 | 6.9 | 6.321 | M3G1L3W35M5P2 | 6.321 (=49*0.129) |
| 8 | 4.1 | 4.386 | M3G3L3W23M2P0 | 4.386 (=34*0.129) |
| 9 | 6.1 | 5.418 | M2G3L3W31M1P2 | 5.418 (=42*0.129) |
| 10 | 5.8 | 5.676 | M3G3L3W31M2P2 | 5.676 (=44*0.129) |

*Table 3: Comparison between results of ergonomic assessment and RULA*

| *Instance* | *Cube model result* | *RULA* | |
|---|---|---|---|
| | | *Score* | *Interpretation* |
| 1 | Conditionally acceptable | 3 | Further investigation |
| 2 | Unacceptable | 6 | Unsafe |
| 3 | Unacceptable | 6 | Unsafe |
| 4 | Conditionally acceptable | 3 | Further investigation |
| 5 | Unacceptable | 6 | Unsafe |
| 6 | Unacceptable | 6 | Unsafe |
| 7 | Conditionally acceptable | 3 | Further investigation |
| 8 | Conditionally acceptable | 3 | Further investigation |
| 9 | Unacceptable | 6 | Unsafe |
| 10 | Unacceptable | 6 | Unsafe |

The results of the case study analysis indicate that the integrated micro motion level simulation approach can be effectively used by process designers and production planners to provide firstly a standard duration for manual tasks and secondly an insight into the level of ergonomic risks associated with a design. The correlation between the actual and simulation time can be used as a benchmark for worker's performance which enables evaluating the efficiency of operations for work process design and redesign purposes.

In future work, the authors will consider linking human motion data to the elements of the developed special purpose simulation template in order to enable more accurate ergonomic evaluation and mitigation of ergonomic risks while redesigning ongoing operations. Furthermore, more data from jobsites will be collected in order to formulate the conversion of standard MODAPTS time to realistic jobsite time by accounting for inefficiencies in carrying out manual tasks. The result can be used to develop efficiency factors for various types of manual activities.

## 4 Conclusion

Simulation modelling is an effective tool for design and planning of operations since a digitalised model representing a physical system allows for prediction of the system performance in the real world, as well as for understanding system behaviour under various conditions. In construction, advances in simulation modelling techniques have been made to estimate the duration of different types of projects and the productivity of various operational systems. However, it still remains challenging to analyse manual operations due to the uncertainties in modelling human movements and estimating the cycle time of manual tasks. Furthermore, although production planners intend to take ergonomic considerations into account while planning operations, there is a lack of tools to enable ergonomic safety assessment during workplace design. In this regard, this study proposes integrating ergonomic analysis into micro-motion-level simulation modelling, which enables simultaneous estimation of the standard duration of manual activities and ergonomic risks inherent in production design. A special purpose simulation template is developed that enables modifying a workplace design in order to achieve both ergonomically safe and productive operation systems and potentially allows for the in-depth understanding of the trade-off between productivity and safety. This study provides industry practitioners with the opportunity to incorporate reliable representations of manual activities into simulation models of operations and take advantage of standard PMTS methods as well as ergonomic assessment tools, without requiring prior knowledge about these systems.

## References

Farrell, J.M.: Predetermined Motion-Time Standards in Rehabilitation, A Review. Work: A Journal of Prevention, Assessment and Rehabilitation 3.2 (1993), pp. 56-72.

Golabchi, A.; Han, S.; Fayek, A. Robinson (2015a) A Fuzzy Logic approach to posture-based ergonomic analysis for field observation and assessment of construction manual operations. Canadian Journal of Civil Engineering (2015). (in review, submitted March 2015)

Golabchi, A.; Han, S.; Seo, J.; Han, S.; Lee, S.; and Al-Hussein, M. (2015b) An automated biomechanical simulation approach to ergonomic job analysis for workplace design. Journal of Construction Engineering and Management 141 (2015) 8, online.

Golabchi, A.; Han, S.; AbouRizk, S.M. (2015c) Integration of predetermined motion time systems into simulation modelling of manual construction operations. In: Proceedings of the 5th International/11th Construction Specialty Conference, Vancouver (Canada), June 2015. (accepted for publication)

Hajjar, D.; AbouRizk. S.: Simphony: An environment for building special purpose construction simulation tools. Proceeedings of the 1999 Winter Simulation Conference. Phoenix (AZ) 1999, pp. 998-1006.

Heyde, G.C.: Modapts. Industrial Engineering (1966), pp. 11-15.

Heyde, G.C.: Concepts and history of ModaptsPlus, Industrial Engineering 30 (1989), pp. 24-29.

ILO: Introduction to work study (3rd Edition). Geneva: International Labour Organization (ILO) 1986.

Kadefors, R.: A model for assessment and design of workplaces for manual welding. In: Marras, W.S., Karwowski, W., Pacholski, L. (eds.): The Ergonomics of Manual Work. London: Taylor & Francis 1993.

Kadefors, R.: An ergonomic model for workplace assessment. In: Proceedings of the IEA'94, Vol. 5. International Ergonomics Association, Toronto (Canada) 1994, pp. 210–212.

Kadefors, R.: Evaluation of working situations using the cube model approach. In: Proceedings of the IEA'97, Vol. 4. International Ergonomics Association, Tampere (Finland) 1997, pp. 174–176.

Laring, J.; Forsman, M.; Kadefors, R.; Örtengren, R.: MTM-based ergonomic workload analysis. International Journal of Industrial ergonomics 30 (2002) 3, pp. 135-148.

Maynard, H.B.; Stegemerton, G.J.; Schwab, J.L.: Methods-time measurement. Boston: McGraw-Hill 1948.

McAtamney, L.; and Corlett, E.N.: RULA: A survey method for the investigation of work-related upper limb disorders. Applied Ergonomics 24 (1993) 2, pp. 91–99.

Minter, A.L.: Modapts. Management Services 27 (1983), pp. 8-11.

Snook, S.H.; Ciriello, V.M.: The design of manual handling tasks: Revised tables of maximum acceptable weights and forces. Ergonomics 34 (1991) 9, pp. 1197–1213.

Zandin. K.: MOST. Work Measurement Systems. New York: Dekker 1980.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Performing Scenario Simulations in Construction by Using Standard Software

## *Durchführung von Szenariosimulationen mit Standardsoftware im Hochbau*

Hans-Joachim Bargstädt, Immo Feine, Bauhaus-Universität Weimar, Weimar (Germany), hans-joachim.bargstaedt@uni-weimar.de, immo.feine@uni-weimar.de

Abstract: The authors address the application of scenario simulation in construction from the application side. It is the point of view of the practitioner, who intends to apply process simulation for analysing the realisation phase of a construction process by using standard simulation software. The authors start from the classical application areas for process simulation in production and logistics and describe how the user in the construction industry can profit from scenario simulations taken from standard environment. Typical transfer problems are described as well as examples of simulation models are given to address practical limitations and chances based on the state-of-the-art software package Enterprise Dynamics.

## 1 Problem Statement

Generally it is expected that construction processes are optimised to a certain extent with regard to economical or time related figures. However, this is optimisation is very difficult because the underlying decisions are commonly based on incomplete information about the processes, their boundary conditions and also the interferences between the individual processes (Fiedler and Peldschus 1989). The parameters that govern construction processes are as a rule stochastical values which reflect the uniqueness of the construction in terms of weather and soil conditions, involved machines or personnel, etc.

In practice and for planning purposes estimates are taken which involve experiences acquired from earlier projects. However, it remains unclear if the results of such a planning fulfil the expectations towards the required degree of optimisation sufficiently. A potential solution to this problem can be seen in the use of simulation tools. In other industries (e.g. stationary industry) these virtual tools are inherent part of production planning. Application areas of simulation in the construction industry, however, have been shown by different authors, e.g. Voigtmann (2014), Berner et al. (2013) and Weber (2007). Whereas on scientific level the use of simulation in construction has been widely discussed, its application in practice is far behind.

Different authors report about the reasons for this situation and mention

- the lack of expertise in the companies to use simulation tools,
- the inadequateness of tools coming from the stationary industry in relation to the specific requirements of the construction industry,
- the enormous effort needed for the creation of adequate simulation models for the construction sector,
- the time needed for any adaptation of the simulation model to changes in the real system and
- the provision of valid input data for the simulation

as the main obstacles for applying process simulation tools in construction (Kochkine 2014; Kugler 2012).

In this paper the authors address the issue from the application side. It is the point of view of the practitioner, who is willing to apply process simulation for analysing the realisation phase of a construction process by using standard simulation software.

## 2 Literature Review

In the scientific literature many of the obstacles for applying process simulation to construction processes have been addressed by researchers. Standard software has been adapted by the provision of special toolboxes containing elements for the modelling of specific construction activities (Voigtmann 2014; Weber 2007). Some researchers developed new simulation tools focusing on individual problems of specific construction projects, e.g. Steinhauer (2007), Al-Hussein et al. (2006), Ruwanpura et al. (2000), Hajjar and AbouRizk (1999). Others attempted to address the question of data provision by linking simulation tools to CAD-software or by importing input data from building information models (Lu and Olofsson 2014; Wang et al. 2014; König et al. 2012; Kugler 2012; Xu et al. 2003). However, many of the scientific proposals have so far not been implemented in any commercial software packages. Especially for providing input data for the process simulation in construction, individual and hand-made solutions like spreadsheet processing programs are used for collecting and preparing the data (Kochkine 2014).

Weber (2007) describes an approach on the use of the simulation software Enterprise Dynamics for the simulation of construction logistical processes. The author uses 3D-CAD data to determine the material quantities of the construction site. These quantities have been connected to construction elements and linked with information of the time schedule. Weber creates a simulation model containing the logistical elements that are necessary to transport the calculated quantities of material to its place of use. By doing so the author uses simulation to test and evaluate different supply and warehousing strategies on a specific building construction site. The simulation runs and gives back throughput, waiting and idle times. The simulation acts as a support tool for decision making in the areas of site layout and logistics. Weber describes that the shortening of construction times by creating more parallel jobs on site requires higher coordination efforts. He argues that simulation tools in construction can be used as a means to increase productivity. The obstacles for the application of simulation tools are poor cost effectiveness, missing data input and others.

Pitsch (2011) reports that the challenges for commercial simulation tools for the application in one-of-a-kind processes are typical for the construction industry. He points out that especially the nature of one-of-a-kind processes is in conflict with standard tools which have more or less been developed for universal applications. Therefore standard simulation tools must be equipped with additional functions that allow an easy adaptation of its preset elements to the unique requirements of the processes to be modelled in construction.

Voigtmann (2014) reports on her research based on the use of the simulation software Plant Simulation (Siemens 2013). This has been expanded by the Simulation Toolkit Shipbuilding (STS package). STS provides predefined simulation objects developed for the modelling of shipbuilding processes. Finishing processes in shipbuilding are to a large extent similar to those of building construction. Voigtmann used a highly complex simulation program to model the logistical and technological construction processes in the finishing phase of a multi-storey office building. The dynamic simulation model that has been created extends the standard functionalities of the simulation software by the use of hard- and soft constraints (Beißert 2012), space management (Marx and König 2010), logistics control and others. Based on the number of parameters in the model and the individual range that each parameter can take the simulation time of the model changes dramatically. Voigtmann concludes that a sensitivity analysis of the impact of single parameters to the simulation time (time until completion of construction) can help limiting the total number of parameters as well as restrict the value range of each parameter. Voigtmann points out that this approach can contribute to some less complex simulation models and thus facilitate the application of simulation in construction.

Kochkine (2014) introduces a concept for process based production planning for optimisation purposes in construction. He argues that the simulation of production processes in construction is still challenging especially in the area of data collection, data import and model creation. Automated approaches which make use of building information models can be regarded as a chance to overcome some of the problems, but only for larger construction elements. Applying simulation to the finishing phase of construction projects which is governed by high construction costs, many small tasks (e.g. technical building services) and a relatively large number of different crafts on site is rather complicated to be suitable for practice.

Berner et al. (2013) discuss the use of simulation in manufacturing planning. They illustrate different applications based on Plant Simulation supplemented by the STS. It might be debatable whether the application of STS can be regarded as standard software because STS contains custom made simulation functionalities which facilitate the modelling in shipbuilding. However, by presenting the technical possibilities of simulation in shipbuilding, it gives a strong impression on how far the use of simulation in construction can go starting from a continuous enrichment of standard simulation tools. The following application areas of simulation in construction have been described:

- Determination of optimum work sections on a daily basis;
- Analysis of alternative operation methods;
- Determination of resource requirements and personnel;
- Analysis of different logistical concepts by considering the dynamic environment.

Considering the above mentioned application fields it shows that the use for process simulation is well described on the scientific level. The benefits of simulation tools have been clearly addressed. In technical respect different simulation programs are in use. By disregarding special simulators for isolated problems the use of conventional simulation tools in construction is discussed controversially. The potential for the application of process simulation as shown in scientific papers is still problematic if transferred to general purpose problems in construction. In order to address this issue a small series of simulation projects has been carried out which will be described in the following.

# 3 Use of Standard Simulation Tools

## 3.1 Introduction

### *3.1.1 Standard Simulation Software*

The term "standard simulation software" as defined by the authors describes publically available software offered on the market, which can be used for the creation of multipurpose simulation models by using standard elements. The software should not be specialised on specific application fields nor does it have branch-specific solutions (Pitsch 2011).

There are two advantages by the use of standard software: 1st it offers broad application fields by providing a large set of standard elements. 2nd the software developing power can be higher because the number of clients from different applications can be considerably stronger. At first glance the software doesn't suit the requirements of a construction company because its prevailing problems are varying construction sites and conditions. Another disadvantage of using standard software is that the creation of individualised simulation models is more time consuming and requires higher adaptation skills of the simulation engineer than by using one-purpose simulation software.

A market review of simulation software for construction purposes did not detect a specific tool that can be classified as a standard in this field and which might be explicitly designed for modelling of construction processes already. In contrary, simulation tools were found that were coming from the stationary industry and which have been used in scientific applications for the modelling of construction processes already. These tools contain predefined elements which support modelling situations in the stationary industry and which show a high affection for adaptation to be applied in the construction industry.

### *3.1.2 Perspective*

Basically a construction company has to take the make or buy decision, when deciding on technical tools and equipment. The company could either hire a consultant who offers simulation services or the company could purchase simulation software and provide it to its engineers as a planning support tool. Since project and production planning belongs to the core tasks of a company, the purchase option is most likely in the long run.

In the following chapters the perspective of the company's engineer is taken that has so far no experience with simulation but whose task it is to introduce it by the use of

standard simulation tools available on the market. The input to these chapters is gained from projects carried out at the Bauhaus-Universität Weimar in the term 2014/2015. The involved master students in civil engineering and construction management courses had so far no or only little theoretical background in simulation. They got the assignment to identify typical on site problems that could be tackled by the application of standard simulation tools such as Enterprise Dynamics. The students had to develop individual and team-oriented ways to get familiar with the software and to develop simulation models that were based on the standard functionalities of the software. Finally they presented the simulation results which were solutions for the identified problems in work planning.

The experiences gained during the project work as well as the results of these projects were thoroughly reported and documented. They contribute to a clearer perspective about the practical application areas and related challenges. The idea is that these project settings represent similar situations in practice by involving typical problems related to the work of construction companies.

## 3.2 Getting Started

After having acquired a simulation package the engineer in charge in the construction company usually finds himself in an isolated software environment which is not linked to the company's standard tools like computer aided design (CAD), building information modelling (BIM), enterprise resource planning (ERP) or network planning. This means that the data provision from construction projects or machinery pool into the simulation model must be performed by using the import/export functionalities of the corresponding software tools, or it must be generated manually with the simulation software. The engineer has to decide which information is needed and by which way (interface) it can be provided for a simulation model. Of course the engineer has to acquire some basic knowledge to learn how to perform a simulation study (theoretical knowledge) and he has to become familiar with the simulation tool as such (practical simulation knowledge) which might also include learning a new script language allowing the adaptation of the standard functionalities.

Furthermore the engineer who is using standard software will have to decide which standard elements can be used for modelling his real problem even if at first glance it appears inappropriate since they have been originally designed for a use in another application environment, most likely the stationary industry. An alternative could be that standard elements must be manipulated or new elements have to be developed and added to the library of the simulation tool.

The challenges of a first start with new simulation software in construction applications can be summarised as follows:

- Availability of qualified personnel to set up the simulation objectives and the modeling requirements (theoretical knowledge);
- Availability of qualified personnel to use a specific simulation software and to perform the simulation (practical knowledge);
- Generation or provision of input data to the simulation tool;
- Adaptation of standard simulation elements to the special needs of modelling construction processes.

Considering all these challenges the overall barrier to introduce simulation and develop a proper basis in a specific construction company for the use of simulation as a planning support tool is comparably high.

## 3.3 Simple Models for Construction Purposes

### *3.3.1 Overview*

The projects that have been carried out at the Bauhaus-Universität Weimar were based on four different topics (table 1). The projects had been assigned each to 2 to 5 persons as a team and included theoretical studies as well as practical work with the provided standard simulation software. Enterprise Dynamics (software version 8.2.5 and 9) was used. During the work it has been individually assessed whether each task could be fulfilled by using and applying the standard elements and to what extent an adaptation of these had to be undertaken. Based on this assessment a recommendation was given whether the topic can be regarded as an appropriate application in construction companies that are willing to apply simulation but so far have not had experience with it.

*Table 1: Simple models for construction purposes*

| No. | Purpose | Adaption of standard functionality by script language | Difficulty for simulation starters | Recommended application for use in construction |
|---|---|---|---|---|
| 1 | Setting up a transport chain for soil work | no | normal | yes |
| 2 | Material laying strategies by the use of 3D-BIM data | yes | high | no |
| 3 | Improving resource combinations | no | medium | yes |
| 4 | Modelling of logistical strategies | yes | medium | yes |

Two of these projects and the experiences gained with them are presented here in detail.

### *3.3.2 Material Laying Strategies by the Use of 3D-BIM Data*

This project focused on the connection between BIM and some standard simulation tool. The aim of the project was to elaborate how the 3D-coordinates of building components from a BIM model can be transferred into the simulation model. Within the simulation model it was intended to use the 3D-coordinates for transportation and material laying purposes.

The project started with the creation of a 3D-BIM model of a parking deck using Autocad Revit 2015 (fig. 1).

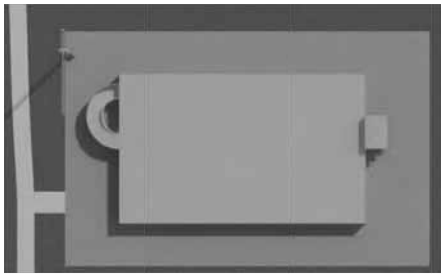


*Figure 1: Parking deck*

It was decided to simulate the asphalt laying on the park deck by identifying the positions of the existing concrete columns. The necessary task was thus the selection of the coordinates of concrete columns and the edges of the asphalt area.

Difficulties arose when trying to select these coordinates in Autocad Revit 2015 in order to export this information into a compatible format for further processing. As an alternative solution the 3D-Model was exported to AutoCAD 2015 where the position of the desired elements was selected (fig. 2) and exported into an MS Excel file.

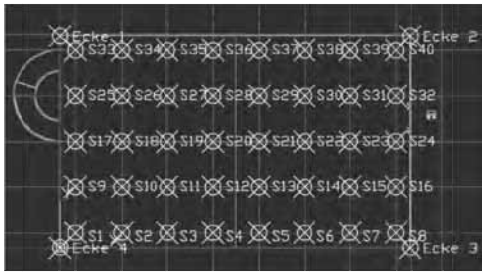



*Figure 2: Selection of columns*

The export into MS Excel was necessary, because the underlying coordinate systems in AutoCAD and in the simulation software were different, so that a conversion of the coordinates had to be performed. In order to improve the visualisation, a background drawing of the parking deck was inserted into the model layout. The creation of the paving material to be layed was simulated by using a source atom of enterprise dynamics library which was adapted in such a way that each square meter of the material to be delivered and laid receives an individual label from the excel file, which gave information to its final position. Figure 3 shows the model layout during the simulation run.

The created model provides a proper basis for a scenario simulation of different laying strategies. However, the efforts needed for the creation of such a model remain high. The experiences from a practitioner's point of view can be summarised as follows:

- Modelling work was time consuming;
- Coupling of different software system was necessary;
- Provision of data flow required comprehensive knowledge of software interfaces (exchange formats);
- The selection process of the required 3D-data need to be improved because up to now it requires a separate manual work step.

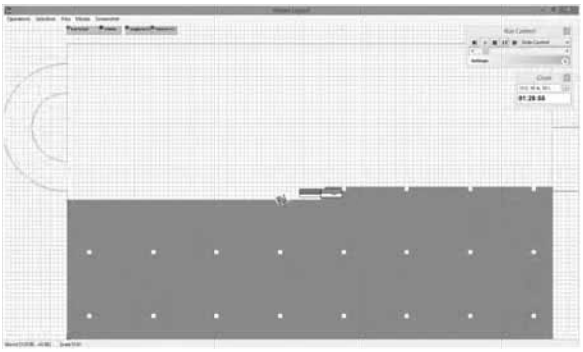


*Figure 3: Visualisation of material laying during simulation*

Considering the efforts needed for the conversion and transfer of 3D-BIM data into the standard simulation software at this stage of development the application cannot be recommended for construction companies that are at the beginning of applying simulation for their daily production planning.

### *3.3.3 Improving Resource Combinations*

The main application of simulation in this project was analysing the combination between material requirements by workers and the corresponding logistical chain on site. It should be analysed how the site logistics must be organised in order to ensure a nearly 100 % productivity rate of bricklayers. In this project a ground plot of the construction site was used at the background picture for modelling purposes. The approach turned out to be suitable for a quick transfer of the site specific characteristics like roads, building structures etc. The project included the modelling of material delivery on site and the following transportation of the material to a smaller storage area near the entrance of the building. The ground plot (fig. 4) has been scaled to ensure the correct recognition of distances in the simulation model. Later the supply paths for transporters and workers have been placed onto the ground plot. These modelling elements are seen in figure 4.

Recommended values for production ratios (ARH-Tables) have been used to determine the material demands of professional workers at the different work locations. Scenario simulation has been carried out to determine the consequences for the logistical chain if the number of bricklayers is increased. By using a table calculation, is has been determined at which combination it is economically favourable to employ logistical personnel (less hourly wage rate) that focuses exclusively on logistical tasks in order to increase productivity of the bricklayers.

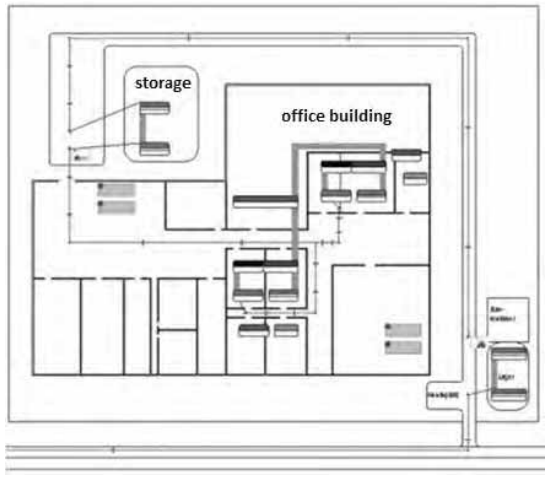



*Figure 4: Ground plot of construction site*

As a result of this application of simulation in the field of resource combinations, the experiences can be summarised as follows:

- Standard software supports this application well in a static system;
- Finding appropriate parameter values to model time realistic processes is time-consuming;
- The separation of logistical and technological (productive) tasks is difficult to model with standard simulation elements;
- The analysis of costs requires a comprehensive side calculation by using different time values measured during the simulation runs.

The use of standard simulation software in this application can be recommended. However, the modelling effort increases drastically when increasing the complexity by considering more types of material, different teams of workers and the dynamical environment of the construction site. Furthermore it turned out that the management of relevant parameters within the simulation model must be considered at an early stage. This can be achieved for example by implementing an interaction form or by hosting the parameters in an easily accessible calculation table.

## 4 Discussion and Conclusion

After having demonstrated various possibilities of applying simulation, Berner et al. (2013) have raised the question how long it will take the construction industry to use simulation for its production planning. Considering the results of the simulation projects that have been shown in this paper and the reviewed literature, there is a clear answer: the application of simulation tools to support production planning in construction will find its way into practice if

a. the entry barriers to use simulation tools have been lowered,
b. the integration of the tools within the corporate data management and IT infrastructure is given;
c. the data provision and data import is enhanced;

d. the time frame for development and verification of simulation models is shortened;
e. the approach of applying these tools in real projects will follow a formalised procedure.

Referring to a) and considering that the construction industry in Germany is highly fragmented by small and medium sized companies and engineering consultants, the economic barriers to use simulation as planning support tools are still fairly high. The task to perform a sound simulation study to support planning processes requires specially trained personnel which must have both, knowledge about simulation and construction works. It is obvious that an investment into simulation and modelling is rather undertaken in larger companies than in traditional small and medium sized companies. This is presumably also the reason why theoretical or prototypical concepts for the use of simulation in construction have not yet been implemented in the available tools by the software industry.

Referring to b) and c), it is not to be expected that simulation tools will be useful when operated in completely separated environments of the company's IT infrastructure. This requires connecting simulation with the existing planning and management tools. Then practitioners will not have to spend additional time for continuously aligning simulation systems with the existing information technology. Simulation tools are a powerful environment for a bidirectional processing together with traditional planning results. Suggestions to improve construction planning have to be transferred back into the traditional planning tools (time planning, BIM, ERP) and can then be considered during execution. A central connection or integration of standard simulation tools is therefor required.

An automated adaption of the simulation model refers to item d). Different changes of resource base or in design must be made possible without exceeding the required modelling time. Furthermore an optimal level of detail must be found that reflects the character of the construction project and shortens simulation time. In fast changing projects am in-depth simulation approach might be outdated because the chosen parameters and assumptions aren't valid anymore.

The application of simulation by practitioners requires additional formalisation of the steps to be executed with simulation tools as referred by item e). This means that a general approach on how to integrate a simulation tool in construction planning should be provided, which also functions as a guideline. This approach should include instructions on how to proceed related to different standard situations.

The one-of-a-kind character of construction projects often requires a thorough evaluation "from scratch" when designing simulation models as planning support tools. The main tasks for further research are to develop concepts for the integration of simulation tools as well as the further development of standard simulation elements that allow an easy adaptation to changed conditions on a specific site. Simulation software must approach itself closer to standard software applications in construction or even become an integral part of it. To overcome the existing hurdles for a successful application of standard simulation tools by construction companies is for many companies on the market currently still regarded as rather high.

## References

ARH-Tables: www.zeittechnik-verlag.de; last access 23.07.2015

Al-Hussein, M.; Athar Niaz, M.; Yu, H.; Kim, H.: Integrating 3D visualization and simulation for tower crane operations on construction sites. Automation in Construction 15 (2006) 5, pp. 554–562.

Beißert, U.: Constraint-basierte Simulation zur Terminplanung von Ausführungsprozessen: Repräsentation baubetrieblichen Wissens mittels Soft Constraints. Weimar: Verl. der Bauhaus-Universität 2012.

Berner, F.; Habenicht, I.; Kochkine, V.; Spieckermann, S.; Väth, C.: Simulation in der Fertigungsplanung von Bauwerken. Bauingenieur (2013) 88, pp. 89–97.

Fiedler, K.; Peldschus, F.: Die Anwendung der Spieltheorie für die Technologie der Bauprozesse. IABSE journal (1989) J-39, pp. 24–40.

Hajjar, D.; AbouRizk, S.: Simphony: an environment for building special purpose construction simulation tools, in: P. A. Farrington: Proceedings of the 31st conference on Winter simulation: Simulation - a bridge to the future - Volume 2, Phoenix, Arizona, USA, ACM, 1999.

Kochkine, V.: Prozessbasierte Fertigungsplanung bei der Optimierung der Baustellenproduktion: Ein Verfahrensmodell. Berlin: Beuth 2014.

König, M.; Koch, C.; Habenicht, I.; Spieckermann, S.: Intelligent BIM-based construction scheduling using discrete event simulation. In: 2012 Winter Simulation Conference, Berlin, Germany, pp. 1–12.

Kugler, M.: CAD-integrierte Modellierung von agentenbasierten Simulationsmodellen für die Bauablaufsimulation im Hochbau. Kassel Univ. Press 2012.

Lu, W.; Olofsson, T.: Building information modeling and discrete event simulation: Towards an integrated framework. Automation in Construction 44 (2014), pp. 73–83.

Marx, A.; König, M.: Simulation von Bauprozessen unter Berücksichtigung zeitvariabler räumlicher Restriktionen. In: Integrationsaspekte der Simulation: Technik, Organisation und Personal. G. Zülch & P. Stock (Hrsg) (2010), pp. 133–140.

Pitsch, H.: Simulation von Unikatprozessen. In: Franz, V. (Hrsg.): Simulation von Unikatprozessen: Herausforderung für kommerzielle Simulationswerkzeuge. Kassel: Kassel Univ. Press 2011, pp. 11–27.

Ruwanpura, J. Y.; AbouRizk, S. M.; Er, K.C.; Fernando, S.: Special Purpose Simulation Templates for Tunneling. In: Walsh, K.D. (Hrsg.): Construction Congress VI, Orlando, Florida, United States, February 20-22, 2000, pp. 97–106.

Siemens PLM Software (2013): Product data sheet Plant Simulation, via: http://www.plm.automation.siemens.com/en_us/products/tecnomatix/plant_design/plant_simulation.shtml, last access 23.07.2015.

Steinhauer, D.: Simulation im Schiffbau und Kooperation mit der Bauhaus-Universität Weimar, in: V. Franz (Hrsg.): 1. IBW Workshop Simulation in der Bauwirtschaft, Kassel, kassel university press, 2007.

Voigtmann, J. K., 2014: Simulation baulogistischer Prozesse. 1. Aufl. Hg. v. H.-J. Bargstädt. Kromsdorf: Bauhaus-Universitätsverlag Weimar.

Wang, W.-C.; Weng, S.-W.; Wang, S.-H.; Chen, C.-Y.: Integrating building information models with construction process simulations for project scheduling support. Automation in Construction 37 (2014), pp. 68–80.

Weber, J.: Simulation von Logistikprozessen auf Baustellen auf Basis von 3D-CAD Daten. Dissertation; Universität Dortmund, 2007.

Xu, J.; AbouRizk, S.M.; Fraser, C.: Integrated three-dimensional computer-aided design and discrete-event simulation models. Canadian Journal of Civil Engineering 30 (2003) 2, pp. 449–459.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# GIL-Baustellensimulation – Ein entscheidungsunterstützendes Planungswerkzeug für Baustellen zur Errichtung kundenindividueller Anlagen

## *GIL-Construction Site Simulation – A Decision Support Planning Tool for Construction Sites for the Erection of Customised Facilities*

Aron Heck, Technische Hochschule Nürnberg Georg Simon Ohm, Nürnberg (Germany), heckaron@gmx.de

Ilka Habenicht, SimPlan AG, Maintal (Germany), ilka.habenicht@simplan.de

Abstract: This document describes the use of simulation in the planning process of construction sites for the erection of Gas Insulated Lines (GIL). Simulation is integrated into a decision support planning tool in order to investigate different execution variants of a GIL project. It includes a data base for saving all relevant project data, a simulation tool kit for modelling individual construction sites and an analysis tool to evaluate different configurations. The integration of the decision support planning tool into the planning process of a GIL project is described. It is already used in first projects in order to determine a configuration to meet installation dates and realise an acceptable utilisation of workers and resources.

## 1 Einführung

Gasisolierte Übertragungsleitungen (GIL) werden eingesetzt, um platzsparend und verlustarm hohe elektrische Leistungen zu übertragen. Eine GIL kann auf Trassen, in Tunneln oder unterirdisch in einem Graben verlegt werden. Bei der Installation solcher Anlagen werden in der Regel Entfernungen von mehreren Kilometern überbrückt, wobei verschiedene Platzverhältnisse zur Installation der Anlage vorgefunden werden. Dies stellt besondere Anforderungen an die Planung und Durchführung der Installation der GIL-Anlagen. Die Siemens AG, Energy Management Division will aus diesem Grund die Planung solcher Projekte mit einem simulationsbasierten Werkzeug unterstützen.

### 1.1 Anwendungsumfeld und Motivation

Für die Errichtung einer GIL-Anlage wird zunächst eine Baustelle eingerichtet. Die notwendigen Komponenten, insbesondere Leiter- und Kapselrohre, Winkelbausteine und Kompensatoren, werden in Containern angeliefert, gelagert, vorgefertigt und installiert. Eine GIL-Anlage kann je nach Anlagengröße an einem oder mehreren

Orten gleichzeitig installiert werden. Bei größeren Anlagen müssen die Installationsarbeitsplätze nach Abschluss der Arbeiten an einem Montageabschnitt möglicherweise an einen anderen Ort versetzt werden. Nach der Installation wird die Leitung mit einem Isoliergas befüllt und in Betrieb genommen.

Die Fertigungsreihenfolge der Montageabschnitte ist normalerweise durch Rahmenbedingungen vorgegeben, teilweise aber auch frei wählbar. Es sind verschiedene Geräte und unterschiedlich qualifiziertes Personal, sowie Verbrauchsmaterial für die Durchführung der einzelnen Arbeitsschritte notwendig. Personal, Geräte und Material müssen in ausreichender Menge und in der richtigen Kombination zur Verfügung stehen. Entsprechend dem Baufortschritt müssen die Komponenten der Anlage in der richtigen Reihenfolge angeliefert und möglicherweise in Lagern gepuffert werden. Die Lagerfläche auf solchen Baustellen ist jedoch in der Regel stark begrenzt. Die GIL-Fertigung ist prinzipiell mit ähnlichen Problemen konfrontiert, wie sie beispielsweise von Ailland und Bargstädt (2008), Kugler und Franz (2008), sowie Voigtmann und Bargstädt (2008) für das Bauwesen beschrieben werden.

Eine Herausforderung bei der Planung von GIL-Baustellen liegt darin, eine möglichst hohe Auslastung der Installationsarbeitsplätze zu erreichen, weil deren Durchsatz maßgeblich für einen zügigen Baufortschritt ist. Das wird möglich, indem die Arbeitsplätze der Vorfertigung die notwendigen Baugruppen für die Installation fristgerecht fertigstellen. Dazu müssen die Liefertermine und die Zusammensetzung der Container mit den anzuliefernden Teilen entsprechend dem Bedarf auf der Baustelle ermittelt werden. Weiterhin müssen die Mitarbeiterzahlen und bereitzustellenden Gerätschaften so geplant werden, dass eine fristgerechte Vorfertigung und Installation erfolgen kann.

## 1.2 Zielsetzung

Um diesen Anforderungen besser gerecht zu werden, soll die Simulation als entscheidungsunterstützendes Instrument in die Planung von GIL-Projekten integriert werden. Auf diese Weise sollen Fehler bei der Auslegung der Baustelle früher erkannt werden als in der Vergangenheit. Ziel ist es, ein Instrument bereitzustellen, welches projektübergreifend bei der Planung der Baustellen eingesetzt werden kann. Dieses Werkzeug stellt die notwendigen Datenstrukturen für die Stammdaten und die projektspezifischen Daten, sowie einen Simulationsbausteinkasten für die Modellierung der Baustellen zur Verfügung.

Der Simulationsbausteinkasten soll so ausgelegt sein, dass geänderte Arbeitsschritte oder neuentwickelte Baugruppen beliebig konfiguriert und simuliert werden können. Der Planer soll in der Lage sein, ohne detaillierte Kenntnisse der Simulationssoftware das Simulationsmodell einer GIL-Baustelle zu erstellen und zu parametrisieren. Dabei soll berücksichtigt werden, dass vorrangig auf Basis der verschiedenen Prozesse auf der Baustelle und deren Anordnungsbeziehungen geplant wird.

Im weiteren Verlauf dieses Beitrages wird der Aufbau der Simulationsumgebung mit seinen Komponenten beschrieben. Schwerpunkte sind die bereitgestellten Datenstrukturen und die Bausteinbibliothek für die Modellierung der Baustellen. Es wird erörtert, wie die Simulation als Instrument in die Planung von GIL-Baustellen integriert werden kann. Abschließend werden die erzielten Projektergebnisse präsentiert und ein Ausblick zur zukünftigen Verwendung und Weiterentwicklung der Simulationsumgebung gegeben.

## 2 Komponenten der GIL-Baustellensimulation

In Abbildung 1 sind die Komponenten zur Verwaltung der Eingangsdaten, die Simulationsumgebung und die drei Komponenten zur Auswertung der Simulationsergebnisse dargestellt. Zusammen bilden sie die GIL-Baustellensimulation.

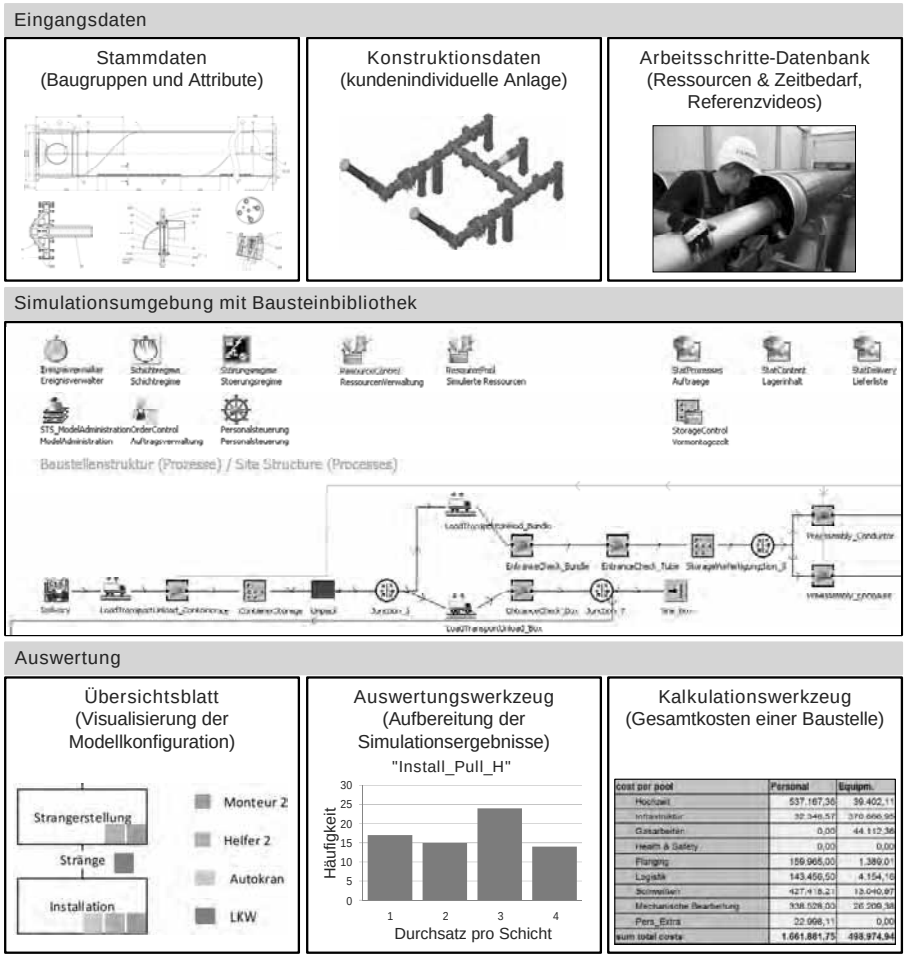



*Abbildung 1: Komponenten der GIL-Baustellensimulation*

Der Datentransfer zwischen den Komponenten erfolgt automatisiert über vorkonfigurierte Excel- und Datenbankschnittstellen. Die Modellierung des Materialflusses und der daraus resultierenden Prozesse basiert auf den Stammdaten mit den Attributen der Baugruppen, den Konstruktionsdaten einer kundenindividuellen Anlage und der Arbeitsschritte-Datenbank.

Die Simulationsumgebung stellt Dialoge und eine Bausteinbibliothek bereit, die es dem Planungsingenieur ermöglichen auf Basis der Eingangsdaten ausführbare Simulationsmodelle einer speziellen GIL-Baustelle zu konfigurieren. Die drei Kom-

ponenten zur Auswertung der Simulationsexperimente bereiten die Konfigurations- und Ergebnisdaten für die Interpretation und Validierung einer simulierten Gestaltungsvariante auf.

Um den Aufwand für die Modellerstellung möglichst gering zu halten, wurde der Simulationsbausteinkasten so ausgelegt, dass

- die Datenerfassung der Baugruppen und Arbeitsschritte pro Projekt inkrementell erfolgt,
- auf Basis der Konstruktionsdaten ermittelt wird, ob und wie oft Arbeitsschritte an einer Station durchgeführt werden und
- für die Konfiguration verschiedener Gestaltungsvarianten alternative Arbeitsschritte aus der Projekthistorie verfügbar sind.

## 2.1 Eingangsdaten der Simulation

Die Struktur der Stammdaten (Baugruppen und Attribute) wurde an dem bestehenden Sachnummernsystem ausgerichtet. Die Attribute der Baugruppen, welche für die Ermittlung der Arbeitsschritte relevant sind, werden auf Basis der technischen Zeichnungen identifiziert und in einer Excel-Mappe gespeichert. Neuentwickelte Baugruppen können durch einen Eintrag in der Excel-Mappe mit den Stammdaten ergänzt werden.

Die Konstruktionsdaten mit dem exakten Aufbau einer kundenindividuellen Anlage werden aus dem Softwaresystem des Anlageningenieurs exportiert, um sie anschließend, zusammen mit den Stammdaten, über eine Excel-Schnittstelle einzulesen. Diese Daten bilden die Basis für die Konfiguration der Installationsaufträge, die innerhalb der Simulation den Materialfluss steuern.

Die dritte Komponente ist eine Datenbank in der die Arbeitsschritte mit deren Zeit und Ressourcenbedarf hinterlegt sind. Sie enthält alle Arbeitsschritte die an Arbeitsplätzen einer GIL-Baustelle durchgeführt werden können und bietet eine Benutzeroberfläche zum Editieren vorhandener und Anlegen neuer Arbeitsschritte. Der Ablauf der Arbeitsschritte wird auf Videos festgehalten, die in der Datenbank mit den Arbeitsschritten verlinkt sind. Das in dieser Datenbank gespeicherte Prozesswissen dient zusätzlich als Ausgangsbasis für Verbesserungen von Arbeitsabläufen oder Fertigungsverfahren.

## 2.2 Bausteinbibliothek für die Modellierung der Baustellen

Die Fragestellungen und Randbedingungen bei der Planung und Durchführung von GIL-Bauprojekten sind vergleichbar mit typischen Problemen bei der Durchführung von Projekten in der Baubranche. Der Simulation-Toolkit-Shipbuilding-Bausteinkasten (STS) wurde bereits erfolgreich für die Modellierung von Prozessen im Bauwesen eingesetzt. Der Bausteinkasten wurde ursprünglich von der Flensburger Schiffbau Gesellschaft für die Modellierung schiffbaulicher Prozesse entwickelt und für die Anwendung im Bauwesen erweitert (vgl. Steinhauer und König 2010). Nachdem sich die simulationstechnischen Anforderungen mit denen im Bauwesen überschneiden, wurden die Bausteine zur Simulation von GIL-Baustellen auf Basis des STS-Bausteinkastens implementiert.

Abbildung 2 gibt einen Überblick über den Simulationsbausteinkasten zur Konfiguration von Simulationsmodellen kundenindividueller GIL-Baustellen. Die gesamte Modellierung einer GIL-Baustelle ist dialoggestützt.

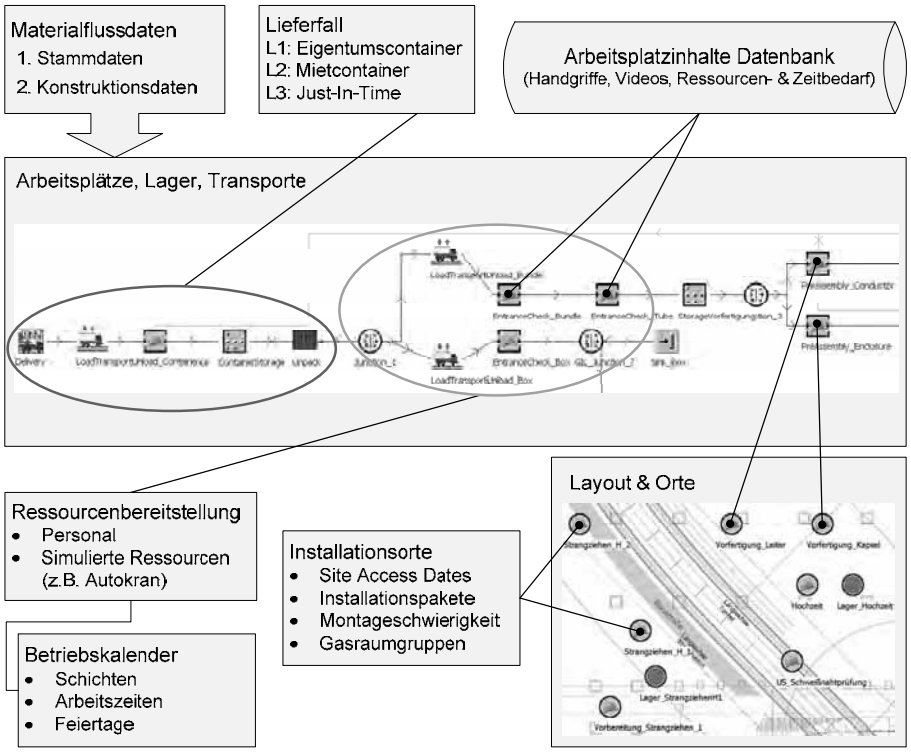



*Abbildung 2: Simulationsbausteinkasten zur Konfiguration eines Baustellenmodells*

Der Simulationsbausteinkasten besteht aus speziellen Bausteinen für Arbeitsplätze, Transporte und Lager, die es dem Planungsingenieur ermöglichen, die individuelle logistische Situation einer Baustelle zu modellieren ohne über tiefergehende Simulationskenntnisse zu verfügen. Die Arbeitsplatz-, Lager- und Transportbausteine bilden in Kombination mit den Material- und Informationsflusskanten die logistische Situation einer Baustelle ab.

Die exakte Sequenz der Werkstücke innerhalb des Materialflusses wird definiert, indem die Konstruktionsdaten der Anlage den Installationsarbeitsplätzen in Form von Aufträgen zugewiesen werden. Entsprechend dieser Auftragssequenz bestellen die Installationsarbeitsplätze entlang der Informationsflusskanten schließlich das Material, welches mit einer GIL-typischen Lieferstrategie der GIL-Baustellenfertigung zugeführt wird. Zu diesem Zweck verfügt der Bausteinkasten über Bausteine zur Modellierung der verschiedenen Logistikvarianten, die durch unterschiedliche Lieferfälle und Verpackungsarten gekennzeichnet sind.

Der Planungsingenieur ordnet jedem Arbeitsplatz, die dort auszuführenden Arbeitsschritte aus der Arbeitsinhalte-Datenbank, zu. Die Arbeitsschritte sind mit Attributen

der Baugruppen verknüpft, die für die Ausführung dieser Arbeitsschritte ursächlich sind. Ausgehend von dieser Konfiguration wird auf Basis der Werkstücke im Materialfluss entschieden, ob und wie oft ein Arbeitsschritt ausgeführt wird.

Ergänzende Randbedingungen wie Installationsreihenfolgen, verfügbare Ressourcen, Schicht- und Betriebskalender, sowie relevante Eingangsdaten zur Definition des Materialflusses und der Arbeitsschritte konfiguriert der Planungsingenieur über Dialoge.

## 2.3 Auswertung der Simulationsergebnisse

Mit Hilfe der Simulation können unterschiedliche Konfigurationen einer Baustelle untersucht und verglichen werden. Auf diese Weise ist eine iterative Verbesserung der Baustellenkonfiguration hinsichtlich der Einhaltung der Installationstermine und Auslastung der Anlagen möglich. Für die Dokumentation der Iterationsschritte sind die Parameter einer Modellkonfiguration mit den Arbeitsplätzen, Lagern und Transporten auf einem Übersichtsblatt dargestellt. Die Parameter sind mit einem Export der Konfigurationsdaten verknüpft, damit sie im Anschluss an einen Simulationslauf mit geringem Aufwand aktualisiert werden können.

Die Komponente zur Aufbereitung der Simulationsergebnisse zu aussagekräftigen Kennzahlen und Diagrammen ermöglicht den Einsatz der von Rabe et al. (2008) beschriebenen Techniken zur Verifikation und Validierung der Daten und Modelle der GIL-Baustellensimulation. Der Planungsingenieur vergleicht Gestaltungsvarianten und bewertet sie nach leistungs- oder serviceorientierten Gesichtspunkten wie bspw. Auslastung, Durchsatz oder Fertigstellungstermine.

Weigert et al. (2010) analysierten vorhandene Definitionen von Kennzahlen zur Bewertung logistischer Systeme. Auf Grundlage dieser Analyse wurde ein Kennzahlenkatalog zusammengestellt, der ein simuliertes Produktionssystem vollständig abbildet und grundlegend unterschiedliche logistische Sachverhalte beschreibt. Das Auswertungswerkzeug (Aufbereitung der Simulationsergebnisse) stellt die folgenden unterschiedlich aggregierten Kennzahlen bereit:

- Durchsatz,
- Bestand,
- Durchlaufzeiten,
- Auslastungen,
- Termine.

Nachdem eine hinreichend günstige, als valide einzustufende Konfiguration einer Baustelle gefunden ist, empfiehlt es sich, diese Gestaltungsvariante monetär zu bewerten und sie mit alternativen Gestaltungsvarianten zu vergleichen. Die Bewertung erfolgt auf Basis der Gesamtkosten einer Baustelle. Das Werkzeug zur Kalkulation der Gesamtkosten greift auf die aktuellen Preise für Equipment und Verbrauchsmaterialien des Hauptlieferanten zurück und kombiniert diese mit den Einsatzzeiträumen der Arbeitsbereiche und des Personals. Final ergänzt der Planungsingenieur Preise für Personal oder extern bezogenes Equipment und überprüft die Ausstattung der Baustelle auf Vollständigkeit. Als Ergebnis erhält der Planungsingenieur die anfallenden Kosten für Equipment und Personal pro Arbeitsbereich der Baustelle. In einer weiteren Tabelle sind die Kosten entsprechend interner Kalkula-

tionsgruppen aggregiert, damit sie von dem Projektmanagement ohne zusätzlichen Aufwand für die Kalkulation der Projektkosten übernommen werden können.

# 3 Integration des Simulationsbausteinkastens als entscheidungsunterstützendes Planungswerkzeug

Die Integration des Simulationsbausteinkastens als entscheidungsunterstützendes Planungswerkzeug setzt die Wiederverwendbarkeit der Daten in allen vorhandenen und neu entwickelten Systemen voraus. Dies wurde im Wesentlichen durch die Implementierung von Datenschnittstellen erreicht, die die Eingangsdaten in bestehender Form einlesen und die Ergebnisse in benötigter Form ausgeben.

Ein weiterer Aspekt ist die Integration des Simulationsbausteinkastens in den bestehenden Angebots- und Auftragsabwicklungsprozess. Hier ändert sich das grundsätzliche Vorgehen während der Baustellenplanung, welches im folgenden Abschnitt genauer erläutert wird.

## 3.1 Entscheidungsunterstützende Simulation als Bestandteil der Planung

Zu Beginn des Angebotsprozesses wird entsprechend der Spezifikation einer Ausschreibung eine kundenindividuelle Anlage konstruiert. Für die Erstellung der Angebotskalkulation müssen im nächsten Schritt die Aufwendungen zur Installation der Anlage abgeschätzt werden. In dieser Phase unterstützen die Simulationsexperimente bei Entscheidungen zur Gestaltung der Baustelle auf Basis getroffener Annahmen.

Diese Annahmen werden im Zuge der Auftragsabwicklung mit dem Kunden und den Lieferanten abgestimmt und allmählich zu Planparametern konkretisiert. Nun wird das nachfolgend erläuterte experimentelle Vorgehen zur simulationsgestützten Planung einer GIL-Baustelle erneut durchgeführt.

Die verschiedenen Gestaltungsvarianten einer Baustelle werden mit dem jeweiligen Stand der Eingangsdaten und den Auswertungen in einer Ordnerstruktur gespeichert. Der Planungsingenieur dokumentiert auf dem Übersichtsblatt die Eingangswerte einer Konfiguration, wie beispielsweise Lieferstrategie, Kapazitäten und Zuordnungen von Personal bzw. Arbeitsinhalten, damit die Simulationsexperimente nachvollziehbar bleiben.

Im Anschluss an einen Simulationslauf werden die Simulationsergebnisse mithilfe eines Auswertungswerkzeuges für die analytische Bewertung und Validierung des Modells zu Kennzahlen und Diagrammen aufbereitet, siehe Kapitel 2.3. In Abbildung 3 ist der Weg von der Leistungsbewertung des ersten Simulationsexperiments hin zur Kostenkalkulation auf Basis der Ergebnisse einer gültigen Gestaltungsvariante dargestellt.

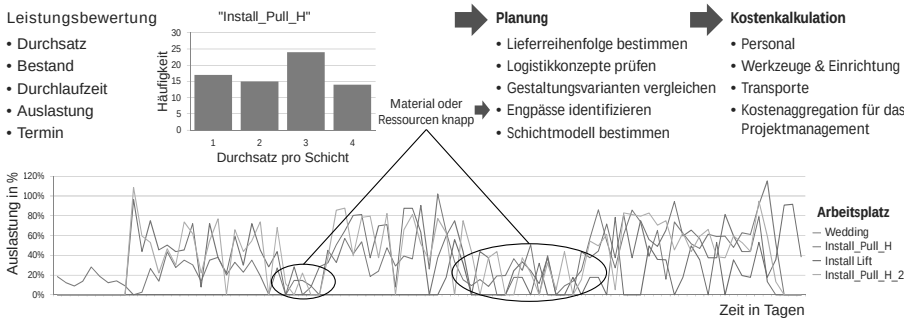



*Abbildung 3: Von der Leistungsbewertung des ersten Experiments zur Kostenkalkulation einer Baustelle*

Anhand der Simulationsexperimente werden alternative Logistikkonzepte, Anlagenvarianten und Ressourcenkonfigurationen zur Errichtung einer Anlage iterativ geprüft, bis eine hinreichend günstige, von allen Projektbeteiligten als valide einzustufende Konfiguration einer Baustelle gefunden ist. Diese ist Grundlage zur Festlegung des Schichtmodells, einer geeigneten Zusammensetzung des Personals und zur Bestimmung des kritischen Pfades auf der Baustelle. Außerdem wird auf Basis des simulierten Baustellenablaufes, eine Liste mit einem Vorschlag zur Verpackung der Anlagenkomponenten und den Terminen zur Bereitstellung dieser Packstücke generiert.

Die Prüfung der Logistikkonzepte umfasst die Konfiguration verschiedener Kapazitäten für realisierbare Lager und Transportmittel. Ein Logistikkonzept ist geeignet, wenn an den Arbeitsplätzen keine Verzögerungen entstehen, die auf eine zu geringe Auslegung der Lager oder Transportmittel zurückzuführen sind.

GIL-Baugruppen sind entweder günstiger in der Anschaffung (z. B. Rohr mit Schweißkante) oder einfacher zu installieren (z. B. Rohr mit Flansch). Eine Anlagenvariante bestehend aus geschweißten Baugruppen wird nur dann gebaut, wenn der Zusatzaufwand zur Installation geringer ist als die Mehrkosten zur Anschaffung der geflanschten Baugruppen. Durch Simulation dieser beiden Anlagenvarianten, wird projektspezifisch der kostengünstigere Aufbau der Anlage unter Berücksichtigung des Arbeitsaufwandes ermittelt.

Eine weitere Anwendung der Simulation ist die Bestimmung von Auswirkungen auf den Baustellenablauf infolge von Planänderungen zur Stärkung der eignen Verhandlungsposition. Als Beispiel für solche Planänderungen seien hier eine kundenseitige Vorgabe zur Änderung der Installationsreihenfolge oder eine lieferantenseitige Verzögerung bei der Zustellung von Material bzw. Equipment genannt.

Über eine Excel-Schnittstelle liefert die Simulation schließlich die Eingangsdaten für das Werkzeug zur Kalkulation der Gesamtkosten dieser Baustellenvariante. In der Kostenkalkulation werden die simulationsbasiert ermittelten Bedarfszeiträume der Werker und Geräte mit den Kostensätzen für Kauf, Miete oder Anstellung kombiniert. Außerdem stellt die Kostenkalkulation dem Planungsingenieur Oberflächen zur Erfassung von Personal und Equipment für nicht simulierte Vorgänge zur Verfügung. Im Rahmen der Kalkulation wird der Plan für die Ressourcenausstattungen der Arbeitsbereiche überprüft und gegebenenfalls vervollständigt.

Durchführbare Techniken der Verifikation und Validierung werden während der Entwicklung entsprechend dem Vorgehensmodell von Rabe et al. (2008) nach jeder Implementierungsetappe durchgeführt, wobei involvierte Projekt- und Bauleiter die Modellgültigkeit nach etwaigen Korrekturen stets bestätigten.

## 3.2 Ergebnisse aus der praktischen Anwendung

Die Entwicklung der GIL-Baustellensimulation erfolgte parallel zu einer realen Baustelle, deren Ablauf mit hinreichender Übereinstimmung zur Realität simuliert werden konnte. Außerdem hat sich mit dem aktuellen Entwicklungsstand der GIL-Baustellensimulation die Flexibilität der Daten und Modelle zur Simulation kundenindividueller Baustellen, insofern als ausreichend herausgestellt, als dass ein geänderter Installationsprozess mit neuentwickelten Baugruppen ohne zusätzlichen Implementierungsaufwand konfiguriert und simuliert werden konnte.

Die Simulation zeigte, in welchem Ausmaß und unter welchen Voraussetzungen der Durchsatz bei einer Umstellung auf den geänderten Installationsprozesses gesteigert werden kann. Bei den Untersuchungen in diesem Projektbeispiel wurde vor allem geprüft, welche Auswirkungen die Erhöhung des Durchsatzes an den Arbeitsplätzen der Vorfertigung auf den Prozess hat. An dieser Stelle dient die Arbeitsschritte-Datenbank als Ausgangsbasis zur Optimierung bestehender Vorfertigungsprozesse (Reduktion der Durchlaufzeiten in der Vorfertigung) und die Simulation als Experimentierumgebung zur optimalen Auslegung der Arbeitsplätze der neuen GIL-Baustellenfertigung. Die Auslegung verfolgt hier zusätzlich das Ziel einer möglichst gleichverteilten Auslastung der Ressourcen während ihrer kostenverursachenden Zeiträume.

Abbildung 4 zeigt, in welchem Ausmaß der Durchsatz am Arbeitsplatz Installation gesteigert werden konnte.

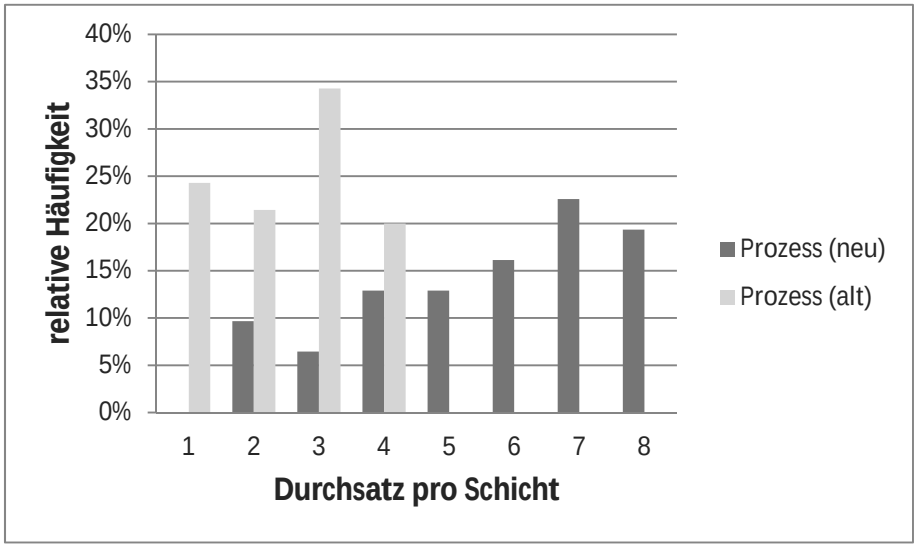



*Abbildung 4: Steigerung des Durchsatzes am Arbeitsplatz Installation*

## 4 Zusammenfassung und Ausblick

In diesem Beitrag wurde ein Werkzeug für die simulationsgestützte Planung von GIL-Baustellen beschrieben. Das Werkzeug umfasst neben einer vollständigen Datenbasis eines Bauprojektes einen Simulationsbausteinkasten zur Modellierung der Prozesse auf der Baustelle und eine Auswertungskomponente. Mit Hilfe der Simulation können bei der Planung verschiedene Gestaltungsvarianten einer Baustelle bewertet und verglichen werden.

Nachdem der Automatisierungsgrad mit Umsetzung neuer Fertigungsprozesse zunimmt, werden in einem weiteren Schritt die Komponenten zur stochastischen Simulation und zur statistischen Auswertung erweitert, wie sie von Helmer (2014) für die Anwendung auf die GIL-Baustellensimulation bereits untersucht wurden.

Die GIL-Baustellensimulation wird während der Angebotserstellung und Auftragsabwicklung zur Planung und Kalkulation von allen zukünftigen GIL-Baustellen eingesetzt. Auf Basis der bei der Anwendung gesammelten Erfahrungen wird die GIL-Baustellensimulation weiter an die Bedürfnisse des Projektmanagements angepasst. So wird beispielsweise eine Komponente entworfen, welche die wichtigsten Ergebnisse automatisch nach dem Simulationslauf in einer aussagekräftigen Form bereitstellt.

## Literatur

Ailland, K.; Bargstädt, H.J.: Tagesgenaues Termincontrolling auf Baustellen mit Hilfe der Simulation. In: Rabe, M. (Hrsg.): Advances in Simulation for Production and Logistics Applications. Stuttgart: Fraunhofer IRB Verlag 2008, 171-178.

Helmer, M.: Erstellung, Verifikation und Validierung von Wahrscheinlichkeitsverteilungen für eine Materialflusssimulation. Masterarbeit Technische Hochschule Nürnberg Georg Simon Ohm, Fakultät Elektrotechnik, Feinwerktechnik und Informationstechnik, 2014.

Kugler, M.; Franz, V.: Einsatz von Simulation zur Effizienzsteigerung von Produktionsprozessen im Bauwesen. In: Rabe, M. (Hrsg.): Advances in Simulation for Production and Logistics Applications. Stuttgart: Fraunhofer IRB Verlag 2008, 151-160.

Rabe, M.; Spieckermann, S.; Wenzel, S.: Verifikation und Validierung für die Simulation in Produktion und Logistik. Berlin, Heidelberg: Springer 2008.

Steinhauer, D.; König, M.: Konzepte zum effektiven Aufbau von Simulationsmodellen für die Unikatproduktion. In: Zülch, G.; Stock, P. (Hrsg.):Integrationsaspekte der Simulation: Technik, Organisation und Personal. Karlsruhe: KIT Scientific Publishing 2010, S. 157-164.

Voigtmann, J.K.; Bargstädt, H.J.: Simulation von Baulogistikprozessen im Ausbau. In: Rabe, M. (Hrsg.): Advances in Simulation for Production and Logistics Applications. Stuttgart: Fraunhofer IRB Verlag 2008, 131-140.

Weigert, G.; Rose, O.; Gocev, P.; Mayer, G.: Kennzahlen zur Bewertung logistischer Systeme. In: Zülch, G.; Stock, P. (Hrsg.): Integrationsaspekte der Simulation: Technik, Organisation und Personal. Karlsruhe: KIT Scientific Publishing 2010, S. 599-606.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Prozesssimulation zur Analyse von Wartungsstrategien im maschinellen Tunnelbau

## *Analysis of Maintenance Strategies in Mechanised Tunnelling Using Process Simulation*

Alena Conrads, Markus Scheffer, Markus König, Markus Thewes, Ruhr-Universität Bochum, Bochum (Germany), alena.conrads@rub.de, markus.scheffer@rub.de, koenig@inf.bi.rub.de, markus.thewes@rub.de

Abstract: In mechanised tunnelling detailed project planning with an accurate performance prediction of the tunnel boring machine (TBM) is needed for an efficient tunnelling progress. Undersized logistical components, disturbances of the supply chain as well as negligence of maintenance schedules reduce the TBM performance and frequently lead to avoidable times of standstill. Process simulation is a useful method to analyse the interaction between production and logistical processes. Performance data and information of the failure rates of technical components extracted from already finished projects are included as input data by applying methods of data mining. Thus, this paper presents the implementation of a performance forecast model for mechanised tunnelling projects focussing on disturbances and maintenance processes. The developed simulation model can easily be used to evaluate different maintenance strategies for a tunnelling project. Thus an improved maintenance strategy for reducing the time of standstill can be found.

## 1 Einführung

Wartung und Inspektion sind essentiell um die Funktionsfähigkeit von Maschinen zu gewährleisten. Unzureichende Wartung kann zu technischen Störungen und Ausfällen einzelner Elemente sowie zum Versagen des Gesamtsystems führen. Dies stellt ein sicherheitstechnisches Risiko dar und kann zusätzlich längere Produktionsausfälle und somit wirtschaftliche Schäden verursachen. In der Literatur finden sich zahlreiche Untersuchungen und Optimierungsansätze für die Wartung von verschiedenen Produktionsprozessen. In Ahmadi et al. (2010) werden z.B. multikriterielle Entscheidungsmethoden zur Auswahl von Wartungsstrategien vorgestellt. Pereira et al. (2010) verwenden einen Particle Swarm Optimization Ansatz (PSO) zur optimalen Planung von präventiven Wartungsintervallen für Kernkraftwerke. Beide Ansätze bedienen sich einer von der Logistik isolierten Betrachtung der Produktionsprozesse und sind somit nur unzureichend auf den maschinellen Tunnelbau übertragbar. Eine Erläuterung der Notwendigkeiten, die Instandhaltung bzw. War-

tung von Produktionssystems nicht getrennt von den logistischen Prozessen zu betrachten, ist in Achermann (2008) zu finden.

Der maschinelle Tunnelbau ähnelt aufgrund seines quasi-stationären Aufbaus einer fabrikartigen Produktion. Aufgrund der Vielzahl der zur Produktion benötigten Maschinenkomponenten sowie deren sensible Abhängigkeiten von den Logistikprozessen ergibt sich ein komplexes und dynamisches Systemverhalten. Eine unterdimensionierte Logistik, Störungen in der Versorgungskette und Ausfall von Produktionselementen aufgrund ineffizienter Wartungsstrategien können die Vortriebsleistung von Tunnelvortriebsmaschinen (TVM) reduzieren und vermeidbare Stillstandszeiten generieren. Aufgrund der komplexen Prozessabläufe einer TVM und sensitiver Interaktionen, können Störungen einzelner Prozesse zur Störung der gesamten Prozesskette führen. Diese Fortpflanzungseffekte erschweren die Identifikation der Störungsursachen und konsequenterweise auch die Entwicklung effizienter Wartungsstrategien. Durch die regelmäßige Durchführung von Wartungsarbeiten kann das Auftreten von technischen Störungen reduziert werden. Allerdings generieren Wartungsarbeiten in der Regel zusätzliche Produktionsstillstände und sollten daher effizient geplant werden. Unscharfe sowie unsichere projektspezifische Eingangsparameter, z.B. geotechnische Randbedingungen des Baugrunds oder Ausfallraten der Maschinenteile, erschweren hierbei eine genaue Planung.

Prozesssimulation ist eine etablierte Methode um komplexe Interaktionen zwischen Produktions- und Logistikprozessen abzubilden und zu analysieren (AbouRizk 2010). Die Umsetzung realistischer Simulationsmodelle erfordert die detaillierte Berücksichtigung der wesentlichen Systemelemente und gut aufbereitete Eingangsdaten, z.B. Prozesszeiten und Ausfallraten (König et al. 2014; Rahm et al. 2015). Die bislang veröffentlichten Ansätze zur simulationsgestützten Leistungsermittlung von TVM berücksichtigen die Ausfallraten von Produktionskomponenten ohne die Betrachtung von Wartungsarbeiten (Scheffer et al. 2014; Rahm et al. 2015). Im Folgenden wird ein Simulationsansatz zur Analyse der Auswirkungen von Wartungsarbeiten auf die Leistungsfähigkeit von TVM vorgestellt, um verschiedene Wartungsstrategien analysieren und vergleichen zu können.

## 2 Grundlagen

Der maschinelle Tunnelvortrieb ist ein sehr sensibles Produktionssystem, welches für einen reibungslosen Betrieb eine intensive Wartung erfordert. Es existiert eine Vielzahl an unterschiedlichen Wartungsarbeiten für einzelne technische Elemente, die bei der Erarbeitung einer Wartungsstrategie berücksichtigt werden müssen. In diesem Abschnitt werden zunächst die Grundlagen für die Planung der Wartung im maschinellen Tunnelbau aufgezeigt.

### 2.1 Maschinentechnik

Beim maschinellen Tunnelbau wird der Tunnel durch den Einsatz einer TVM aufgefahren. Ein weit verbreiteter Maschinentyp ist das Erddruckschild (EPB: Earth Pressure Balance), welcher besonders in bindigen Böden eingesetzt wird (Maidl 2012).

Abbildung 1 zeigt den Aufbau einer EPB-Schildmaschine. Mit Hilfe des Schneidrads (1) wird der anstehende Boden abgebaut und gelangt durch Schneidrad-

öffnungen in die Abbaukammer (2). Dort wird er unter Hinzugabe eines Konditionierungsmittels zu einem plastischen Erdbrei verarbeitet, sodass er als Stützmedium dem anstehenden Erddruck entgegenwirkt.

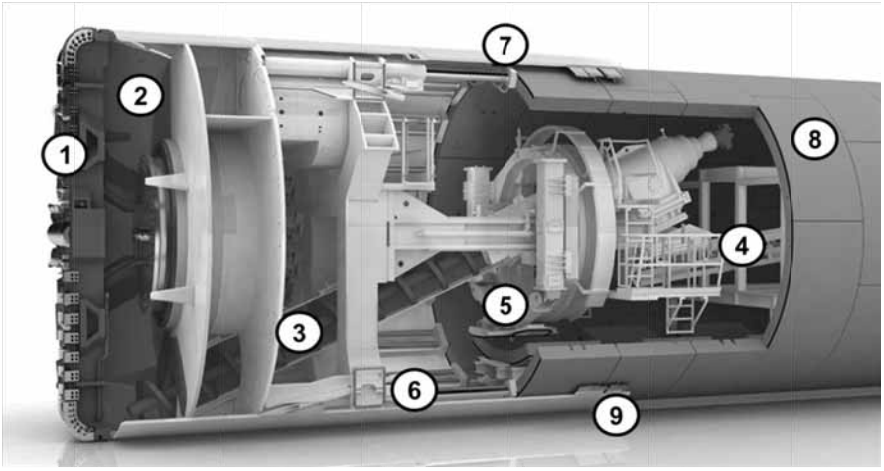



*Abbildung 1: Aufbau einer Erddruckschildmaschine (Herrenknecht 2015)*

Anschließend wird der Abraum über eine Förderschnecke (3) aus der druckbeaufschlagten Abbaukammer in den unter atmosphärischen Druck stehenden Teil des Tunnels gefördert. Dort wird der Abraum über ein Fließband (4) oder durch Tunnelfahrzeuge (z.B. Züge) an die Oberfläche transportiert. Die Vortriebskraft wird durch Hydraulikpressen (6) erzeugt, welche sich am Ende des Tunnelausbaus abstützen und die Maschine weiter voran in das Erdreich drücken.

Im Schutz des Schildschwanzes (7) wird der Tunnel durch Ringe aus Stahlbetonfertigteilsegmenten (8), sogenannten Tübbingen, ausgebaut. Diese werden vom Erektor (5) einzeln zusammengesetzt. Die Segmente werden in einem Fertigteilwerk hergestellt und mit LKWs zur Baustelle transportiert und zwischengelagert. Später werden die Segmente von einem Kran auf die Tunnelfahrzeuge verladen und zum Nachläufer der TVM transportiert.

Der Ringspalt, der zwischen dem Tübbingring und dem anstehenden Boden entsteht, wird während des Vortriebs mit Mörtel verpresst (9). Dieser wird auf der sich über Tage befindlichen Baustelle hergestellt und gemeinsam mit den Tübbingen von den Tunnelfahrzeugen zur TVM transportiert. Alternativ kann der Mörtel auch durch Rohrleitungen bis zur TVM gepumpt werden. Am Nachläufer wird dieser dann in einem Tank gelagert.

Der Abbau des Baugrundes und der Einbau eines Tübbingringes erfolgen immer im Wechsel. Im Idealfall schließen beide Vorgänge direkt aneinander an. Diese alternierende Arbeitsweise hat den Vorteil, dass für bestimmte Prozesse planmäßige Arbeitsunterbrechungen entstehen und nicht jeder Ausfall eines Elements direkt zu einer Verzögerung im Vortrieb führt.

## 2.2 Wartungsstrategien

Wartung bzw. Instandhaltung ist laut DIN 13306 (2001) die „Kombination aller technischen und administrativen Maßnahmen und des Maßnahmenmanagements während des Lebenszyklus einer Einheit zur Erhaltung des funktionsfähigen Zustandes oder der Rückführung in diesen, so dass sie die geforderte Funktion erfüllen kann“. Im Allgemeinen gibt es verschiedene mögliche Strategien zur Wartung einer Maschine.

Die planerisch einfachste Strategie ist die korrektive oder fehlerbehebende Wartung. Beim Auftreten einer Störung wird die Funktionsfähigkeit des Elements wiederhergestellt. Da die Wahrscheinlichkeit für das Auftreten technischer Störungen stochastischen Verteilungsfunktionen folgt, ist die fehlerbehebende Wartung stets zu berücksichtigen.

Eine weitere Wartungsstrategie ist die intervallbasierte Wartung. Nach Ablauf eines zuvor definierten Intervalls wird die Maschine gewartet. Dieses Intervall ist entweder von der Zeit (z.B. Betriebsstunden) oder der Produktion (z.B. gefahrene km) abhängig.

Kann der Zustand eines Elements visuell oder messtechnisch erfasst werden, sollten die Wartungsintervalle daran angepasst werden. In diesem Zusammenhang wird von einer zustandsorientierten Wartung gesprochen.

Wird ein Element gewartet bzw. ausgetauscht, obwohl sein Zustand noch keine Wartung erfordert, handelt es sich um eine präventive Wartung. Diese wird vor allem bei schlechter Zugänglichkeit zu einem Element sowie bei kritischen Abschnitten im Produktionsverlauf eingesetzt, um größere Ausfälle zu vermeiden.

## 2.3 Wartung im maschinellen Tunnelbau

Die Wartung der meisten Maschinenelemente einer TVM erfolgt in der Praxis intervallbasiert auf Grundlage der Vorgaben des TVM-Herstellers. Diese umfassen das Reinigen und Schmieren der beweglichen Maschinenelemente, aber auch die Zustandserfassung um weitere Wartungsvorgänge planen zu können. Diese Arbeiten erfolgen größtenteils begleitend zum Vortrieb oder Ringbau und haben somit nur geringe Auswirkungen auf die Leistung der Maschine.

Den größten Einfluss auf die Vortriebsleistung hat die Wartung der Abbauwerkzeuge am Schneidrad. Ist die Verschleißgrenze der Abbauwerkzeuge erreicht, müssen diese ausgetauscht werden. Zusätzlich können Verklebungen des Schneidrads durch den Boden oder das Auftreten von Findlingen im Boden einen Einstieg in die Abbaukammer erforderlich machen. Das Schneidrad ist bei Schildvortrieben mit flüssigkeits- bzw. erdbreigestützter Ortsbrust, während des Vortriebs nicht zugänglich. Da die Ortsbrust zu jeder Zeit gestützt werden muss, um ein Einbrechen des anstehenden Bodens zu vermeiden, kann die Wartung des Schneidrads nur unter Druckluft erfolgen. Das erforderliche Ein- und Ausschleusen der Arbeiter sowie die stark begrenzten Arbeitszeiten führen zu langen Stillstandszeiten bei Durchführung der Wartungsarbeiten.

Die Abbauwerkzeuge und das Schneidrad werden möglichst zustandsorientiert inspiziert und gewartet. Daher ist es sinnvoll, das Verschleißverhalten der Abbauwerkzeuge im Vorhinein abzuschätzen. Köppl (2014) hat dazu ein empirisches Prognosemodell zur Abschätzung des Verschleißes der Abbauwerkzeuge im Locker-

gestein entwickelt. Allerdings sind zusätzliche Randbedingungen wie der Baugrund oder das anstehende Grundwasser zu berücksichtigen. Je nach geotechnischen Randbedingungen und Oberflächenbebauung ist ein Einstieg in die Abbaukammer der TVM nicht oder unter erschwerten Bedingungen möglich. Um korrektive Wartungsarbeiten an diesen kritischen Punkten zu vermeiden, ist eine präventive Wartung unabdingbar.

Eine ganzheitliche Analyse des komplexen Zusammenspiels von Produktions- und Logistikprozessen unter Berücksichtigung von dynamischen Randbedingungen, unscharfen Ausfallzeiten und variierendem Wartungsmanagement ist mit den in der Praxis verfügbaren Planungskonzepten derzeit nicht möglich.

# 3 Simulationsmodell

Der Einsatz von simulationsgestützten Analysen findet in den letzten Jahren zunehmend Verbreitung im Bauwesen. Existierende Simulationsansätze für den maschinellen Tunnelbau beschäftigen sich bislang meist mit einer isolierten Betrachtung der Produktions- oder Logistikprozesse einer Baustelle (Al-Bataineh et al. 2013; Ebrahimy et al. 2011a; Ebrahimy et al. 2011b). Detaillierte Modelle zur Analyse von Prozessinteraktionen existieren bislang nur für einzelne Projekte (Li et al. 2010). Solche projektspezifischen Simulationsmodelle lassen sich jedoch nur sehr aufwendig auf andere Tunnelbauprojekte mit teilweise sehr unterschiedlichen Logistiksystemen übertragen.

## 3.1 Aufbau des Modells

In vorherigen Arbeiten wurden die Produktions- und Logistikkomponenten des maschinellen Tunnelbaus formalisiert und in der Simulationsumgebung *AnyLogic* implementiert (König et al. 2014; Scheffer et al. 2014). Zur formalen Beschreibung des Systems wurde die *System Modeling Language* (SysML) verwendet. Die entwickelten Systemelemente sowie die hierarchische Struktur wurden durch Verwendung von Blockdefinitionsdiagrammen (*bdd-Diagram*) realisiert. In Abbildung 2 ist exemplarisch die Struktur des unterirdischen Baustellenbereichs abgebildet. Durch eine formale Beschreibung der ablaufenden Prozesse und Prozessabhängigkeiten durch Sequenzdiagramme lassen sich einzelne Simulationskomponenten flexibel austauschen und konfigurieren.

Abbildung 2 zeigt die hierarchische Anordnung der einzelnen Komponenten. Zunächst wird das Gesamtsystem aufgeteilt in eine oberirdische und eine unterirdische Komponente. Sämtliche Kommunikation der beiden Hauptkomponenten erfolgt über das Logistikmanagement, d. h. Bestellungen für Materialien sowie Störfallmeldungen werden vom Logistikmanagement an die betroffenen Systemelemente weitergeleitet. Die formale Systemstruktur ist entsprechend dem Blockdefinitionsdiagramm hierarchisch implementiert (Abb. 3).

Die Prozesse der einzelnen Elemente sind durch Zustandsdiagramme modelliert. Die Kommunikation der Systemelemente ist durch das Senden und Empfangen von Signalen nach dem Observer-Design-Pattern implementiert (Gamma et al. 2014). So wird z. B. nach Beendigung des Vortriebs von dem Schneidrad das Signal *AdvanceFinished* gesendet. Parallel wird vom Segmentzuführer durch das Signal *SegmentAvailable* signalisiert, dass ein Segment zum Einbau bereit liegt. Der

Erektor startet nach dem Empfang dieser Signale mit dem Ringbau. Nach dem Aufnehmen des Segments durch den Erektor wird das Signal *SegmentRemove* gesendet. Falls möglich stellt der Segmentzuführer ein neues Segment zum Einbau bereit. Diese Abfolge wiederholt sich, bis der gesamte Ringbau beendet ist. Das Ende des Ringbaus wird dem Schneidrad über das Signal *RingbuildFinished* übermittelt.

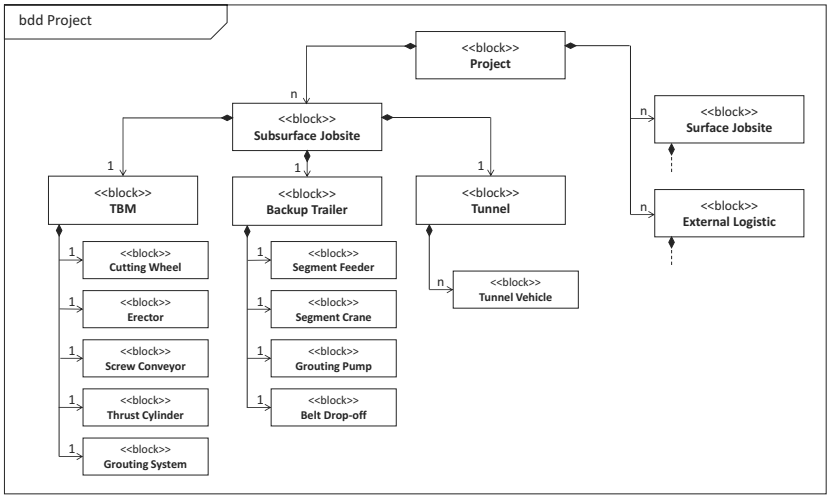



*Abbildung 2: SysML Block-definition Diagramm der unterirdischen Komponenten*

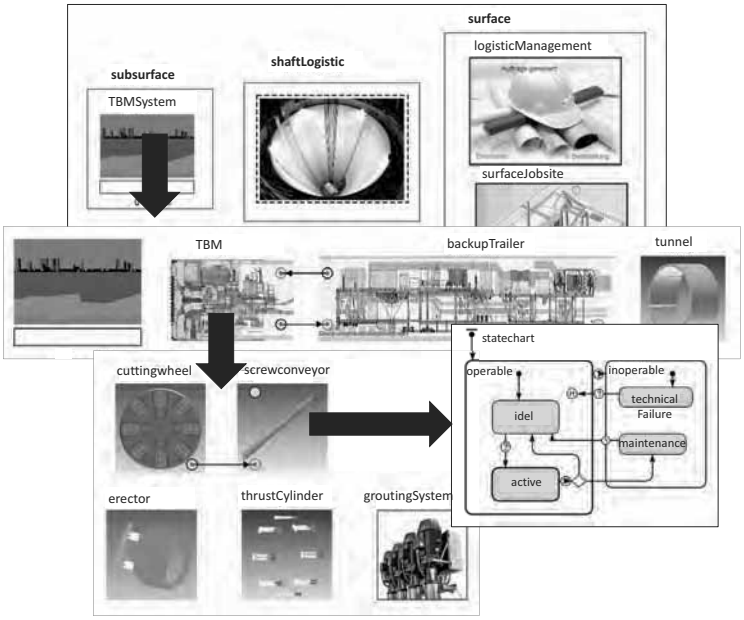



*Abbildung 3: Hierarchische Struktur des Simulationsmodells*

Die Störungen der einzelnen Elemente können aufgrund der implementierten Prozessabhängigkeiten an die jeweils betroffenen Elemente weitergegeben werden (Rahm et al. 2015). Analog zur Dauer der Produktionsprozesse für den Vortrieb und dem Ringbau, kann die Dauer bis zum nächsten Eintritt einer Störung (Time Between Failure, TBF), sowie die Dauer zur Behebung (Time To Repair, TTR) über Verteilungsfunktionen abgebildet werden. Diese wurden auf Basis von Daten aus bereits abgeschlossenen Projekten durch statistische Auswertungsmethoden ermittelt. Die Tabelle 1 zeigt beispielhaft die Verteilungsfunktionen des Produktionsprozess Ringbau und der TBF des Erektors.

*Tabelle 1: Verteilungsfunktionen des Produktionsprozesses Ringbau und der Time Between Failure des Erektors*

| | Verteilungsfunktion | Parameter |
|---|---|---|
| Produktionsprozess Ringbau | Weibull-Verteilung | loc = 3.19<br>α = 1.12<br>β = 1.25<br>Ø = 30,7 min/ring |
| TBF des Erektors | Beta-Verteilung | p = 0.46<br>q = 3.10<br>min = 9.25<br>max = 54209.84<br>Ø = 7012,55 min |

Der Materialfluss des abgebauten Bodens sowie des verwendeten Mörtels wird mit Hilfe systemdynamischer Simulationskomponenten abgebildet (s. Abbildung 4) und mit den ereignis-diskreten Zustandsdiagrammen verknüpft. Die Menge des Materialflusses ist dabei von der variablen Vortriebsgeschwindigkeit abhängig.

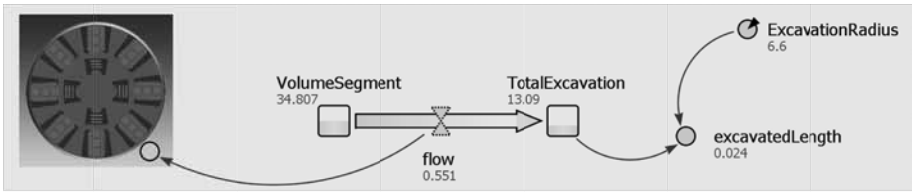



*Abbildung 4: Materialfluss des abgebauten Bodens innerhalb des Schneidrads*

## 3.2 Integration der Wartungsprozesse

Die Wartung eines Elementes wird, neben den Produktionszuständen, als ein weiterer Zustand dem ereignis-diskreten Modell hinzugefügt. Wie auch bei den technischen Störungen befindet sich das Element während der Wartungsarbeiten im Zustand *inoperable*. Abbildung 5 zeigt beispielhaft das Zustandsdiagramm des Elements Erektor.

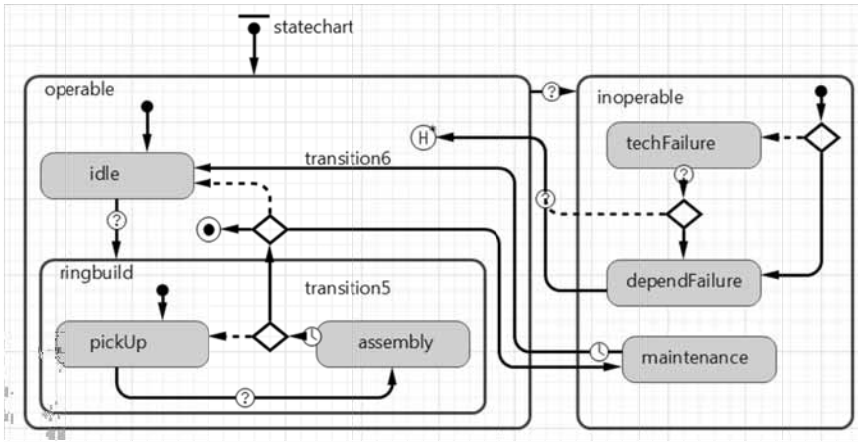



*Abbildung 5: Zustandsdiagramm des Erektors inklusive der Wartungsprozesse*

Da die Wartung im Gegensatz zu technischen Störungen geplant eintritt, unterbricht sie nicht den aktuellen Arbeitsschritt, sondern schließt sich unmittelbar danach an. Durch den asynchronen Ablauf von Ringbau und Vortrieb führen geplante Wartungen nicht zwangsläufig zu einem Stillstand des Gesamtsystems. Durch die implementierten Abhängigkeiten der Prozesse können sich Störungen oder zeitintensive Wartungsarbeiten jedoch auch direkt auf die Produktionsprozesse auswirken. Wartungsarbeiten im maschinellen Tunnelbau werden in der Regel nach dem Ringbau durchgeführt. Ist der Ringbau beendet und das Zeitintervall bis zur nächsten Wartung überschritten, wechseln die Komponenten in den Zustand *maintenance*. Der aktuell durchgeführte Prozess wird dabei nicht unterbrochen, sondern vollständig ausgeführt.

Für die Zeit eines Wartungsintervalls sowie der eigentlichen Wartungsdauer wird ein fester Wert gesetzt. Die Wartung beinhaltet dabei die Inspektion des Elements. Ist die Zeit bis zur nächsten zu erwartenden technischen Störung (TBF) zu gering, wird das Element repariert bzw. ausgetauscht. In diesem Fall wird eine neue TBF bestimmt. Ist die ermittelte TBF kürzer als das Wartungsintervall, kommt es zu einer technischen Störung. Da die TBF stochastisch ermittelt wird, sind technische Störungen auch durch eine Wartung nicht vollständig vermeidbar. Zusätzlich zu den häufig durchgeführten kurzen Wartungsarbeiten werden in größeren Abständen längere Wartungsarbeiten, wie z.B. der Austausch verschlissener Abbauwerkzeuge am Schneidrad, durchgeführt.

## 4 Anwendungsbeispiel

Im folgenden Anwendungsbeispiel wird der Einfluss einer präventiven Wartung mit festgelegten Wartungsintervallen unter Berücksichtigung von stochastischen Eingangsdaten auf die Vortriebsdauer untersucht. Um eine möglichst robuste Aussage über die zu erwartende Vortriebsleistung treffen zu können, wird die Vortriebsdauer mit Hilfe von Monte-Carlo Simulationen ermittelt und anschließend analysiert. Es wurden je Szenario 1000 Simulationsläufe durchgeführt, um eine statistische Auswertung der Ergebnisse zu ermöglichen. Als wesentliche Bewertungsgrundlage wird das 95 %-Quantil der Projektdauer verwendet. Bei dem untersuchten Beispiel

handelt es sich um einen Tunnel mit einer Länge von 3000 m und einem Durchmesser von 6,6 m. Ein Tunnelring wird aus 6 Segmenten zusammengesetzt.

### 4.1 Szenario 0: Keine Störungen

Im ersten Szenario wird die theoretische Leistung des Gesamtsystems ermittelt, die sich bei einem Vortrieb ohne technische Störungen und ohne Wartungsarbeiten ergeben. Die Vortriebsgeschwindigkeit ist abhängig von dem anstehenden Baugrund und wird als stochastische Größe mit einbezogen. Eventuelle Stillstände resultieren aus Einschränkungen durch die Baustelleneinrichtung und der logistischen Versorgung der TVM. Die Projektdauer beträgt bei diesem Szenario mit einer 95 % Wahrscheinlichkeit maximal 374,7 Tage.

### 4.2 Szenario 1: Korrektive Wartung

Durch die Implementierung von Störungen wird die Leistung des Gesamtsystems vermindert. Die Störungen können die Produktionsprozesse direkt betreffen oder sich auf Supportprozesse, z.B. der Nachläuferlogistik, auswirken. Diese indirekten Störungen der Supportprozesse führen nicht zwangsläufig zu einer Unterbrechung der Produktionsprozesse. Tritt eine Störung ein, wird das betroffene Element korrektiv gewartet und die Störung wird behoben. Durch den Ausfall einzelner Elemente erhöht sich das 95 %-Quantil der Projektdauer auf 467,2 Tage.

### 4.3 Szenario 2: Intervallbasierte Wartung

Um die Ausfallzeiten durch die technischen Störungen zu reduzieren, wird eine intervallbasierte Wartung durchgeführt. Es wird angenommen, dass alle 11 Stunden für eine Dauer von ca. 1 Stunde eine Inspektion bzw. Wartung der Elemente durchgeführt werden muss. Dies entspricht bei einem 2-Schicht-System eine Stunde Wartung pro Schicht. Zusätzlich werden einmal im Monat für etwa 12 Stunden größere Wartungsarbeiten, wie z.B. der Wechsel der Abbauwerkzeuge am Schneidrad, durchgeführt. Durch die Wartung in den gewählten Wartungsintervallen entstehen geplanten Stillstandszeiten von insgesamt ca. 43,5 Tagen. Die gesamte Stillstandsdauer, verursacht durch Wartungsarbeiten und technischen Störungen, ist jedoch um ca. 25 % geringer als in Szenario 1 und das 95 %-Quantil der Projektdauer reduziert sich auf 445,4 Tage. Abbildung 6 zeigt die Häufigkeitsverteilung der zu erwartenden Projektdauer. Es ist ersichtlich, dass trotz kurzer Wartungsintervalle und der daraus resultierenden Produktionsausfälle die Leistung des Gesamtsystems durch die Wartungsarbeiten verbessert wird.

## 5 Zusammenfassung und Ausblick

Es wurden die Einflüsse von Störungen sowie der Wartung auf die Leistung eines gekoppelten Produktions- und Logistikmodells aufgezeigt. Technische Störungen reduzieren die Leistung des Gesamtsystems und verursachen deutlich längere Projektdauern. Mit Hilfe von regelmäßigen Wartungsarbeiten kann die Häufigkeit des Auftretens technischer Störungen reduziert werden.

Das Beispiel zeigt, dass bereits mit intervallbasierter Wartung die Leistung des Projektes verbessert wird. Eine Anpassung von Anzahl und Dauer der Wartungsarbeiten unter Berücksichtigung einer zustandsbasierten Bewertung ermöglicht eine weitere

Verringerung von Produktionsausfällen. Elemente, deren Zustand visuell oder messtechnisch nicht erfassbar ist, können in weiteren Arbeiten durch den Einsatz geeigneter Ersatzsysteme in einer optimierten Wartungsplanung berücksichtigt werden.

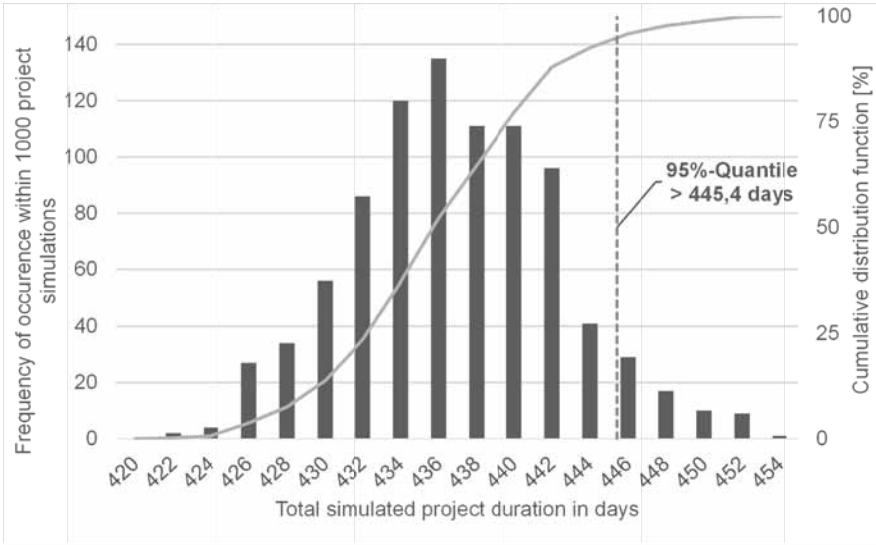



*Abbildung 6: Histogramm der Projektdauer des Szenarios 2 bei 1000 Simulationsläufen*

## Danksagung

Die Autoren danken der Deutschen Forschungsgesellschaft (DFG) für die finanzielle Förderung dieser Arbeit im Rahmen des Teilprojekts C3 des Sonderforschungsbereichs SFB 837 „Interaktionsmodelle für den maschinellen Tunnelbau“.

## Literatur

AbouRizk, S.: Role of Simulation in construction engineering and management. Journal of Construction Engineering and Management 136 (2010) 10, S. 1140-1153.

Achermann, D.: Modelling, simulation and optimization of maintenance strategies under consideration of logistic processes. Dissertation Swiss Federal Institute of Technology, Zurich, 2008.

Ahmadi, A.; Gupta, S.; Karim, R; Kumar. U.: Selection of maintenance strategy for aircraft systems using multi-criteria decision making methodology. International Journal of Reliability, Quality and Safety Engineering 17 (2010) 3, S. 223-243.

Al-Bataineh, M.; AbouRizk, S.; Parkis, H.: Using simulation to plan tunnel construction. Journal of Construction Engineering and Management 139 (2013) 5, S. 564-571.

Deutsches Institut für Normung: DIN EN 13306: Maintenance Terminology; Trilingual version EN 13306: 2001.

Ebrahimy, Y.; AbouRizk, S.; Fernando, S.; Mohamed, Y. (2011a) Simphony Supply Chain Simulator: a simulation toolkit to model the supply chain of construction projects. Simulation 87 (2011) 8, S. 657-667.

Ebrahimy Y.; AbouRizk, S.; Fernando, S.; Mohamed, Y. (2011b) Simulation modeling and sensitivity analysis of a tunneling construction project's supply chain. Engineering, Construction and Architectural Management 18 (2011) 5, S. 462-480.

Gamma, E., Helm, R., Johnson, R., & Vlissides, J.: Design patterns – elements of reusable object-oriented software. Reading (MA): Addison-Wesley 2014.

Herrenknecht 2015: https://www.herrenknecht.com/de/produkte/kernprodukte/tunnelling/epb-schild.html, abgerufen am 12.05.2015.

König, M.; Thewes, M.; Rahm, T.; Scheffer, M.; Sadri, K.; Conrads, A.: Prozesssimulation von maschinellen Tunnelvortrieben – Verfügbarkeitsanalysen der Leistungsprozesse unter Berücksichtigung von Stillständen. Bauingenieur 89 (2014) 11, S. 467-477.

Köppl, F.: Abbauwerkzeugverschleiß und empirische Verschleißprognose beim Vortrieb mit Hydroschild TVM in Lockergesteinen. Dissertation, TU München 2014.

Li, L.; Li, J.; Ma, W.; Xue, Y.; Liu, Y.: Virtual prototype of the tunnel boring machine and movement simulation in DIVISION MOCKUP2000i2. In: Huang, G.Q.; Mak, K.L.; Maropoulos, P.G. (Hrsg.): Proceedings of the 6th CIRP-Sponsored International Conference on Digital Enterprise Technology. Berlin, Heidelberg: Springer 2010, S. 525-532.

Maidl, B.; Herrenknecht, M.; Maidl, U.; Wehrmeyer,G.: Mechanised shield tunneling. Berlin: Ernst & Sohn 2012.

Pereira, C.; Lapa, C.; Mol, A.; da Luz, A.: A Particle Swarm Optimization (PSO) approach for non-periodic preventive maintenance scheduling programming. Progress in Nuclear Energy 52 (2010) 8, S. 710-714.

Rahm, T.; Scheffer, M.; Duhme, R.; Thewes, M.; König, M.: Evaluation of disturbances in mechanized tunneling using process simulation. Computer-Aided Civil and Infrastructure Engineering (Mai 2015, early view).

Scheffer, M.; Rahm, T.; Duhme, J.; Thewes, M.; König, M.: Jobsite logistic simulation in mechanized tunneling. In: Proceedings of the 2014 Winter Simulation Conference, Savannah 2014, S. 1843-1854.

The AnyLogic Company: http://www.anylogic.com//overview, 12.05.2015.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Simulationsgestützte Planung und Bewertung der Energieeffizienz für Produktionssysteme in der Automobilindustrie

## *Simulation-based Planning and Evaluation of Energy Efficiency for Production Systems in Car Manufacturing*

Tim Peter, Sigrid Wenzel, Universität Kassel, Kassel (Germany),
tim.peter@uni-kassel.de, s.wenzel@uni-kassel.de

Abstract: This paper gives a short overview of some current approaches towards including energy aspects into material flow simulation and presents the simulation approach developed in the SimEnergy project which couples discrete event material flow simulation with continuous energy simulation. The implemented solution interconnects the discrete event simulation tool Plant Simulation and MATLAB Simulink as continuous simulation tool. However, the solution is flexible and can be easily adapted to other simulation tools. A use case in the automotive industry is used as a basis for the simulation runs which show dependencies of temperature based deactivation of production machines on the throughput of the production systems.

## 1 Einleitung

Mit der gestiegenen Bedeutung des Energieverbrauchs in der Produktion hat auch die Berücksichtigung energetischer Einflussfaktoren in der Produktionsplanung erheblich an Stellenwert gewonnen. Das zeigt sich ebenfalls im Bereich der simulationsgestützten Planung und der Forschung in diesem Umfeld. So beziehen sich aktuelle Arbeiten verstärkt auf die Integration energetischer Aspekte in die ereignisdiskrete Materialflusssimulation. Ein Integrationsansatz wurde im Rahmen des Forschungsprojektes „SimEnergy“ (Hessen Agentur Projekt Nr. 369/13-08) vom Fachgebiet für Produktionsorganisation und Fabrikplanung der Universität Kassel in Kooperation mit den Simulationsdienstleistern SimPlan AG, Maintal, und Limón GmbH, Kassel, sowie dem Anwendungspartner Volkswagen AG entwickelt. Der folgende Beitrag soll den Stand der Technik auf dem Gebiet der Integration energetischer Einflussgrößen in die Simulation in Produktion und Logistik aufzeigen, den im Rahmen des Projektes entwickelten Lösungsansatz vorstellen, die Anwendung an einem Praxisbeispiel in der Automobilproduktion demonstrieren und die erzielten

Simulationsergebnisse zusammenfassen. Abschließend wird ein Ausblick auf weitere Anwendungsfälle und Entwicklungsaufgaben gegeben.

## 2 Stand der Technik

Während bei neuen Produktionsanlagen „wie selbstverständlich“ bereits „energiesparende Verfahren und Technologien“ eingesetzt werden (VDI 2012, S. 12) ist die Berücksichtigung der energetischen Bewertung bei der Planung und dem Betrieb von Produktions- und Logistikprozessen noch Forschungsgegenstand. Die Bewertung der Energieeffizienz orientiert sich dabei zum einen an den heutigen Kennzahlen für die energetische Systembewertung wie Energieverbrauch oder $CO_2$-Emissionen aber auch an klassischen Logistikkennzahlen wie Durchsatz, Auslastungsgrad, Beständen und deren Reichweite.

Ein Bewertungsansatz ist die sogenannte Energiewertstrommethode, die auf der klassischen Wertstromanalyse aufsetzt und diese um energetische Betrachtungen erweitert (Erlach und Westkämper 2009, S. 17–22). Der ermittelte Energiewertstrom lässt sich mit Hilfe von Kennzahlen (Energieintensität und Effizienzgrad) bezüglich seiner Energieeffizienz bewerten (Erlach und Westkämper 2009, S. 32, 33). Diese Betrachtungen bleiben jedoch statisch und können keine dynamischen Wechselwirkungen oder stochastischen Einflüsse berücksichtigen, wie es in der Simulation von Produktions- und Logistikprozessen der Fall ist. Gerade die Betrachtung von dynamischen Wechselwirkungen oder stochastischen Einflüsse stellen aber eine Grundvoraussetzung für eine gesicherte Planung dar (Wenzel et al. 2008; Rabe et al. 2008). Auch bei der Analyse energetischer Einflussgrößen in einem Produktions- oder Logistiksystem wird Simulation eingesetzt, wenn sich das Verhalten der Energieströme in Abhängigkeit von der Zeit verändert (Hesselbach 2012; Mendes et al. 2001; Herrmann et al. 2011). Sobald sich allerdings Produktions- und Logistikprozesse und energetische Einflussgrößen gegenseitig beeinflussen und somit Wechselwirkungen auftreten, stellt sich die Frage, wie diese bewertet und auch in der Simulation adäquat berücksichtigt werden können. Heute existieren bereits verschiedene Ansätze zur Integration von energetischen Analysen in die Simulation in Produktion und Logistik, die sich in ihrem Implementierungsaufwand und den erzielbaren Ergebnissen unterscheiden. Eine mögliche Lösung ist eine direkte Einbindung energetischer Bewertungsbausteine in ein ereignisdiskretes Simulationswerkzeug, wie sie in Plant Simulation ab Version 11 implementiert ist (Siemens Industry Software GmbH & Co. KG 2015). Bei dieser Lösung handelt es sich um eine Energieverbrauchsberechnung in Abhängigkeit von den jeweiligen Anlagenzuständen. Dabei wird nur eine Energieform (hier: elektrische Energie) betrachtet, die Energiezustände sind vordefiniert und Veränderungen können nur an der Leistungsaufnahme und an den Übergangszeiten zwischen zwei Zuständen vorgenommen werden. Eine detaillierte Abbildung von Leistungskurven ist nicht möglich, da die Berechnungen direkt in den entsprechenden Objekten erfolgen. Dementsprechend werden auch keine Wechselwirkungen zwischen der Energiebereitstellung und den Materialflussobjekten abgebildet (Schlegel et al. 2013, S. 189).

Auch Weinert (2010) entwickelt einen Integrationsansatz, der eine Betrachtung energetischer Einflussfaktoren mit hohem Detaillierungsgrad ermöglicht, indem er sogenannte Energy Blocks aus Lastprofilen von Maschinen erzeugt und diese in einer Datenbank ablegt, die mit einer ereignisdiskreten Materialflusssimulation

gekoppelt wird. Hierbei sind zwar detaillierte Auswertungen über Energieverbräuche möglich, Rückkopplungen des Materialflusses beispielsweise auf die Umgebungstemperatur können aber nicht abgebildet werden (Weinert 2010).

Bei dem von Junge (2007) entwickelten Integrationsansatz werden mehrere spezialisierte Simulationswerkzeuge (SIMFLEX 3D für die Materialflusssimulation und TRNSYS für die thermische Gebäudesimulation) über eine TCP/IP Schnittstelle gekoppelt. Die Materialflusssimulation steuert die Simulation der Stoff- und Energieflüsse. Dabei werden Temperaturen, Energiebedarf, Wärmeabgabe und Emissionen betrachtet. Bei Beginn eines Bearbeitungsschrittes, der die Stoff- und Energieflüsse verändert, wird ein Signal von der Materialflusssimulation an die Gebäudesimulation übertragen und diese speichert den Zeitpunkt des Arbeitsbeginns. Zum Zeitpunkt des Bearbeitungsendes, der ebenfalls als Signal von der Materialflusssimulation übertragen wird, wird nun rückwirkend die aufgenommene Leistung berechnet. Die integrierten Wärme- und Emissionsmengen bilden nach Abschluss des Simulationslaufs die Basis für die Berechnung des Temperaturverlaufs (Junge 2007, S. 82). Der Ansatz von Junge (2007) bietet die Möglichkeit, Temperaturverläufe und Energieverbräuche detailliert abzubilden. Durch die implementierte Form der Vernetzung ist es allerdings nicht möglich, dass der Temperaturverlauf innerhalb eines Simulationslaufs Auswirkungen auf die Produktion hat.

Im Gegensatz zu den obigen Integrationsansätzen bietet der im Forschungsprojekt SimEnergy entwickelte Lösungsansatz die Möglichkeit, unter Einsatz von spezialisierten Simulationswerkzeugen bidirektionale Wechselwirkungen zwischen Materialfluss und energetischen Größen zu simulieren. Diese technische Umsetzung bildet die Voraussetzung, um energetische Aspekte bei der Planung und Steuerung der Produktionsabläufe zu berücksichtigen. Für das Gesamtverständnis wird daher zunächst der entwickelte Lösungsansatz kurz erläutert.

## 3 Gekoppelte Simulation von Material- und Energiefluss

Im Rahmen des Forschungsprojektes wird die Kopplung unter Verwendung des ereignisdiskreten Simulationswerkzeugs Plant Simulation von Siemens Tecnomatix umgesetzt, da dieses in der Automobilindustrie das am häufigsten eingesetzte ereignisdiskrete Simulationswerkzeug ist (Wenzel und Peter 2013, S. 252). Für die Simulation der Energieströme wird MATLAB Simulink als kontinuierlicher Simulator verwendet. Dieser dominiert den Markt der kontinuierlichen Simulation von physikalischen Vorgängen (Borshchev und Filippov 2004). Abbildung 1 zeigt das entwickelte Architekturkonzept zur Integration von Materialfluss- und Energiesimulation. Neben den beiden Simulationswerkzeugen wird die Software SimAssist der SimPlan AG, welche aus den von der Hessenagentur geförderten Projekten AssistSim (HA-Projekt-Nr.: 185/09-15) und EDASim (HA-Projekt-Nr.: 260/11-06) hervorgegangen ist, als Kommunikationsplattform eingesetzt. Sie weist eine gute Erweiterbarkeit und Anpassbarkeit auf (Gutfeld et al. 2014, S. 3430), wodurch es möglich ist, ein Plug-in zu entwickeln, das die Projektverwaltung, den Datenaustausch zwischen den Modellen und die Zeitsynchronisation steuert. Außerdem können die gekoppelten Modelle über diese Plattform parametrisiert (z. B. Anschlussleistungen von Maschinen oder die Dauer eines Simulationslaufs) werden.

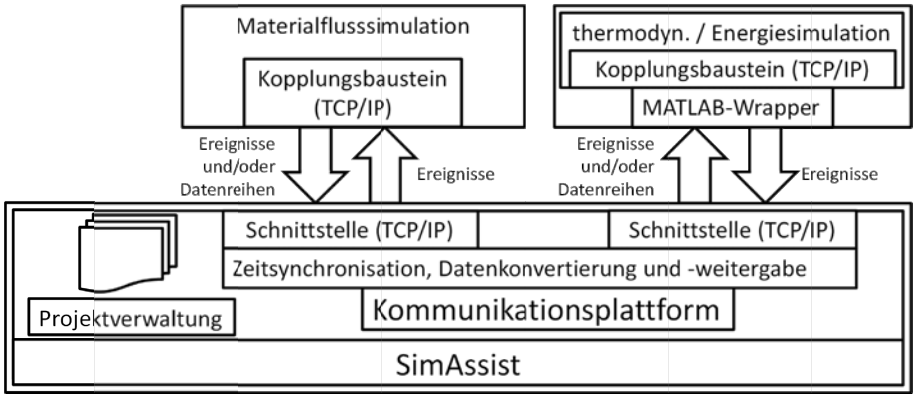



*Abbildung 1: Architekturkonzept der SimEnergy-Integrationslösung*

Die Zeitsynchronisation stellt sicher, dass die Simulationszeiten der beiden zu koppelnden Modelle aufeinander abgestimmt werden. Wird ein Simulationslauf von der SimAssist-Bedienoberfläche aus gestartet, läuft die Materialflusssimulation bis zum Ende des ersten Zeitschrittes; im Anschluss erfolgt die Kommunikation über die Kommunikationsplattform an die Energiesimulation. Hat die Energiesimulation denselben Zeitschritt erreicht, werden die ermittelten Ergebnisse über die Plattform an die Materialflusssimulation weitergereicht. Dieser Vorgang wiederholt sich für jeden Zeitschritt bis zu dem vom Anwender festgelegten Ende der Simulationsdauer. Beispiele für die im Rahmen des Kommunikationsprotokolls auszutauschenden Daten sind Ereignisse zu Änderungen von Maschinenzuständen oder Temperaturen. Nach Abschluss eines Simulationslaufs bietet die SimAssist-Plattform ergänzende Möglichkeiten zum Ergebnisexport, zur grafischen Aufbereitung der Simulationsergebnisse sowie zur Archivierung. Die Datenübertragung zwischen den Modellen und der SimAssist-Plattform ist unter Verwendung eines TCP/IP-Protokolls umgesetzt. Auf der Seite von Plant Simulation kommt der dort angebotene Socket-Baustein zum Einsatz, der Funktionalitäten für die Netzwerkkommunikation bereitstellt. Selbst programmierte Methoden legen fest, welche Daten gesendet und wie empfangene Daten innerhalb des Modells weiterverarbeitet werden. Zur Anbindung von MATLAB Simulink über TCP/IP wird als Rahmenapplikation ein sogenannter „MATLAB-Wrapper“ entwickelt, welcher die Codierung von zu sendenden Daten und die Decodierung von empfangenen Daten übernimmt. Um das Modell über den Wrapper anzusprechen, wird es als .dll kompiliert und in die Wrapper-Applikation eingebunden. Die Wrapper-Applikation ist außerdem in der Lage, das MATLAB Simulink Modell zu parametrisieren, und bezieht die Werte hierfür aus einer Konfigurationsdatei. Diese kann vom Simulationsanwender geändert werden, ohne dass er in das MATLAB-Modell eingreifen muss, so dass eine vereinfachte Durchführung von Parametervariationen in einem Simulationsexperiment gegeben ist. Die Gesamtarchitektur ist unabhängig vom verwendeten Simulationswerkzeug konzipiert, solange dieses über eine Schnittstelle zur TCP/IP Kommunikation verfügt.

# 4 Anwendung

Die in Kapitel 3 erläuterte Softwarelösung wird unter Verwendung der beiden Simulationswerkzeuge an einem Praxisbeispiel validiert. Hierzu werden in Zusammenarbeit mit Fachexperten aus Planung, Fabrikbetrieb und technischer Gebäudeausstattung des Anwendungspartners Volkswagen zunächst unterschiedliche Anwendungsfälle dahingehend bewertet, ob sie zum einen eine hinreichende bidirektionale Wechselwirkung zwischen Materialfluss und energetische Einflussgrößen besitzen und zum anderen ausreichend Daten zur Modellbildung vorliegen. Im Folgenden wird ein Anwendungsfall, der diese Vorgaben erfüllt, beschrieben.

## 4.1 Anwendungsfall in der Automobilindustrie

Als Anwendungsbeispiel wird ein Abschnitt aus einem Fertigungsbereich des Anwendungspartners gewählt, in dem Dreh- und Schleifprozesse auf verschiedenen Maschinen durchgeführt werden. Die Bearbeitungsmaschinen sind mit Schaltkästen ausgestattet, in denen die Maschinensteuerung untergebracht ist. Abhängig von der Leistungsaufnahme der Maschinen erwärmen sich diese Schaltkästen, die mit Wasser gekühlt werden, das wiederum über einen Kühlturm und entsprechende Rohrleitungen bereitgestellt wird. In diesem Anwendungsbeispiel hat die Produktion also einen Einfluss auf die Temperatur des Kühlwassers. Umgekehrt führen die Schaltschränke eine Notabschaltung durch, wenn eine Grenztemperatur von 44 °C (Lufttemperatur im Schaltschrank) überschritten wird. Ab dieser Grenztemperatur kann eine zuverlässige Kühlung nicht mehr sichergestellt werden, da die Temperaturen an den Chips bereits erheblich höher liegen. Obwohl die Kühlung der Schaltschränke großzügig ausgelegt ist, treten die beschriebenen Fälle an heißen Sommertagen auf. Mit einer Abschaltung hat der energetische Zustand der Schaltschränke somit eine Wirkung auf die Produktion, die im Falle einer Notabschaltung der Schaltschränke ebenfalls angehalten würde. Dadurch käme es zu einem Abkühlen der Schaltschränke. Eine bidirektionale Wechselwirkung ist also – wie gefordert – gegeben.

## 4.2 Entwickelte Modelle

Für die unter 4.1 beschriebene Problemstellung wurden zwei Simulationsmodelle entwickelt, die im Folgenden kurz beschrieben werden. Auf eine grafische Darstellung der Modelle muss aus Geheimhaltungsgründen verzichtet werden. Insgesamt werden vier Maschinen aus dem Gesamtsystem berücksichtigt. Es handelt sich dabei um drei Maschinen zum Drehen und eine Anlage zum Schleifen. Der Teiletransport zwischen den Maschinen erfolgt über verkettete Stetigförderer. Unter Verwendung des Bausteinkasten des Verbandes der Automobilindustrie (VDA) werden in Plant Simulation die Bearbeitungsmaschinen, die Fördertechnik sowie die Stationen zum Beladen und Entnehmen der Werkstücke auf die Werkstückträger modelliert. Die einzelnen Bearbeitungsstationen werden mit seitens der Fachexperten vorgegebenen Bearbeitungszeiten parametrisiert. Der in Kapitel 3 erläuterte Socket-Baustein für die anwendungsspezifische Modellkommunikation ist so parametrisiert, dass er die Betriebszustände der Maschinen auslesen und an die SimAssist-Plattform senden kann. Zusätzliche in Plant Simulation programmierte Methoden sorgen dafür, dass Maschinen abgeschaltet werden, sobald von MATLAB Simulink das Signal kommt,

dass eine Abschaltung aufgrund einer Überschreitung der definierten Grenztemperatur vorzunehmen ist.

Mittels MATLAB Simulink ist ein Modell erstellt, das die Schaltschränke der vier Maschinen als Raummodelle abbildet, die über eine Wärmekapazität (Außenwände sowie Luft) und über eine Wärmequelle (Chips) verfügen. Über die aufgenommene Leistung und einen Faktor für den Wärmeeintrag wird die Erwärmung des Schaltschrankes berechnet. Luftströmungen innerhalb des Schaltschrankes werden dabei ebenso wie die Wärmeübertragung durch Strahlung, Konvektion und Transmission (werkstoffabhängige Wärmeleitung in den Bauteilen) hinsichtlich eines praxistauglichen Detaillierungsgrades modellhaft abgebildet. Die Simulation berücksichtigt die Wärmeentwicklung der elektrischen Bauteile (Wärmequelle), die Wärmeübertragungsprozesse von der Wärmequelle an die Luft im Schaltschrank sowie von der Luft an das Kühlwasser und somit an die Umgebungsluft (Außenluft). Aus der Erwärmung und der Abkühlung wird kontinuierlich eine momentane Schaltschranktemperatur berechnet, die in den vom Simulationsanwender festgelegten Zeitabständen an die SimAssist-Plattform übermittelt wird. Die Hallentemperatur beeinflusst das Erwärmungsverhalten ebenfalls. Vom Anwendungspartner liegen Messdaten zur Hallentemperatur vor, die in Abhängigkeit von der Tages- und Jahreszeit schwanken und somit Einfluss auf die Simulationsergebnisse nehmen können. Die entsprechenden Hallentemperaturwerte sind in Plant Simulation als Tabelle hinterlegt. Je nach Simulationszeit wird der entsprechende Wert an die SimAssist Plattform gesendet, die ihn an das MATLAB Simulink Modell übergibt, in dem er als Eingangswert für den nächsten Simulationszeitschritt dient.

Die Validierung der Einzelmodelle erfolgt im Gespräch mit den Fachexperten des Anwendungspartners. Dabei kann beim Energiemodell der Schaltschränke auf umfangreiche Messdaten zurückgegriffen werden. Die Ergebnisse der thermischen Simulation decken sich weitgehend mit den tatsächlichen Messwerten. Das Materialflussmodell wird im Rahmen einer Face Validation (vgl. Rabe et al. 2008) in Zusammenarbeit mit Fertigungsingenieuren des Anwendungspartners validiert. Nach Validierung der Einzelmodelle erfolgt eine Überprüfung der gekoppelten Modelle in Bezug auf Korrektheit und Vollständigkeit der auszutauschenden Daten. Auf Basis der Einzelmodelle und ihrer Kopplung werden Simulationsläufe durchgeführt, die im Folgenden beschrieben werden.

## 4.3 Simulationsläufe und Ergebnisse

Im Rahmen der bisher durchgeführten Simulationsläufe enthält das Plant Simulation Modell keine zufallsabhängigen Parameter, um zunächst nur die zeitlichen Abhängigkeiten der Ereignisse durch die wechselnden Temperaturen im Tagesablauf zu untersuchen. Auf Wiederholungen von Simulationsläufen mit unterschiedlichen Startwerten kann somit verzichtet werden, da die Ergebnisse davon nicht beeinflusst werden. Auf Basis der entwickelten Modelle werden unterschiedliche Simulationsläufe durchgeführt, bei denen die Abschalttemperatur der Maschinen variiert und die Auswirkungen auf die Ausbringungsmenge untersucht werden. Im ersten Schritt wird ein Simulationslauf durchgeführt, bei dem keine Abschaltung in Abhängigkeit der Temperatur stattfindet. Im Anschluss werden als Abschalttemperatur 44 °C festgelegt; die Maschinen werden wieder in den Betriebszustand versetzt, sobald die Schaltschränke auf 40 °C abgekühlt sind. Im dritten Lauf wird als Abschalttemperatur 40 °C definiert und der Betrieb wird bei Erreichen von 36 °C

Schaltschranktemperatur wieder aufgenommen. Die Simulationsdauer beträgt jeweils 24 Stunden. Die Temperatur der Schaltschränke hängt von der Hallentemperatur ab, die wiederum vom Wetter beeinflusst wird. Für die Durchführung der Simulationsläufe sind Extremwerte am interessantesten; daher wird ein Tag im Juni mit einer jahreszeittypischen hohen Hallentemperatur gewählt. Abbildung 2 zeigt die Temperaturentwicklung der vier betrachteten Schaltschränke im Zeitverlauf.

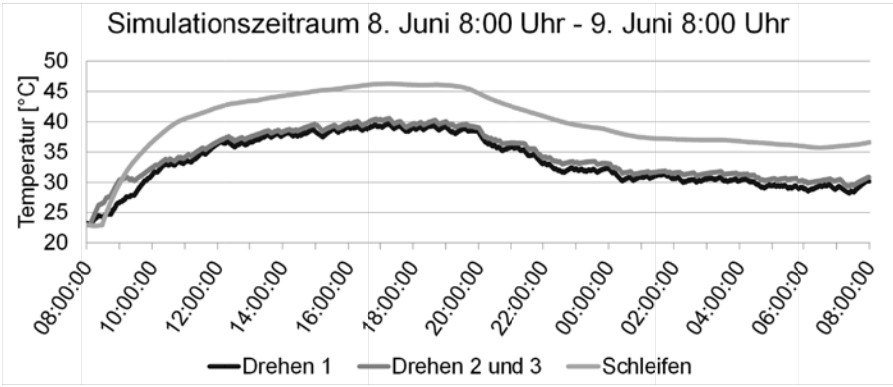



*Abbildung 2: Temperaturentwicklung Schaltschränke – Lauf 1*

Die Temperaturen der Schaltschränke „Drehen 2 und 3“ sind zusammengefasst, da sich die Temperaturwerte nur geringfügig unterscheiden und separate Liniendarstellungen im Diagramm nicht unterscheidbar gewesen wären. Zu erkennen ist, wie die Temperatur vom Maschinenzustand abhängt (Schwankungen bei den Anlagen „Drehen 1“ und „Drehen 2 und 3“) und dass die Temperatur in der Nachtschicht durch die geringere Außen- und damit Hallentemperatur absinkt.

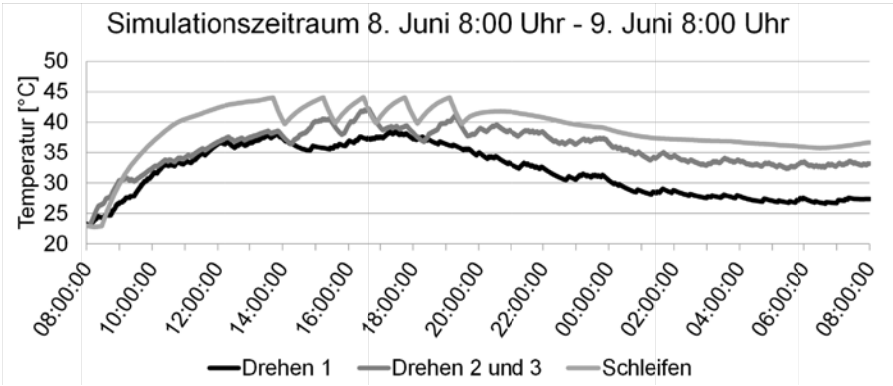



*Abbildung 3: Temperaturentwicklung Schaltschränke – Lauf 2*

In einem weiteren Simulationslauf (vgl. Abbildung 3) werden jeweils Bedingungen für das Ab- und Anschalten der Maschinen definiert. Das MATLAB Simulink Modell sendet ein Abschaltsignal für die jeweilige Bearbeitungsstation an das Plant

Simulation Modell, sobald die Temperatur des entsprechenden Schaltschrankes 44 °C erreicht. Hierbei ist zu erkennen, dass in diesem Szenario nur die Anlage „Schleifen“ von einer temperaturbedingten Abschaltung betroffen ist. Zum Zeitpunkt der Abschaltung kühlt sich der Schaltschrank ab, bis die Temperatur für das Wiederanschalten (hier 40 °C) erreicht ist. Die Temperatur der Schaltschränke „Drehen 2 und 3“ schwankt ebenfalls zeitlich versetzt zu der Temperatur der Anlage „Schleifen“. Dies ist dadurch zu erklären, dass sich die Teile im Materialflussmodell vor der Anlage „Schleifen“ stauen und die Drehmaschinen nicht produzieren, wenn der Puffer vor der Schleifmaschine voll ist. Der Durchsatz an Teilen liegt in diesem Lauf 25 % niedriger als in Lauf 1.

In einem dritten Simulationslauf wird die erlaubte Maximaltemperatur auf 40 °C gesenkt, um die Auswirkungen auf das Abschaltverhalten und damit auf die erreichbare Stückzahl zu untersuchen. Abbildung 4 zeigt die Temperaturverläufe für die betrachteten Anlagen. Erneut ist von einer temperaturbedingten Abschaltung nur die „Anlage Schleifen“ betroffen, die anderen Anlagen erreichen die definierte Grenztemperatur nicht. Bei den Anlagen „Drehen 2 und 3“ kommt es wiederum zu Abschaltungen, die nicht temperaturbedingt sind. Die Ursache ist, wie zuvor, ein voller Puffer, der ein Weiterarbeiten der Drehmaschinen verhindert; dieser Zusammenhang ist im Plant Simulation Modell erkennbar. Die Abschaltung tritt zeitlich versetzt ein, da es eine gewisse Zeit dauert, bis die Pufferstrecken voll sind. Das beobachtete Verhalten wird im Gespräch mit den Fachexperten des Anwendungspartners bestätigt. In diesem dritten Simulationslauf liegt die zugehörige Ausbringungsmenge in 24 Stunden 35 % niedriger als im ersten Simulationslauf ohne Abschaltungen.

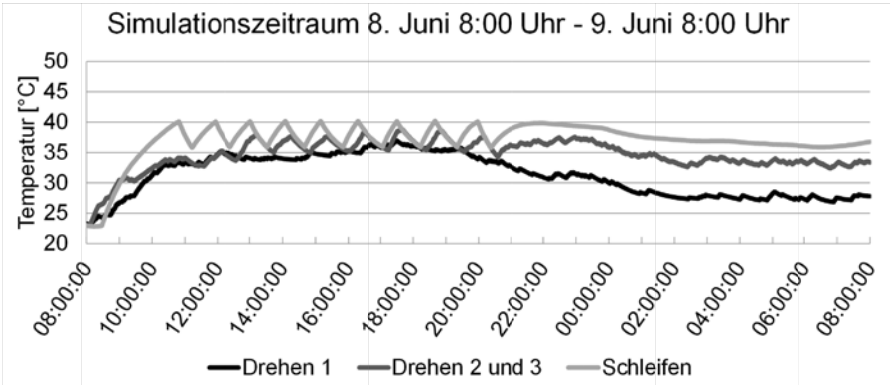



*Abbildung 4: Temperaturentwicklung Schaltschränke – Lauf 3*

Zusammenfassend ist festzuhalten, dass die Kopplung technisch erfolgreich umgesetzt ist. Die SimAssist Plattform protokolliert die gesendeten und empfangenen Daten; die Datenkommunikation ist vollständig transparent. Des Weiteren wird mit dem Anwendungsbeispiel gezeigt, dass bestehende Wechselwirkungen zwischen Produktion und Energiefluss auf Basis gekoppelter Simulationsmodelle nachvollzogen und bewertet werden können. Die Ergebnisse sind im Gegensatz zu den in Kapitel 2 dargestellten Integrationsansätzen im Falle der Temperaturkurven – bedingt durch das verwendete Simulationswerkzeug MATLAB Simulink – sehr

genau. Funktionalität und Anwendbarkeit der bidirektionalen Kopplung sind somit umfassend gezeigt und in Zusammenarbeit mit dem Anwendungspartner validiert worden.

## 5 Ausblick

Die entwickelte Integrationslösung bietet die Möglichkeit, bidirektionale Wechselwirkungen in ereignisdiskreten Produktionsmodellen und kontinuierlichen Energiemodellen simulativ zu untersuchen und Maßnahmen zur Bewertung und Verbesserung der Energieeffizienz von Produktionssystemen vorab zu überprüfen und abzusichern. Bei der Suche nach einem geeigneten Anwendungsfall in der Automobilindustrie für die entwickelte Integrationslösung hat sich allerdings gezeigt, dass auch heute noch in der Produktion die Erreichung der Sollstückzahl das wichtigste Ziel darstellt und Anpassungen an der Produktionssteuerung zugunsten des Energieverbrauchs nur zulässig sind, wenn sie nicht zu Lasten der Stückzahl gehen. Des Weiteren sind Energiebereitstellungsanlagen oftmals so dimensioniert, dass sie die Anforderungen aufgrund hoher Sicherheitszuschläge bewältigen können. Eine Betrachtung von möglichen Wechselwirkungen zwischen Energieverbrauch und Produktion erscheint daher zum gegenwärtigen Zeitpunkt allein aus Anwendersicht wenig relevant. Allerdings sorgen zurzeit energieintensive Fertigungsverfahren in Verbindung mit steigenden Energiepreisen für einen höheren Leidensdruck in den Unternehmen und für eine größere Bereitschaft Energie- und Medienverbrauch als gleichrangiges Planungsziel neben der Stückzahl zu akzeptieren. Aus diesem Grund soll die Lösung auch auf andere Anwendungsfälle außerhalb der Automobilindustrie übertragen werden. Denkbar wären Anwendungsfälle in Unternehmen der Gießereibranche, der Lebensmittelindustrie, Druckereien oder in Verzinkereien. Diese haben den Umstand gemeinsam, dass die Produktion stark von den klimatischen Bedingungen (Temperatur, Luftfeuchtigkeit) abhängt und somit Wechselwirkungen zwischen Produktion und Energie- und Medienbereitstellung bestehen. Des Weiteren hat sich in der Zusammenarbeit im Forschungsprojekt gezeigt, dass einzelne Fachabteilungen mit ihren jeweiligen spezifischen Kompetenzen fachliche Unterstützung benötigen, wann der Einsatz einer gekoppelten Simulation notwendig und sinnvoll ist. Auch an dieser Stelle sind weitere Forschungsarbeiten notwendig.

## Danksagung

Die Ergebnisse entstanden im Rahmen eines gemeinsamen Forschungsvorhabens der SimPlan AG, der Limón GmbH, der Volkswagen AG und des Fachgebietes Produktionsorganisation und Fabrikplanung der Universität Kassel. Das vom 01.04.2013 bis 31.03.2015 laufende Projekt mit dem Titel "Simulationsgestützte Planung und Bewertung der Energieeffizienz für Produktionssysteme in der Automobilindustrie" (HA-Projekt-Nr.: 369/13-08) wurde im Rahmen von Hessen ModellProjekte aus Mitteln der Energietechnologieoffensive Hessen gefördert.

## Literatur

Borshchev, A.; Filippov, A.: From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools. In: Kennedy, M.; Winch,

G.W.; Langer, R.S.; Rowe, J.I.; Yanni, J.M. (Hrsg.): Proceedings of the 22nd International Conference of the System Dynamics Society, Oxford 2004.

Erlach, K.; Westkämper, E. (Hrsg.): Energiewertstrom. Der Weg zur energieeffizienten Fabrik. Stuttgart: Fraunhofer-Verlag 2009.

Gutfeld, T.; Jessen, U.; Wenzel, S.; Laroque, C.; Weber, J.: A technical concept for plant engineering by simulation-based and logistic-integrated project management. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): Proceedings of the Winter Simulation Conference. IEEE, Omnipress 2014, S. 3423-3434.

Herrmann, C.; Thiede, S.; Kara, S.; Hesselbach, J.: Energy oriented simulation of manufacturing systems – Concept and application. CIRP Annals - Manufacturing Technology 60 (2011) 1, S. 45–48.

Hesselbach, J.: Energie- und klimaeffiziente Produktion. Grundlagen, Leitlinien und Praxisbeispiele. Wiesbaden: Vieweg+Teubner 2012.

Junge, M.: Simulationsgestützte Entwicklung und Optimierung einer energieeffizienten Produktionssteuerung. Kassel: Kassel Univ. Press. Produktion & Energie. Bd. 1. 2007.

Mendes, N.; Oliveira, G.H.; de Araújo, H.X.: Building thermal performance analysis by using MATLAB/SIMULINK. In: Lamberts, R.; Negrao, C.; Hensen, J. (Hrsg.): Proceedings of Seventh International IBPSA Conference, Rio de Janeiro (Brazil) 2001, S. 473-480.

Rabe, M.; Spieckermann, S.; Wenzel, S.: Verifikation und Validierung für die Simulation in Produktion und Logistik. Vorgehensmodelle und Techniken. Berlin: Springer. 2008.

Schlegel, A.; Stoldt, J.; Putz, M.: Erweiterte Integration energetischer Betrachtungen in der Materialflusssimulation. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): Simulation in Produktion und Logistik 2013. Paderborn: Heinz-Nixdorf-Inst. Univ. Paderborn 2013, S. 187-196.

SIEMENS INDUSTRY SOFTWARE GMBH & CO. KG 2015. *Plant Simulation* [online] [Zugriff am: 11. Mai 2015]. Verfügbar unter: http://www.plm.automation.siemens.com/de_de/products/tecnomatix/plant_design/plant_simulation.shtml

VDI: Produktion und Logistik in Deutschland 2025. Trends, Tendenzen, Schlussfolgerungen. Düsseldorf: VDI 2012.

Weinert, N.: Vorgehensweise für Planung und Betrieb energieeffizienter Produktionssysteme. Berichte aus dem Produktionstechnischen Zentrum Berlin. Stuttgart: Fraunhofer-Verlag 2010.

Wenzel, S.; Peter, T.: Simulation zur Sicherstellung der Wettbewerbsfähigkeit – Ergebnisse einer Umfrage zur Simulation in Produktion und Logistik. In: Friedewald, A.; Lödding, H. (Hrsg.): Produzieren in Deutschland. Wettbewerbsfähigkeit im 21. Jahrhundert. Berlin: Gito 2013, S. 243-264.

Wenzel, S.; Weiß, M.; Collisi-Böhmer, S.; Pitsch, H.; Rose, O.: Qualitätskriterien für die Simulation in Produktion und Logistik. Planung und Durchführung von Simulationsstudien. Berlin: Springer 2008.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Energieeffizienzsteigerung im Tagebaubetrieb durch simulationsgestützte Untersuchung konkurrierender Flexibilitäten

## *Improving Energy Efficiency in Open-cast Mining Operations through Simulative Studies of Competing Flexibilities*

Christin Fanghänel, Andreas Schlegel, Johannes Stoldt, Matthias Putz, Fraunhofer Institut für Werkzeugmaschinen und Umformtechnik IWU, Chemnitz (Germany), christin.fanghaenel@iwu.fraunhofer.de, andreas.schlegel@iwu.fraunhofer.de, johannes.stoldt@iwu.fraunhofer.de, matthias.putz@iwu.fraunhofer.de

Hans Rüdiger Lange, Thomas Woldt, Vattenfall Europe Generation AG, Cottbus (Germany), hans-ruediger.lange@vattenfall.de, Thomas.woldt@vattenfall.de

Abstract: Coal mining operations are not just a primary source for the generation of electricity but also a large consumer thereof. Hence, the optimisation of the power station and mine tandem is a prime concern to retain profitability despite the surging market shares of renewable energy sources. This paper describes the use of discrete event simulation (DES) for modelling a lignite open-cast mine to identify and analyse potentials for energy efficiency therein. In particular, the characteristics of the real system as well as the procedure for creating the model are presented. By means of the developed simulation model influenceable operational procedures, feasible margins for change, as well as associated effects can be determined.

## 1 Einleitung

Im Forschungsfeld der Energieeffizienz wurde Ablaufsimulation bereits verschiedentlich eingesetzt. Nachstehend werden die Motivation und der Ansatz hinter der in diesem Beitrag vorgestellten Anwendung in einem Tagebau diskutiert.

### 1.1 Motivation

In der von der Politik gestalteten Energiewende bilden der Ausbau der erneuerbaren Energien und die Steigerung der Energieeffizienz die wesentlichen Schwerpunkte. Infolgedessen ist der Anteil der erneuerbaren Energien an der Bruttostromerzeugung in Deutschland in den letzten Jahren stetig gewachsen und zum bestimmenden Element des Strommarktes geworden. Zugleich wird jedoch circa ein Viertel des erzeugten Stroms (≈ 161 TWh) aus Braunkohle gewonnen (BMWi 2014). Damit gewährleistet die konventionelle Energieerzeugung aus Braunkohle in wind- und

sonneschwachen Zeiten eine zuverlässige Grundversorgung und bei hoher Einspeisung erneuerbarer Energien die erforderliche Flexibilität zu deren Integration in das Stromnetz. Zugleich existieren in der Gewinnung und Verstromung von Braunkohle als stromintensive Produktion Potenziale für die Steigerung der Energieeffizienz. Aufgrund der bereits langjährig betriebenen Optimierung und der technisch-physikalischen Grenzen ist damit zu rechnen, dass Effizienzgewinne an der Anlagentechnik zunehmend in den Hintergrund treten werden. Vielmehr wird das Zusammenwirken einzelner Systembestandteile Ausgangspunkt für Steigerungen der Energieeffizienz sein. Hierfür ist es erforderlich, die Flexibilitäten der Teilsysteme systematisch zu erschließen, um den Wechsel zwischen Last und Erzeugung sowie zwischen konventionellen und erneuerbaren Energiequellen prozessübergreifend zu erleichtern.

Um die Voraussetzungen für eine Steigerung der systemischen Energieeffizienz in den Kraftwerken und Tagebaubetrieben zu schaffen, wurden von Vattenfall bereits vielfältige Maßnahmen der Flexibilisierung im Erzeugungs- und Lastmanagement umgesetzt (Lange et al. 2014). Diese Maßnahmen sind Ausgangspunkt für systematische Analysen zu den Prozessstrukturen und zum Energieeinsatz im Tagebau.

Dieser Beitrag stellt dar, wie Ablaufsimulation zur Identifikation und Abschätzung der diversen Flexibilitäts- und Optimierungspotenziale in einem, in den Strommarkt und die Stromerzeugung eingebundenen Tagebaubetrieb, genutzt werden kann. Resultierend sollen hierdurch die Steigerung der systemischen Energieeffizienz durch zeitaufgelöste Optimierung der Energieintensität (d. h. Realisierung einer vorgegebenen Ausbringungsmenge durch einen zeitoptimalen Energieeinsatz), die damit verbundene Kostensenkung sowie insgesamt eine Verbesserung der Prozessqualität durch eine präzisere Beherrschung der zunehmend flexibleren Produktionsprozesse erreicht werden.

## 1.2 Konzept der Simulationsstudie

Um die angestrebten Ziele zu realisieren, erfolgte die modellhafte Abbildung des Tagebaus mit dem Simulationswerkzeug Tecnomatix Plant Simulation von Siemens. Die dynamische ereignisdiskrete Materialflusssimulation soll dazu beitragen, die Bewirtschaftung eines Tagebaubetriebs unter Beachtung konkurrierender Flexibilitäten und Einflüsse (Personaleinsatz, Elektroenergiebezug, Just-in-Time-Kohlelieferung, Witterungsrisiken, Störungen, Instandhaltung) zu verbessern. Vorarbeiten zur Simulation von Erdbau- und Minenprozessen, allerdings ohne Energiebetrachtungen, finden sich etwa bei Basu und Baafi (1999) sowie Ji et al. (2010).

Der Ansatz der Nutzung von Materialflusssimulation erscheint hierbei besonders geeignet, da mit diesem die Vielzahl stochastischer Einflüsse, welche im System wirken, modelliert und in die Untersuchungen einbezogen werden können. Zugleich ist festzustellen, dass im betrieblichen Alltag bereits eine Diskretisierung bei der Datenerhebung vorgenommen wird, etwa bei der Erfassung abgetragener Bodenarten. Vereinfacht lässt sich ein Tagebau daher mit einem verketteten Logistiksystem vergleichen, wie es vielmals im Fahrzeugbau vorgefunden wird. Demzufolge wird der Tagebau als getaktete Fertigung mit variablen Produktions-anforderungen interpretiert und bewährte Ansätze der Flexibilisierung aus dem Automobilbau werden im Sinne eines „cross learning“ auf den Energiesektor übertragen. Die angestellten Untersuchungen haben das Ziel, diese Annahme zu bestätigen und übertragbare Ansätze zur Betriebsverbesserung abzuleiten.

Um insbesondere die Flexibilitäts- und Optimierungspotenziale im Hinblick auf den Ressourceneinsatz identifizieren zu können, ist es notwendig, dass neben dem Material- auch der Energiefluss für die Simulation modelliert wird. Zur Abbildung der energetischen Verbräuche in Abhängigkeit unterschiedlicher Betriebszustände wird der generische Simulationsbaustein eniBRIC verwendet (Schlegel et al. 2013). Dieser Baustein ist branchenneutral konzipiert und eignet sich für die Modellierung unterschiedlicher Anlagen und Betriebsmittel. Insbesondere seine Fähigkeit gezielt in Betriebszustände geschaltet zu werden, welche sich auch auf den Materialfluss auswirken, qualifiziert ihn für Untersuchungen, welche Verbesserungen im betrieblichen Energieeinsatz mittels betriebsorganisatorischer Ansätze zum Ziel haben. Hiermit erweiterte Simulationsmodelle erlauben folglich die gleichzeitige Betrachtung von Material- sowie Energieflüssen in einem Softwarewerkzeug und dienen zur digitalen Absicherung und Bewertung identifizierter Potenziale.

## 2 Charakterisierung des Realsystems

Betrachtungsgegenstand der simulationsgestützten Untersuchungen ist der Braunkohletagebau Welzow-Süd. Um die Komplexität der zu simulierenden Produktionsprozesse in einem Tagebau zu verdeutlichen, werden in den folgenden Abschnitten die Prozessstrukturen und -elemente sowie die Material- und Energieflüsse erläutert.

### 2.1 Produktionsprozesse im Tagebau

Der Tagebau Welzow-Süd ist durch seine kontinuierliche Tagebautechnik charakterisiert. Bei dieser wird das abzutragende Material kontinuierlich aus dem Gebirgsverband (d. h. überlagernde Erdschichten sowie Kohleflöz) gelöst, geladen, transportiert, verstürzt bzw. zur Verladung gebracht (Stoll et al. 2009). Das Verhältnis des zu bewegenden Abraumvolumens (m³) zum gewinnbaren Kohlevorrat (t) beläuft sich auf etwa 6:1. Genauer müssen für die jährliche Förderung von 20 Millionen Tonnen Rohbraunkohle etwa 110 Millionen m³ Abraum bewegt werden.

Des Weiteren zeichnet sich der betrachtete Tagebau durch ein Direkt-Versturz-System aus. Dabei wird mittels einer Abraumförderbrücke das gewonnene Material über den offenen Tagebau transportiert und anschließend direkt auf eine Kippe verstürzt. Ein derartiger Verband ist sehr leistungsfähig und weist niedrigere spezifische Produktionskosten auf. Folglich wird im Tagebaubetrieb beabsichtigt die maximal mögliche Menge an Abraum mit dem Brückenverband zu fördern. Im Vorschnitt I und II arbeiten zwei Bagger der Abraumförderbrücke voraus und tragen die Erdschichten ab, die durch den Brückenverband (bestehend aus Förderbrücke und zwei angeschlossenen Eimerkettenbaggern) nicht abgebaut werden können. Die Abraumbandanlagen transportieren das Vorschnittmaterial auf die ausgekohlte Kippenseite des Tagebaus. Hier wird der Abraum durch Absetzer verkippt. Die Freilage des Kohleflözes erfolgt durch den Brückenverband. Unterhalb der Abraumförderbrücke befindet sich die Grube, in der der eigentliche Abbau der Rohbraunkohle erfolgt. In diesem Betriebsbereich wird die Kohle selektiv entsprechend ihrer Qualität durch fünf Bagger im Hoch- und Tiefschnitt gefördert und über eine Bandanlage entweder direkt zur Zugverladung oder zum Grabenbunker transportiert. Der Tagebauproduktionsprozess sowie die Einteilung in die fünf Betriebsbereiche werden in Abbildung 1 verdeutlicht.

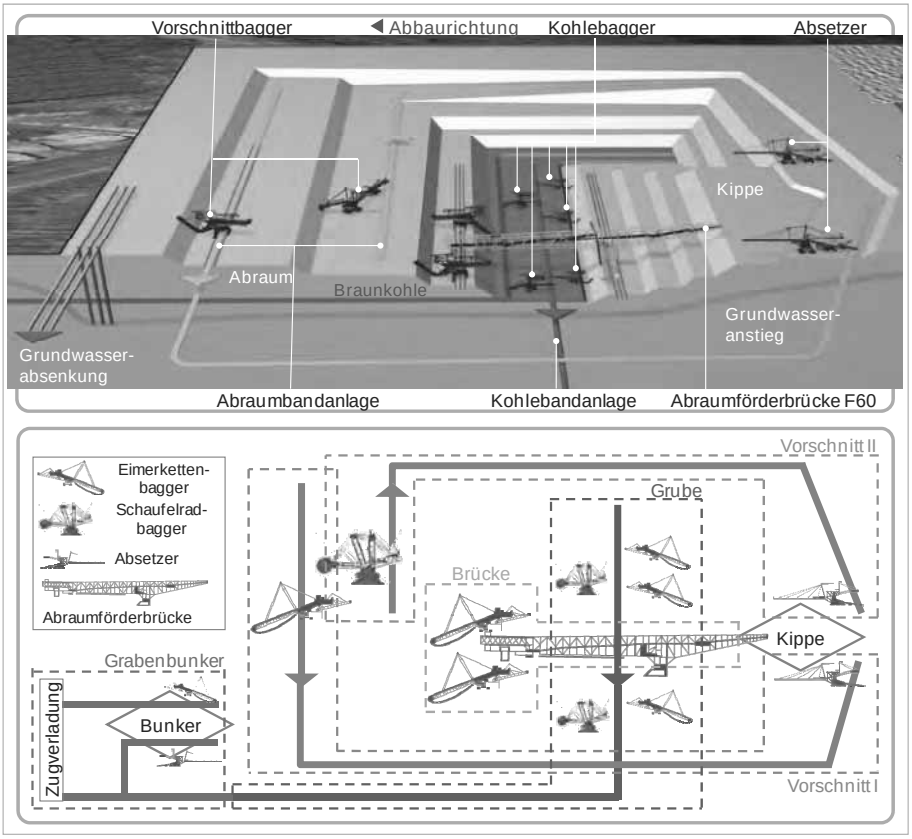



*Abbildung 1: Tagebauproduktionsprozess Welzow-Süd*

Im Grabenbunker wird die Kohle schließlich mittels eines Schüttgerätes auf Halden verkippt, die qualitätsseitig unterschieden werden. Die Gesamtkapazität beträgt ca. 170.000 t. Am Wochenende erfolgt keine direkte Zugverladung aus der Grube. Stattdessen werden die Züge über den Grabenbunker verladen. Darüber hinaus kann der Lagerplatz auch für Engpässe bei der Kohlelieferung verwendet werden. Ein Füllstand von 90.000 t ist jedoch nicht zu unterschreiten.

Die einzelnen Betriebsbereiche lassen sich als eigenständige Prozesslinien auffassen, die auch unabhängig voneinander bewirtschaftet werden können. Allerdings existieren hierbei Einschränkungen. Beispielsweise muss die vom Kraftwerk angeforderte Kohle auch vom Brückenverband freigelegt werden. Ebenso ist ein bestimmter Vorlaufbereich, d. h. die bisher abgetragene Erdschicht, zum nachfolgenden Betriebsbereich einzuhalten. Die Verwendung dieser Flächen als Flexibilitäten ist nicht vorgesehen, da die Organisation der maximalen Betriebszeit stets Priorität hat. Diese Vorgehensweise begründet sich in den häufig auftretenden Störungen. Bezugnehmend auf die Vorschnittbagger wird die verfügbare Betriebszeit eines Monats jeweils um bis zu 20 % durch ungeplante Stillstände beansprucht. Aufgrund dessen können mögliche Eingriffe in den laufenden Betrieb zur Steigerung der Energieeffizienz auch nur im Kurzzeitbereich erfolgen. Abschaltkonzepte zur Reduzierung

des Ressourcenbedarfes, wie sie durch Knafla und Richter (2013) im Karosseriebau eingesetzt werden, sind daher nicht ohne Weiteres auf den Tagebau übertragbar.

## 2.2 Energieeinsatz

Der Tagebau Welzow-Süd hat einen jährlich schwankenden Elektroenergiebedarf von 300 – 350 GWh. Etwa 50 % davon entfallen auf die Vorschnitte und ca. 30 % auf den Brückenverband. Die verbleibenden 20 % verteilen sich auf Grube und Bunker sowie auf Nebenanlagen, wie z.B. die Oberflächenentwässerung. Um die Dimensionen weiter zu verdeutlichen sind in Tabelle 1 die stündlichen Durchschnittsverbräuche und Energiekosten ausgewählter Systemelemente sowie des gesamten Tagebaus aufgeführt. Der Stromverbrauch bezieht sich dabei rein auf Betriebszeiten, d. h. Grundlasten in Stillstandszeiten wurden hier nicht einbezogen.

*Tabelle 1: Stdl. Stromverbrauch und Energiekosten für ausgewählte Systemelemente*

| | Ø-Stromverbrauch MWh | Ø-Energiekosten (€/h) |
|---|---|---|
| Schaufelradbagger (VS II) | 3,7 | 109,09 |
| Abraumbandanlage (VS II) | 13,77 | 406,72 |
| Absetzer (VS II) | 2,71 | 80,40 |
| Eimerkettenbagger (Brücke) | 4,31 | 144,71 |
| Abraumförderbrücke | 8,04 | 271,22 |
| Schaufelradbagger (Grube) | 0,9 | 28,22 |
| Σ Welzow-Süd | 65,33 | 2.085,32 |

Im Energiesystem nimmt der Tagebau prinzipiell zwei Rollen ein. Einerseits ist er Lieferant für die Kraftwerke und somit für die Stromerzeugung, andererseits ist er Kunde dieser. Für die Steigerung der systemischen Energieeffizienz ist sowohl die Rolle als Kohlelieferant als auch die des Elektroenergiekunden von wesentlicher Bedeutung. Aus diesem Grunde wurden der Energieeinsatz im Tagebau sowie mögliche Einflüsse für eine detaillierte Untersuchung mittels Simulation ausgewählt.

Der Energiebedarf in einem Braunkohletagebau wird nicht alleinig durch die Betriebszeiten der Systemelemente bestimmt, sondern auch maßgeblich durch das Produktionsumfeld beeinflusst. Stromverbrauch und effektive Förderleistung sind vor allem von den geologischen Gegebenheiten vor Ort abhängig. In den abzutragenden Erdschichten variiert der Boden beispielsweise hinsichtlich des Feuchtigkeitsgehaltes, was sich unmittelbar auf die Transporteigenschaft des geförderten Materials auswirkt und eine angepasste Fahrweise erfordern kann. In diesem Betriebszustand verringert sich die Leistungsaufnahme, aber auch die Förderleistung. Weiterhin wird der Elektroenergiebedarf erheblich durch die zu bewältigenden Längen- und Höhenunterschiede der Bandanlagen beeinflusst. Diese können im Vorschnitt bis zu 14 km lang sein und müssen auf dem Weg zum Bunker bis zu 80 Höhenmeter bewältigen. Da die Geschwindigkeit der Gurtförderer konstant ist, weisen die Bandanlagen zudem eine massenabhängige, d. h. bodenabhängige, Leistungsaufnahme auf.

# 3 Vorbereitungen der Simulationsuntersuchungen

Um die Untersuchung hinsichtlich bestehender Flexibilitäten und Optimierungspotenziale vornehmen zu können, wurde das betrachtete System Tagebau in Vorbereitung späterer Simulationen modelliert. Dies ist nachfolgend detailliert.

## 3.1 Systemgrenzen und Abstraktion

Im Rahmen der Modellierung wurden zunächst Systemgrenzen definiert, welche die zu betrachtenden Elemente und Ströme eingrenzen. Für das betrachtete System wurde festgelegt, dass lediglich die Systemelemente betrachtet werden, welche unmittelbar mit den Materialflüssen von Kohle und Abraum in Verbindung stehen sowie der Fluss elektrischer Energie. Weitere indirekte Bereiche des Tagebaus, etwa die Wasserwirtschaft, der Betrieb von Hilfsgeräten, die Vorfeldberäumung oder die Rekultivierung, finden keine Beachtung. Die Kohleverladung in die Züge stellt die äußere räumliche Grenze des zu simulierenden Systems dar. Das Kraftwerk fungiert in den Untersuchungen lediglich als innerbetrieblicher Auftraggeber und initiiert die Fertigungsaufträge. Dabei stellen angeforderte Kohlemengen die primären und die zuvor abzutragenden Erdschichten die sekundären Produktionsaufträge dar.

Wie in Abschnitt 2.2 beschrieben, beeinflusst der Aufbau der Lagerstätte inklusive Kohleflöz und überlagerndem Deckgebirge (geologisches Profil) den Elektroenergiebedarf sowie die effektive Förderleistung. Um diese Abhängigkeiten in die Modellierung einbeziehen zu können, ist das geologische Profil zu abstrahieren. Die neun vorkommenden Bodenarten werden hierzu auf die wesentlichen reduziert und durch eine mittlere Zusammensetzung charakterisiert. In dieser Weise beschriebene Erdschichten werden in durchschnittliche Volumenblöcke eingeteilt, welche in fünf Minuten Normalbetrieb gefördert werden. Resultierend entstehen Abtragungsreihenfolgen für die Vorschnitte sowie für den Brückenverband. Analog wird das Flöz qualitätsseitig in Brikettier-, Förder- und Kesselkohle unterschieden und die entsprechende Abfolge für die Gewinnung erzeugt.

## 3.2 Datenanalyse

Vattenfall verfügt über ein Echtzeitdatensystem für die Elektroenergieproduktion, welches zurzeit um Elemente der Tagebauproduktion erweitert wird, um eine integrierte Optimierung von Markt-, Netz-, Erzeugungs- und Produktionsdaten in einem einheitlichen Steuerungssystem zu ermöglichen. Ausgehend von den umfangreichen Realdaten erfolgte eine Analyse der zeitlichen Prozessabläufe und der assoziierten Energiebedarfe, welche zunächst einen Ist-Stand des Tagebaus liefern. Dessen Abstraktion stellt den Ausgangspunkt für die Modellierung der Betriebsmittel sowie der Produktionsprozesse dar.

Im Fokus der Datenanalyse standen die Identifizierung der Betriebszustände (energetische Lastfälle, welche sich auf den Materialfluss auswirken) sowie der Nachweis der im Abschnitt 2.2 beschriebenen Einflüsse auf die Leistungsaufnahme. Um eine realistische Zuordnung von Ressourcenbedarf und Förderleistung zu den entsprechenden Betriebszuständen vornehmen zu können, wurden mehrere Datensätze und -quellen sowie das Fachwissen bei Vattenfall einbezogen. Die energetischen Betriebszustände wurden entsprechend definiert und hinsichtlich ihres Prozesscharakters beschrieben. Für den Eimerkettenbagger im Vorschnitt I ergeben sich beispielsweise die fünf Betriebszustände Aus, Gestört, Fahren, Hoch-/Tiefschnitt-

wechsel, Lastgradanpassung und Normalbetrieb, wohingegen ein Kohlebagger nur die drei Betriebszustände Aus, Anfahren und Normalbetrieb aufweist. Um die jeweiligen Verbräuche und Förderleistung zu ermitteln, wurden Berichtsdaten, zeitaufgelöste Messdaten zu Stromverbrauch und Förderleistung sowie die geplanten und ungeplanten Stillstandzeiten miteinander in Verbindung gesetzt.

## 3.3 Modellierung

Um das Prozessverständnis der zunehmend flexibleren Tagebauproduktion zu verbessern, wurden alle Betriebsbereiche in einem Netzwerk (d. h. Teilmodell) abgebildet. Wie in Abbildung 2 dargestellt, wurde das Layout des Simulationsmodells dem realen des Tagebaus nachempfunden. Auf eine Visualisierung der Lagerstätte wurde hingegen verzichtet.

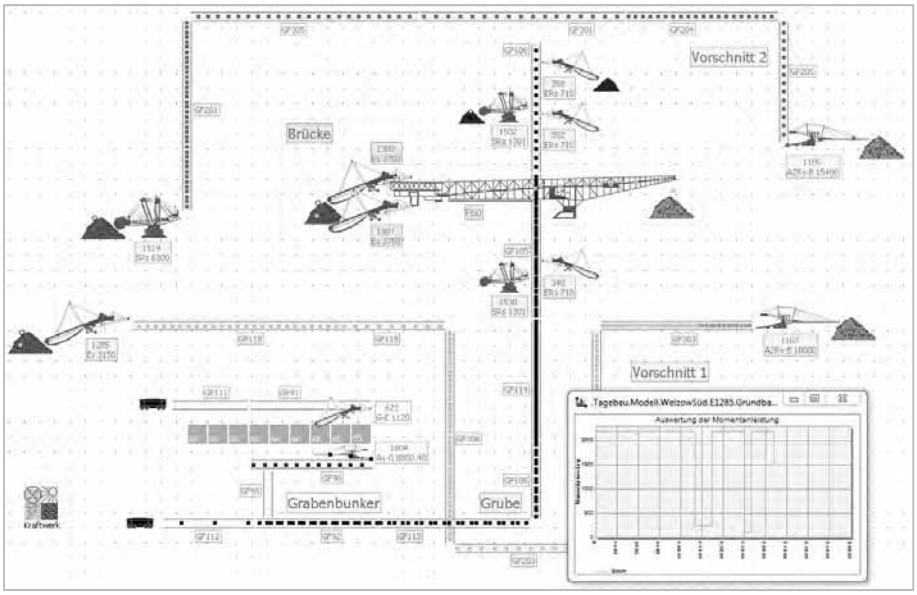



*Abbildung 2: Screenshot Simulationsmodell*

Für die Modellierung des geologischen Profils wurden Klassen beweglicher Elemente (BE-Klassen) für die verschiedenen Boden- und Kohlearten abgeleitet und in geeigneter Weise attribuiert (z. B. Volumen, Leistungsbedarf, Strompreis). Anhand der zuvor definierten Volumenblöcke (siehe Abschnitt 3.1) wurden Abfolgen für die zu fördernden Erd- und Kohleschichten, d. h. für die in die Simulation eintretenden BE-Instanzen (BEs), festgelegt. Die BEs werden in jedem Betriebsbereich durch eine Quelle in der entsprechenden Reihenfolge generiert. Die Grube stellt dabei eine Ausnahme mit zwei Quellen dar, da hier prozess- und qualitätsseitig Unterschiede zwischen Hoch- und Tiefschnitt bestehen.

Alle Bagger wurden auf Grundlage einer Objektklasse modelliert und weisen somit eine identische interne Struktur sowie Funktionsweise auf. Sie unterscheiden sich einzig hinsichtlich ihrer modellierten Energieflüsse. Während der Simulation

nehmen sie die 5-Minuten-Volumenpakete aus den Quellen auf und generieren aus diesen neue, kleinere Volumenpakete. Dieser Schritt wird notwendig, da eine kleingliedrige Modellierung der Lagerstätten mit einem nichtzielführenden Aufwand verbunden wäre. Neben dieser Umwandlungsfunktion protokollieren die Bagger-Modellelemente die austretenden BEs, um relevante Leistungsindikatoren (z. B. stündliche Förderleistung) zu berechnen.

Die Abraumförderbrücke, die Bandanlagen sowie die Absetzer sind in ihrer Funktionalität identisch als Förderstrecken modelliert. Um insbesondere die volumenabhängigen Verbräuche der Bandanlagen realistisch abbilden zu können, wird das jeweils auf der Förderstrecke befindliche Gesamtvolumen während der Simulation erfasst. Werden zuvor definierte Grenzen über- oder unterschritten, wird der Simulationsbaustein eniBRIC durch eine implementierte Logik in den entsprechenden energetischen Lastfall, d. h. Betriebszustand, geschaltet.

Zu Beginn der Simulation werden notwendige Eingangsdaten, wie die Bandparameter, geplante und ungeplante Stillstandzeiten sowie Betriebszustands-wechsel, zur Parametrierung des Modells eingelesen. Störungen, Wartungen und Betriebszustandswechsel werden durch Trigger ausgelöst und mittels zugehöriger Methoden gesteuert. Die Anfahrprozesse nach Stillstandszeiten erfolgen ebenfalls über die so implementierte Steuerungslogik.

Die zeitlichen Prozessabläufe in Grube und Grabenbunker werden über zusätzliche Methoden realisiert. Die Kohlenachfrage durch Kraftwerk und Veredlungsbetrieb, gehen zu Beginn als Eingangsinformationen in das Modell ein und repräsentieren die zu erfüllenden Vorgaben. Die Verlade-, Ab- und Aufhaldevorgänge im Bunker sowie die Kohleförderung im Hoch- bzw. Tiefschnitt der Grube werden auf Grundlage dieser Bedarfe gesteuert.

## 3.4 Integration von eniBRIC

Um die Energieflüsse und somit den Einsatz von Elektroenergie im Tagebau realistisch abbilden zu können, wurde der generische Simulationsbaustein eniBRIC (Schlegel et al. 2013) verwendet. Dieser wurde in jedem energetisch zu betrachtendem Element des Tagebaus instanziiert. Zu diesem Zweck wurden die zugrundeliegenden Objektklassen in der Klassenbibliothek um den Baustein erweitert und mit der jeweiligen Steuerungslogik der Materialflusselemente verknüpft.

In den vorgesehenen Parametertabellen für die Energieaufnahme bzw. -abgabe der einzelnen Instanzen wurden die zuvor identifizierten Betriebszustände (vgl. Abschnitt 3.2) eingetragen und die jeweiligen Ressourcenverbräuche zugeordnet. Für das entwickelte Simulationsmodell wurde festgelegt, dass Wechsel des Betriebszustands alleinig durch eine steuernde Instanz, d. h. durch eine Logik außerhalb von eniBRIC, realisiert werden. Weiterhin wurde in den Parametertabellen definiert, in welchen Betriebszustand das betrachtete Element einen Materialfluss zu realisieren vermag. So kann beispielsweise in den Zuständen „Fahren“ oder „Hoch-/Tiefschnittwechsel“ kein Material gefördert werden.

Als Quelle für elektrische Energie wurde zudem eine Instanz von eniBRIC geschaffen. Diese besteht alleinig aus dem Erweiterungsmodul und lässt sich im Weiteren nutzen, um etwa den Einfluss flexibler Energieverfügbarkeit, abgebildet durch verschiedene Betriebszustände, zu untersuchen. Überdies wurden die zu eniBRIC gehörigen Konfigurations- und Auswertebaustein im Modell instanziiert, um

notwendige Einstellungen vorzunehmen und Ad-hoc-Visualisierungen sowie -Auswertungen des Energieverbrauchs vornehmen zu können.

### 3.5 Validierung

Die Validierung des Simulationsmodells erfolgt auf der Grundlage von realen Mess- und Berichtsdaten eines kompletten Monats. Störungen, Wartungen und Betriebszustandswechsel werden mit Zeitpunkt und Dauer direkt aus den entsprechenden Berichten entnommen. Eine zufällige Verteilung für Dauer und Abstand zwischen ungeplanten Stillständen wird für die Validierung zunächst nicht angenommen. Stattdessen wird die Zeitreihe des Referenzmonats abgearbeitet. Die stochastische Betrachtung des Störverhaltens erfolgt dann in den Simulationsexperimenten. Die Abtragungsreihenfolgen für Abraum und Kohle werden anhand von Zeitreihen, in denen die zeitliche, mengen- und bodenspezifische Abfolge festgehalten ist, erstellt und in die entsprechenden Reihenfolgetabellen übertragen.

In einer abschließenden Auswertung werden die Berichtsdaten sowie die zeitlich aufgelösten elektrischen Verbräuche und die Förderleistung den geloggten Daten gegenübergestellt.

## 4 Einsatz des Simulationsmodells

Für die Gewinnung und Verstromung von Braunkohle verliert die Kennzahl Energieeffizienz, als das Verhältnis der erzielten Wirkung zur eingesetzten Energie zunehmend, an Aussagekraft. Bei einem Überangebot von Elektroenergie aus erneuerbaren Energiequellen kann es etwa systemeffizient sein, temporär geringere Energieeffizienz in einem Teilprozess zu tolerieren. Dadurch ließe sich im Substitutionsfall oder bei Lastverschiebung der Verbrauch ressourcen-intensiverer Energiequellen beispielsweise zu anderen Zeitpunkten mindern. Diese veränderte Perspektive auf die Effizienz des Energieeinsatzes im Produktionszusammenhang ist Ausgangspunkt der Untersuchungen zu den konkurrierenden Flexibilitäten.

Daher liegt neben der realistischen Abbildung des Tagebaus Welzow-Süd der Schwerpunkt des vorgestellten Simulationsprojektes auf der Identifizierung von Flexibilitäten, die zur Optimierung der Betriebsstrategie herangezogen und in der Simulation validiert werden können. Diese sind grundsätzlich in der Tagebauproduktion vorhanden, aber durch komplexe Zwangsbedingungen nur sehr begrenzt aktivierbar. Zudem unterliegen einzelne Betriebsbereiche, wie z. B. die Vorschnitte, (lärm-)emissionsrechtlichen Einschränkungen, die eine Produktion an den Wochenendtagen nicht erlauben. Eine Verschiebung oder Ausweitung von Betriebszeiten ist aufgrund dessen keine Option.

Mithilfe des entwickelten Simulationsmodells kann daher eruiert werden, wo sich beeinflussbare betriebliche Abläufe, Spiel- und Belastungsräume befinden und welche Effekte sich aus deren Ausnutzung ergeben. Der Beitrag dieser Potenziale zur Steigerung der systemischen Energieeffizienz sowie die Auswirkungen externer Veränderungen auf das Gesamtsystem werden zu diesem Zweck in „Was-wäre-wenn-Experimenten“ qualitativ und quantitativ untersucht. Die Variation der Parameter und Eingangsdaten erfolgt dabei anhand von Expertenwissen (bzgl. des Tagebaubetriebs) und Analysen der Simulationsergebnisse des validierten Modells.

## 5 Zusammenfassung und Ausblick

In diesem Beitrag wurde verdeutlicht, dass die Materialflusssimulation im Sinne eines „cross learning“ geeignet ist, um die Produktionsprozesse eines Braunkohletagebaus modellhaft abzubilden. Hierzu ist der branchenübergreifend einsetzbare Simulationsbaustein eniBRIC für die Modellierung von Energieflüssen zur Anwendung gekommen. Die gleichzeitige Betrachtung von Material- und Energieflüssen in einem Simulationswerkzeug ermöglicht die Durchführung von Untersuchungen zur Steigerung der systemischen Energieeffizienz.

Ausgehend vom hier beschriebenen Modell für den Tagebau Welzow-Süd wird in künftigen Arbeiten beabsichtigt geeignete Flexibilitäten mittels simulationsbasierter Untersuchungen zu identifizieren und deren Potenziale qualitativ und quantitativ abzuschätzen. Insbesondere die Auswirkungen dynamischer Rahmenbedingungen auf die Produktion, wie Stromangebot und -preis, Produktnachfrage sowie Witterung müssen in die Betrachtungen einbezogen werden, so dass neue optimierte Betriebsweisen im Hinblick auf Netz-, Strompreis- oder Versorgungssicherheit entwickelt werden können.

## Literatur

Basu, A.J.; Baafi, E.Y.: Discrete event simulation of mining systems: Current practice in Australia. International Journal of Surface Mining, Reclamation and Environment. 13:2 (1999), p. 79-84.

BMWi: Energiedaten: Gesamtausgabe. http://www.bmwi.de/BMWi/Redaktion/PDF/E/energiestatistiken-grafiken, 2014

Ji, Y.; Borrmann, A.; Wimmer, J.; Günthner, W.A.: Bidirectional Coupling of Macroscopic Optimization and Microscopic Simulation of Earthwork Processes. 27th Conference on Applications of IT in the AEC Industry, 2010.

Knafla, F.; Richter, M.: Einfach mal abschalten – Karosseriebau: Effizienter produzieren. IEE 58 (2013), Nr. 5, S. 66-68.

Lange, H.R; Woldt, T.; Köhler, U.: Von „just in time“ zu „just to energy“? Ein Praxisbericht zur Integration von Last- und Erzeugungsmanagement im industriellen Maßstab (Bergbau). Institutskolloquium am Institut für Industriebetriebslehre und Industrielle Produktion (IIP), Karlsruhe 2014.

Schlegel, A.; Stoldt, J.; Putz, M.; Erweiterte Integration energetischer Betrachtungen in der Materialflusssimulation. In: Dangelmaier. W. (Hrsg.) Simulation in Produktion und Logistik, S. 187-196; ISBN: 978-3-942647-35-9, 2013.

Stoll, R.D; Niemann-Delius, C.; Drebenstedt, C.; Müllensiefen, K. (Hrsg.): Der Braunkohletagebau – Bedeutung, Planung, betrieb, Technik, Umwelt. Berlin, Heidelberg: Springer 2009.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Ein Simulationsmodell zur Abbildung und Prognose der elektrischen Leistungsaufnahme in Seehafen-Containerterminals

## *A Discrete Event-driven Simulation Model to Analyse Power Consumption Processes in Container Terminals*

Nico Grundmeier, Norman Ihle, Axel Hahn, Universität Oldenburg, Oldenburg (Germany), nico.grundmeier@uni-oldenburg.de, norman.ihle@uni-oldenburg.de, axel.hahn@uni-oldenburg.de

Abstract: This article presents a simulation model (GreenPort) which combines logistic processes and electricity consumption in container terminals using electric transport vehicles and a battery exchange station. The model can be used to evaluate new logistic strategies in consideration of energy consumption and to provide a short-term energy demand load curve forecast. If the energy consumption can be forecasted precisely, it is possible to benefit from several use cases. In order to keep the container handlings of electric vehicles on a high level, it is necessary to develop intelligent strategies for battery exchange times. In this article it is shown how Greenport can be used to develop and test new strategies to resolve the disadvantage of frequent battery exchanges that these battery-powered vehicles have in comparison to their diesel-powered versions.

## 1 Einführung

In Containerterminals gibt es einen deutlichen Trend zur Elektrifizierung von schwerem Gerät, welches für Containerumschläge und deren Transport zuständig ist. Als Beispiel sind in der jüngeren Vergangenheit neben Diesel-betriebenen Transportfahrzeugen auch Batterie-elektrisch betriebene Fahrzeuge für den Einsatz im Containerterminal entwickelt worden. Weltweit war das HHLA Containerterminal Altenwerder (CTA) in Hamburg das erste Terminal, welches Prototypen eines automatischen Transportfahrzeugs (AGV) mit einem voll-elektrischen Antriebsstrang eingesetzt hat (Gottwald Port Technology GmbH et al. 2011). Mit der Einführung solcher Elektrofahrzeugflotten im Terminalbetrieb und der Liberalisierung des Energiemarkts eröffnen sich neue Möglichkeiten für das Energiemanagement eines Terminals (Wawer 2007; Runge et al. 2014). Dies kann z. B. durch die Nutzung eines Batteriewechselkonzepts für die Elektrofahrzeuge erreicht werden, bei welchem Wechselbatterien für die Antriebsenergie bezugspreisoptimiert

geladen oder Lastverschiebungspotenziale vermarktet werden. Um als Akteur am Strommarkt agieren zu können und somit von Energiekostensenkungen zu profitieren, ist es notwendig, Kenntnisse über den eigenen Energieverbrauch zu besitzen und die zukünftige Leistungsaufnahme über einen kurzfristigen Zeitraum zu prognostizieren.

In dieser Arbeit wird ein integriertes Simulationsmodell (GreenPort) vorgestellt, welches die kombinierte Betrachtung von Logistikprozessen und Energieverbrauchsprozessen zum Ziel hat. Dieses Modell ermöglicht die Bewertung und Validierung von neuen Betriebsstrategien unter Berücksichtigung der Leistungsaufnahme der beteiligten Umschlaggeräte. Durch die Simulation in der Zukunft liegender Logistikvorgänge kann das Modell dazu genutzt werden, eine Prognose der Leistungsaufnahme des Gesamtterminals anzufertigen und somit eine Interaktion mit dem Energiemarkt zu ermöglichen. Eine Beschreibung, wie eine solche Interaktion dabei konkret aussehen kann, erfolgt in (Ihle et al. 2014). Eine kurzfristige Prognose der Leistungsaufnahme über einen Zeitraum von 48 Stunden ermöglicht die Nutzung verschiedener Anwendungsfälle, welche sich dazu eignen, eine Energiekostensenkung zu erreichen. Beispiele hierfür sind Lastgangvoranmeldung beim Energieversorger, Bereitstellung von Ausgleichs- oder Regelleistung, Lastgangglättung und Lastverschiebung (Grundmeier et al. 2014).

## 2 Wissenschaftliche Abgrenzung und Darstellung von möglichen Anwendungen

Bisher wurden Planungsprobleme im Containerterminal zumeist aus logistischer Perspektive betrachtet und energietechnische Aspekte oftmals ausgeblendet (Xin et al. 2014). So zielen Optimierungsmöglichkeiten traditionell auf die Erhöhung des Containerumschlagvolumens und eine Reduzierung der Schiffsliegezeiten ab, ohne jedoch die Energieversorgung der Umschlagsgeräte mit einzubeziehen. Stahlbock und Voß (2008) geben einen Überblick über aktuelle Modelle und Verfahren zur Optimierung von Betriebsstrategien unterschiedlicher Planungsbereiche im Containerterminal. Gemeinsam haben die vorgestellten Modelle, dass Energieverbräuche der beteiligten Umschlagsgeräte nicht betrachtet werden, welche gerade bei der Nutzung von elektrisch angetriebenen Transportfahrzeugflotten von großer Bedeutung sind. Durch eine verkürzte Einsatzzeit bis zum nächsten Energiebedarf von Elektrofahrzeugen müssen Batterieladeoperationen wesentlich häufiger durchgeführt werden, als dieses bei diesel-elektrisch angetrieben Fahrzeugen für Tankvorgänge der Fall ist (Gottwald Port Technology GmbH et al. 2011). Um diesen Umstand zu kompensieren, sollten neue Betriebsstrategien im Bereich der Energieversorgung und -bereitstellung entwickelt werden, um eine nahtlose Integration von Elektrofahrzeugen in die Logistikprozesse zu ermöglichen. Erst in jüngster Vergangenheit wurden Arbeiten veröffentlicht, welche sich mit den elektrischen Leistungsaufnahmen in Containerterminals befassen. Eine Arbeit, welche sich ausführlich mit dem Energieverbrauch von Containerterminals beschäftigt, wurde von Tran (2012) veröffentlicht. Im Rahmen dieser Arbeit wurde die Leistungsaufnahme der einzelnen Teilverbraucher des Melbourner Containerterminals über einen Zeitraum von einem Jahr aufgezeichnet und analysiert. Als Ergebnis wird darauf basierend eine analytische Methode zur Bestimmung der zu erwartenden maximalen Leistungsaufnahme vorgestellt. Auch wurden in jüngster Zeit Simulationsmodelle entwickelt, welche die

Leistungsaufnahme von Containerkränen abbilden (Johanson 2010; Le 2012). Eine ganzheitliche Betrachtung der Leistungsaufnahme im gesamten Terminalbereich, insbesondere eine Prognose der Leistungsaufnahme, ist nach Kenntnis der Autoren noch nicht erfolgt. Neben Modellen zur Abbildung von Logistikprozessen in Containerterminals gibt es unterschiedliche, etablierte Methoden zur Energiebedarfsprognose. Hierzu zählen etwa Trendverfahren, Hochrechnungsverfahren, Regressionsverfahren oder die Verwendung von Neuronalen Netzen (Feinberg und Genethliou 2005). Auch wenn mit diesen Methoden teilweise sehr gute Ergebnisse im Bereich der Lastgangprognose erzielt werden können, haben sie jedoch alle einen gemeinsamen Nachteil: Einflüsse unterschiedlicher Betriebsstrategien auf den Energieverbrauch lassen sich durch diese Methoden nicht bewerten und auch lassen sich keine Erkenntnisse über die Energieverbrauchsprozesse der Einzelkomponenten gewinnen. Aus diesem Grund wird eine Simulation als geeignete Methode gewählt. Hiermit wird ein universelles Werkzeug zur Verfügung gestellt, um neue Betriebsstrategien zu bewerten und gleichzeitig die Auswirkungen auf die Leistungsaufnahme im Terminal verfolgen und prognostizieren zu können

# 3 Greenport

Bei einem Seehafen-Containerterminal handelt es sich um eine Anlage, in welcher Container zwischen Containerschiffen und Landverkehrsmitteln umgeschlagen werden. Hierzu werden an der Kaimauer liegende Schiffe von Containerbrücken be- und entladen. Die Container werden von Transportfahrzeugen in terminaleigene Lagerbereiche transportiert, von wo aus der Weitertransport per LKW oder Bahn erfolgt. Hierbei finden hoch dynamische und stochastische Prozesse statt, welche eine vorausschauende Planung der Logistikabläufe erschweren. Im Regelfall werden für die einzelnen Planungsprobleme, wie etwa der Transportauftragszuordnung, Online-Planungsstrategien eingesetzt. Diese arbeiten mit einem Planungshorizont von nur wenigen Minuten (Lehmann 2006). Die Nutzung von Online-Strategien wird auch durch die Tatsache gefördert, dass Containerattribute wie etwa Positionen und Gewicht oftmals erst mit der Ankunft eines Schiffes zur Verfügung stehen. Aus diesem Grund ist die Hauptanforderung an ein Modell zur Prognose der Leistungsaufnahme die Logistikprozesse abzubilden, ohne dabei tatsächliche Containerattribute wie das Gewicht oder die exakte Position des Containers im Schiff zu berücksichtigen. Die Grundidee des Ansatzes ist daher, durch Simulation der Logistikprozesse, basierend auf den geplanten Schiffsankünften, die Leistungsaufnahme zu prognostizieren.

## 3.1 Terminallayout

Das Layout eines Containerterminals wird durch einen gerichteten Multigraphen mit Kantengewichten abgebildet. Abbildung 1 zeigt ein beispielhaftes Layout eines kleinen Terminals mit zwei Blocklagern auf der linken und einer Kaimauer auf der rechten Seite sowie einem Wegenetzwerk, auf welchem sich Transportfahrzeuge bewegen können. Weitere Terminalkomponenten wie etwa Blocklager, Containerbrücken oder Batteriewechselstationen können so nicht direkt auf dem Layout platziert werden, sondern müssen in dieses integriert werden. Hierzu müssen für alle auf dem Terminalgelände befindlichen Objekte Übergangsknoten definiert werden, welche die Anbindung an das Wegenetzwerk der Transportfahrzeuge darstellen. Im aufgeführten Beispiel sind die Knoten $E_1$ und $E_2$ Übergabeknoten für das Blocklager

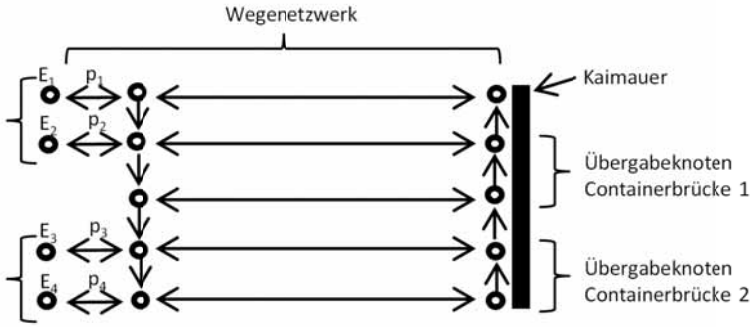



*Abbildung 1: Beispielhaftes schematisches Containerterminallayout*

1. Es ist zu beachten, dass sich Übergabeknoten von Containerbrücken dynamisch ändern, sobald diese neue Positionen an der Kaimauer einnehmen.

## 3.2 Verbraucher

Relevante Verbrauchsgruppen sind nach Grundmeier et al. (2014) die Beleuchtung des Terminals, Kühlcontainer, Containerbrücken und Portalkräne, Ladeinfrastruktur zur Energieversorgung der Transportfahrzeugflotte sowie der Verbrauch von IT-Anlagen, Bürogebäuden und Werkstätten (Grundlast).

Der Beleuchtung wird ein konstanter Energieverbrauchswert zugeordnet. Es wird dabei zwischen dem Verbrauch am Tage und dem Verbrauch in der Nacht unterschieden, der höher ist als am Tage. Zum Sonnenauf- bzw. Sonnenuntergang wird der Energieverbrauchswert innerhalb einer Dreiviertelstunde um drei Stufen gesenkt bzw. erhöht. Der obere und untere Leistungsaufnahmewert ist dabei vorgegeben. Die Zeiten eines Sonnenauf- bzw. Sonnenuntergangs werden analytisch bestimmt.

Die Leistungsaufnahme der Kühlcontainer hängt von vielen Faktoren ab, welche zum Durchführungszeitraum der Simulation nicht zur Verfügung stehen. Aus diesem Grund wird eine durchschnittliche Leistungsaufnahme angenommen. In der durchschnittlichen Leistungsaufnahme werden auch statistische Ein- und Abschaltwahrscheinlichkeiten der Kühlaggregate berücksichtigt. Mit der Einlagerung eines Kühlcontainers in den Lagerbereich des Terminals beginnt auch die damit verbundene Leistungsaufnahme. Sobald ein Kühlcontainer eingelagert wurde, wird von der Instanz des entsprechenden Lagers ein Einlagerungssignal an einen *ReeferManager* übermittelt. Bei der Auslagerung eines Kühlcontainers wird ein entsprechendes Auslagerungssignal übermittelt. Nach einem Ein- oder Auslagerungssignal und der Erhöhung oder Verminderung des Kühlcontainerbestandes wird die Gesamtleistungsaufnahme aller in den Lagern befindlicher Kühlcontainer bestimmt.

Containerbrücken und Portalkräne verfügen über eine von Containerumschlägen unabhängige Leistungsaufnahme, welche bezogen wird, sobald sich ein Kran in Betrieb befindet. Die Basisleistungsaufnahme wird durch Beleuchtung, Klimaanlage, Hydraulikpumpen und Flutlichter erzeugt (Johanson 2010). Für die Leistungsaufnahme eines Containerumschlags wird zusätzlich ein fester Energieverbrauchswert in kWh ($pC_{avr}$) angenommen, wobei für diesen zwischen 20 und 40 Fuß Container unterschieden wird. Die Umschlagszeit (t) wird durch eine diskrete Wahr-

scheinlichkeitsverteilung abgebildet, welche aus historischen Aufzeichnungen von Containerumschlagszeiten gewonnen wird. Dafür werden aufgezeichnete Umschlagszeiten pro Containerbewegung in Klassen ($K_i$) von je fünf Sekunden eingeordnet. Die Wahrscheinlichkeit, eine Umschlagszeit t(x) aus einer der Klassen zu ziehen, ist dann durch Formel 1 gegeben.

$$t(x) = \frac{|\, M = \{x \,|\, x \in K_i\}\,|}{Anz_{Gesamt}} \tag{1}$$

$Anz_{Gesamt}$ beschreibt die Gesamtanzahl der beobachteten Containerumschläge. Bei Containerschiffen wird zwischen Feedern und Seeschiffen unterschieden. Da die Umschlagszeiten eines Containers bei diesen beiden Schiffsklassensignifikant voneinander abweichen, sollten hier unterschiedliche Wahrscheinlichkeitsverteilungen genutzt werden. Die elektrische Leistungsaufnahme ($P_c$) für einen Containerumschlag ist abhängig von der Zeit und wird definiert durch:

$$P_c = \frac{pC_{avr}}{t} * 3600 \tag{2}$$

Zu beachten ist, dass an Containerbrücken Schichtwechsel der Brückenbesatzung durchgeführt werden. Zu diesen Zeiten werden von den Containerbrücken keine Umschläge durchgeführt und somit nur die Basisleistung aufgenommen.

Für ein AGV werden die grundlegenden physikalischen Zusammenhänge zwischen Geschwindigkeit (v) und benötigter Leistung ($P_v$) nach Guzzella und Sciaratta abgebildet (Guzzella und Sciaretta 2013), wobei $F_L$ den Luftwiderstand und $F_R$ den Rollwiderstand beschreibt.

$$P_v = F_L + F_R * v \tag{3}$$

Jedes AGV verfügt über ein Batteriewechselfach, welches eine Batterie aufnehmen kann. Batterieaufladungen werden durch Ladefächer in einer Batteriewechselstation in der Batteriewechselstation durchgeführt, wodurch Logistik- und Batterieladeprozesse zeitlich entkoppelt werden.

## 3.3 Prozesse

In der sogenannten Segelliste (SL) sind Datensätze aller zukünftigen und zurückliegender Schiffsankünfte aufgelistet. Ein Datensatz einer Schiffsankunft enthält dabei die in Abbildung 2 aufgeführten Informationen. Die Angaben über die an Bord befindlichen Import- und Exportcontainer beziehen sich auf die Container, welche im betrachteten Terminal be- und entladen werden sollen und nicht auf die Gesamtanzahl der an Bord befindlichen Container.

Abbildung 3 zeigt den generellen Simulationsablauf. Nach der Ankunft eines Schiffes wird diesem ein Liegeplatz (*Berth allocation*) zugewiesen und Containerbrücken (Quay Cranes, QC) zugeordnet (*QC assignment*). Die zugewiesenen Containerbrücken bewegen sich entlang der Kaimauer Richtung Zielschiff *(QC starts moving*). Im Anschluss wird die Ladereihenfolge der Container (*QC dispatch*) bestimmt. Eine Containerbrücke wartet (*QC waiting*) solange bis ein YardTruck be-

reitsteht. Nach der Umlagerung eines Containers wird ggf. die Reihenfolge der nächsten umzulagernden Container angepasst (*QC dispatch*) und die Containerbrückenzuweisung (*QC assignment*) erneut ausgeführt. Dies ist notwendig, um einer Containerbrücke, nachdem alle erreichbaren Container an ihrer Position umgeschlagen wurden, eine neue Aufgabe zuordnen zu können. Auch kann es in bestimmten Situationen sinnvoll sein, Containerbrücken von einem aktiven Auftrag abzuziehen. Zeitgleich mit dem Start der Containerbrückenbewegung wird die Jobliste erstellt (*Create QC job list*), welche Informationen über alle von einem Kran umzuschlagenden Container enthält.

| | |
|---|---|
| 1. $CS_{JSNR}$ = Schiffs ID | 6. $CExp_{40}$ = Anzahl 40f Exportcontainer |
| 2. $CS_{Name}$ = Schiffsname | 7. $Cimp_{20}$ = Anzahl 20f Importcontainer |
| 3. $Cs_{arr}$ = Ankunftszeit | 8. $CImp_{40}$ = Anzahl 40f Importcontainer |
| 4. $Cs_{dep}$ = Abfahrtszeit | 9. $CImp_{Ref}$ = Anzahl Import Kühlcontainer |
| 5. $CEXP_{20}$ = Anzahl 20f Exportcontainer | 10. $CExp_{Ref}$ = Anzahl Export Kühlcontainer |

*Abbildung 2: Schematischer Aufbau einer Segelliste*

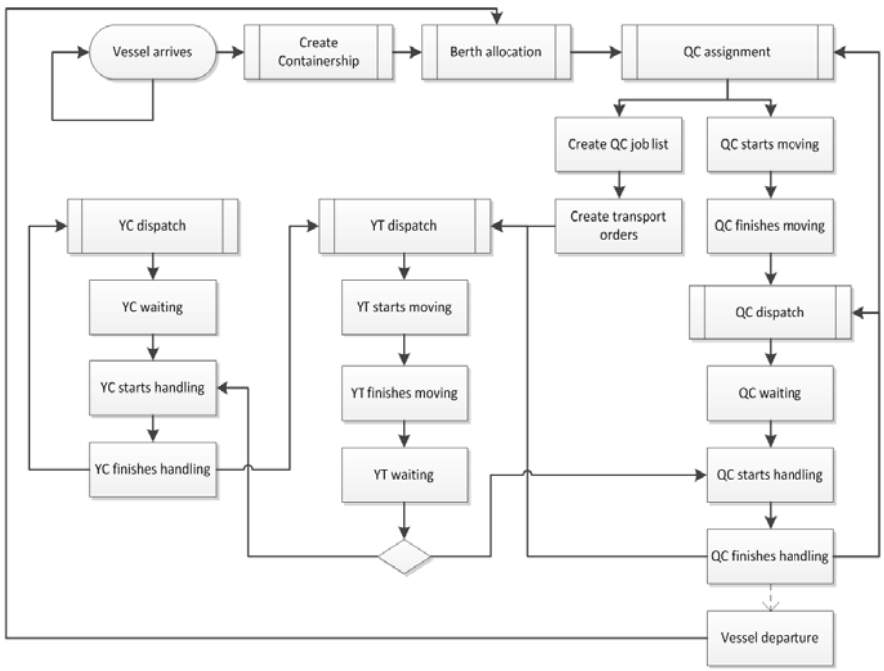



*Abbildung 3: Prozessdiagramm des allgemeinen Simulationsablaufs*

Mit diesen Informationen werden Transportaufträge erstellt (*Create Transport orders*) und einem Yard Truck zugewiesen. Die Zuweisung eines Yard Trucks und der Durchführungszeit zu einem Transportauftrag erfolgt in der Transportauftragszuordnung (*YC dispatch*). Wurde einem Yard Truck ein Transportauftrag zugewiesen begibt sich dieser Richtung Zielort (*YC starts moving*). Mit der Ankunft

am Zielort (*YC finishes moving*) wartet der YardTruck (*YT waiting*). Je nachdem ob es sich um einen Import- oder Exporttransportauftrag handelt, startet nun der entsprechende Kaikran oder die Containerbrücke den Umschlagvorgang eines Containers (*QC starts handling, YC starts handling*). Sind sämtliche Container eines Schiffes umgeschlagen, verlässt dieses das Terminal (*Vessel departure*).

Bei Änderung der Leistungsaufnahme eines Verbrauchers wird die neue Leistungsaufnahme einem *EnergyAnalyzer* Modul übermittelt. Dieses berechnet aus den übermittelten Änderungen der Leistungsaufnahmen die verrichtete elektrische Arbeit. Die aggregierte Leistungsaufnahme aller Verbraucher ergibt die Gesamtleistungsaufnahme des Terminals.

Abbildung 4 zeigt die Prozesskette, welche durch eine Schiffsankunft ausgelöst wird. Zum Zeitpunkt t wird eine Instanz *Containership* (CS) erzeugt, anschließend werden Import- und Exportcontainer erzeugt.

*Abbildung 4: Prozessablauf einer Schiffsankunft*

Ein Containerschiff verfügt über eine feste Anzahl an Bays (Stellplätzen/Reihen), in welchen Container verstaut werden. Entscheidend ist die Anzahl der Bays ($Nbr_{Bays}$), welche von der Länge des Schiffes $Bay_{Length}$ abhängt. Es gilt:

$$Nbr_{Bays} = \left\lfloor \frac{CS_{length}}{Bay_{Length}} \right\rfloor \geq 1 \tag{4}$$

Jede Bay verfügt über eine Startposition, welche die Position einer Bay relativ zur Schiffslänge angibt. Nach der Instanziierung eines Containerschiffs werden die zu entladenden (Import-)Container als Objekte generiert und zufällig auf die Containerbays verteilt, wobei eine gleichmäßige Aufteilung der Container auf die Bays erfolgt. Dies bedeutet, dass sich in jeder Bay ungefähr die gleiche Anzahl an Containern befindet, wobei durchaus in einer Bay mehr 20 Fuß Importcontainer als 40 Fuß Importcontainer oder Kühlcontainer vorkommen können. Die Gesamtzahl der an Bord befindlichen Anzahl von Importcontainer ($CIMP_{total}$) ist die Summe aus 20 Fuß, 40 Fuß und für den Import bestimmte Kühlcontainer. Im letzten Schritt erfolgt die Erzeugung der Exportcontainerobjekte, welche dem gleichen Schema wie die Erzeugung der Importcontainer folgt. Wie bei Importcontainern sind die Exportcontainer gleichmäßig auf Containerlager zu verteilen. Die Zuweisung der Container zu ihrem Lagerplatz erfolgt nach einem nutzerspezifischen System, welches auf das abzubildenden Containerterminal angepasst ist.

Der vollautomatische Wechsel einer Batterie nimmt in etwa 10 Minuten in Anspruch. Nach der Entnahme einer Batterie aus einem Fahrzeug wird diese in ein Batterieladefach eingelagert und gemäß der hinterlegten Ladestrategie (*Plan charging*) geladen. GreenPort nutzt ein Ladeverfahren mit einer konstanten Laderate C, wobei je nach Fragestellung beliebige Ladeverfahren implementiert werden können. Abbildung 5 stellt den Prozess eines Batteriewechsels dar.

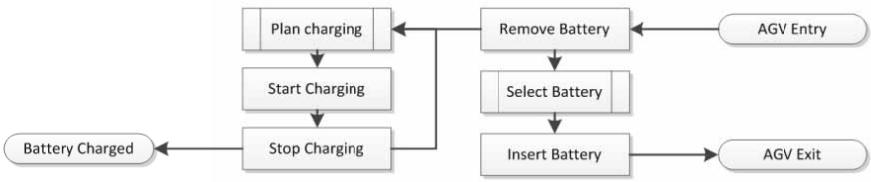



*Abbildung 5: Prozessablauf eines Batteriewechsels*

Zuletzt müssen Ankünfte von Containern per Bahn und LKW abgebildet werden. Hierzu müssen die zu erwartenden stündlichen Containeranlieferungen und Abholungen bekannt sein. Aus diesen Informationen werden Containeranlieferungen innerhalb einer Stunde zu einem zufälligen Zeitpunkt generiert. Zur Bestimmung der exakten Liefer/Abholzeit kommt eine Normalverteilung zum Einsatz.

## 3.4 Ergebnisse

Um die Validität des Modells zu bewerten, wurden Simulationsdurchläufe mit historischen Segellistendaten des CTA aus dem Jahr 2013 durchgeführt und die tatsächlichen Lastverläufe mit simulierten Ergebnissen verglichen. Dabei wurde die Güte der Prognose in Abhängigkeit der Abweichung zwischen realen und simulierten Schiffsabfertigungszeiten betrachtet. Tabelle 1 zeigt die Ergebnisse der Prognose anhand der Gütemaße „Mittlere prozentualer Fehler“ (MPE), „Mittlerer absoluter prozentualer Fehler“ (MAPE) und „Mittlerer quadratischer prozentualer Fehler“ (RMSPE) in Relation zu den Differenzen zwischen tatsächlicher und simulierter Abfertigungszeit der Containerschiffe (DAZ). Abbildung 6 zeigt exemplarisch die reale Leistungsaufnahme eines Tages und die dazugehörige Leistungsaufnahme der Simulation.

*Tabelle 1: Gütemaße der Lastgangprognose*

| Gütemaß/DAZ | < 30 min | < 60 min | > 60 min |
|---|---|---|---|
| MPE | 1,14 | 1,15 | 3,04 |
| MAPE | 7,18 | 8,18 | 9,70 |
| RMSPE | 10,32 | 11,04 | 13,6 |

In den beschriebenen Simulationsdurchläufen zur Bestimmung der Gesamtlastgangprognose wurden Energieverbräuche der Transportfahrzeugflotte noch nicht berücksichtigt, da diese in den historischen Energieverbrauchsdaten nicht erfasst wurden.

Um den Einfluss batteriebetriebener Transportfahrzeuge auf den Containerdurchsatz zu ermitteln, wurden Simulationsläufe mit Elektrofahrzeugen durchgeführt und der sich verändernde Containerdurchsatz mit dem verglichen, welcher durch eine reine Dieselfahrzeugflotte in der gleichen Zeit bewerkstelligt wurde. Der Containerdurchsatz wurde unter einer Vollauslastung des Terminals über acht Tage bestimmt. Dabei kamen 66 Batterie-AGVs mit einem Fahrverbrauch von 47 Ampere und einem Basisverbrauch von 1,39 Ampere zum Einsatz. Eine Batteriewechselstation mit 37 Wechselbatterien mit je einer Ladekapazität von 200 Amperestunden wurde

*Abbildung 6: Realer Lastgang und Lastgangprognose eines Beispieltages*

im Terminallayout integriert. Die Laderate eines Batterieladefachs betrug 1/7 C. Dies bedeutet, dass die benötigte Zeit für eine Vollladung 7 Stunden beträgt. Die benötigte Zeit für einen Batteriewechsel wurde auf 10 Minuten festgelegt. Batteriefahrzeuge führen immer dann einen Batteriewechsel aus, wenn der Ladezustand unter 30 % des Maximalwerts fällt. Die Ergebnisse sind:

- Durchgeführte Containerumschläge mit Diesel-AGVs: 55142
- Containerumschläge B-AGVs: 48752

Der verringerte Containerdurchsatz beim Einsatz von B-AGVs ist zum einen auf die verkürzte Einsatzzeit der Fahrzeuge zurückzuführen, zum anderen ist jedoch aufgefallen, dass durch die ungeplanten Batteriewechselaufträge es immer wieder zu Stauungen an der Batteriewechselstation kommt. So warten im Extremfall 19 Fahrzeuge gleichzeitig auf einen Batteriewechsel und stehen nicht für Containertransporte zur Verfügung. Die durchschnittliche Wartezeit auf einen Batteriewechsel betrug 48 Minuten, wobei es Stunden gab, in welchen überhaupt keine Batteriewechsel vorgenommen wurden. Es wurde eine weitere Batteriewechselstrategie (gesteuerte Batteriewechsel) implementiert und die Simulation wiederholt. Hierbei wird unabhängig vom tatsächlichen Ladestand alle 20 Minuten ein Batteriewechsel bei dem Fahrzeug mit dem aktuell geringsten Ladezustand durchgeführt. Unabhängig davon wurde ein Batteriewechsel auch beim Unterschreiten der 30% Grenze durchgeführt. Tabelle 2 zeigt die minimale, maximale und durchschnittliche Wartezeit für Batteriewechsel, wobei die Zeiten auch den eigentlichen Batteriewechsel umfassen.

*Tabelle 2: Vergleich der minimalen, maximalen und durchschnittlichen Wartezeit für einen Batteriewechsel.*

| Wartezeiten Ungesteuertes Laden | | | Wartezeiten Gesteuerte Batteriewechsel | | |
|---|---|---|---|---|---|
| Min | Max | Durchschnitt | Min | Max | Durchschnitt |
| 11 min | 190 min | 48 min | 11 min | 34 min | 22 min |

Durch das geplante Wechseln von Batterien sank die durchschnittliche Wartezeit eines AGVs für einen Batteriewechsel, wodurch die Containerumschläge auf 52.550 durchgeführte Umlagerungen gesteigert werden konnte.

## 4 Zusammenfassung

Durch die Einführung der elektrischen Leistungsaufnahme an Containerumschlagsgeräten ist es möglich, Wechselwirkungen zwischen Logistikprozessen und der

Energieversorgung der Transportfahrzeugflotte zu identifizieren. Auf dieser Grundlage können Strategien entwickelt werden, den Containerdurchfluss unter Berücksichtigung der Energieversorgung zu erhöhen. Durch Simulation der zukünftigen Logistikprozesse ist es zudem möglich, den Energiebedarf zu prognostizieren.

## Literatur

Feinberg, E. A., Genethiliou, D.: Load forecasting. In: Chow, Joe H., Wu, Felix F., Momoh, James A. (eds.): Applied Mathematics for Restructured Electric Power Systems. New York: Springer Science+Business Media 2005, S. 269-285.

Gottwald Port Technology GmbH, HHLA Container-Terminal Altenwerder GmbH, Institut für Energie- und Umweltforschung Heidelberg, RWTH Aachen – Institut für Kraftfahrzeuge: Forschung und Entwicklung batteriebetriebener Schwerlastfahrzeuge (AGV) und deren Erprobung in einem Feldversuch im Container-Terminal Altenwerder in Hamburg. Düsseldorf 2011.

Grundmeier, N., Hahn, A., Ihle, N., Runge, S. Meyer-Barlag, C.: A simulation based approach to forecast a demand load curve for a container terminal using battery powered vehicles. In: International Joint Conference on Neural Networks (IJCNN), Bejing 2014, S. 1711-1718.

Guzella, L.; Sciarretta, A.: Vehicle propulsion systems: Introduction to modeling and optimization. Berlin: Springer 2013.

Ihle, N., Runge, S., Grundmeier, N., Meyer-Barlag, C., Appelrath, H. J.: An IT-architecture to support energy efficiency and the usage of flexible loads at a container terminal. In: Proceedings of the 28th EnviroInfo Conference, Oldenburg 2014, S. 357-364.

Johanson, F.: Efficient use of energy in container cranes. In: Port Technology International 48 (2010), S. 51-54.

Le, K.: Analysing electric yard cranes with simulation. In: Port Technology International 53 (2012), S. 60-61.

Lehmann, M.: Einsatzplanung von Fahrerlosen Transportsystemen in Seehafen-Containerterminals. Dissertation, Technische Universität Berlin 2006.

Runge, S., Ihle, N., Meyer-Barlag, C., Grundmeier, N., Hahn, A., Appelrath, H.-J.: Herausforderungen für Demand-Side-Integration mit Elektrofahrzeugen in geschlossenen Logistiksystemen („Battery-to-Market“). VDE-Kongress, Frankfurt 2014.

Stahlbock, R., Voß, S.: Operations research at container terminals: A literature update. OR Spectrum 30 (2008) 1, S. 1-52.

Tran, T. K.: Study of electrical usage and demand at the container terminal. Dissertation, Deakin University Melbourne 2012.

Wawer, T.: Förderung erneuerbarer Energien im liberalisierten deutschen Strommarkt. Dissertation, Universität Münster 2007.

Xin, J.,Negenborn, R., Lodewijks, G.: Energy-aware control for automated container terminals using integrated flow shop scheduling and optimal control. Transportation Research Part C: Emerging Technologies 44 (2014), S. 214-230.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Allokation von Emissionswerten auf Behälterebene in multimodalen Umschlagsanlagen mittels Simulation

## *Allocation of Greenhouse Gas Emissions for Containers in Multimodal Transshipment Facilities Using Simulation*

Jan Kaffka, Uwe Clausen, Zoran Miodrag, Institut für Transportlogistik, TU Dortmund, Dortmund (Germany), kaffka@itl.tu-dortmund.de

Holger Pitsch, INCONTROL Simulation Solutions, Wiesbaden (Germany), holger.pitsch@incontrolsim.com

Abstract: The assessment of emissions caused by logistics operations in general and their allocation to individual customers is a major challenge for logistics service providers. Presently, numerous standards and guidelines exist (e.g. ISO 14064-1, ISO 14065, DIN EN 14040, DIN EN 14044, DIN EN 16258, PAS 2050) for the calculation of Greenhouse Gas (GHG) emissions caused by logistics processes. To support container terminal operators to determine their emissions, we use simulation to allocate overall emissions in a container terminal to a single container handling. To approach this goal, at first this paper describes the measurement of energy consumptions from the handling equipment. After that, the paper shows the simulation approach to allocate the emissions.

## 1 Einführung

Der Klimawandel und die damit einhergehenden negativen Entwicklungen, wie steigende Temperaturen, Extremwetterereignisse und weitere witterungsbedingte Einflüsse, sind auf den Ausstoß von Treibhausgasen (THG) und darunter insbesondere Kohlenstoffdioxid ($CO_2$) zurückzuführen. Um diesen Auswirkungen entgegenzuwirken und die THG-Emissionen zu reduzieren, werden verstärkt sowohl nationale als auch internationale Maßnahmen getroffen.

Der Verkehrssektor zeigte sich bisher jedoch nicht in der Lage, Maßnahmen umzusetzen, die eine ähnliche Emissions-Reduzierung erwirken konnten, wie z. B das verarbeitende Gewerbe oder die Energiewirtschaft. Der $CO_2$-Ausstoß des Güterverkehrs ist im Zeitraum von 1995 bis 2010 sogar um 13 % gestiegen (Thomas 2012). Da der Verkehr mit einem Anteil von 20 % an den emittierten THG ein gewichtiger Sektor ist, ist die Umsetzung erfolgreicher Maßnahmen in diesem Bereich erforderlich.

Ein Grund für die im Verkehr schwer zu reduzierenden THG-Emissionen ist der anhaltende Trend der steigenden Transportleistung sowie des steigenden Transportaufkommens. Die ergriffenen Maßnahmen zur Reduzierung wurden so überkompensiert (Clausen 2008).

Eine Möglichkeit mit Potentialen zur Reduktion der THG im Güterverkehr ist die Verkehrsverlagerung von der Straße auf andere Verkehrsträger. Da jedoch kein anderer Verkehrsträger auch nur eine annähernd gleich große Netzabdeckung bietet wie die Straße, ist eine Verknüpfung der verschiedenen Verkehrsträger hierfür in der Regel erforderlich. Eine solche Verknüpfung innerhalb eines Transportes bietet der Kombinierte Verkehr (KV), bei dem das Frachtgut von der Quelle zur Senke in einer Ladeeinheit durchgehend mit mehreren Verkehrsträgern, die als Transportkette hintereinander geschaltet sind, befördert wird. (Arnold et al. 2008)

Um Maßnahmen entwickeln zu können, die an den emissions-intensivsten Stellen angreifen können, ist zunächst die Möglichkeit zur möglichst exakten Bilanzierung und Allokation von THG-Emissionen zu schaffen. Diese besteht zum einen für Transporte, die mit einem einzelnen Verkehrsträger durchgeführt werden. Zum anderen können die THG-Emissionen für eine Transportkette ermittelt werden, indem sie für jeden Transportabschnitt allokiert werden. Die Emissionen der dazugehörigen Umschlagprozesse können bisher nur anhand des Gesamtenergieverbrauchs der logistischen Anlage bestimmt werden. Die Bestimmung der THG-Emissionen anhand des Gesamtenergieverbrauchs für den Umschlag entspricht jedoch einer sehr ungenauen Methode und ist somit für die Ausweisung auf Behälterebene, in Abhängigkeit von verschiedenen Einflussfaktoren, nicht ausreichend. In multimodalen Umschlagsanlagen existieren hingegen eine große Anzahl von Einflussfaktoren auf den Umschlag (Lagerfüllgrad, Anzahl Umlagervorgänge je Auftrag, Priorisierung von Verkehrsträgern, Lastfahrt, Leerfahrt usw.), die es unabdingbar machen, eine feinere Unterscheidung und somit eine exaktere Allokation zu ermöglichen. Hierzu wird in diesem Beitrag die Methode der Simulation genutzt, um eine exaktere Allokation von stationär ermittelten Verbrauchswerten auf die Ladeeinheit je nach Einflussfaktor durchzuführen und somit Referenzwerte für unterschiedliche Terminalarten und Einflussfaktoren zu liefern.

Der Beitrag ist wie folgt aufgebaut: Nach der Einleitung folgt im Stand der Forschung eine Abgrenzung des Themas gegenüber den Arbeiten anderer Autoren. Daraufhin werden in Kapitel 3 die Ermittlung der Messwerte und deren Integration in die Simulationsumgebung beschrieben, bevor in Kapitel 4 die Simulationsstudie zur Allokation der Verbrauchswerte auf Behälterebene erfolgt. Der Beitrag schließt mit einem Fazit.

# 2 Stand der Forschung

## 2.1 $CO_2$ Bilanzierung

Die Bilanzierung der THG-Emissionen (d.h. $CO_2$, $CH_4$, $N_2O$, HFC‘s, PFC‘s, $SF_6$) wird CARBON FOOTPRINT genannt und stellt einen Ausschnitt aus einer Ökobilanz dar. Die ÖKOBILANZ wiederum ist ein Instrument, das Umweltbelastungen unterschiedlicher Produkte (Waren und Dienstleistungen) berechnet und miteinander vergleichbar macht. Die Normen DIN EN ISO 14040 und DIN EN ISO 14044

beinhalten die Prinzipien und den Untersuchungsrahmen für Ökobilanzen sowie die methodischen Anforderungen zur Durchführung einer Ökobilanzierungsstudie. Die Ökobilanz bezieht sich auf die Umweltaspekte und potenziellen Umweltwirkungen, die im Verlauf des Lebensweges eines Produktes ausgehend von der Rohstoffgewinnung über Produktion, Anwendung, Abfallbehandlung, Recycling bis zur endgültigen Beseitigung (d.h. »cradle-to-grave«) auftreten. Bei einer Ökobilanz werden je nach angewendeter Methode zur Wirkungsabschätzung bis zu 14 Wirkungskategorien betrachtet, zu denen auch das THG-Potential gehört. Nachteilig bei der Erstellung einer Ökobilanz ist, dass selbst einfache Systeme einen erheblichen Datenbedarf und damit einen hohen Erfassungsaufwand aufweisen. Um diesen Erfassungsaufwand möglichst gering zu halten, müssen die Systeme sehr stark vereinfacht werden, wodurch wiederum eine realitätsgetreue Abbildung und Bewertung nicht möglich ist.

Ein CARBON FOOTPRINT in der Logistik kann sowohl für ein Produkt (Ware und Dienstleistung), ein Unternehmen, einen Standort, eine Organisation oder eine Transportdienstleistung ermittelt werden. Derzeit existieren zahlreiche Methoden, Normen, Leitfäden und Studien (Kranke et al. 2011; Schmied und Knörr 2011; Mc Kinnon and Piecyk 2010), die einen branchenspezifischen Rahmen für die Bewertung von Logistikprozessen und -systemen ermöglichen. Eine Norm über die Berechnung und die Berichterstattung von Energieverbrauch und THG-Emissionen von Transportdienstleistungen (Güter- und Personenverkehr) wurde Ende 2008 vom europäischen Normenausschuss CEN TC 320/WG 10 initiiert. Die 2013 publizierte Norm DIN EN 16258 enthält eine standardisierte Vorgehensweise zur Ermittlung der THG-Emissionen von Transportdienstleistungen aller gängigen Verkehrsträger (Straßen-, Schienen-, Schiffs- und Luftverkehr) in der gesamten Versorgungskette (vom Verlader über Spediteure und Frachtführer bis hin zum Subdienstleister) und entsprechenden Vorgaben für die Ergebnisdokumentation. Darüber hinaus werden zur Ermittlung der Datengrundlagen Empfehlungen gegeben. Letztendlich kann aber der Anwender frei entscheiden, welche Auswahl (d. h. individuelle Messwerte, fahrzeug-/routenspezifische Durchschnitts- oder Flottenwerte des Transportdienstleisters, Vorgabewerte aus Datenbanken) getroffen wird, auch wenn die Ergebnisse auf unterschiedlichem Detaillierungsniveau basieren.

Nicht in der Norm enthalten sind jedoch administrative bzw. unterstützende Tätigkeiten (z. B. Produktionsplanungs- und Steuerungssysteme, Wartung, Entsorgung) und stationäre Prozesse (z. B. innerbetrieblicher Umschlag), so dass wesentliche Bestandteile in den Bilanzergebnissen nicht enthalten sind.

In der Regel entstehen $CO_2$-Emissionen stationärer Logistikprozesse in multimodalen Logistikknoten z. B. durch

- Stromverbrauch der Umschlageinrichtungen, Terminals, Lager und Büros
- Wärmeenergieverbrauch für Terminals, Lager und Büros
- Weitere verwendete Energieträger (z. B. Gas, Diesel) für zusätzliches Equipment, wie beispielsweise Reachstacker, andere Terminalfahrzeuge und Gabelstapler.

Da stationäre Prozesse bei der aktuell vorliegenden Norm DIN EN 16258 explizit ausgeschlossen wurden, werden nach dem Greenhouse Gas Protocol die direkten THG-Emissionen, die beispielsweise bei der Verbrennung von Energieträgern sowie durch den Bezug von Strom und Wärme entstehen, ermittelt. Letztendlich wird fol-

gende Formel zur Ermittlung der THG-Emissionen sämtlicher stationärer Prozesse herangezogen: THG-Emissionen = Energieverbrauch x Emissionsfaktor. Aufgrund fehlender valider Energieverbrauchsdaten wesentlicher Elemente einer Umschlagsanlage des kombinierten Verkehrs muss auf anerkannte Quellen und Abschätzungen von Energieverbrauchsdaten zurückgegriffen werden. Hierzu können Normen, Richtlinien, Literatur-, Hersteller- sowie Unternehmensangaben herangezogen werden. Allerdings sind nicht für alle Elemente einer multimodalen Umschlagsanlage Energieverbrauchsdaten vorhanden, und für vorhandene Daten werden relevante Einflussfaktoren, wie beispielsweise Geländebeschaffenheit oder Terminallayout, nicht berücksichtigt. Daher wird in diesem Beitrag aufgezeigt, wie Energieverbrauchswerte mittels punktueller Verbrauchsmessungen erhoben werden und mittels energetischer Simulation auf Behältereben allokiert werden.

## 2.2 Energetische Simulation im kombinierten Verkehr

Zur Untersuchung und Bewertung des Verhaltens im Zeitablauf komplexer multimodaler Umschlagknoten und im Containerhandling wird vielfach auf das Hilfsmittel der Simulation zurückgegriffen (z. B. Lampe 2006; Lee et al. 2006; Duinkerken et al. 2006). Simulationslösungen zur ganzheitlichen Abbildung von Anlagen im kombinierten Verkehr wurde von den Autoren in vorhergegangen Arbeiten entwickelt. Diese Simulationslösungen basieren auf dem Materialflusssimulator Enterprise Dynamics und beinhalten sowohl eine Simulationslösung zur Feinplanung von Terminalsystemen als auch eine Low-Level-of-Detail-Lösung zur Grobplanung von multimodalen Umschlagsanlagen. Die Lösungen beinhalten neben Bausteinen zur Abbildung von Containerterminals mathematische Heuristiken zur optimierten Steuerung der Umschlagmittel, Ladestellenzuordnung und Yardmanagement (Kaffka et al. 2010; Kaffka und Mest 2010; Clausen et al. 2012; Clausen und Kaffka 2012; Kaffka et al. 2014). Eine simulative Betrachtung des Carbon Footprints von Supply Chains und großen Logistik Netzwerken erfolgte durch Rabe und Deininger (2012) und Rabe et al. (2014). Wie diese Arbeiten zeigen, wird in derzeitigen Forschungsarbeiten und Lösungsansätzen zur energetischen Simulation der Fokus auf die Betrachtung des Transportes gelegt. Eine Betrachtung der Umschlagsanlagen erfolgt hierbei entweder gar nicht oder auf einer stark abstrahierten Ebene. Da jedoch nach Kranke et al. (2011) ca. 25 % der $CO_2$-Emissionen eines Transportes auf stationäre Prozesse entfallen, ist eine derartige Betrachtung von hoher Bedeutung. Derzeitig existierende Simulationsumgebungen bieten jedoch keine Möglichkeiten, THG-Emissionen in Umschlagsanlagen zu ermitteln. Vereinzelte Simulationsumgebungen ermöglichen zwar die Angabe eines durchschnittlichen Energieverbrauches der Umschlagmittel, wie z. B. bei elektronischen Flurförderzeugen. Dies dient jedoch lediglich zur Berechnung der Reichweite des Umschlagmittels. Auf diese Weise werden Ladezyklen mit in den Materialfluss integriert, so dass die Planung in dem nötigen Detaillierungsgrad erfolgen kann. Hierbei wird jedoch nur mit Durchschnittswerten gerechnet, eine Unterscheidung von Umgebungseinflüssen, wie z. B. Steigungen und Gefälle oder Last- und Leerfahrt erfolgt nicht. Des Weiteren wird dies nur für batteriebetriebe Umschlaggeräte einer Umschlagsanlage eingesetzt; die Methode ermittelt keine THG-Emissionen. Dementsprechend existiert ein hoher Bedarf an der Integration von Messdaten unterschiedlicher Umgebungseinflüsse und THG-Ermittlungsmethoden in die bestehende Simulationsumgebung für multimodale Umschlagsanlagen.

## 3 Integration von Messwerten in die Simulationsumgebung

Für Containerumschlagsmittel ist derzeit keine verlässliche Datenbasis für Energieverbrauchswerte vorhanden, so dass im Rahmen der vorgestellten Arbeit in einem ersten Schritt eine Datenbasis erzeugt werden musste. Zu diesem Zweck erfolgten Energieverbrauchsmessungen an unterschiedlichen Kranmodellen mit unterschiedlichem Alter, mit dem Ziel diese in das Kranmodul einer Container Terminal Simulation einzufügen. Zusätzlich zu den Energieverbrauchsmessungen erfolgte eine detaillierte und auf die Energieverbrauchsmessung synchronisierte Prozessaufnahme, um die gemessenen Energieverbräuche den einzelnen Teilprozessen eine Kranspiels zuordnen zu können. Ein Kranspiel unterteilt sich dabei in die in Abbildung 1 aufgezeigten Prozessschritte.

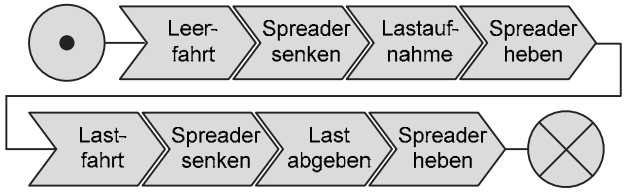



*Abbildung 1: Einzelprozesse eine Kranspiels*

Die Prozesskette beginnt mit einer Leerfahrt zur nächsten umzuschlagenden Ladeeinheit. Nachdem der Spreader über der Ladeeinheit positioniert wurde, wird dieser gesenkt. Durch Schließen der Twist-Locks oder der Greifarme wird die Ladeeinheit verriegelt und aufgenommen. Anschließend wird die Last durch Anheben des Spreaders hochgezogen und zur Zielposition verfahren. Sobald die Endposition erreicht wurde, wird der Spreader gesenkt, die Ladeeinheit abgestellt und die Twist-Locks geöffnet. Anschließend hebt der Kran den Spreader wieder an und steht für einen weiteren Umschlag(-auftrag), beginnend mit der Leerfahrt, zur Verfügung. Die so aufgenommenen Wirkleistungen (Abb. 2) konnten sekundengenau und sehr feingliedrig einzelnen Prozessschritten zugeordnet und in die jeweiligen Ereignisse des Kranmoduls der Simulationsumgebung integriert werden.

Als Simulationsumgebung wird TerminalSim, welche von Kaffka et al. (2014) entwickelt wurde, eingesetzt. Die Simulationsumgebung basiert auf der Simulationssoftware Enterprise Dynamics 8 der Firma INONTROL Simulation Solutions. TerminalSim ist eine Low-Level-of-Detail Simulationslösung für Containerterminals. Sie besteht aus parametrisierbaren Bausteinen, mit denen alle Funktionsbereiche einer multimodalen Umschlagsanlage abgebildet werden können. Um eine schnelle Modellierung und Auswertung unterschiedlicher Simulationsszenarien zu gewährleisten, ist in TerminalSim eine standardisierte Datenbankanbindung integriert, über die Simulationsszenarien automatisch parametrisiert, durchgeführt und ausgewertet werden. Die Kranumschlagsbausteine von TerminalSim wurden um die Ergebnisse der Energieverbrauchsmessungen der unterschiedlichen Krane erweitert. Die Messergebnisse haben gezeigt, dass es Unterschiede in den Verbräuchen der Krane bezüglich ihres Alters gibt. Neue und moderne Krane mit oder ohne Energierückgewinnung haben niedrigere Verbräuche, als Krane in einem Alter zwischen

einem und zehn Jahren. Krane die älter als zehn Jahre alt sind, weisen einen nochmals schlechteren Energieverbrauch auf. Die Kranumschlagsbausteine lassen sich dementsprechend parametrisieren.

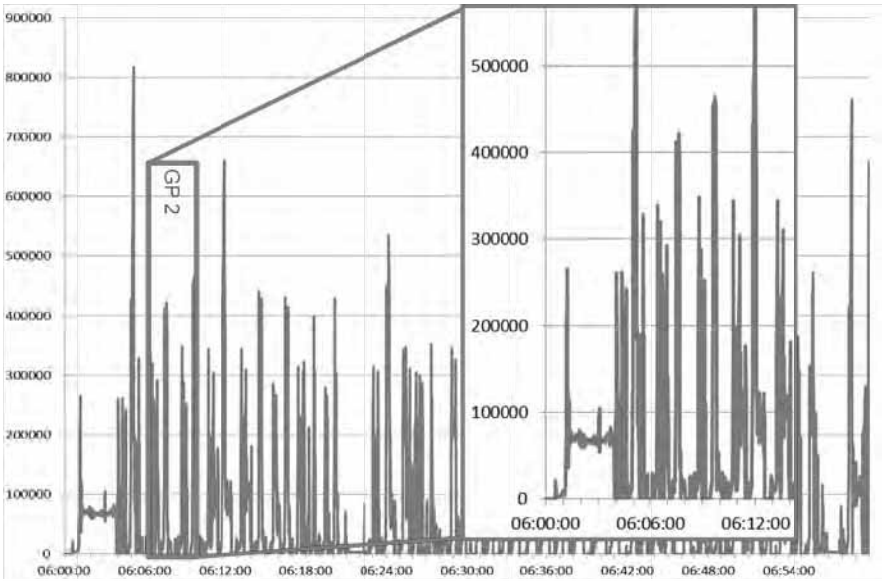



*Abbildung 2: Aufgenommene Wirkleistung P [W]*

## 4 Verbrauchsallokation mittels Simulation

Im Rahmen der vorgestellten Arbeit sollen Gesamtverbräuche eine multimodalen Umschlagsanlage in Abhängigkeit unterschiedlicher Einflussparameter mittels Simulation auf Containerebene allokiert werden. Hierzu werden Simulationsmodelle mit den entsprechenden Einflussparametern gebildet und Verbrauchswerte für einzelne Umschlagssituationen abgeleitet, da durch die Prozessaufnahme nur kurze Momentaufnahmen geschaffen wurden, welche keine allgemeinen Rückschlüsse bezüglich Containeraufkommen und –struktur erlauben. Zwar können die real gemessenen Energieverbräuche übernommen werden; der Gesamtverbrauch über ein Jahr mit zeitgleicher Allokation auf Containerebene kann auf diese Weise jedoch nicht bestimmt werden. Aus diesem Grund wird eine Simulationsstudie mit der Laufzeit von einem Jahr durchgeführt.

In diesem Beitrag werden die Ergebnisse der ersten betrachteten Umschlagsanlage aufgezeigt. Das betrachtete Terminal ist ein trimodale Umschlagsanlage in der Containerumschläge mittels Kran und Reach Stacker durchgeführt werden. Die Umschlagsfläche hat eine Länge von 420m und eine Breite von 135m und wird komplett von einem Portalkran überspannt. Unterhalb des Kranes befindet sich der Kaiumschlagsplatz für Schiffe, das Containerdepot, 4 Zugladegleise mit 350m Länge und die Ladeplätze für die LKW. Alle Containerumschläge von und auf das Schiff werden ausschließlich von dem Kran durchgeführt. Die Umschläge von und auf den Zug und den LKW werden hauptsächlich von den Reach Stackern

durchgeführt. Hier assistiert der Kran, wenn er freie Kapazitäten hat oder der Verladeschluss erreicht werden muss. Insgesamt schlägt das Terminal 150.000 TEU im Jahr um. Im Rahmen dieses Beitrages werden nur die Kranumschläge ausgewertet und betrachtet. Eine Einbeziehung von Reach-Stacker-Umschlägen erfolgt zu einem späteren Zeitpunkt des Projektes. Die Handlingszeiten innerhalb des Terminals für Containeraufnahme und –abgabe des Kranes, die Rangierzeiten der LKWs und Züge, die Bereitstellungszeit der Züge und die Anlege- und Ablegezeit der Schiffe wurden eine Woche lang von den Terminalmitarbeitern gemessen. Aus diesen Messdaten wurde mittels MiniTab jeweils eine stochastische Verteilungsfunktion abgeleitet, welche in das Simulationsmodell als Zeitverbrauch für diese Prozessschritte integriert wurde.

Basierend auf den realen Vergangenheitswerten des Containerterminals wurde für das Jahr 2014 ein Eingangsdatensatz erzeugt. Dies geschah, um saisonale Einflüsse, wie z. B. das Weihnachtsgeschäft, mit in die Auswertung mit einzubeziehen. Auf Basis dieser Eingangsdaten wurden 50 Simulationsläufe für das Jahr 2014 durchgeführt. Die Validierung des Simulationsmodells erfolgte durch Expertengespräche vor Ort.

Bei der Auswertung der Mess- und der Simulationswerte wurde schnell deutlich, dass die unterschiedlichen Containergewichte keinen Einfluss auf die Energieverbräuche haben, da die meiste Energie dafür gebraucht wird, um den Kran mit einer Masse von ca. 300t in Bewegung zu setzen. Hierbei hat ein Unterschied von max. 20t in den Containergewichten keinen großen Einfluss auf die Verbräuche, der sich bemerkbar macht. Dahingegen hat die Bewegungsentfernung einen sehr großen Einfluss auf die Verbräuche gezeigt. Hierzu wurden, basierend auf einer ABC-Analyse, folgende 4 Entfernungsklassen gebildet:

- Klasse E1 von 0m bis 20m
- Klasse E2 von 21m bis 60m
- Klasse E3 von 61m bis 100m
- Klasse E4 über 100m

Abbildung 3 zeigt das Ergebnis für die einzelnen Entfernungsklassen. Die Verbräuche steigen mit größer werdender Entfernung.

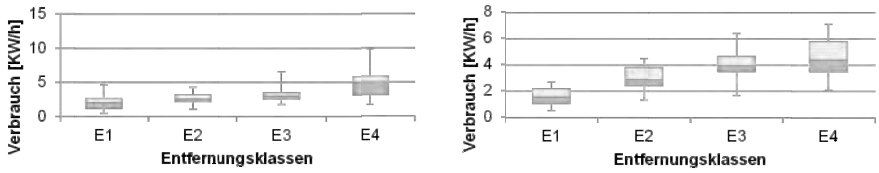



*Abbildung 3: Verbrauch pro Entfernungsklasse für Container mit 40 Fuß(links) und 20 Fuss (rechts)*

Auf Basis der Simulationsergebnisse mit Einbeziehung von Umwelteinflüssen und vor allem Umlagerungen eines Containers konnten die in Tabelle 1 allokierten Verbrauchswerte für die einzelnen Containergrößen ermittelt werden.

*Tabelle 1: Allokierte Verbrauchsdaten auf Containerebene*

| Container | Ø-Zeit | Ø-KWh | Ø-€/KWh | Ø-KG $CO_2$/KWh | Ø-Wegstrecke[m] |
|---|---|---|---|---|---|
| 20 Fuß Leer | 0:02:58 | 3,65 | 0,40 | 1,56 | 88,84 |
| 40 Fuß Leer | 0:02:47 | 3,09 | 0,34 | 1,32 | 66,68 |
| 20 Fuß Leicht | 0:03:00 | 4,02 | 0,44 | 1,72 | 124,14 |
| 40 Fuß Leicht | 0:03:30 | 3,48 | 0,38 | 1,49 | 84,60 |
| 20 Fuß Mittel | 0:02:50 | 3,14 | 0,35 | 1,34 | 180,35 |
| 40 Fuß Mittel | 0:05:51 | 3,63 | 0,40 | 1,55 | 72,06 |
| 20 Fuß Schwer | 0:03:00 | 4,19 | 0,46 | 1,79 | 95,23 |
| 40 Fuß Schwer | 0:02:42 | 3,82 | 0,42 | 1,63 | 123,99 |

Insgesamt kostet in diesem ersten Fall ein Containerumschlag im Durchschnitt 0,40 € und emittiert durchschnittlich 1,55 KG CO2, was dem Wert der Ökobilanz entspricht und zeigt, dass das Simulationsmodel valide ist. Mittels der neuen Methode ist es jedoch nun möglich, für einzelne Containertypen genauere Verbrauchswerte zuzuordnen, was so vorher nicht möglich war Dies kann von den Terminalbetreibern sowohl zu Marketingzwecken als auch zu betrieblichen Optimierungen oder zur differenzierteren Abrechnung genutzt werden.

## 5 Zusammenfassung und Ausblick

Im Rahmen des Forschungsvorhabens konnten erfolgreich Energieverbräuche und Einflussfaktoren auf diese in multimodalen Umschlagsanlagen ermittelt werden. Die Ergebnisse haben vor allem gezeigt, dass die Umschlagsentfernungen ausschlaggebend für hohe Verbräuche sind. Um Energie und somit Kosten zu sparen ist es wichtig, die Umschlagsentfernung soweit wie möglich zu reduzieren. Des Weiteren sind Umlagerungen zu vermeiden, da jegliche Zusatzbewegung weitere Verbräuche erzeugt. Die Ergebnisse sollen in weiteren Schritten mit weiteren Terminallayouts auf der einen Seite und mit Verbräuchen der mobilen Umschlaggeräte und nicht umschlagsspezifischen Anlagen (Verwaltung, Lichtmasten, usw.) auf der anderen Seite erweitert werden. Des Weiteren werden die Ergebnisse in ein auf Microsoft Excel basierendes $CO_2$-Bilanzierungstool integriert werden.

## Acknowledgement

Diese Ergebnisse wurden im Forschungsprojekt "$CO_2$-Bilanz – Multimodale Logistikknoten: Erweiterung des Methodenbaukastens $CO_2$-Bilanz zur exakteren Ermittlung und Zuordnung von Umwelteffekten in Multimodalen Logistikknoten" erzielt. Das Forschungsvorhaben wird gefördert durch das Bundesministerium für Wirtschaft und Energie mittels der AiF im Auftrag der BVL (Bundesvereinigung Logistik) unter dem Grant Agreement 17961 N/1.

## Literaturverzeichnis

Arnold, D.; Isermann, H.; Kuhn, A.; Tempelmeier, H.; Furmans, K.: Handbuch Logistik, 3. Aufl. Berlin, Heidelberg, u.a.: Springer 2008.

Clausen, U.: Maßnahmen im mobilen Bereich – Ergebnisse von StaMoLo und deren Umsetzung in der Praxis. Klimaschutz in der Logistik: Treibhausgase mindern – Kosten senken. Bonn. 2008, ohne Seitenzahlen

Clausen, U. und Kaffka, J.: Modeling of handling task sequencing to improve crane control strategies in container terminals. In: Laroque,C.; Himmelspach, J.; Pasupathy, R.; Rose, O.; Uhrmacher, A.M. (Hrsg.): Proceedings of the 2012 Winter Simulation Conference. Piscataway: IEEE 2012, S. 3104-3113.

Clausen, U.; Kaffka, J.; Meier, F.: CONTSIM - Container terminal management with simulation. In: Proceedings of EWGT2012 - 15th Meeting of the EURO Working Group on Transportation, September 2012, Paris; Procedia - Social and Behavioral Sciences 54 (2012); S. 332-340

DIN EN ISO 14040 – Norm DIN EN ISO 14040, 10/2006: Umweltmanagement - Ökobilanz - Grundsätze und Rahmenbedingungen.

DIN EN ISO 14044 – Norm DIN EN ISO 14044, 10/2006: Umweltmanagement - Ökobilanz - Anforderungen und Anleitungen.

DIN EN 16258 – Norm DIN EN 16258, 03/2013: Methode zur Berechnung und Deklaration des Energieverbrauchs und der Treibhausgasemissionen bei Transportdienstleistungen (Güter- und Personenverkehr).

Duinkerken, M. B.,Dekker, R., Kurstjens, S. T., Ottjes, J. A., Dellaert, N. P.: Comparing transportation systems for inter-terminal transport at the Maasvlakte container terminals. In: OR Spectrum 28 (2006), S. 469-493.

Kaffka, J.; Clausen, U.; Chmielewski, A.; Schütte, T.: Optimization and simulation of operating strategies for container terminals. In: General Proceedings of the 12th WCTR, Lissabon, 2010, ohne Seitenzahlen,

Kaffka, J.; Mest, L.: Optimierung von Binnenhafencontainerterminals mittels Simulation. In: Schifffahrt und Technik 29 (2010) 1, S. 85.

Kaffka, J; Clausen, U; Stein, S.: Revealing gaps in the material flow of inland port container terminals alongside the Danube with simulation. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley,S.; Miller, J.A. (Hrsg.): Proceedings of the 2014 Winter Simulation Conference. Piscataway: IEEE 2014, S. 1807-1818.

Kranke, A.; Schmied, M.; Schön, A.D.: CO2-Berechnung in der Logistik. Datenquellen, Formeln, Standards. München: Vogel 2011.

Lampe, H. (2006) Untersuchung von Dispositionsentscheidungen in Umschlagterminals des Kombinierten Verkehrs Schiene/Straße. Dortmund: Praxiswissen 2006.

Lee, B.K.; Jung, B.J.; Kim, K.H.; Park, S.O.; Seo, J.H.: A simulation study for designing a rail terminal in a container port. In: Perrone, L.F.; Wieland, F.P.; Liu, J.; Lawson, B.G.; Nicol, D.M.; Fujimoto, R.M. (Hrsg.): Proceedings of the 2006 Winter Simulation Conference. Piscataway: IEEE 2006, S. 1388-1397.

McKinnon, A.; Piecyk, M.: Measuring and managing $CO_2$-emissions of European chemical transport. Edited by Cefic -The European Chemical Industry Council. Heriot-Watt University, Logistics Research Centre. Edinburgh. Available online at www.cefic.org. 2010

Rabe, M.; Deininger, M.: State of art and research demands for simulation modeling of green supply chains. International Journal of Automation Technology 6 (2012) 3, S. 296-303.

Rabe, M.; Gutenschwager, K.; Fechteler, T.; Sari, U.: A data model for carbon footprint simulation in consumer goods supply chains. In: Pasupathy, R.; Kim, S.-H.; Tolk, A.; Hill, R.; Kuhl, M.E. (Hrsg.): Proceedings of the 2013 Winter Simulation Conference. Picataway: IEEE 2013, S. 2677-2688.

Schmied, M.; Knörr, W.: Berechnung von Treibhausgasemissionen in Spedition und Logistik. Begriffe, Methoden, Beispiele. Edited by Deutscher Speditions- und Logistikverband e.V. Öko-Institut e.V.; ifeu - Institut für Energie- und Umweltforschung Heidelberg GmbH. Köln, 2011

Thomas, J.: Umweltökonomische Analyse des Güterverkehrs 1995 bis 2010. In: Statistisches Bundesamt (Hrsg.) Wirtschaft und Statistik – Juni 2012, Wiesbaden, S. 503-514.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Potentiale in der Reduzierung des Gesamtenergieverbrauchs einer Werkstattfertigung in der Maschinenbelegungsplanung

## *Potential of Reducing the Total Energy Consumption by Scheduling a Jobshop Production System*

Maximilian Selmair, Frank Herrmann, OTH Regensburg, Regensburg (Germany),
maximilian.selmair@gmail.com, frank.herrmann@oth-regensburg.de

Thorsten Claus, Enrico Teich, TU Dresden, Dresden (Germany),
thorsten.claus@tu-dresden.de, enrico.teich@tu-dresden.de

Abstract: This research paper has evaluated the potential of reducing energy consumption by scheduling jobshop production systems with machines using standby modes in free times. The initial introduction of the planning issue is followed by a description of the approach to assess the available potential. Subsequently, the research procedure by means of simulation, a detailed discussion of the results and a perspective on future research is given. Moreover, the notion that scheduling influences the energy consumption in jobshop production systems is supported. The presented simulation research also documents that there is no direct correlation between energy consumption and total lead time, although this was presumed usually. Finally, this paper provides a forecast for a possible optimisation model as well as an exemplary model with an energy-optimised schedule plan.

## 1 Einführung

Produktionssysteme, die sich aus Werkzeugmaschinen zur Zerspanung und Umformung konstituieren, wurden in der Vergangenheit bei Untersuchungen zu Energieeinsparungen in der Industrie nicht berücksichtigt (Neugebauer 2008). Trotz des immensen Energiebedarfs solcher Systeme, ist deren Energieverbrauch erst seit einigen Jahren zum Augenmerk von sowohl Maschinennutzern als auch -herstellern geworden (ebenda). Grund hierfür sind sicherlich die steigenden Energiepreise der vergangenen Jahre (Wagner 2012), aber auch ein gestiegenes Bewusstsein vieler Unternehmen für das Thema Nachhaltigkeit und der damit verbundenen Intention der Ressourcenschonung. Abbildung 1 zeigt die Strompreisentwicklung für die Industrie seit 1998.

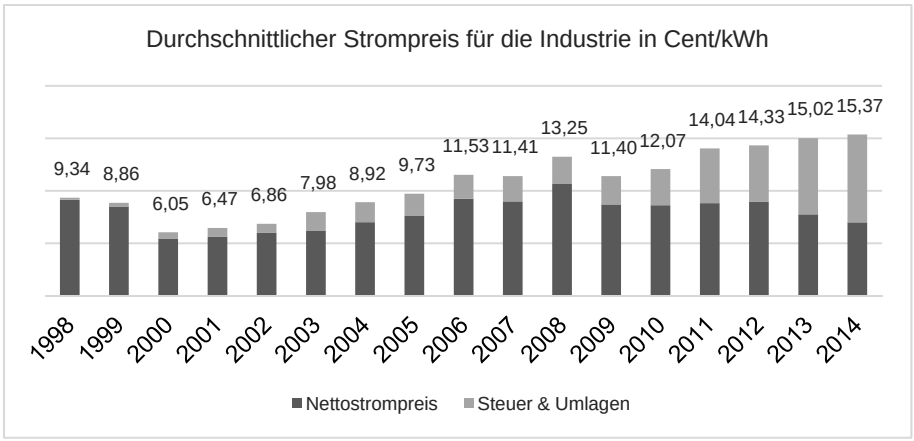



*Abbildung 1: Entwicklung der Energiepreise für die Industrie in Deutschland von 1998 bis 2014 (Bauernhansl et. al 2013, BDEW 2014)*

Anhand des Werkzeugmaschinenbestandes in Deutschland und der im Rahmen einer Untersuchung der Fraunhofer Gesellschaft ermittelten Jahresenergieverbräuche dieser Maschinen lässt sich abschätzen, dass deren Anteil am gesamten industriellen Energiebedarf nicht zu vernachlässigen ist. Die zuletzt ermittelte Anzahl an Werkzeugmaschinen in der deutschen Industrie wurde 1995 vom Verein Deutscher Werkzeugmaschinenfabriken e.V. mit 1,04 Millionen Einheiten festgestellt (Neugebauer 2008). Ein deutlicher Großteil von 83 Prozent, 863.000 Stück, entfiel dabei auf spanende Maschinen. Die im Rahmen einer Untersuchung der Fraunhofer Gesellschaft ermittelten Jahresenergieverbräuche solcher Maschinen liegen zwischen 5.000 kWh und 150.000 kWh (ebenda). Obwohl aufgrund der hohen Varianz des Energieverbrauchs der einzelnen Maschinen eine exakte Hochrechnung nicht möglich ist, kann dennoch festgestellt werden, dass spanende Maschinen einen beträchtlichen Anteil am gesamten industriellen Energiebedarf haben.

Dass dies wenigstens in einigen speziellen Industriebereichen so ist, verdeutlicht das folgende Beispiel eines Motorenwerkes eines deutschen Automobilherstellers, dessen vorwiegend spanende Produktion 66 Prozent des Gesamtenergiebedarfs des Werkes ausmacht (Tab. 1).

In Abbildung 2 ist die tatsächliche Leistungsaufnahme einer Senkrecht-Drehmaschine für die Bearbeitung von bis zu 70 Tonnen schweren Komponenten für Dampfturbinen mit einer Anschlussleistung von 150 kW dargestellt. Gut zu erkennen ist die hohe Grundlast mit einem Leistungsabruf von 11 kW bis 15 kW über etwa drei Stunden.

Untersuchungen der Technischen Universität Chemnitz haben ergeben, dass die Grundleistung einen beträchtlichen Teil des Gesamtenergiebedarfs spanender Werkzeugmaschinen einnimmt. Es wird geschätzt, dass bis zu 30 Prozent des Gesamtenergiebedarfs auf den Energieverbrauch im Standby-Betrieb entfällt.

*Tabelle 1: Verteilung des Elektroenergieverbrauchs in einem Motorenwerk (ebenda)*

| Bereich | Anteil in % |
|---|---|
| Produktion | 66 |
| Druckluft | 11 |
| Kälteanlagen | 8 |
| Emulsionsanlagen | 7 |
| Beleuchtung | 4 |
| Kühlturm | 2 |
| Heizhaus | 1 |
| Ölaufbereitung | 1 |
| Gesamt | 100 |

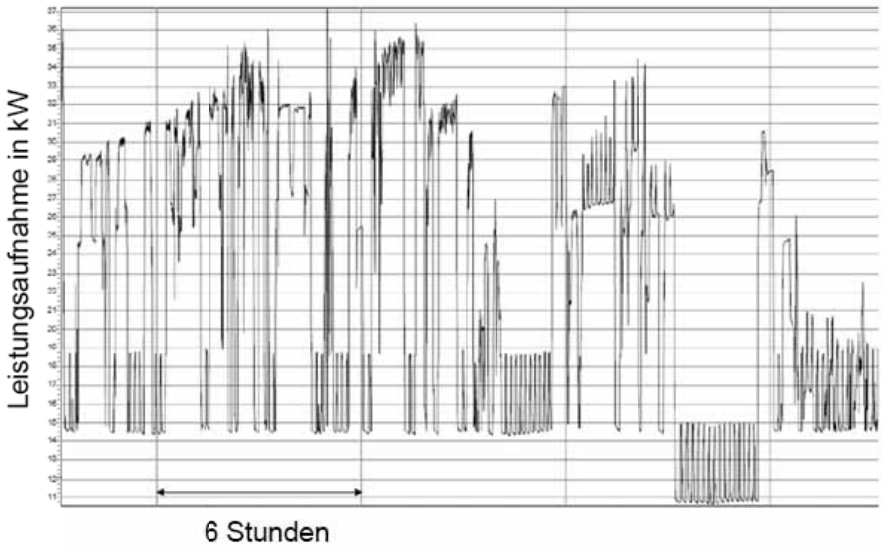



*Abbildung 2: Leistungsaufnahme einer Senkrecht-Drehmaschine (Neugebauer 2008)*

Einen detaillierten Einblick gibt hierzu eine Studie der RAVEL-Dokumentationen (Neugebauer 2008). Untersucht wurde die Verteilung der Betriebszeiten spanender Werkzeugmaschinen. So wird in der Großserienproduktion lediglich in 38 Prozent der Betriebszeit tatsächlich Span abgehoben. Der übrige Teil fällt auf Umrüstvorgänge und Wartezeiten. Noch kleiner fällt der Anteil bei der Produktion von Kleinserien aus. Hier sind es nur 15 Prozent (Abb. 3).

*Tabelle 2: Zeitliche Auslastung von Werkzeugmaschinen in der Klein- und Großserie*

| | Großserie | Kleinserie |
|---|---|---|
| Spanabhebung | 38 % | 15 % |
| Wartezeit | 36 % | 55 % |
| Werkzeugwechsel, Einstellungen | 26 % | 30 % |

Für die Produktionsplanung ergeben sich unmittelbar Handlungsfelder im Bereich der Ressourcenbelegungsplanung. Auf den ersten Blick erscheinen kurze Leerzeiten bzw. eine hohe Maschinenauslastung zielführend zu sein. Damit ließe sich der Energieverbrauch im Standby-Betrieb reduzieren. Möglich sind aber auch lange Leerzeiten, die eine vollständige Abschaltung von Maschinen gestatten. Für diese Option müsste das rechtzeitige sichere Hochfahren und damit die reibungslose Wiederaufnahme der Fertigung durch die Planung gewährleistet werden.

Das zweite Handlungsfeld, ebenfalls in der Produktionsplanung angesiedelt, liegt im Bereich der Losgrößenplanung. Aus größeren Losen und einem damit kleineren Produktmix je Planungsperiode ergeben sich unmittelbar weniger Rüstvorgänge und damit kürzere Wartezeiten, in denen die Maschinen mit Grundlast fahren.

## 2 Untersuchungsgegenstand

Die Maschinenbelegungsplanung (auch Feinplanung, Ablaufplanung oder Reihenfolgeplanung) legt fest, wann und in welcher Reihenfolge Aufträge an zur Verfügung stehenden Ressourcen bearbeitet werden – stets unter der Einhaltung bestimmter Restriktionen. Diese können beispielsweise vorgegebene Start- und Endtermine oder auch Rüst- und Betriebszustände der Maschinen sein.

Zur Lösung solcher Probleme wurden zahlreiche Entscheidungsmodelle mit unterschiedlichen Annahmen und Zielen entwickelt. Derartige Modelle werden in der Praxis dazu genutzt, vorhandene Ressourcen unter Betrachtung der geforderten Restriktionen und Zielsetzungen bestmöglich zu nutzen. Neben solchen, meist sehr rechenaufwendigen und damit zeitintensiven Constraint-basierten Verfahren, gibt es auch eine Reihe von heuristischen Verfahren zur Maschinenbelegungsplanung. Diesbezüglich ist der Einsatz von Prioritätsregeln in der Unternehmenspraxis vielfach anzutreffen. Bekannte Regeln dieser Art sind beispielsweise:

- FIFO: First-In-First-Out
- LIFO: Last-In-First-Out
- KOZ: Kürzeste Operationszeitregel
- LOZ: Längste Operationszeitregel
- GRB: Größte Restbearbeitungszeitregel
- KRB: Kürzeste Restbearbeitungszeitregel
- EDD: Earliest-Due-Date

Mit Hilfe der Simulationssoftware Plant Simulation 12 wird in diesem Beitrag untersucht, wie sich die Anwendung solcher Prioritätsregeln in der Maschinenbelegungsplanung auf den Gesamtenergieverbrauch und damit auf die entstehenden Energie-

kosten einer Werkstattfertigung auswirkt. Im Mittelpunkt steht hierbei der von der Produktion nicht direkt verursachte Energiebedarf, der beispielsweise im Standby-Modus und beim Hoch- bzw. Herunterfahren von Anlagen auftritt (siehe beispielhaft Abb. 2).

Ziel der Untersuchung ist es, die Relevanz der Betrachtung von Energiebedarfen und der damit verbundenen Kosten im Bereich der Feinplanung simulativ zu beurteilen.

# 3 Simulationsuntersuchung

## 3.1 Aufbau der Simulationsuntersuchung

Das modellierte Produktionssystem besteht aus neun Maschinen mit jeweils einer Warteschlange (Abb. 3).

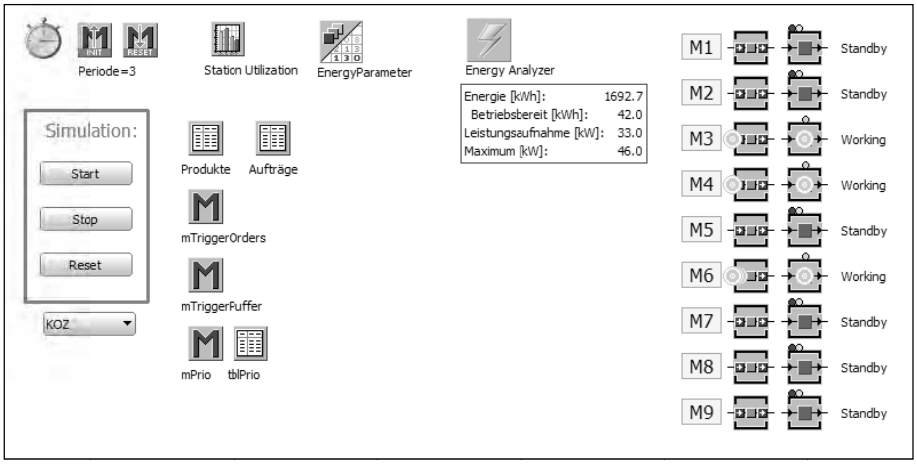



*Abbildung 3: Aufbau des Simulationsnetzwerks*

Alle Maschinen sind hinsichtlich ihrer Energieparameter und Bereitschaftszeiten gleich beschaffen (Abb. 4).

Die Überwachung sämtlicher Leistungsaufnahmen gewährleistet der Baustein *Energy Analyzer*. Er kumuliert die einzelnen Verbräuche in den verschiedenen Betriebszuständen jeder Maschine einzeln und protokolliert diese.

Mit Auftragsstart werden Aufträge zunächst nach FIFO-Reglement den Warteschlangen der jeweiligen Maschinen zugewiesen. Je nach Prioritätsregel wird dann die Reihenfolge der Aufträge in der Warteschlange bestimmt und diese entsprechend sortiert bevor der nächste Auftrag auf die Maschine bearbeitet wird.

Jede Maschine wird erst dann hochgefahren, wenn sie von einem Auftrag benötigt wird und erst dann wieder heruntergefahren, wenn kein Auftrag mehr für die Maschine vorgesehen ist. Wird die Maschine in der Zwischenzeit nicht benötigt, nutzt sie einen Standby-Modus um Energie und Zeit für das Hoch- und Herunterfahren zu sparen. Dieses Verhalten dient in der Praxis zur Absicherung der

Prozessstabilität. Ein häufiges Hoch- und Herunterfahren von Anlagen würde im beschriebenen Beispiel die Gesamtdurchlaufzeit stark verlängern.

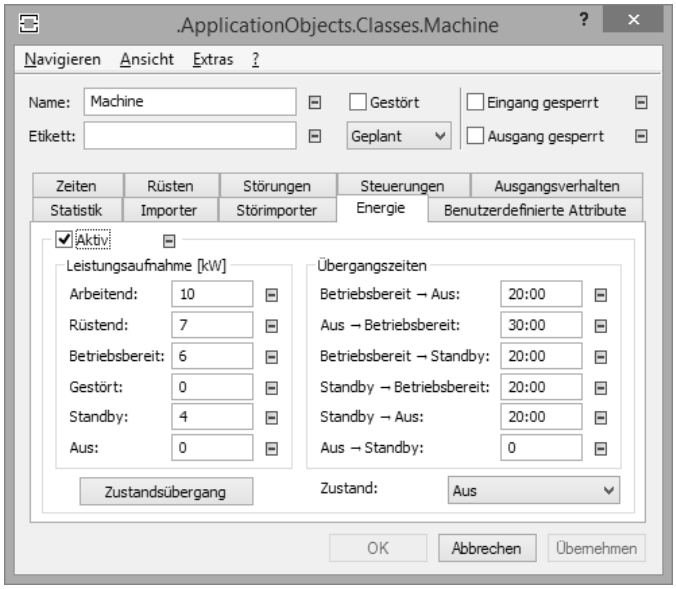



*Abbildung 4: Energieparameter der Maschinen*

Hat ein Auftrag eine Maschine durchlaufen, wird er in die Warteschlange der nächsten Station gelegt und dort erneut nach Anwendung der gewählten Prioritätsregel auf der Folgestation bearbeitet. Fertige Aufträge verlassen das System.

Aufgrund der Seitenlimitierung dieser Arbeit können die Aufträge und Arbeitspläne der einzelnen Produkte nicht mit aufgeführt werden. Für diese Untersuchung ist die Struktur der Aufträge und Arbeitspläne jedoch nicht direkt relevant, da lediglich geprüft wird, ob signifikante Unterschiede in der Leistungsaufnahme bei Anwendung verschiedener Planungsstrategien auf denselben Auftragspool bestehen.

Um die Experimente bewerten zu können, werden neben allen energetisch relevanten Daten zu den Maschinen auch die Durchlaufzeiten und die Auslastung aller Ressourcen und deren Betriebszustände protokolliert.

## 3.2 Ergebnisse der Simulationsuntersuchung

### *3.2.1 Gesamtenergiebedarf versus Gesamtdurchlaufzeit*

Bei der Bewertung der einzelnen Simulationsläufe sollte insbesondere der Gesamtenergiebedarf aller Maschinen und die Gesamtdurchlaufzeit aller Aufträge miteinander verglichen werden. Da die Rüst- und Bearbeitungszeiten für alle Experimente gleich sind und jede Maschine auch nur zu Fertigungsbeginn und -ende hoch- bzw. heruntergefahren wird, kann festgehalten werden, dass längere Durchlaufzeiten mit längeren Standby-Zeiten einhergehen. Da der Standby-Betrieb dieser Maschinen nicht zu vernachlässigen ist, wirkt sich dies signifikant auf den Gesamtenergiebedarf des Produktionssystems aus.

Zunächst erscheint es schlüssig, dass bei einer längeren Durchlaufzeit mehr Energiebedarf auf den Standby-Betrieb entfällt und damit auch der Gesamtenergiebedarf größer ist. Dass zwischen diesen beiden Größen kein linearer Zusammenhang besteht, zeigt der Vergleich der jeweiligen Auswertung in Abbildung 5 und Abbildung 6.

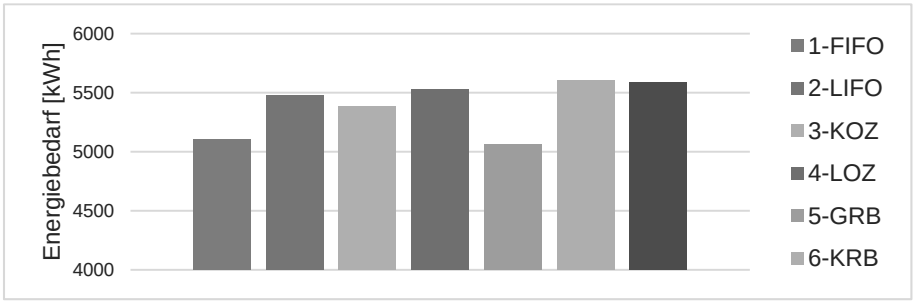



*Abbildung 5: Gesamtenergiebedarf nach Prioritätsregel*

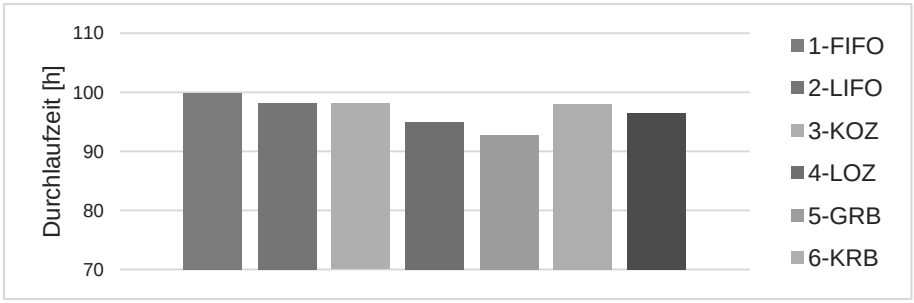



*Abbildung 6: Summe der Durchlaufzeit aller Aufträge nach Prioritätsregel*

Die beste Prioritätsregel hinsichtlich Durchlaufzeit und Energieverbrauch ist die Regel nach der größten Restbearbeitungszeit (GRB). Den geringsten Energiebedarf benötigt die FIFO-Regelung mit 51 kW/h, was auf einen erhöhten Anteil an Standby-Betrieb zurückzuführen ist. Selbstverständlich ist die Prioritätsregel nach der größten Restbearbeitungszeit nicht als Patentlösung für eine sowohl schnelle als auch ressourcenschonende Bearbeitung in der Werkstattfertigung zu werten. Welche Prioritätsregeln zielführend sind hängt von der Auftragsstruktur und unternehmensspezifischen Zielsetzungen ab.

### *3.2.2 Flexible Energiebedarfe*

Nachfolgende Abbildung 7 erläutert den Energiebedarf in den verschiedenen Betriebszuständen nach angewandter Prioritätsregel. Zu erkennen sind deutlich die zwischen 470 kWh und 1011 kWh schwankenden Standby-Energiebedarfe der einzelnen Szenarien.

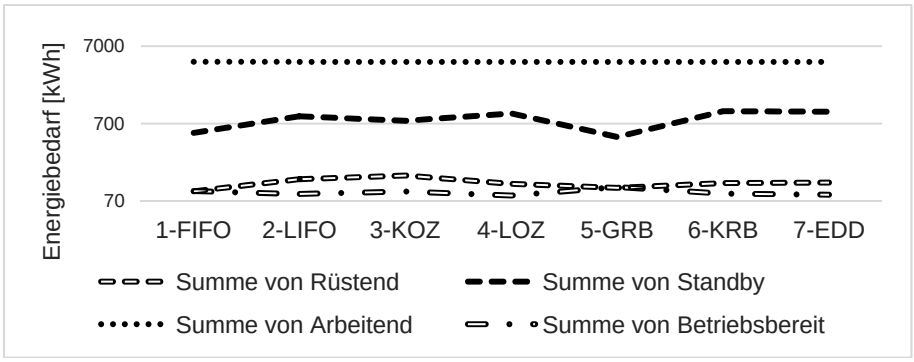



*Abbildung 7: Anteil am Gesamtenergiebedarf je Betriebszustand nach Prioritätsregel*

### 3.2.3 Erkenntnisse

Die Simulationsergebnisse haben gezeigt, dass es einen Bezug zwischen der Maschinenbelegungsplanung und der aufzuwendenden Energie in einer Werkstattfertigung gibt und sich dieser nicht linear zur Betriebszeit bzw. Gesamtdurchlaufzeit verhält. Infolgedessen existiert ein Optimierungspotential, das sich wie folgt erläutern lässt.

Die Abbildungen 8 und 9 zeigen in beiden Fällen einen gültigen Belegungsplan für dasselbe Szenario ohne Verspätungen. Der erste Belegungsplan nach FIFO nutzt infolge der späteren Starttermine für Job 3 bis 5 den Standby-Betrieb um die Prozessstabilität zu gewährleisten.

Geht man von einer Leistungsaufnahme wie in Tabelle 3 aufgeführt aus, so benötigt der Ablaufplan nach Abbildung 8 einen Gesamtenergiebedarf von 292 kWh im Planungszeitraum.

Im zweiten Belegungsplan (Abbildung 9) wird die zeitliche Flexibilität von Job 1 und 2 dazu genutzt, die Standby-Zeiten zu verkürzen. Alle Maschinen werden wesentlich später hochgefahren, um alle Aufträge ohne Pausen zu bearbeiten. Dieser Plan benötigt im Planungszeitraum einen Gesamtenergiebedarf von 245 kWh. Im Vergleich zur Planung nach FIFO ergibt sich eine Energieeinsparung von 16 %. Der Standby-Energiebedarf für dieses Beispiel ist im Vergleich zur industriellen Praxis gering. Enthalten Maschinen Kühlaggregate oder Laser, so sind Leistungsabrufe während des Standby-Betriebs von 10 kW/h oder mehr keine Seltenheit.

*Tabelle 3: Beispielhafte Leistungsaufnahme für fünf Maschinen in kW/h*

| | $M_1$ | $M_2$ | $M_3$ | $M_4$ | $M_5$ |
|---|---|---|---|---|---|
| Standby | 4 | 4 | 4 | 4 | 4 |
| Hochfahrend | 10 | 10 | 10 | 10 | 10 |
| Herunterfahren | 5 | 5 | 5 | 5 | 5 |
| Rüstend | 10 | 10 | 10 | 10 | 10 |
| Bearbeitend | 10 | 10 | 10 | 10 | 10 |

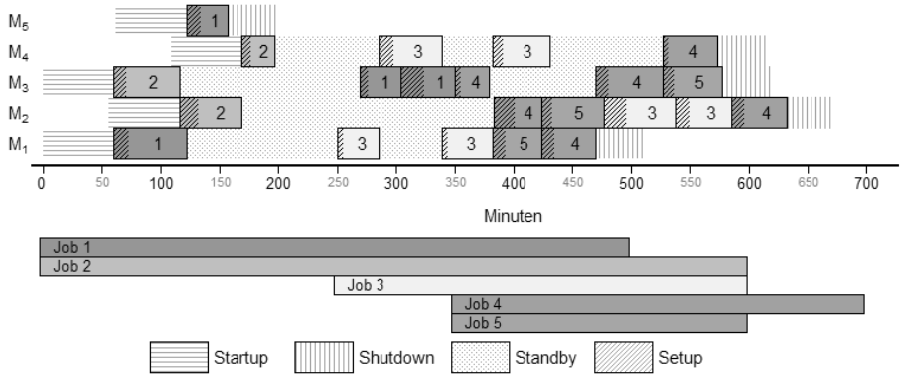



*Abbildung 8: Gültiger Maschinenbelegungsplan ohne Verspätungen*

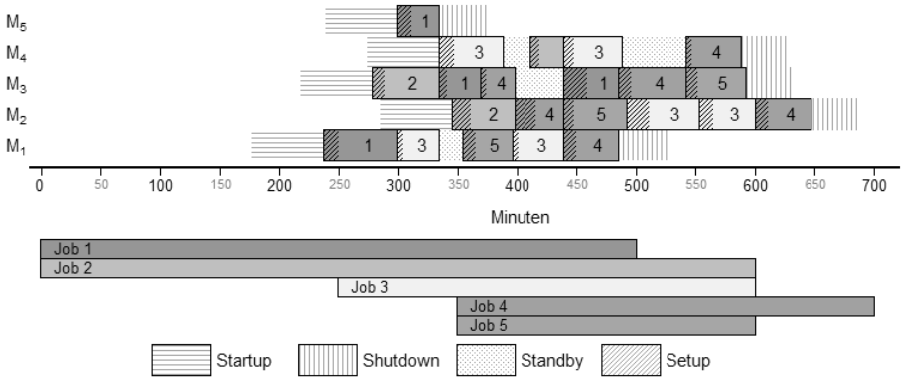



*Abbildung 9: Nach Energiebedarf optimierter Maschinenbelegungsplan ohne Verspätungen*

## 4 Ausblick

Grundsätzlich müsste ein Optimierungsverfahren zur nachhaltigen Maschinenbelegungsplanung die nicht von der Produktion selbst verursachten Energiebedarfe berücksichtigen: das Hoch- und Herunterfahren und den Standby-Betrieb von Maschinen. Bei Maschinen mit einem höheren Standby-Verbrauch würden dann verstärkt Leerzeiten vermieden werden. Maschinen könnten außerdem geplant herunter- und wieder hochgefahren werden, um in langen Wartezeiten Energie einzusparen. Dies ist dann der Fall, wenn der Standby-Verbrauch über den Wartezeitraum höher ist, als der Energiebedarf zum Herunter- und wieder Hochfahren.

Um den Energiebedarf für ein Optimierungsverfahren korrekt mit Kosten zu bewerten, müsste ein flexibler Energiepreis über den Planungszeitraum angenommen werden, da produzierende Unternehmen in der Regel Sonderverträge abschließen, welche individuell abgestimmte Energietarife enthalten. Die üblichen Bestandteile des so festgelegten Energiepreises sind nach Bönig (2013) in der Regel:

- Arbeitspreis: Darunter ist der Preis für eine bezogene kWh zu verstehen. Er fällt für die tatsächlich genutzten Kilowattstunden elektrischer Arbeit an und kann in Abhängigkeit des Tages variieren.
- Leistungspreis: Er wird für die Bereitstellung der elektrischen Arbeit berechnet. Die maximale Leistung, also die Lastspitze, innerhalb eines Abrechnungszeitraums ist hier der Multiplikator für die Berechnung.
- Verrechnungspreis: Der Verrechnungspreis beinhaltet den Abschlag für die vom Energieversorger installierten Zähler und Messgeräte.
- Blindleistungsanteil: Unternimmt ein Unternehmen keine Blindleistungskompensation, so muss infolge dessen ein Blindleistungsanteil abgetreten werden.
- Hinzu kommen noch Nutzungsentgelte, Abgaben, Umlagen und Steuern.

## Literatur

Bauernhansel, T.; Mandel, J.; Wahren, S.; Kasprowicz, R.; Miehe, R.: Energieeffizient in Deutschland. Stuttgart: Institut für Energieeffizient in der Produktion 2013

BDEW Bundesverband der Energie- und Wasserwirtschaft e.V.: Industriestrompreise, 28.04.2014.

Bönig, C.: Entwicklung einer Methode zu energiekostenorientierten Belegungsplanung. Schlussbericht der AiF-Forschungsvereinigung, 11.11.2013

Neugebauer, R.: Untersuchung zur Energieeffizienz in der Produktion. München: Fraunhofer Gesellschaft 2008

Wagner, J.; Schäfer, M.; Kohl, D.; Thide, H.; Grüning, C.; Hesselbach, J.; Böhm, S.: Energieeffizienz in der Industrie. Productivity Management 17 (2012) 4, S. 47-50.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Offshore-Windenergie: Kostensenkung durch Logistiksimulation

## *Offshore Wind Energy: Cost Reduction Through Logistics Simulation*

Torsten Münsterberg, Carlos Jahn,
Fraunhofer-Center für Maritime Logistik und Dienstleistungen CML,
Hamburg (Germany), torsten.muensterberg@cml.fraunhofer.de,
carlos.jahn@cml.fraunhofer.de

Abstract: Electricity production costs of offshore wind farms are too high. They have to be reduced to make offshore wind more competitive. The operational costs, especially logistics costs, have a great potential for cost reduction. In this article the author develops a modular simulation model for the operation phase of offshore wind farms by using the software Enterprise Dynamics. The model is able to represent offshore and onshore based logistics concepts. The model results are logistics and opportunity costs. Thus, the model evaluates the cost effectiveness of different logistics concepts. In multiple simulation experiments the model shows that for the selected scenario, an offshore wind farm located 100 km away from shore with 90 wind turbines, an offshore based mother vessel concept without helicopter support provides the best performance. Further research should focus on the impact of worsening weather conditions and increasing failure rates.

## 1 Einführung und Problemstellung

Die Stromgestehungskosten für Offshore-Windenergie variieren momentan zwischen 119 €/MWh und 194 €/MWh. Um Offshore-Windenergie wirtschaftlich zu nutzen, sollen die Stromgestehungskosten für Offshore-Windparks (OWP) ab 2020 auf unter 100 €/MWh gesenkt werden (DONG 2013; Kost et al. 2013). Der Betriebskostenanteil (OPEX) beträgt in etwa 25 – 30 % an den Stromgestehungskosten (Pieterman 2012). Die OPEX eröffnen somit bedeutende Kostensenkungspotenziale (Dinwoodie et al. 2012).

Das größte Potenzial fällt laut einer Industrieumfrage bei den Transport- und Logistikkosten an (Krüger et al. 2012). Auf Grund nur geringer Betriebserfahrungen mangelt es aktuell an standardisierten Lösungen. Betroffen ist besonders die Planung und Gestaltung der maritimen Logistikketten sowie die Auswahl und Disposition von Transportmitteln (Vojdani et al. 2012). Für den nachhaltigen wirtschaftlichen

Erfolg von OWP ist die Entwicklung eines effizienten Logistikkonzepts von fundamentaler Bedeutung, da Investitionen in logistische Mittel direkt in die Betriebskosten eingehen.

In diesem Beitrag wird die Analyse und Bewertung der Leistungsfähigkeit von Logistikkonzepten für den Betrieb von Offshore-Windenergieanlagen (OWEA) dargestellt. Zur Untersuchung dieser komplexen Zusammenhänge und der stochastisch verteilten Betriebsereignisse wurde ein Simulationsmodell in Enterprise Dynamics entwickelt. Mit dem Modell können sowohl onshore- als auch offshore-basierte Logistikkonzepte untersucht und mittels zahlreicher Variationen von Eingangsparametern umfassend analysiert und bewertet werden.

## 2 Stand der Wissenschaft und Technik

In diesem Kapitel werden zunächst die in der Praxis verwendeten und geplanten Logistikkonzepte beschrieben. Im zweiten Abschnitt wird das entwickelte Simulationsmodell abgegrenzt.

### 2.1 Logistikkonzepte

Abbildung 1 stellt die unterschiedenen Logistikkonzepte für den Betrieb von OWP grob dar (Phillips et al. 2013; Pistol 2011; Rehfeldt 2012).

*Abbildung 1: Logistikkonzepte*

Onshore-basierte Konzepte zeichnen sich durch Stützpunkte an Land aus, die als Ausgangsbasis für direkte Transporte von Personal und Material zu den OWEA dienen. Für Instandhaltungsmaßnahmen bilden Servicehäfen den Startpunkt für Routineeinsätze (Wartung und Instandhaltung) und kurzfristig anfallende Reparaturen. Onshore-basierte Konzepte, die nur auf Schiffe setzten, sind in ihrer Leistungsfähigkeit begrenzt. Ab einer bestimmten signifikanten Wellenhöhe ist kein sicherer Übergang zur OWEA mehr gewährleistet. Bei zunehmender Küstenentfernung vergrößert sich zudem die Transferzeit zu Lasten der verfügbaren Arbeitszeit. Daher werden bei größeren Entfernungen und insbesondere für die Durchführung reaktiver Instandhaltungsmaßnahmen und Inspektionen bevorzugt Helikopter eingesetzt. Dies hat den Zweck der Verringerung der Transferzeiten und der Gewährleistung des Zugangs bei starkem Seegang.

Mit offshore-basierten Konzepten sind die Erfahrungen noch beschränkt. Sie beinhalten stationäre oder mobile Einrichtungen (Wohnplattform, Hotelschiff, Mutter-

schiff) auf See für die Unterbringung von Personal sowie für die Lagerung von Material am Standort des OWP (Jahn et al. 2013; Neulinger et al. 2013). Das Personal verbringt dabei bis zu zwei Wochen am Stück auf See. Kurze Reaktionszeiten nach einem unvorhergesehenen OWEA-Ausfall erhöhen die Verfügbarkeit. Durch die geringere Transferzeit kann die produktive Arbeitszeit des Personals gesteigert werden (Pistol 2011).

## 2.2 Abgrenzung zu existierenden Modellen

Bisherige Methoden zur Untersuchung und Bewertung von Logistikkonzepten für den Betrieb von OWP setzten meist auf statisch-analytische Lösungen, die weder hindcast Wetterdaten (i. d. R. statistische Wetterdaten) noch konkurrierende Ressourcen (Schiffe, Helikopter) bei einem gleichzeitigen Ausfall zweier oder mehrerer OWEA berücksichtigen. Damit werden diese Ansätze der Komplexität des Systems OWP nicht gerecht. Wenn existierende Ansätze Simulation als Methode verwenden, handelt es sich aus kommerziellen Gründen meist um eine nicht nachvollziehbare Blackbox. Der hier verfolgte Ansatz bietet nachvollziehbare (visuelle) Prozesse. Das Modell hat eine strategische Ausrichtung, bildet aber ebenfalls nachvollziehbare operative Prozesse ab. Nach Kenntnis der Autoren existiert ein solch umfangreicher und transparenter Simulationsansatz zum jetzigen Zeitpunkt nicht. Mit dem entwickelten Modell sollen systematisch Zusammenhänge zwischen verschiedenen Einflussparametern und der Leistungsfähigkeit eines OWP untersucht werden. Abbildung 2 stellt den modellierten Prozess vereinfacht dar.

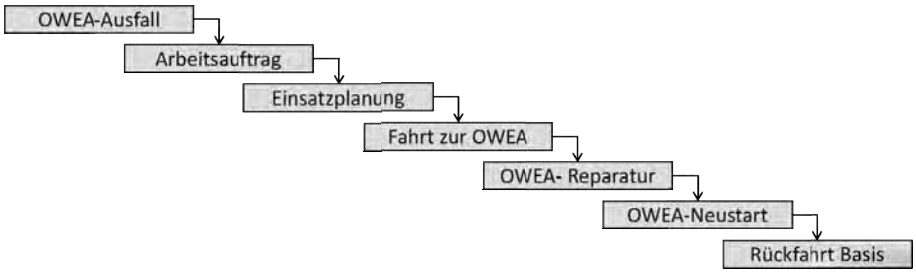



*Abbildung 2: Modellierter Prozess*

Betrachtet werden nur Anlagenteile über der Wasseroberfläche. Weitere Strukturen eines OWP (wie Umspannplattformen) werden nicht berücksichtigt. Das Modell bildet die Transportaktivitäten zwischen einem Service-Stützpunkt und dem OWP auf See ab.

# 3 Simulationsmodell

Dieses Kapitel befasst sich zunächst mit den Eingabe- und Ausgabeparametern des Simulationsmodells. Darauf folgen der grundsätzliche Aufbau, das Layout des Modells sowie eine Beschreibung der Datenverarbeitung im Modell.

## 3.1 Eingabe- und Ausgabeparameter

Das im Folgenden vorgestellte Simulationsmodell wurde in der Software Enterprise Dynamics von der Firma Incontrol entwickelt. Die wesentlichen Eingabeparameter (siehe Abb. 3) sind Informationen zum OWP und den OWEA, das eingesetzte Equipment für das Logistikkonzept, die geplante Instandhaltungsstrategie sowie die Umweltdaten im Windpark. Bis auf die Umweltdaten, welche sich aus Performancegründen bereits im Modell befinden, werden die übrigen Eingabeparameter per ActiveX-Schnittstelle aus MS Excel in das Modell geladen und dort in Tabellen gespeichert.

Zum OWP werden u. a. die Anzahl, Leistung und Ausfallraten der enthaltenen OWEA in das Modell übertragen. Das generierte Logistikkonzept wird in MS Excel näher spezifiziert. Je nachdem kann so die Anzahl der Transferschiffe und Helikopter variiert werden. Das Modell ist in der Lage onshore- und offshore-basierte Konzepte abzubilden. In Bezug auf die Instandhaltungsstrategie kann die Priorität der Ausfälle aber auch die Einsatzreihenfolge der zur Verfügung stehenden Fahrzeuge variiert werden.

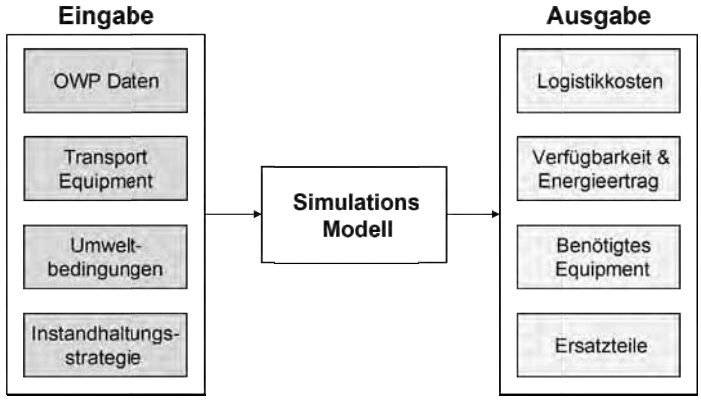



*Abbildung 3: Eingabe- und Ausgabeparameter*

Die wesentlichen Ausgabedaten des Modells sind die Gesamtlogistikkosten inklusive der Opportunitätskosten durch entgangenen Stromertrag. Daneben berechnet das Modell die Verfügbarkeit sowie weitere statistische Kennzahlen (u. a. Fehlerhäufigkeit, Kapazitätsfaktor, Stromproduktion) der einzelnen OWEA und des gesamten OWP. Der Vergleich verschiedener Logistikkonzepte gibt Aufschluss darüber, welches Konzept und Transportequipment eingesetzt werden sollte, um die logistischen Kosten für den Betrieb des OWP zu minimieren. Daneben kann aus den Ergebnissen abgeleitet werden, welche Ersatzteile benötigt und welcher Platz bzw. welche Fläche dafür auf dem Schiff bzw. im Hafen bereitgehalten werden muss.

## 3.2 Modelllayout

Abbildung 4 zeigt das Layout des Simulationsmodells in Enterprise Dynamics. Das Modell besteht dabei aus verschiedenen Objekten und Modulen, wie den OWEA, Schiffen, Helikopter, Hafen, Personal und Arbeitsaufträgen. Daneben existiert ein

Funktioneneditor, der Funktionen zur Steuerung des Modells enthält. Zudem enthält das Modell verschiedene Tabellen zur Datenspeicherung. Die Daten werden teilweise zur Auswertung automatisch nach dem Simulationslauf in MS Excel exportiert.

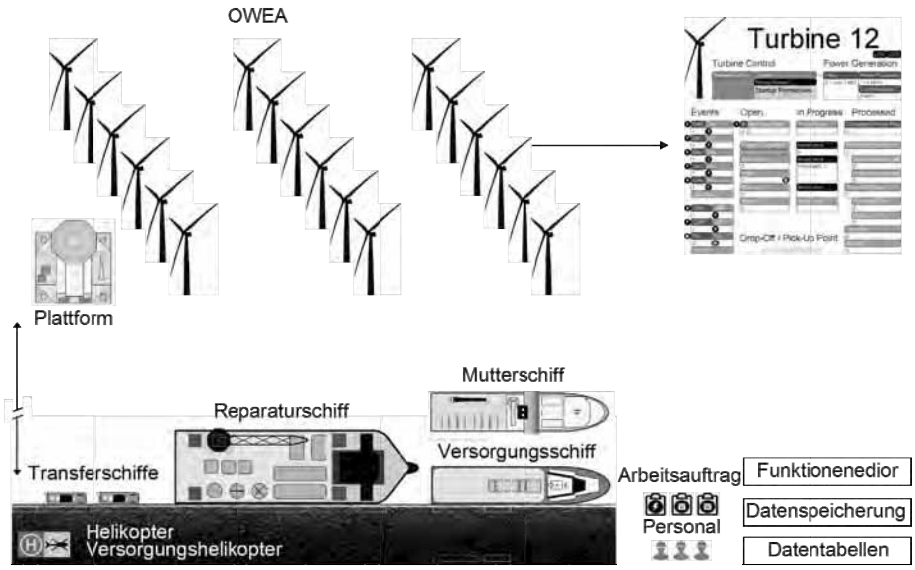



*Abbildung 4: Modelllayout*

Grundsätzlich kann im Modell zwischen onshore-basierten und offshore-basierten Konzepten gewählt werden. Dafür stehen verschiedene Fahrzeuge zur Auswahl. Transferschiffe und Helikopter werden für den Transfer zu den OWEA eingesetzt. Als Reparaturschiff wird ein selbstangetriebenes Hub-Schiff mit Bordkran verwendet. Es wird für Missionen mit hohen Anforderungen an die Transport- und Hebekapazität eingesetzt. Für offshore-basierte Konzepte kann zwischen einem Mutterschiff und einer bemannten Plattform gewählt werden. Der Personalwechsel und die Materialversorgung des Mutterschiffs bzw. der Plattform werden von einem Versorgerschiff bzw. Helikopter übernommen.

Bei den Fahrzeugen handelt es sich um modifizierte Advanced Transporter, die um einige Funktionen und Attribute erweitert wurden. Die Funktionen werden bei bestimmten Ereignissen ausgelöst. Erweiterte Attribute sind u. a. die zurückgelegte Strecke, Einsatzzeit und Kosten.

Jede OWEA besitzt verschiedene Fehlerklassen, die einer Komponentengruppe entsprechen. Den Fehlerklassen können unterschiedliche Wahrscheinlichkeiten für das Auftreten eines Fehlers zugewiesen werden. Je nach Art des Fehlers kann nur ein bestimmtes Fahrzeug zur Reparatur eingesetzt werden. Helikopter oder Transferschiffe werden bspw. für kleinere Reparaturen eingesetzt. Für größere Reparaturen wird ein Transferschiff oder das Mutterschiff verwendet. Für den Großkomponentenaustausch kann nur das Reparaturschiff herangezogen werden.

Das OWEA-Modul besteht hauptsächlich aus einer Kombination von Quellen, Senken und Servern. Die verschiedenen Fehlerklassen haben jeweils eigene Quellen mit Produkten, Senken und Servern, die parallel laufen. Die Zeit bis zu einem Ausfall wird durch die Serverdurchlaufzeit abgebildet. Die Durchlaufzeit (Zeit bis zum nächsten Ausfall) wird als Label auf dem Produkt gespeichert. Sobald ein Produkt z. B. den Server für korrektive Instandhaltung durchlaufen hat, fällt die OWEA aus und sendet einen Arbeitsauftrag an eine zentrale Auftragsplanung. Diese löst einen Einsatz mit Personal zur OWEA aus. Der Reparaturvorgang wird mittels eines Multiservers auf der OWEA simuliert. Im Anschluss verlässt das Personal die OWEA und diese wird wieder neugestartet.

## 3.3 Datenverarbeitung

Abbildung 5 zeigt schematisch, wie die verschiedenen Eingabeparameter in das Modell eingehen und dort verarbeitet werden.

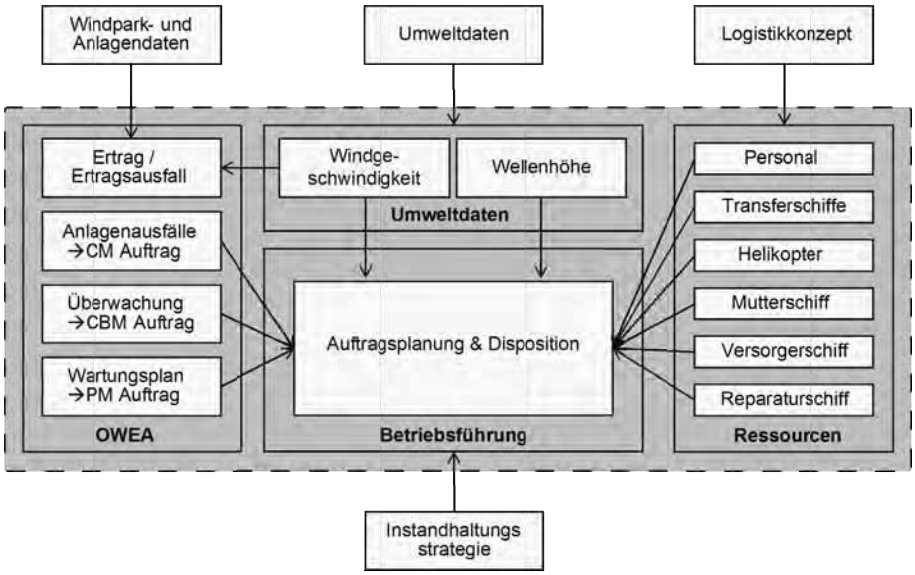



*Abbildung 5: Datenverwendung*

Die Anlagendaten umfassen Ausfallwahrscheinlichkeiten und gehen direkt zum OWEA-Modul im Modell, um dort Ereignisse für korrektive Instandhaltung (CM), zustandsbasierte Instandhaltung (CBM) und geplante Instandhaltung (PM) auszulösen. Diese Ereignisse sind negativ exponentiell verteilt. Die Wahrscheinlichkeit für einen Ausfall innerhalb eines bestimmten Zeitraums beschreibt Gleichung 1 (Dinwoodie et al. 2012):

$$U(t) = 1 - e^{-\lambda t} \tag{1}$$

$\lambda$ = Ereignisrate

t = Zeit

Die OWEA greift gleichzeitig auf die aktuelle Windgeschwindigkeit zu, um so in Kombination mit ihrer Leistungskurve die Stromproduktion zu berechnen bzw. den Stromausfall (während eines Störfalls der OWEA).

Dabei wird anhand der Leistungskurvenfunktion die entsprechende Leistung für die aktuelle Windgeschwindigkeit berechnet und im OWEA-Modul kumuliert gespeichert.

Die Umweltdaten (stündliche Werte für Wind und Welle) werden zudem bei der Auftragsplanung verwendet. Sobald eine OWEA ausfällt oder geplante Instandhaltung ansteht, wird ein Arbeitsauftrag generiert. Das Auftragsplanungsmodul sucht dann innerhalb der Umweltdaten mittels verschiedener Schleifen nach einem passenden Wetterfenster, in dem die Limits von Wellenhöhe und Windgeschwindigkeiten für einen sicheren Überstieg zur OWEA nicht überschritten werden. Ist ein Wetterfenster gefunden wird der Arbeitsauftrag, je nach Instandhaltungsstrategie einem bestimmten Fahrzeug des Logistikkonzeptes zugewiesen. Die Auftragsplanung löst ein Ereignis zum Start eines entsprechenden Einsatzes für Instandhaltung bzw. Instandsetzung zur OWEA aus.

## 4 Experimente

Die Fähigkeiten des Modells sollen anhand eines Experimentes verdeutlicht werden. In dem Experiment wird ein Szenario mit 90 OWEA bei einer Entfernung von 100 km zur Küste gewählt. In Bezug auf die OWP in der deutschen Nordsee handelt es sich um einen großen und weit von der Küste entfernten OWP. Die verwendeten Umweltdaten enthalten stündliche Werte für Wellenhöhe und Windgeschwindigkeit im Bereich der deutschen OWP in der Nordsee. Zunächst soll analysiert werden, welches Logistikkonzept die geringsten Gesamtkosten (Logistikkosten und Ertragsausfälle) ergibt. Dabei werden für onshore- und offshore-basierte Konzepte mit Helikoptereinsatz die Anzahl der Transferschiffe von 0 bis 3 bzw. 1 bis 4 variiert. Im Anschluss wird der Einfluss des Helikopters genauer untersucht, indem in einem weiteren Simulationsexperiment auf seinen Einsatz verzichtet wird. Auf diesem Wege kann die Robustheit der Konzepte gegen Helikopterausfälle untersucht werden.

In den Experimenten wird davon ausgegangen (Maples et al. 2013), dass jede OWEA pro Jahr 9,5 Mal ausfällt, wobei es sich in 5 Fällen davon um „Remote Resets“ handelt, die keinen Einsatz vor Ort nötig machen. In ca. der Hälfte der übrigen 4,5 Vorfälle kann ein Helikopter eingesetzt werden. Für die restlichen Vorfälle muss ein Transferschiff oder in ganz weniger Fällen (< 0,2) ein großes Reparaturschiff eingesetzt werden.

Tabelle 1 gibt einen Überblick über die eingesetzten Fahrzeuge und Annahmen zu deren wichtigsten Kennzahlen.

Bis auf das Reparaturschiff werden die Fahrzeuge im Ein-Schicht-Betrieb eingesetzt. Die Schicht darf dabei maximal 12 Stunden inklusive Reisezeit dauern.

*Tabelle 1: Fahrzeugkennzahlen*

| | Helikopter | Transferschiff | Mutterschiff | Reparaturschiff |
|---|---|---|---|---|
| PAX (Techniker) | 3 | 12 | bis zu 51 | 6 |
| Max. $H_S$ in m | - | 1,5 | 3 | 2 |
| Max. $V_W$ in m/s | 17 | 12 | 17 | 10 |
| Geschwindigkeit in kn | 128 | 20 | 11 | 10 |
| Jährliche Kosten in € | 2.160.000 | 900.000 | 7.500.000 | - |
| Tägliche Kosten in € | - | 1.200 | 4.000 | 150.000 |
| Stündliche Kosten in € | 1.200 | - | - | - |
| Mobilisierungskosten in € | - | - | - | 400.000 |

## 5 Ergebnisse

Bei den getroffenen Annahmen ergeben die Simulationsläufe ein Kostenminimum beim Mutterschiffkonzept ohne Transferschiff. Die jährlichen Kosten, bestehend aus Logistikkosten (Schiffe und Personal) und Ertragsausfällen, beziffern sich auf ca. 30 Mio. € (siehe Abb. 6). An zweiter Stelle folgt das onshore-basierte Konzept mit zwei Transferschiffen. Dieses Konzept führt zu Kosten von ca. 31 Mio. € pro Jahr. Das günstigste Plattform-Konzept kommt mit einem Transferschiff auf ca. 34 Mio. € pro Jahr. Es sei darauf hingewiesen, dass die Untersuchung nur auf „harten" Faktoren wie Kosten basiert. Die Berücksichtigung „weicher" Faktoren (wie z. B. permanente Parküberwachung, Sicherheitspuffer bei Fahrzeugausfall oder Seekrankheit) könnten daher auch zur Auswahl eines anderen Konzeptes führen. Der Einsatz eines weiteren Transferschiffes könnte unter Einbezug weicher Faktoren ebenfalls als vorteilhaft angesehen werden.

Bei der Untersuchung der Robustheit der jeweils besten Ausprägung (Anzahl Transferschiffe) der drei Konzepte zeigt sich ein heterogenes Bild. Beim onshore-basierten Konzept mit zwei Transferschiffen nehmen die jährlichen Kosten um 8 Mio. € (ca. 25 %) zu. Beim offshore-basierten Plattformkonzept führt der dauerhafte Wegfall des Helikopters dazu, dass dieses Konzept nicht mehr in der Lage ist die jährliche Wartung ausreichend durchzuführen sowie die korrektive Instandhaltung zeitnah vorzunehmen. Die Kosten steigen um mehr als 100 % auf ca. 74,5 Mio. €. Der größte Teil hiervor entfällt auf Ertragsausfälle.

Beim Mutterschiffkonzept führt der dauerhafte Wegfall des Helikopters zu einem gegenteiligen Effekt. Die Kosten nehmen ab. Der Mehrwert des Helikoptereinsatzes wiegt demnach nicht die deutlichen Mehrkosten für seinen Einsatz auf. Das Mutterschiff ist also in der Lage, die meisten Einsätze des Helikopters auch selber zu übernehmen.

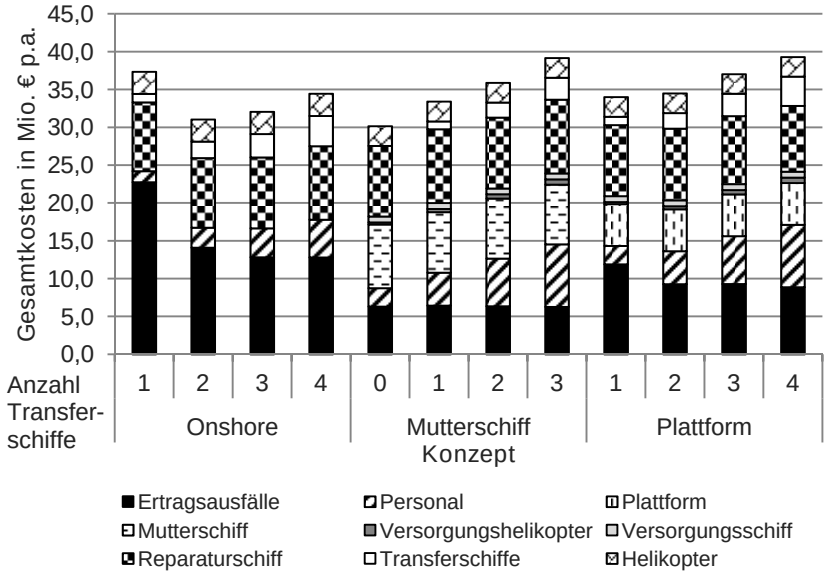



*Abbildung 6: Gesamtkosten verschiedener Logistikkonzepte*

## 6 Fazit

Das vorgestellte Simulationsmodell zur Bewertung von Logistikkonzepten für OWP ist in der Lage diverse Konzepte monetär zu vergleichen und zu bewerten. Im untersuchten Szenario mit 90 OWEA bei einer Küstenentfernung von 100 km liefert ein offshore-basiertes Mutterschiffkonzept die beste Performance. Die Robustheit der Konzepte in Bezug auf die Helikopterunterstützung variiert sehr stark. Das favorisierte Konzept ist ohne den Helikoptereinsatz noch kostengünstiger. Das onshore-basierte Konzept wird deutlich teurer ohne Helikopter und das offshore-basierte Plattformkonzept mit einem Transferschiff wird untauglich ohne den Helikoptereinsatz.

Zukünftig sollten weitere Effekte auf die Logistikkonzepte, wie z. B. schlechteres Wetter oder erhöhte Ausfallraten getestet werden. Ein weiteres interessantes Feld ist die unterjährige Analyse der Konzepte. Diese kann Aufschluss darüber geben, in wie weit ein Konzept im Laufe des Jahres angepasst werden sollte, um dauerhaft optimal für einen OWP zu sein.

## Literatur

Dinwoodie, I. A.; Mcmillan, D.; Quail, F.: Analysis of offshore wind turbine operation & maintenance using a novel time domain Meteo-Ocean modeling approach. In: ASME Turbo Expo 2012, Kopenhagen, 11.-15.06.2012. Strathprints Institutional Repository, S. 847-857.

DONG: Offshore wind cost-of-Electricity. Information package February 2013. Dong Energy 2013.

Jahn, C.; Münsterberg, T.: Simulation-based design and evaluation of O&M logistics concepts for offshore wind power plants. In: Grabe, J. (Hrsg.): Proceedings of the Conference on Maritime Energy, Hamburg 21.-22.5.2013, S. 285–296.

Krüger, D.; Bichler, M.; Bartholl, C.; Hellich, P.; Knote, C.: Europäische Wind-Service-Studie. Woher der Wind weht. Deloitte, TaylorWessing 2012, online www.taylorwessing.com/uploads/tx_siruplawyermanagement/Europäische_Wind-Services-Studie_Taylor_Wessing_DE.pdf

Maples, B.; Saur, G.; Hand, M. et al.: Maintenance strategies to reduce the cost of offshore wind energy installation, operation, and maintenance strategies to reduce the cost of offshore wind energy. Golden, USA: National Renewables Energy Laboratory (NREL) 2013.

Neulinger, R.; Clark, R.; Perinic, K. et al.: Betriebserfahrungen und Betriebskosten. In: Böttcher, Jörg (Hrsg.) Handbuch Offshore-Windenergie. Rechtliche, technische und wirtschaftliche Aspekte. München: Oldenbourg Wissenschaftsverlag 2013, S. 432–454.

Pieterman, R.P.v.d.: Data collection & analysis as input for O&M cost modelling. In: Windpower Monthly Wind Farm Data Management & Analysis Forum, Hamburg, 20.-21.11.2012. Energy Research Centre of the Netherlands (ECN) 2012.

Pistol, B.: „Konkretisierung des Hafenkonzeptes Offshore-Häfen Nordsee SH. Endbericht“. Hamburg: UNICONSULT Universal Transport Consulting, The Interreg IVB North Sea Region Programme 2011.

Phillips, J.; Fitch-Roy, O.; Reynolds, P. et al.: A guide to UK offshore wind operations and maintenance. London: GL Garrad Hassan, Scottish Enterprises, The Crown Estate 2013.

Rehfeldt, K.: Bedeutung von nachhaltigen O&M Konzepten für die Offshore Windenergie. Hamburg: Stiftung Offshore Windenergie, WindGuard 2012.

Vojdani, N.; Lootz, F.: Life-cycle-orientierte Gestaltung der logistischen Prozesse für Offshore-Windenergieanlagen. In: Jahrbuch Logistik 2012. Korschenbroich: free beratung GmbH, S. 146–152.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Vorstudie zur energetischen Modellierung automatischer Parksysteme

## *Preliminary Study for Energetic Modelling of Automated Parking Systems*

Katrin Sinha, Bernd Künne, TU Dortmund, Dortmund (Germany),
katrin.sinha@tu-dortmund.de, bernd.kuenne@tu-dortmund.de

Abstract: In view of decreasing fossil fuel, energy efficiency is one of the most important keys to all industry sectors. By contrast, the number of cars utilised all over the world is increasing which leads to a growing stationary traffic. From this it follows a demand to develop a system creating more parking spaces. An automated parking system provides the reduction of space requirements concerning conventional car parks. In this case, the parking space is doubled at the same cubature. The aim of the project is the analysis of the energy consumption of automated parking systems and the improvement of the energy efficiency. In this case different simulation models of automated parking systems are going to be developed. The bases of these models are theoretical backgrounds of the propulsion technology and the system architecture of available parking systems. This paper demonstrates the results of the simulation preliminary study.

## 1 Einleitung

Neben regenerativen Energien, angepassten Netzstrukturen und Elektromobilität übernimmt die Steigerung der Energieeffizienz eine wichtige Rolle bei der erfolgreichen Umsetzung der Energiewende (Stadt Dortmund Umweltamt 2011). Stetig steigender Energiekonsum, bedingt durch den hohen Automatisierungsgrad heutiger Prozesse, erzwingt eine Effektivitätssteigerung von Anlagen und Maschinen. Energieeffiziente Anlagen und Produkte haben im Hinblick auf eine nachhaltige Energieversorgung daher eine große ökonomische und ökologische Bedeutung.

Steigender Individualverkehr und die dazugehörige Unterbringung des ruhenden Verkehrs sind ein immer größeres Problem. Untersuchungen zeigen, dass in Spitzenzeiten Parksuchverkehre bis zu 80% des Verkehrsaufkommens verursachen (BAGMG 1990). Park-Such-Situationen sind mit Stausituationen gleichzusetzen. Verglichen mit dem fließenden Verkehr werden hier drei- bis fünfmal so viel Schadstoffe produziert (Fittinghoff 2001). Die Forderung nach mehr Stellplätzen ist demnach nicht nur eine ökonomische, sondern auch eine ökologische Lösung. Dadurch

bedingt rücken automatische Parksysteme (APS) bei der Stadt- und Verkehrsplanung immer mehr in den Fokus und finden z. B. in neuen sogenannten Smart-City-Konzepten Anwendung. Eine energetische Optimierung von APS erfordert sowohl die Abstimmung von Produktivität als auch zeitlicher Effizienz.

## 2 Stand der Technik

Im folgenden wird ein Überblick über den Stand der Technik automatischer Parksysteme gegeben sowie Anwendungsbereiche der Simulation im Kontext logistischer Systeme mit energetischen Fragestellungen präsentiert. Der Schwerpunkt liegt dabei auf der Betrachtung von Simulationen, die als Ziel die energieeffiziente Untersuchung von Intralogistischen Anlagen anstreben.

### 2.1 Automatische Parksysteme

Ein Parksystem verfügt über mindestens zwei Stellplätze, einen Zubringer, der als Straße oder Fahrgasse ausgebaut ist, und die zugehörigen Subsysteme, sodass das Parken von mindestens zwei Pkw möglich ist (Kirchmann 1993). Eine Systematisierung von Parksystemen nach dem Automatisierungsgrad differenziert Anlagen in konventionelle (herkömmliche) und automatisierte Systeme. Während in konventionellen Systemen ausschließlich die Antriebskraft des fahrzeugeigenen Motors zum Erreichen des Stellplatzes genutzt wird, übernehmen in APS technische Mittel oder Förderanlagen den Parkvorgang entweder teilweise oder vollständig. Grundsätzlich besteht ein APS aus einer Übergabezone/-station, Fördertechnik mit Antriebsaggregaten und einem Lagermittel, das zur Speicherung der Pkw dient.

Die Entwicklung von APS begann in den Anfängen des 20. Jahrhunderts (Hamelink 2011). Heute sind auf dem weltweiten Markt unterschiedliche Systemvarianten mit verschiedenen Fördertechniken verbreitet. Die Hersteller setzen dabei auf die von ihnen entwickelte und seit Jahren erprobte Technik. Einheitliche System- oder Fördertechnik finden herstellerübergreifend keine Anwendung.

*Abbildung 1: Multifunktionales APS in Containerbauweise. Prototyp des FG ME der TU Dortmund. Links: Vorderansicht; Einfahrtsrampe und geöffnete E/A-Station. Rechts: Innenansicht; vorne das in der Gasse verfahrende RBG, dahinter (r.) die mit der Öffnung nach innen gerichteten Container (Künne und Sinha 2013a)*

Energetische Betrachtungen standen bislang nicht im Vordergrund der Forschung. Obwohl ein hoher Energiebedarf als Nachteil in der Literatur aufgeführt wird, fehlen wissenschaftliche Untersuchungen und Optimierungsansätze (Söhnchen 1995; Bellmann 1997). Entsprechend sind energetische Verlustmechanismen und Einflusspara-

meter nur unzureichend analysiert und deren Einfluss auf den Energiebedarf wenig erforscht. Auch die unterschiedlichen Förderaggregate (Regalbediengeräte (RBG), Shuttle-Lift-System (SLS)), Stellplatzanordnungen (Turm- oder Gassensysteme) und Systemvarianten (mit und ohne Ladehilfsmittel (LHM)) sind bislang noch nicht im Bezug auf den Energiebedarf vergleichend gegenübergestellt worden.

Seit einigen Jahren wird eine Kombination aus automatischem Park- und Lagersystem, in Form multifunktionaler Park- und Lagersysteme, diskutiert (Kuhn 2004). Im Hinblick auf ein steigendes Verkehrsaufkommen bietet die zeitgleiche oder versetzte Mehrfachnutzung logistischer Systeme zur Lagerung von Pkw und Waren eine Möglichkeit mit hoher Flexibilität und multipler Funktion in innerstädtischen Bereichen und Ballungsgebieten, dezentralen Park- und Lagerraum zu schaffen. Im Rahmen eines Verbundprojektes wurde vom Fachgebiet Maschinenelemente (FG ME) gemeinsam mit Industrie- und Hochschulpartnern ein neuartiges multifunktionales APS in Containerbauweise entwickelt (Abb. 1).

## 2.2 Simulation in der Logistik

Der Einsatzbereich von Simulation in der Logistik ist vielfältig. Ein Indikator für den Einsatz der Simulation in diesem Bereich ist u. a. die Zeitanhängigkeit der dort betrachteten Systeme. Eine detaillierte Übersicht, die unterschiedliche Simulationsansätze in logistischen Systemen zeigt, geben Blecker et al. (2014) in *Innovative Methods in Logistics and Supply Chain Management*.

Ein wichtiges Ziel in der aktuellen logistischen Forschungslandschaft sind u. a. Maßnahmen zum Umweltschutz. Beispielsweise erfolgt die Erfassung von $CO_2$-Parametern im Kontext der Green Supply Chain (Rabe et al. 2012). Aber auch die Steigerung der Energieeffizienz intralogistischer Anlagen wird erforscht. Die energetische Optimierung eines Kleinteilelagers ist im Forschungsprojekt *Eneff Lager - Energieeffiziente Lagerstrategien und Lastverteilung* vorgenommen worden (Zadek 2014). Die Verwendung der Simulation als Analysetool ist hierbei gängige Praxis. Forschungsarbeiten beziehen sich auf den Aufbau und die Simulation von Modellen, die Ergebnisse geben z. B. Aufschluss über das Potenzial von Leichtbau-Werkstoffen in RBG (Günthner et al. 2011) oder spiegeln den Einfluss kinematischer Parameter auf den Energiebedarf von RBG in Kleinteilelagern wider (Schulz et al. 2012).

Jedes intralogistische System verfügt über elektrische Antriebe. Im Zusammenhang einer Energieeffizienzanalyse von Antriebskomponenten kann die Simulation Anwendung für die Bewertung von elektrischen Antrieben finden, und energieexakte Simulationsmodelle können aufgebaut werden (Blank und Roth-Stielow 2011).

Vermehrt findet die Simulation nun auch im Kontext mit APS Anwendung. Die Modellierung eines APS nahm Schrepfer (2000) vor und untersuchte diese u. a. hinsichtlich der logistischen Kapazität. Weiterhin sind Vorarbeiten aus der Forschung bekannt, die das Ein- und Ausfahrtverhalten der Kunden modellieren, damit Rückstauungen auf öffentliche Verkehrswege vermieden werden (Bernhard und Fender 2003). Die Untersuchung alternativer Fördermitteltechnik in APS erforschte Eller (2010). Das Ziel dieser Arbeit ist, die Reduktion des Lastpendelns beim Einsatz von Kranen als Fördermittel in APS zu erreichen, dafür werden unterschiedliche Steuerungsstrategien von Krananlagen simuliert.

Die energetische Optimierung von APS ist bislang nicht wissenschaftlich untersucht worden. Dementsprechend fehlen Ansätze sowie quantitative belegbare Aussagen.

Die vorliegenden Forschungsarbeiten dienen dazu, die Forschungslücke zu schließen und ein Simulationsmodell aufzubauen, um den Energiebedarf unterschiedlicher APS-Systemvarianten sowie mögliche Energieeinsparpotentiale aufzuzeigen.

# 3 Vorstudien für den Simulationsaufbau

Für den Aufbau eines Simulationsmodells, das den Energiebedarf von APS abbildet, ist eine Systemanalyse zwingend erforderlich. Im Zuge der Forschungsarbeiten werden APS der Kategorie Parkregale mit feststehenden Regalen unter Variation verschiedener Fördermittel einer Energieeffizienzanalyse unterzogen. Deswegen wird an dieser Stelle für eine detaillierte Beschreibung anderer Parksystemvarianten auf die Literatur verwiesen (Fittinghoff 2011; Irmscher et al. 2013; Kirchmann 1993).

Um die energetische Optimierung von APS zu erreichen, müssen die Einflussgrößen auf den Energiebedarf bekannt sein. Entsprechend erfolgt vorab eine Systemanalyse, die sowohl den allgemeinen Aufbau als auch die Funktionsweise von APS umfasst.

## 3.1 Allgemeiner Aufbau und Funktionsweise von APS

APS sind in der VDI-Richtlinie 4466 definiert als „Systeme mit fördertechnischen Einrichtungen, die in einer Übergabekabine ein dort geparktes Fahrzeug aufnehmen und automatisch zu einem Lagerfach (Stellplatz) transportieren, um es dort einzulagern. Nach einer variablen Lagerzeit (Parkzeit) wird das Fahrzeug auf Anforderung vom Stellplatz automatisch ausgelagert und in eine Übergabekabine befördert, wo es dem Benutzer wieder zur Verfügung steht“. Der allgemeine Aufbau und die Funktionsweise von APS entsprechen grundlegend dem Aufbau eines in der Intralogistik etablierten automatischen Hochregallagers.

Der Parkvorgang endet aus des Nutzers Sicht in einem APS nach Abgabe des Fahrzeuges in der Übergabestation. Systemseitig jedoch beginnt nun der eigentliche Parkprozess. Es folgt der automatische Pkw-Transport durch den Einsatz entsprechender Fördermitteltechnik zu einem freien Stellplatz. Nach Beendigung der Parkdauer kann das Fahrzeug in der Übergabezone durch den Nutzer abgerufen werden. Nach Transport zur Nutzerschnittstelle ist der Vorgang aus System- und Betreibersicht abgeschlossen. Der Pkw ist wieder im Verantwortungsbereich des Nutzers.

## 3.2 Ergebnisse der energetischen Parameterstudie

Parameter, die Einfluss auf den Energiebedarf von APS haben, sind in allgemeingültige, energetische Parameter der Antriebskomponenten und systemspezifische Parameter zu gliedern. Allgemeingültige Einflussparameter sind in jedem Parksystem in ähnlicher Weise vorhanden. Einen Einfluss auf den Energiebedarf des Systems haben beispielsweise die Antriebsgeschwindigkeit, die interne Prozessführung, die Nutzungsfrequenz und Taktzeiten sowie die Massen der unterschiedlichen Ladeeinheiten. Jedoch auch die Massen der Fördermittel und Hilfsaggregate beeinflussen die benötigte Motorleistung. Auf der antriebstechnischen Seite hängen die Einflussparameter von den jeweils verwendeten Antriebskomponenten ab. Dazu zählen u. a. die Art der Motoren, deren Auslegungs- und optimaler Betriebspunkt, die verwendeten Führungselemente und Getriebe sowie die Lagerungsart. Systemspezifische Parameter sind bestimmt vom Aufbau des Parksystems, d. h. abhängig von Etagen- und Spaltenanzahl, Lage der E/A-Station, Taktspielanzahl, Stellplatzanordnung, -tiefe,

-breite und -anzahl sowie der Fördermitteltechnik. Bei gleicher Stellplatzanzahl hat beispielsweise ein Turmsystem einen anderen Energiebedarf als ein Gassensystem. Wieterhin ist die verwendete Fördermitteltechnik ausschlaggebend für den Energiebedarf von APS. SLS transportieren während der Erschließung der horizontalen Verfahrwege, konstruktiv und durch die Arbeitsweise bedingt, geringere Massen als RBG. Systeme mit Palettenwechseleinheit arbeiten mit mehr Taktspielen, dafür aber zeitlich schneller als Systeme ohne ein solches Aggregat.

### 3.3 Erste Projektergebnisse

Erste Untersuchungen und theoretische Betrachtungen bezüglich eines energieeffizienten Betriebs von APS am fachgebietseigenen Demonstrator zeigen, dass der Energiebedarf der Anlage (a) vom Gewicht der einzulagernden Güter vgl. hierzu (Künne und Sinha 2013b; Sinha 2013) und (b) bedingt durch den Anteil der kinetischen Energie von der Verfahrgeschwindigkeit der Förderaggregate abhängt.

Messungen des Energiebedarfs des RBGs in Abhängigkeit von der Nutzlast bei einer Fahrt von Ebene 0 in Ebene 2 wurden zu Testzwecken durchgeführt. Die Lastabhängigkeit der Motorenleistung wird bei einer Beladung von ca. 1200 kg durch einen um ca. 15% erhöhten Energiebedarf im Vergleich zu einer Leerfahrt bestätigt (Künne und Sinha 2013b). Prozessführungen mit höheren Geschwindigkeiten oder unterschiedlicher Priorisierung der einzulagernden Gütermassen beeinflussen nachweislich den Energiebedarf der Antriebskomponenten (vgl. hierzu Sinha 2013). Entsprechend an die Gewichtsklassen der Nutzlast angepasste Prozessführungen würden somit den Energieverbrauch des Systems senken (Sinha und Künne 2012).

## 4 Simulationsaufbau

Die energetische Optimierung von APS ist u. a. durch die Abhängigkeit der energetischen Einflussparameter problematisch. Ein Ziel der Forschungsarbeiten ist herstellerunabhängig die energetisch sowie zeitlich effizienteste APS-Konfiguration zu generieren. Dafür ist eine herstellerübergreifende Analyse und Bewertung von APS notwendig. Einerseits ist es aufgrund der Vielzahl möglicher Konfigurationen notwendig, die Simulation als Analyseinstrument für die verschiedenen Simulationsszenarien einzusetzen (vgl. hierzu März et al. 2011). Andererseits fehlen reale Anlagen zu Testzwecken. Die Analyse unterschiedlicher Fördermittel ist z. B. nicht ohne monetären Aufwand und komplizierte Umbaumaßnahmen zu prüfen. Ferner ist betreiberseitig ein Anlagenstillstand nicht zumutbar.

### 4.1 Simulationsprogramm *Plant Simulation*

Zur Untersuchung der vorliegenden Problematik wird das in der Industrie weit verbreitete Simulationsprogramm *Plant Simulation* der Firma *Siemens Industry Software GmbH & Co. KG* (2015), das zum Produktportfolio von *Tecnomatix* gehört, verwendet. Mit *Plant Simulation* können Produktionssysteme modelliert und simuliert werden. Weiterhin können beispielweise der Materialfluss, die Ressourcennutzung und die Logistik auf allen Stufen der Anlagenplanung optimiert werden.

Aufgrund der Analogie von APS zu intralogistischen Anlagen ist *Plant Simulation* ein guter Kompromiss bzgl. Einsatzbreite und Implementierungsunterstützung. Zudem bietet *Plant Simulation* die Möglichkeit, das Verhalten von Simulationsmodel-

len genauer zu beschreiben, da diese mit der zugrunde liegenden Skriptsprache „*Sim Talk*“ zu steuern sind (Eley 2012).

## 4.2 Allgemeiner Aufbau der Simulation von APS

In den Simulationsaufbau gehen sowohl die Inputparameter, die Einfluss auf den Energieverbrauch eines APS nehmen, als auch die entsprechend zu modellierenden Simulationselemente ein. Abbildung 2 zeigt schematisch den allgemeinen Aufbau der Simulation. Die Simulationselemente können aufgrund ihrer Eigenschaften nicht mit Hilfe von zugewiesenen Parametern gestaltet werden. Exemplarisch hierfür sind die Stellplatzkonfigurationen, für die aufgrund der verschiedenen Anordnungen unterschiedliche Simulationsmodelle erstellt werden müssen, anzuführen. Gleichermaßen sind die Fördermittel als Element zu modellieren. Es ist z. B. nicht möglich, die Verfahrrichtung durch zugewiesene Attribute in Listen abzubilden.

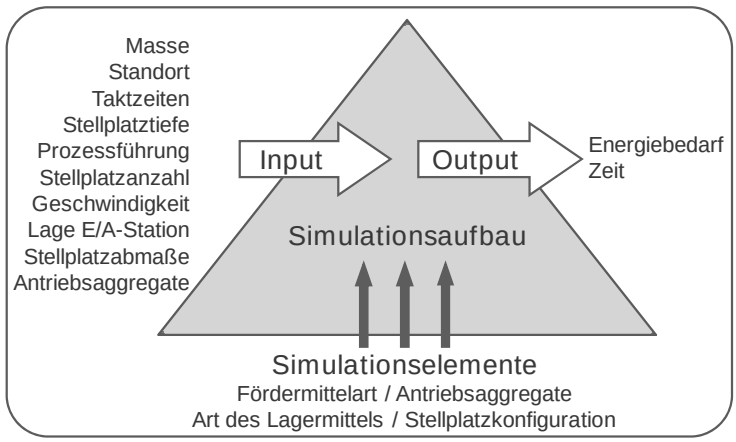



*Abbildung 2: Allgemeiner Aufbau der unterschiedlichen Simulationsmodelle*

APS sind mit intralogistischen Systemen gleichzusetzen. Deshalb muss als Outputparameter sowohl der Energiebedarf als auch die benötigte Zeit für einen Prozessschritt ausgegeben werden. Erst durch beide Parameter ist die energetische Optimierung in Abhängigkeit von Produktivität und zeitlicher Effizienz zu realisieren.

## 4.3 Simulationselemente und Simulationsmodelle

Die energetische Optimierung erfordert Aufbau und Überprüfung mehrerer Simulationsmodelle unterschiedlicher Parksystemkonfigurationen. Durch Variation des Simulationsaufbaus sind neue Simulationsmodelle konfigurierbar. Unterschiedliche Simulationsszenarien werden unter Variation der Inputparameter generiert. Nach Analyse der Ergebnisse erfolgt eine entsprechende Optimierung. Abbildung 3 zeigt einige zu modellierende Simulationselemente mit Stell- und Einflussgrößen. Die Stellplatzkonfiguration bestimmt den Aufbau des APS. Aufgrund unterschiedlicher Arbeitsweisen können auch Fördermittel nicht durch einfaches Zuweisen von Attributen in Listenform abgebildet werden. Die spezifischen Eigenschaften von SLS und RBG müssen über Modelle festgelegt werden. Zur Überprüfung verschiedener

Antriebskonzepte ist ebenso eine Modellierung erforderlich. Aufgrund des Kennlinienverhaltens und der Anhängigkeit unterschiedlicher Parameter voneinander kann eine Zuweisung im Listenformat nicht erfolgen.

**Stellplatzkonfiguration**

- Bauweise des Regals
  - Turmbauweise
  - Gassensystem
- Stellplatzanordnung
  - radial
  - tangential
  - senkrecht zur Gasse
  - parallel zur Gasse
- Etagenhöhe
- Etagenanzahl
- Anzahl der Spalten
- Gesamtstellplatzzahl
- Stellplatzbreite

**Fördermittelkonfiguration**

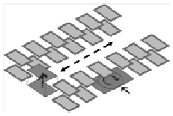
- Shuttle-Lift-System
  - Trennung vertikalen & horizontalen Transfers
  - Anzahl der Shuttle pro Ebene
  - Anzahl der Lifte
  - mit/ohne Palettenwechsler
  - mit/ohne LAM
- RBG
  - Diagonaltransfer möglich
  - Trennung vertikalen & horizontalen Transfers
  - mit/ohne Palettenwechsler
  - mit/ohne LAM

**Antriebskomponenten**

- Antriebsart
  - elektrisch
  - hydraulisch
  - Leistungsklassen
  - Kennlinienverhalten
  - Anfahrverhalten
- Antriebsmotor
  - Synchronmaschine
  - Asynchronmaschine
- Energierückgewinnung
- Lagerungsart
  - Gleit-/Wälzlager
- Führungen

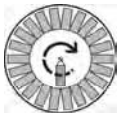
*Abbildung 3: Ausgewählte Simulationselemente, die zum Aufbau eines Simulationsmodells zur energetischen Analyse und Optimierung von APS identifiziert sind*

**Modell 1**



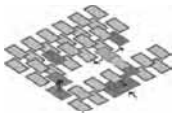
- Gassensystem
- Stellplatzanordnung: senkrecht
- RBG
- LHM
- Palettenwechsler
- einfachtief
- E/A-Station
- Energierückgewinnung

**Modell 2**



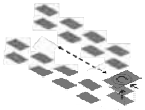
- Turmbauweise
- Stellplatzanordnung: radial
- RBG
- LHM
- Palettenwechsler
- einfachtief
- E/A-Station

**Modell 3**



- Gassensystem
- Stellplatzanordnung: senkrecht
- SLS
- LHM
- Palettenwechsler
- mehrfachrief
- E/A-Station

**Modell 4**



- Gassensystem
- Stellplatzanordnung: parallel
- RBG
- ohne LHM
- ohne Palettenwechsler
- einfachtief
- E/A-Station

*Abbildung 4: Exemplarische Auswahl an Simulationsmodellen durch Variation der Simulationselemente (Darstellung der APS nach Hamelink 2011)*

In Abbildung 4 wird eine Auswahl an Simulationsmodellen, die unterschiedliche Parksystemtypen abbilden, mit zugehörigen Systemspezifikationen, gezeigt. Zu jeder Systemvariante ist eine schematische Abbildung dargestellt. Diese verdeutlicht die Stellplatzkonfiguration mit Anordnung und Tiefe der Stellplätze, die Systembauweise, die Lage der E/A-Station und spiegelt das verwendete Fördermittel wider.

## 4.4 Konkretes Simulationsmodell am Beispiel des multifunktionalen APS in Containerbauweise

Die Erläuterung des konkreten Simulationsmodells erfolgt anhand des Beispiels eines realen Simulationsszenarios für das vorhandene multifunktionale APS.

| Variationsparameter | Beschreibung | Wertebereich |
|---|---|---|
| Gassenseitenanzahl | Stellplatzseiten pro Gasse | 1 - 2 |
| Stellplatzanordnung | Ausrichtung der Stellplätze zur Gasse | 1 / 2 |
| Stellplatzbreite | Breite der Stellplätze in [m] | 2 - 3 |
| Stellplatzhöhe | Höhe der Stellplätze in [m] | 1,4 - 2,5 |
| Stellplatztiefe | Anzahl der Stellplätze pro ISO-Container [Fuss] | 20 / 40 |
| Stellplatzanzahl S. 1 | Stellplätze pro Seite | 1 - 6 |
| Stellplatzanzahl S. 2 | Stellplätze pro Seite | 1 - 9 |
| Etagenanzahl S. 1 | Anzahl *X* der Etagen der Stellplatzmatrix S. 1 | 1 - 3 |
| Etagenanzahl S. 2 | Anzahl *Y* der Etagen der Stellplatzmatrix S. 2 | 1 - 3 |
| Spaltenanzahl S. 1 | Anzahl an Stellplätzen pro Etage: Gassenseite 1 | 1 - 3 |
| Spaltenanzahl S. 2 | Anzahl an Stellplätzen pro Etage: Gassenseite 2 | 1 - 2 |
| Lage der E/A-Station | Lage der E/A-Station im System | S. 1 / 2; X; Y |

*Abbildung 5: Wertebereich der Parameter des Simulationselementes Stellplatzkonfiguration am Beispiel des multifunktionalen APS in Containerbauweise*

Abbildung 5 zeigt den entwickelten Wertebereich der Parameter des Simulationselements *Stellplatzkonfiguration* und die damit verbundenen Variationsmöglichkeiten des fachgebietseigenen Demonstrators. An dieser Stelle wird deutlich, dass die Anwendung der Simulation als Analysetool aufgrund der Vielzahl an Variationen zwingend erforderlich ist.

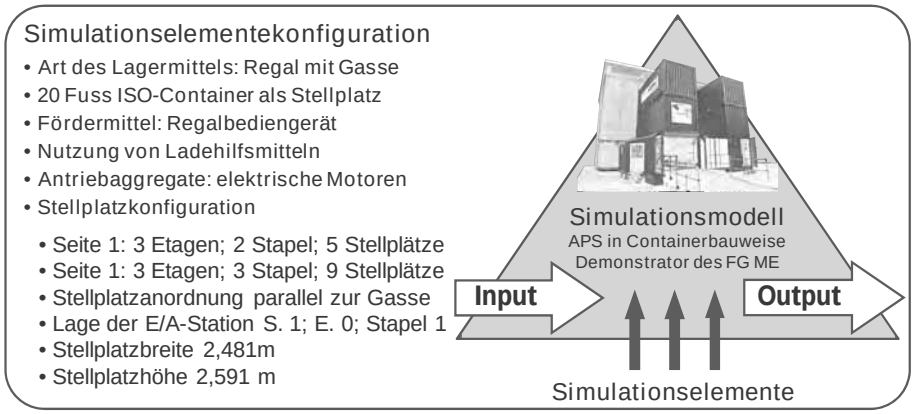



*Abbildung 6: Darstellung des konkreten Simulationsmodellaufbaus am Beispiel des fachgebietseigenen Demonstrators eines multifunktionalen APS in Containerbauweise mit quantitativer Angabe des Simulationselements Stellplatzkonfiguration.*

Abschließend zeigt Abbildung 6 die Konkretisierung des Simulationsmodellaufbaus des multifunktionalen automatischen Parksystems in Containerbauweise in Analogie zum allgemeinen Simulationsaufbau aus Abbildung 2. Der schematisch dargestellte Modellaufbau entspricht der Versuchsanlage des Fachgebiets Maschinenelemente der TU Dortmund und wird in den Simulationen als Referenzobjekt dienen.

## 5 Zusammenfassung und Ausblick

Die Simulationsvorstudie automatischer Parksysteme zeigt, dass die Einflussparameter auf den Energiebedarf von APS in einer wechselseitigen Beziehung zueinander stehen. Dadurch bedingt muss eine genaue Charakterisierung unter Variation dieser energetischen Größen erfolgen. Da die Auswirkung der Parameter auf den Energiebedarf von APS bislang unerforscht ist, ermöglicht erst diese Spezifikation die energetische Optimierung von APS. Die erste Phase der Forschungsarbeiten hat bereits gezeigt, dass ein Potential bezüglich einer energetischen Optimierung besteht und durch beispielsweise eine entsprechende Prozessführung zu erreichen ist (vgl. Kap. 3.3). Eine herstellerübergreifende Optimierung ermöglicht eine Vielzahl von APS-Konfigurationsmöglichkeiten. Aus diesem Grund ist der Einsatz der Simulation als Analyseinstrument der unterschiedlichen Simulationsszenarien erforderlich.

In der aktuellen Forschungsphase wird mit *Plant Simulation* der vorgestellte Prototyp des multifunktionalen Parksystems in Containerbauweise realisiert. Dazu werden die in Kapitel 4.3 vorgestellten Simulationselemente in das Simulationsmodell übertragen. Nach erfolgter Validierung des Modells werden im Anschluss unterschiedliche Simulationsszenarien abgebildet, und eine erste Analyse des Systems erfolgt anhand dieser Ergebnisse. Im Verlauf der Forschungsarbeiten sind weitere Modelle mittels der modellierten Simulationselemente zu konfigurieren, um verschiedene APS abzubilden. Unter Variation der energetischen Parameter werden obige Simulationsszenarien analog durchlaufen und die Ergebnisse ausgewertet. Abschließend werden weitere energetische Einsparpotentiale identifiziert, und eine herstellerunabhängige Variante eines energieeffizient gestalteten APS zu realisieren sein.

## Literatur

BAGMG: Parken und Handel. 2. Auflage, Bundesarbeitsgemeinschaft für Mittel- und Großbetriebe des Einzelhandels e.V. (Hrsg.), Köln 1990.

Bellmann, S.: Mechanische Parkhäuser. Bochum, Ruhr-Universität Bochum, Lehrstuhl für Verkehrswesen, Fakultät für Bauingenieurwesen, Diplomarbeit, 1997.

Bernhard, J.; Fender T.: Experimentelle Anwendung statistischer Verfahren und Visualisierungsmethoden zur Gewinnung ausgesuchter Eingangsdaten im Kontext von GNL. SFB 559, Teilprojekt M9, Dortmund, 2003, online verfügbar https://eldorado.tu-dortmund.de/bitstream/2003/5500/1/03014.pdf, zuletzt geprüft am 10.07.2015.

Blank, F.; Roth-Stielow, J.: Energieexakte Simulationsmodelle zur Bewertung der Energieeffizienz. ETG-Fachbericht 130: Internationaler ETG-Kongress 2011 (mit CD-ROM). Vorträge der Fachtagungen 1 - 5 des Internationalen ETG-Kongresses, 8. - 9.9. 2011, Würzburg, Berlin, Offenbach: VDE-VERLAG, 2011

Blecker, T.; Kersten, W.; Ringle, C. M.: Innovative Methods in Logistics and Supply Chain Management – Current Issues and Emerging Practices. epubli GmbH, Berlin: 2014.

DIN 30781: Teil 1 - Transportkette, Grundbegriffe. Deutsches Institut für Normung (Hrsg.), Berlin, Köln: Beuth 1983.

Eley, M.: Simulation in der Logistik: Einführung in die Erstellung ereignisdiskreter Modelle unter Verwendung des Werkzeuges "Plant Simulation". Berlin, Heidelberg: Springer, 2012.

Fittinghoff, M.: Beitrag zur Leistungsbetrachtung von Parksystemen. Dortmund, Universität Dortmund, Fakultät Maschinenbau, Dissertation, 2001.

Günthner, W. A.; Schipplick, M.; Ertl, R.; Hahn-Woernle, P.: Wettbewerbsfaktor Energieeffizienz– Teil I. Simulationsmodell für Entwicklung und Betrieb von automatischen Regalbediengeräten. Fördern und Heben 6 (2011), S. 276-279.

Hamelink, L.: The mechanical parking guide. CreateSpace Independent Publishing Platform, 2011.

Irmscher, I; Kosarev, I.; Schiefenhövel, A.: Handbuch und Planungshilfe – Parkhäuser und Tiefgaragen. Berlin: DOM Publishers 2013.

Kirchmann, J.: Gestaltung und Betrieb leistungsstarker automatischer Parksysteme. Dortmund, Universität Dortmund, Fakultät Maschinenbau, Dissertation, 1993.

Kuhn, M.: Beitrag zur Gestaltung multifunktionaler Lagersysteme. Dortmund, Universität Dortmund, Fakultät Maschinenbau, Dissertation, 2004.

Künne, B.; Sinha, K.: Energieeffiziente Gestaltung automatischer Parksysteme. Parken aktuell 90 (2013a), S.18 - 21.

Künne, B.; Sinha, K.: Energieeffiziente Gestaltung automatischer Parksysteme - Einsparpotentiale und Einflussgrößen auf den Energiebedarf automatischer Parksysteme. Parken aktuell 91 (2013b) S.17.

März, L.; Krug, W.; Rose, O.; Weigert, G.: Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen. Berlin, Heidelberg: Springer 2011.

Rabe, M.; Horvath, A.; Spieckermann, S.; Fechteler, T.: An Approach for increasing flexibility in green supply chains driven by simulation. Winter Simulation Conference, Berlin: IEEE, 2012, S. 3144 - 3155.

Richtlinie VDI 4466: Automatische Parksysteme – Grundlagen (Blatt 1). Verein Deutscher Ingenieure (Hrsg.), Düsseldorf: Beuth Verlag, 2001.

Schulz, R.; Monecke, J.; Zadek, H.: Der Einfluss kinematischer Parameter auf den Energiebedarf eines Regalbediengerätes. Logistics Journal: Proceedings, ISSN 2192-9084, 2012.

Schrepfer L.: Modellierung und Simulation von Systemen dynamischer Struktur. München: Herbert Utz Verlag GmbH, 2000.

Siemens Industrie Software GmbH & CO. KG: Plant Simulation http://www.plm. automation.siemens.com/de_de/products/ tecnomatix/ plant_ design/plant_ simulation.shtml, zuletzt geprüft am 10.04.2015.

Sinha, K.: Energieeffizienz automatischer Parksysteme. Bundesverband Parken e.V., Tagung am 13.11.2013, Würzburg: Kompetenzforum Parken, Vortrag, 2013.

Sinha, K.; Künne, B.: Energiemanagement multifunktionaler automatischer Containerlagersysteme, Dortmund, TU Dortmund, FG ME, Forschungsschule für energieeffiziente Produktion und Logistik, Zwischenbericht, 2012.

Söhnchen, A.: Innovative Technologie für den ruhenden Verkehr in Stadtzentren und verdichteten Wohngebieten. Wirtschaftspolitische Diskurse, 65, Friedrich-Ebert-Stiftung, Bonn: FES Library, 2002, online verfügbar http://library.fes.de/ fulltext/fo-wirtschaft/00363toc.htm, zuletzt geprüft am 10.05.2015.

Stadt Dortmund Umweltamt: Handlungsprogramm Klimaschutz 2020. Abschlussbericht, Dortmund, 2011.

Zadek, H.: Eneff Lager-Energieeffiziente Lagerstrategien u. Lastverteilung. Schlussbericht, Magdeburg, 2014, online verfügbar http://www.bvl.de/files/ 441/481/522/578/Schlussbericht_17389_Eneff_Lager_2014-06-15.pdf, geprf. 10.07.2015.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Geführte Assistenz für Ablaufsimulationsprojekte

***Assistance for Material Flow Simulation Projects***

Gottfried Mayer, Marielouise Mieschner, BMW AG, München (Germany),
gottfried.mayer@bmw.de, marielouise.mieschner@bmw.de

Abstract: This article describes the development of an assistance system for material flow simulation experts. The way was paved by two research projects, called AssistSim and EDASim. Resulting from those, a development partnership between two automotive manufacturers and a simulation company was founded. It ended with the launch of an official software product called SimAssist. The use of SimAssist made it necessary to create a new process model and a new saving concept for simulation studies. Furthermore, the basic functionality, specific features like the ones for the VDA (Verband der Automobilindustrie) group material flow simulation and a strategy to use the open source library D3JS for visualisation are shown.

## 1 Einleitung

Aktuell werden keine großen Entscheidungen in der Industrie mehr getroffen, ohne diese vorher durch Simulationen absichern und bewerten zu lassen. Die Ablaufsimulation ist hier Basis für nahezu alle Entscheidungen im Umfeld der automatisierten Produktion. Um die vermehrten Anfragen für Simulationen bewältigen zu können, ist es nötig, den Prozess des Aufbaus und der Durchführung von Studien zu optimieren und dadurch zu verkürzen.

An vorderster Stelle stehen hier Vorgehensmodelle für Simulationsstudien, wie sie beispielsweise vom VDI (VDI 2014, S. 19) oder von der ASIM (Rabe et al. 2008, S. 4-8) beschrieben werden. Für den Einsatz in der Automobilindustrie hat die Arbeitsgruppe Ablaufsimulation des Verbands der Automobilindustrie (VDA) die Anwendung dieser Vorgehensmodelle in einer Ausführungsanweisung weiter konkretisiert (VDA 2013).

Doch wenngleich Simulationsvorgehensmodelle seit vielen Jahren bekannt sind und in der Literatur diskutiert werden, wie z. B. die Übersicht in Rabe et al. (2008, S. 27-32) zeigt, hängen auch bei der Verwendung eines Simulationsvorgehensmodells viele Entscheidungen während einer Simulationsstudie maßgeblich vom Wissen und von der Erfahrung der an der Studie Beteiligten ab. Zwar gibt es für die Erstellung von ausführbaren Modellen zahlreiche kommerzielle Simulationswerkzeuge und ergänzende Objektbibliotheken (Mayer und Pöge 2010), die diesen Teil einer Studie

deutlich vereinfachen. Für die Arbeiten im Rahmen einer Studie, die nicht unmittelbar mit der Erstellung und Verwendung des ausführbaren Modells zusammenhängen, gibt es jedoch kein durchgängig eingesetztes Softwarewerkzeug. Natürlich kommen sämtliche Werkzeuge aus der „Office-Welt“ von der Textverarbeitung über die Tabellenkalkulation bis zum Präsentationsprogramm zum Einsatz. Ein strukturierter Einsatz dieser Werkzeuge ist jedoch eine weitgehend individuelle Entscheidung.

Hier kommen Simulationsassistenzsysteme zum Einsatz, die den Ablauf von der anfänglichen Datensammlung und Bewertung über eine teilautomatische Modellgenerierung, des Experimentdesigns sowie dessen Durchführung bis hin zur Auswertung der Ergebnisse unterstützen.

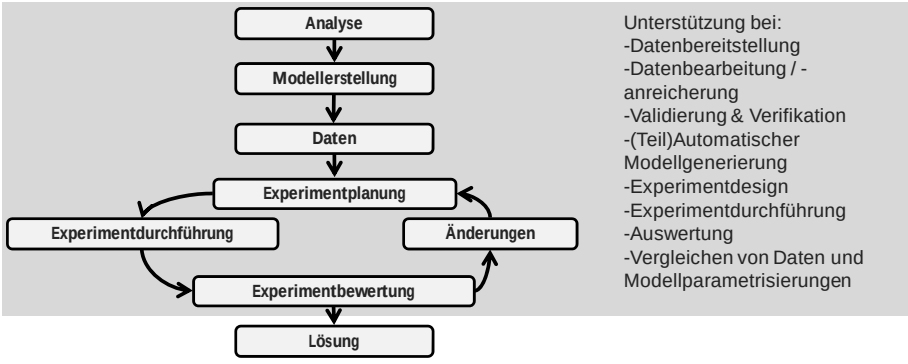



*Abbildung 1: Ziele der Unterstützung anhand eines Vorgehensmodells*

Dieser Beitrag soll die Erstellung eines solchen Systems beleuchten. Die erste Idee für das System liegt über fünf Jahre zurück und die ersten Untersuchungen wurden durch zwei Forschungsprojekte durchgeführt. Das erste Projekt, mit Namen AssistSim, hatte das Ziel, den Rahmen für ein Assistenzsystem zu erstellen. Das zweite, EDASim, sollte AssistSim um Methodiken zu Evaluierung der Daten, sowohl der Eingangs- als auch der Ergebnisdaten, erweitern.

Die Ergebnisse der beiden Projekte wurden durch die Forschungspartner als so gut bewertet, dass aus Teilen der Partner eine Entwicklungskooperation entstand. Mittlerweile ist die Umsetzung des Produktes abgeschlossen und steht als Kaufprodukt zur Verfügung.

Der Ansatz, der schon am Anfang der Forschungen getroffen wurde, eine flexible Oberfläche für die Belange der Ablaufsimulation zu gestalten, hat sich hier als richtig erwiesen. Die beliebige Anpassbarkeit durch ein Plugin- Konzept hat es ermöglicht, sowohl einen allgemeinen Funktionsumfang als auch einen für den Automobilbau spezifischen Umfang zu verwirklichen. Somit ist es möglich, auch sehr spezielle Anforderungen in einem System zu bündeln.

# 2 Von der Idee zur Umsetzung

## 2.1 AssistSim

Die erste Idee für die Forschung an einem Assistenzsystem kam Ende des Jahres 2008 auf. Nach einiger Vorbereitungszeit wurde das Forschungsprojekt AssistSim im Juni 2009 bei der Hessen Agentur aufgesetzt. Kooperationspartner waren die Universitäten Frankfurt und Kassel, die Firmen Incontrol und SimPlan, sowie als Industrievertreter die UST Umweltsensortechnik GmbH und die Arbeitsgruppe Ablaufsimulation des Verbands der Automobilindustrie VDA.

Ziel von AssistSim war es, ein System zu schaffen, das den Ablauf einer Simulationsstudie aktiv unterstützt. Hierzu wurde der Prototyp eines Systems geschaffen, das über mehrere Datenbanken, eine Benutzeroberfläche sowie eine Client- Server-Anwendung es ermöglicht, Simulationsstudien und Experimente zu definieren und diese gezielt durchzuführen.

Hierzu wurde das Vorgehen in eine Planungsassistenz und eine Durchführungsassistenz aufgeteilt.

In der Planungsassistenz war es möglich, die Untersuchungsziele zu definieren und zu dokumentieren. Als zweiten Schritt konnte man die Stellgrößen definieren, d. h. die Werte innerhalb des Simulationsmodells, an denen AssistSim Änderungen vornehmen darf. Als letzter Schritt ließen sich die Messgrößen festlegen, also die Ergebniswerte, anhand derer die Simulationsstudie bewertet wird. Somit war eine Aufgabenstellung ausreichend beschrieben.

In der Durchführungsassistenz definierte man anschließend das Experiment und die dazugehörigen Simulationsläufe. Es war sowohl möglich, Experimente frei zu definieren als auch automatisch über vorgegebene Schrittweiten der Stellgrößen.

Waren alle Experimente definiert, konnte man die Experimentdurchführung starten, durch eine Client- Server Anwendung über mehrere Simulationsrechner verteilt. Den aktuellen Stand sowie die Zwischenergebnisse konnte man in der Benutzeroberfläche einsehen (Abb. 2).

Das Projekt wurde im November 2010 abgeschlossen (Lattner 2011) und 2011 mit dem 2. Platz beim hessischen Kooperationspreis ausgezeichnet.

## 2.2 EDASim

Nach Abschluss des AssistSim Projektes war man sich innerhalb des Teams einig, dass die Idee des Assistenzsystems an sich der richtige Weg war, aber sowohl die Implementierung der Software als auch die Evaluierung der Simulationsdaten noch keine ausreichende Qualität hatten. Aus diesem Grund wurde das zweite Forschungsprojekt mit Namen EDASim (Mayer et al. 2012) ins Leben gerufen. Das Team war hier annähernd das, das auch schon im Projekt AssistSim geforscht hatte, nur UST verließ das Projekt. Dafür kamen die Continental AG und die Universität Trier mit ins Boot.

Ziel des Projektes war es, Methoden zur Evaluierung, Validierung und Verifikation (Rabe 2008) von Simulationsdaten zu finden und den Simulationsexperten über eine geeignete Oberfläche zur Verfügung zu stellen. Als Basis wurden hier die Ergebnisse von AssistSim verwendet.

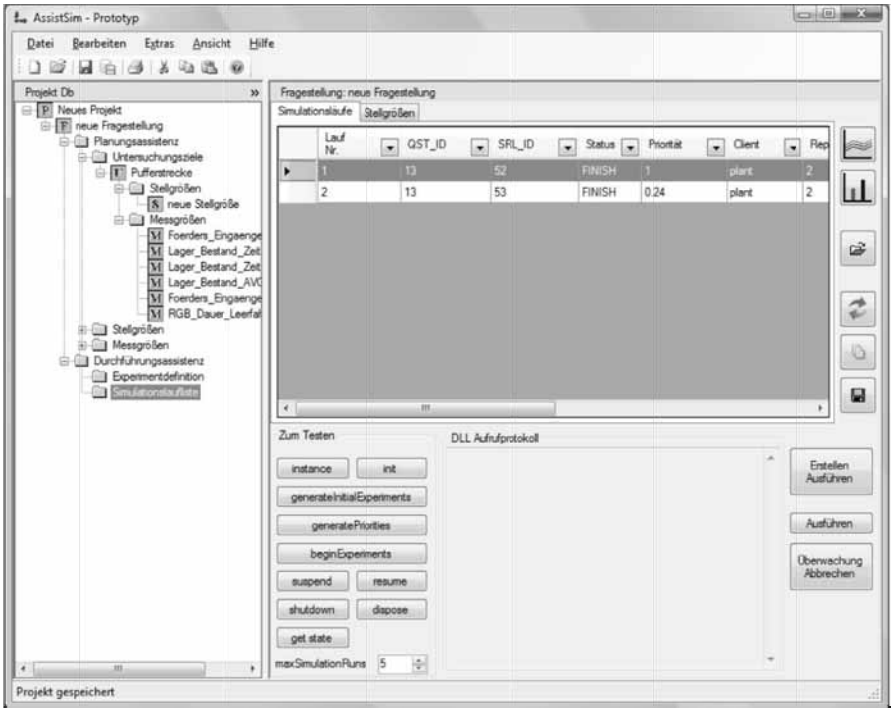



*Abbildung 2: Benutzeroberfläche AssistSim*

Als Grundlage für die Softwareentwicklung wurde auch das Grundkonzept von AssistSim überarbeitet. Die Architektur wurde so umgestellt, dass es ein Hauptprogramm gab, das nur die wichtigsten Funktionen, wie Datenkommunikation oder Ausgebesteuerung beinhaltet. Die eigentlichen Funktionen wurden in sogenannte Plug-Ins gekapselt. Das hat sowohl den Vorteil, dass unterschiedliche Funktionen leichter von einzelnen Partnern entwickelt werden konnten, als auch, dass das Konzept beliebig erweiterbar gemacht wurde.

Das Ergebnis des Forschungsprojektes war eine funktionierende Benutzeroberfläche mit den entwickelten Funktionalitäten aus AssistSim und EDASim (Abb. 3).

Neben Funktionalitäten zur Verknüpfung von beliebigen Datenquellen wurden diverse Statistikfunktionen, wie z. B. eine Kategorisierungsfunktion oder ein Signifikanztest implementiert. Darüber hinaus wurden noch visuelle Auswertemöglichkeiten, wie ein Gantt-Diagramm oder eine Kuchen-Darstellung implementiert.

## 2.3 SimAssist

Da sich die Ergebnisse der beiden Forschungsprojekte sowohl für die Simulationsdienstleister als auch für die Automobilindustrie als zielführend herausgestellt hatten, wurde Ende 2012 beschlossen, die Ergebnisse weiter zu nutzen und über eine Entwicklungskooperation zu einem Produkt zu formen.

Über einen Zeitraum von zwei Jahren wurde basierend auf den Forschungsergebnissen ein Produkt mit Namen SimAssist erstellt. Das Ergebnis wird im nachfolgenden Kapitel näher beleuchtet.

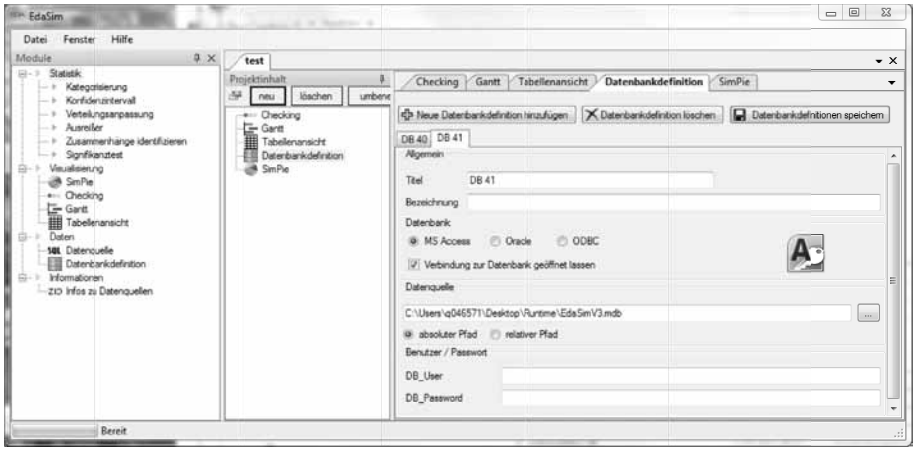



*Abbildung 3: EDASim Oberfläche*

# 3 Aktueller Stand

## 3.1 Allgemeine Funktionalitäten

Als Grundstock für die allgemeinen Funktionen wurde das Ergebnis des EDASim Forschungsprojektes verwendet. Die dort entstandene, Plug-in-fähige Hauptanwendung wurde überarbeitet und an aktuelle Benutzeroberflächen angepasst.

So wurde z. B. das Modul für die Abfragen von Datenbanken über eine flexible Filterfunktion durch eine Definition von Variablen erweitert. Somit ist es möglich, Abfragen-Sets zu erstellen und für unterschiedliche Projekte weiterzuverwenden.

Ein weiterer Punkt ist die Massenverarbeitung und Überprüfung von Daten aus Datenbanken. Das „Checking"-Modul ermöglicht es, beliebig große Datenmengen mit einfachen arithmetischen Funktionen zu überprüfen, wie z. B. „gibt es in den 400 definierten Karosseriebaustationen eine Station, deren Verfügbarkeit kleiner als 90% ist?".

Die statistischen Funktionen aus EDASim wurden ebenso wie die Experimentfunktionen aus AssistSim übernommen.

Die wichtigste Erweiterung war die Portierung der Funktionalitäten des Produktes SimView in SimAssist. Mit dieser Funktionalität ist es möglich, Diagramme automatisch aus Datenquellen zu erzeugen und diese, sowie die Ergebnisse aller anderen Funktionalitäten, in eine Textverarbeitung (Report) oder in Präsentationssoftware zu exportieren.

## 3.2 Erweiterungen der VDA-Arbeitsgruppe Ablaufsimulation

Innerhalb der VDA-Arbeitsgruppe Ablaufsimulation existierten schon vor dem Assistenzsystem gemeinsame Softwareentwicklungen im Umfeld des VDA Automotive Bausteinkastens. Hier sind die Verfügbarkeitsdatenbank (VDB) und die generische Simulationslösung Fördertechnik (GSL-FT) zu nennen. Beide Systeme stellen Schnittstellen in die Ablaufsimulation dar und verfügen über eine Benutzeroberfläche. Als erster Schritt wurden diese Anwendungen in SimAssist portiert.

Als zweiter Schritt wurde eine Datenbank erstellt, die zur Speicherung von Simulationseingangsdaten dient, die sogenannte SIM-DB. Um die Struktur innerhalb dieser SIM-DB bearbeiten und erweitern zu können, wurden ebenfalls Plug-ins geschaffen, die dies ermöglichen.

Des Weiteren wurden Werkzeuge zur Analyse von Zählpunktinformationen geschaffen. Zählpunkte sind Statuspunkte, die sowohl in Simulationsmodellen als auch in der realen Fertigung erzeugt werden können.

Als letzter Schritt wurde die Möglichkeit geschaffen, die Visualisierungsmethodik aus dem Big- Data Umfeld in SimAssist einzubinden. Hierfür wurde die JavaScript Bibliothek D3JS gewählt und in die SimAssist Oberfläche mit eingebunden. D3JS ist eine Sammlung von Visualisierungsdarstellungen, die als freie Software (Open Source) erhältlich ist. Um dies umsetzen zu können, musste als erster Schritt ein Web-Browser in SimAssist eingebunden werden. Hierfür wurde die ebenfalls freie Software Firefox gewählt. Der zweite Schritt war nun, die Daten aus SimAssist an den Browser zu übergeben. Dies geschieht über die Formate CSV (Comma-Separated Values) und über JSON (JavaScript Object Notation). Dadurch ist es nun möglich, nahezu beliebige grafische Auswertungen von Daten in SimAssist zu bewältigen.

## 3.3 SimAssist im Simulations-Vorgehensmodell

Nachdem die Funktionaltäten des Simulationsassistenzsystems definiert waren, war durch die Vielfalt der Funktionalität nicht mehr erkennbar, wie nun genau der Ablauf einer Simulationsstudie verbessert werden sollte. So wurde in der VDA Arbeitsgruppe ein neues Vorgehensmodell entwickelt, das diese Lücke schließt (Abb. 4).

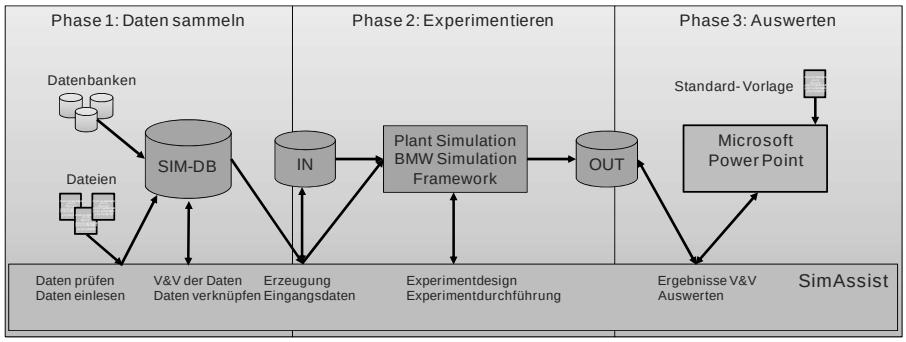



*Abbildung 4: Aktuelles Vorgehensmodell mit SimAssist*

Das erstellte Modell beinhaltet am Anfang den Punkt der Datensammlung. Hier wurde berücksichtigt, dass Daten meist sehr heterogen vorliegen. Durch diverse Imports und Eingaben ist es möglich, alle benötigten Daten in der Simulationsdatenbank zu sammeln.

Sind alle benötigten Daten vorhanden, bietet das Assistenzwerkzeug die Möglichkeit, die Daten zu überprüfen und zu verknüpfen. Hierzu dienen sowohl die statistischen Funktionen, als auch die visuellen Möglichkeiten, die implementiert sind.

Um die Daten reproduzierbar zu halten, ist es nötig, einen Schnappschuss der Daten anzufertigen („IN“ in Abb. 4). Ein weiterer Vorteil dieser Vorgehensweise ist die Unabhängigkeit zur zentralen Simulationsdatenbank.

Mit diesem Datenabzug und geeigneten Bausteinen ist es nun möglich, Teile des Simulationsmodells automatisch aufzubauen und zu parametrieren.

Ist das Modell lauffähig und valide, können im Assistenzsystem Untersuchungen, Experimente und Replikationen geplant und dann verteilt durchgeführt werden. Alle Ergebnisse der Simulatoren werden in einer weiteren Datenbank („OUT“ in Abb. 4) gesammelt.

Als letzter aktiver Schritt erfolgt die Auswertung der Simulationsergebnisse. Dies geschieht wieder durch die Funktionalitäten des Assistenzsystems. Sind die Experimente und Untersuchungen beendet, ist es möglich, die Dokumentation komplett aus den in SimAssist importierten Ergebnissen zu erstellen. Dies funktioniert sowohl als Report in Form eines Textes mit Bildern, als auch als Präsentation. Hierfür lassen sich Vorlagen einbinden, die das jeweilige Firmen-Layout wiedergibt. Die Dokumentation ist updatefähig, d. h. das Assistenzsystem merkt sich die Position der einzelnen Objekte und kann diese bei Datenänderung aktualisieren.

Aus diesem neuen Vorgehen ergibt sich auch eine neue Definition der Dokumentation und der Archivierung der Simulationsuntersuchung. Nimmt man den Datenabzug der Eingangsdaten, die Ergebnisdaten sowie das Modell und das Projekt des Assistenzsystems, sind alle Daten durchgängig archivierbar. Ergänzt man das Ganze noch um die erstellte Dokumentation, so ist das Simulationsprojekt komplett reproduzierbar sicherbar (Abb. 5).

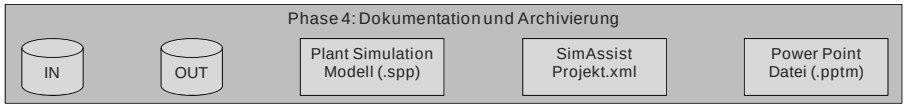



*Abbildung 5: Dokumentation und Archivierung*

## 4 Zusammenfassung und Ausblick

Assistenzsysteme bieten die Möglichkeit, die Durchführung einer Simulationsstudie zu unterstützen und zu verkürzen. Das Vorgehen mit SimAssist zeigt, dass hier viele Möglichkeiten gegeben sind, Simulationsprojekte aktiv zu unterstützen und zeitlich bzw. qualitativ verbessern.

Die flexible Erweiterbarkeit des Systems bietet alle Optionen, neue Erkenntnisse wieder in die Umgebung einfließen zu lassen.

Die VDA Arbeitsgruppe Ablaufsimulation hat mit den Bestrebungen, Open Source Produkte, wie D3JS als Visualisierungsplattform, nutzen zu lassen, gezeigt, dass Weiterentwicklungen kaum Grenzen gesetzt sind. Dieser Ansatz der Weiterverwendung wird das System SimAssist auch in Zukunft Wettbewerbsfähig halten.

## Literatur

Lattner, A.D.; Pitsch, H.; Timm, I.J.; Spieckermann, S.; Wenzel, S.: AssistSim – Towards automation of simulation studies in logistics. Simulation Notes Europe 21 (2011) 3-4, S. 119-128.

Mayer, G.; Pöge, C.: Auf dem Weg zum Standard – Von der Idee zur Umsetzung des VDA Automotive Bausteinkastens. In: Zülch, G.; Stock, P. (Hrsg.): Integrationsaspekte der Simulation: Technik, Organisation, Personal. Karlsruhe: KIT Scientific Publishing 2010, S. 29-36.

Mayer G., Spieckermann S., Wenzel S.: Steigerung der Produktivität in Simulationsstudien mit Assistenzwerkzeugen. ZWF 107 (2012) 3, S. 174-177.

Rabe M., Spieckermann S., Wenzel S.: Verifikation und Validierung für die Simulation in Produktion und Logistik. Berlin: Springer 2008.

VDA: VDA-Empfehlung 4811 „Qualitätskriterien für Simulationsstudien der Ablaufsimulation“. Berlin: Verband der Automobilindustrie 2013.

VDI: Richtlinie 3633 Blatt 1 „Simulation von Logistik-, Materialfluss- und Produktionssystemen“. Berlin: Beuth 2014.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Aufbereitung von unvollständigen Daten für die simulationsbasierte Produktionsplanung im Schiffbau

## *Processing Incomplete Data for Simulation-based Production Planning in Shipbuilding*

Dirk Steinhauer, Flensburger Schiffbau-Gesellschaft mbH & Co. KG, Flensburg (Germany), steinhauer@fsg-ship.de

Hermann Lödding, Axel Friedewald, Jan Niklas Sikorra, Moritz Alexander Haux, Technische Universität Hamburg-Harburg, Hamburg (Germany), loedding@tu-harburg.de, friedewald@tu-harburg.de, sikorra@tuhh.de, moritz.haux@tuhh.de

Abstract: This paper presents an approach on the early use of simulation-based production planning in shipbuilding. In the early stages of the planning process in shipbuilding, the construction of the ship is not yet complete, while an early scheduling of the production already has to be done. Due to the lack of the construction the planning lacks input data, which leads to an uncertainty in the production planning. The uncertainty reduces the accuracy of planning and therefore has a significant impact on its quality and robustness. The planner requires detailed information about this uncertainty and construction details. Therefore a concept to describe and to model the uncertain data as well as a tool to generate construction details from historical data has been created.

## 1 Einleitung

Um die Vorteile der simulationsbasierten Produktionsplanung in der schiffbaulichen Unikatproduktion maximal ausschöpfen zu können, werden detaillierte Planungsinformationen benötigt. Zur Analyse der Abhängigkeiten im komplexen Wirkungsgefüge einer Werft verarbeiten die Simulationsmodelle Eingangsdaten zum Produkt im aktuellen Konstruktionsstand, zu Produktionsprozessen und ihren Bedingungen, zu logistischen Randbedingungen und verfügbaren Ressourcen. Diese Daten liegen in der frühen Planungsphase allerdings nur in einem geringen Umfang vor. Daher kann die simulationsbasierte Produktionsplanung derzeit oft nur punktuell und eher spät im Planungsprozess eingesetzt werden. Es besteht jedoch schon in der Angebotsphase oder direkt nach Auftragseingang die Notwendigkeit, eine erste Produktionsplanung durchzuführen, um die vereinbarte Lieferzeit zu prüfen bzw. mögliche Engpässe und Probleme zu identifizieren. Termintreue und geringe Durchlaufzeit

sind zentrale Zielgrößen der Werften, da Lieferverzug zu empfindlichen Vertragsstrafen führt und die Durchlaufzeit sehr mit den Kosten korreliert. In der frühen Planungsphase ist die Konstruktion jedoch noch nicht abgeschlossen bzw. hat gerade erst begonnen, und es liegen daher nur unvollständige bzw. unscharfe Planungsgrundlagen vor. Diese Unschärfen sind ein charakteristisches Merkmal der Produktionsplanung in der Unikatfertigung (Gruß 2010). Sie reduzieren jedoch die Planungsgenauigkeit und haben somit einen signifikanten Einfluss auf die Qualität und Robustheit der Produktionsplanung. Die Bestimmung von Produktionsdauern und Belegungsplanungen ist für die Werften jedoch essentiell, um die vorhandenen Ressourcen optimal ausnutzen zu können.

Für die Planer ist es daher sehr wichtig, die in den unterschiedlichen Planungsphasen auftretenden Unschärfen bewerten zu können, um geeignete Detaillierungen zu veranlassen oder ausreichende Puffer bei der Planung vorzuhalten. Zur Analyse der vorhandenen Planungsdaten bzw. zur Umsetzung einer Produktionsplanung in frühen Phasen können simulationsbasierte Konzepte sehr gut verwendet werden. Jedoch müssen dazu zum einen die Simulationsmodelle mit unscharfen Planungsdaten umgehen können, und zum anderen müssen Konzepte vorhanden sein, um die vorhandene Datenqualität bewerten und verbessern zu können.

Eine weitere Herausforderung bei der Verwendung von Planungsdaten in frühen Planungsphasen ist, dass viele Informationen eher unstrukturiert vorliegen, wenig auswertbar und über mehrere IT-Systeme verteilt sind. Es ist eine Vielzahl von unterschiedlichen IT-Systemen im Einsatz. Dies führt häufig zu Inkonsistenzen und unvollständigen Informationen, die häufig manuell behoben werden müssen. Der Einsatz von geschätzten Erfahrungswerten füllt diese Lücke, liefert aber keine Betrachtung inwiefern diese Werte das Ergebnis der Planung beeinflussen.

An den beschriebenen Herausforderungen wird seit Oktober 2013 im Forschungsvorhaben HEPP („Hocheffiziente Produktionsplanung für Prototypenkompetenz“) gearbeitet. Das vom BMWi unter dem Förderkennzeichen 03SX365 geförderte Vorhaben hat zum Ziel, Konzepte zur Produktionsplanung von komplexen Unikaten zu entwickeln. Partner in HEPP sind neben der Flensburger Schiffbau-Gesellschaft mbH & Co. KG (FSG) und der Technischen Universität Hamburg-Harburg (TUHH) die Meyer Werft in Papenburg, die SimPlan AG, die Ruhr-Universität Bochum mit dem Lehrstuhl für Bauinformatik sowie das Center of Maritime Technologies e.V. (CMT).

Im Rahmen des Beitrags werden Ansätze zur Modellierung, Klassifizierung und Zusammenführung von unscharfen Planungsdaten für die Produktionsplanung im Schiffbau vorgestellt. Die Klassifizierung der Unschärfen erfolgt auf Basis verschiedener Merkmale (z.B. Aktualität, historische Daten, Erfahrungswerte, Planungsphase, etc.). Auf Basis einer Ontologie werden die einzelnen Planungsdaten systematisch verknüpft. Durch die Zuordnung von Unschärfen zu einzelnen Planungsdaten und die Verwendung die Ontologiebeziehungen können Auswirkungen von Unschärfen auf einzelne Planungsdaten analysiert werden. Das so entstandene Wirkmodell schafft Transparenz über die Unschärfen im Planungsprozess und befähigt den Anwender, gezielt Unsicherheiten zu begegnen.

## 2 Beschreibung und Modellierung unscharfer Daten

Unschärfen sind ein charakteristisches Merkmal der Produktionsplanung in der Unikatfertigung (Gruß 2010). Sie reduzieren die Planungsgenauigkeit und beeinflussen die Qualität und Robustheit der Produktionsplanung signifikant (Werners und Thorn 2002). In der frühen Planung sind viele Produkt- und Prozessdaten mit Unschärfen behaftet. Erst detaillierte Informationen über die Unschärfen befähigen den Planer, diesen gezielt zu begegnen. Deshalb sind folgende Informationen von besonderer Relevanz:

- Welche Planungsdaten sind unscharf?
- Wie stark sind diese Unschärfen ausgeprägt?
- Was sind die Ursachen für die Unschärfen?
- Welche Auswirkungen haben diese auf die Zielgrößen?

Das vorgestellte Konzept zur Beschreibung und Modellierung unscharfer Daten unterstützt den Planer dabei, diese Informationen zu erlangen. Es gliedert sich in drei Schritte: Unscharfe Planungsdaten identifizieren, Unschärfen quantifizieren und Auswirkungen der Unschärfen modellieren (Abb.1). Das Vorgehen wird nachfolgend beschrieben und exemplarisch anhand des Planungsdatums der *Auftragszeit für die Montage eines Blocks* veranschaulicht:

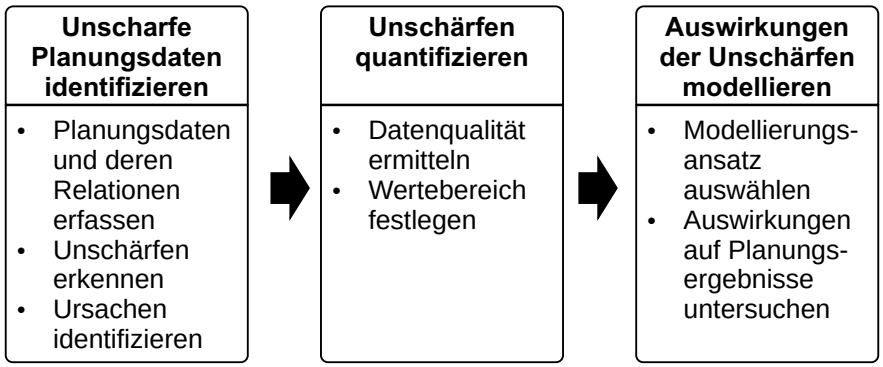



*Abbildung 1: Vorgehen zur Modellierung unscharfer Daten*

*1. Schritt: Unscharfe Planungsdaten identifizieren:* Zunächst werden alle für die Planung relevanten Daten und die Relationen zwischen den einzelnen Daten erfasst und in einer Ontologie festgehalten. Anschließend werden mit Hilfe von Experteninterviews und Vergangenheitsdaten unscharfe Planungsdaten identifiziert. Die unscharfen Planungsdaten werden in die Kategorien Produkt, Prozess, Ressource unterteilt. Ein unscharfes Planungsdatum (Prozess) ist die *Auftragszeit für die Montage eines Blocks*, diese wird u.a. beeinflusst von der Orientierung und Zugänglichkeit der Bauteile (Produkt) und der Schweißgeschwindigkeit (Prozess). Sind die Bauteile beispielsweise nur schwer zugänglich, erhöht sich die Auftragszeit signifikant. Zum Zeitpunkt der Grobterminierung liegen diese Daten aus der Arbeitsvorbereitung nicht oder nur unvollständig vor (Vgl. Kuhlmann 1994). Darauf aufbauend kann ein Wirkmodell entwickelt werden, das die Unschärfen und deren Wechselwirkungen im Planungsprozess abbildet. Der Anwender kann sowohl die

Ursachen der Unschärfen identifizieren als auch die Auswirkungen von unscharfen Planungsdaten auf andere Planungsdaten qualitativ analysieren. Zudem kann ergeeignete Simulationsexperimente ableiten.

*2. Schritt: Unschärfen quantifizieren:* Die Unschärfen in den Planungsdaten sind unterschiedlich stark ausgeprägt. Dieser Umstand spiegelt sich in der Größe des möglichen Wertebereichs wider. Von Interesse für den Planer ist der Wertebereich der unscharfen Daten. Die Qualität der Daten beeinflusst die Ausprägung der Unschärfe. Die Datenqualität erfasst, inwieweit die Planungsdaten die Anforderungen der Planung hinsichtlich ausgewählter Merkmale erfüllen (Vgl. Hinrichs 2002). Relevante Merkmale zur Bewertung der Datenqualität sind die Aktualität, die Glaubwürdigkeit und die Objektivität der Planungsdaten. Der Planer bestimmt den Wertebereich, indem er die Qualität der Aussage linguistisch beschreibt und diese mittels einer Fuzzy-Modellierung in eine Kennzahl der Datenqualität überführt. Der Wertebereich wird als Funktion der Datenqualität und des Planwertes berechnet. Eine mögliche Funktion ist in Gleichung 1 gegeben:

$$Wertebereich = Planwert \pm (1 - Datenqualität) \times Planwert \qquad (1)$$

Am Beispiel der *Auftragszeit für die Montage eines Blocks* wird dieser Schritt verdeutlicht (Tab. 1). Der Planer gibt eine Auftragszeit von 5 Stunden an, bewertet die Merkmale und berechnet eine Datenqualität von 0,8. Die tatsächliche Auftragszeit wird zwischen 4 und 6 Stunden liegen.

*Tabelle 1: Beispiel Auswirkung der Datenqualität auf den Wertebereich*

| Auftragszeit | Aktualität | Glaubwürdigkeit | Objektivität | Datenqualität | Wertebereich |
|---|---|---|---|---|---|
| 5 Stunden | Hoch | Mittel | Hoch | 0,8 | 4-6 Stunden |
| 5 Stunden | Hoch | Gering | Hoch | 0,6 | 3-7 Stunden |
| 5 Stunden | Hoch | Hoch | Hoch | 1 | 5 Stunden |

*3. Schritt: Auswirkungen der Unschärfen modellieren:* Wenn der Planer eine Auftragszeit mit 5 Std angibt, nimmt er an, dass die tatsächliche Auftragszeit bspw. zwischen 4 und 6 Std liegen wird. Jede Abweichung vom Planwert wirkt sich unmittelbar auf die Kosten aus. Neben dieser einfachen Wirkbeziehung resultieren aus dem komplexen Produktionsprozess weitere Wirkbeziehungen, die nur mittels der Simulation abgebildet werden können. Dafür sind die unscharfen Planungsdaten in das Simulationsmodell einzubinden. Für die Integration der unscharfen Planungsdaten in die Simulation wurden zwei Modellierungsansätze ausgewählt: Verteilungsfunktionen und Fuzzy-Zahlen. Abebe et al. (2000) haben diese in einem anderen Kontext bereits angewendet. Für die simulationsgestützte Planung soll untersucht werden, welcher Ansatz besser für die Modellierung der Unschärfen geeignet ist. Bei beiden Ansätzen ist die Datenqualität ein Eingangsparameter. Bei dem ersten Modellierungsansatz werden die unscharfen Daten durch Verteilungsfunktionen abgebildet. Mit Hilfe der Monte-Carlo-Technik wird in verschiedenen Simulationsläufen untersucht, wie sich die Unschärfen auf die Zielgrößen auswirken. Bei dem

zweiten Modellierungsansatz werden die unscharfen Planungsdaten durch Fuzzy-Zahlen (Zugehörigkeitsfunktionen) dargestellt. Hier wird die alpha-cut-Methode eingesetzt, um die Auswirkungen der Unschärfe auf die Zielgrößen zu untersuchen. Die Evaluation des Konzeptes ist Inhalt zukünftiger Arbeiten. Mit dem vorgestellten Konzept kann der Planer die Auswirkungen der Unschärfen ermitteln und daraus Maßnahmen ableiten. Beispielsweise kann er Puffer einplanen oder durch Datengeneratoren gezielt die Datenqualität verbessern. Ein Konzept zur Datengenerierung wird im folgenden Abschnitt vorgestellt.

## 3 Erzeugung von Konstruktionsdaten zur Reduzierung planerischer Unschärfen

Bereits in der frühen Projektphase nach dem Vertragsabschluss muss die Produktionsplanung im Schiffbau den Durchlauf des Schiffes durch die Werft bzw. die zu vergebenen Arbeiten definieren. Abgangstermine und Durchführungszeit für die einzelnen Arbeiten sind zu diesem frühen Zeitpunkt Schätzwerte. Auf Basis der geschätzten Abgangstermine und einer geschätzten Durchführungszeit für die Blöcke auf den Bauplätzen findet bereits zu einem frühen Planungszeitpunkt eine Festlegung der Bauplätze statt, obwohl noch keine detaillierten Konstruktionsdaten vorliegen (vgl. Storch 1995). Diese erste Planung ist aufgrund der fehlenden Konstruktionsdaten mit großen Unschärfen behaftet und die Ursache der Schätzfehler ist im Nachhinein nur schwer nachvollziehbar.

Die Reduzierung identifizierter Unsicherheiten erhöht die Planungssicherheit. Das Ziel, die Reduzierung der Schätzfehler, kann durch den Einsatz der Simulation oder neuer Schätzverfahren erfolgen. Im Projekt HEPP ist ein Werkzeug konzipiert worden, welches auf Basis von Vergangenheitsdaten die Berechnung von Konstruktionsdetails, also von Teilen unter Berücksichtigung spezifischer Metadaten ermöglicht. Dieser basiert auf einem im Projekt GeneSim entwickelten Datengenerator zur Erzeugung stahlbaulicher Teile (Profile und Platten) (vgl. Steinhauer et al. 2011). Der GeneSim-Generator ermöglichte bereits die Bestimmung der in einfachen Bauwerken vorhandenen Teile auf Stücklistenebene. Das in HEPP konzipierte Werkzeug ermöglicht die Auswahl ähnlicher Strukturen auf Basis einer Vergleichsdatenbank. Nach der Auswahl eines ähnlichen Moduls parametriert der Planer eine Vorlage (Template). Er legt dazu bauwerksspezifische Parameter wie Länge, Breite und Höhe fest. Außerdem kann er schiffspezifische Parameter wie Blechdicken und Profilabstände vorgegeben.

Das in HEPP konzipierte Werkzeug zur Datenaufbereitung erzeugt auf Basis der gewählten Vorlage, die aus Vergangenheitsdaten erstellt wurde, unter Berücksichtigung der eingegebenen Parameter die Konstruktionsdetails – stahlbauliche Konstrukte auf Blockebene. Auf Basis der übergebenen Parameter wird so eine Konstruktion des komplexen Bauwerks Block erzeugt. Jedem Bauteil dieser stahlbaulichen Konstruktion sind Metadaten wie Länge, Breite und Höhe zugeordnet. Jedes Teil besitzt weiterhin eine Position. Aus der Position der Teile und deren Metainformationen lassen sich deren Relationen berechnen. Aus der Teilegröße und der Relation der Teile zueinander können die Schweißstöße sowie deren Länge bestimmt werden. Die Verknüpfung der Schweißstöße mit den bekannten Baumethoden ermöglicht die Berechnung von Fertigungsaufwänden im Blockbau. Durch die Zuordnung des Aufwands zu Ressourcen kann durch die in das

Simulation Toolkit Shipbuilding integrierte automatische Parameterschätzung eine Durchführungszeit ermittelt werden. Die im Rahmen des Schätzverfahrens ermittelte Durchführungszeit erhöht die Transparenz der Planung da ihre Berechnung nachvollziehbar ist. Bei Abweichungen zwischen den geschätzten Durchführungszeiten eines Bauprojekts und den später in der Fertigung aufgetretenen Durchführungszeiten sind die Ursachen der Abweichungen identifizierbar.

Um ein genaueres Abbild der Prozesse in der Fertigung in die Berechnung der Durchführungszeiten einzubeziehen, wird die Simulation in der frühen Phase der Produktionsplanung eingesetzt. Durch die Verknüpfung der geschätzten Daten mit den Ressourcen liefert die Simulation für die spezifischen Blöcke Durchführungszeiten auf den Bauplätzen des Blockbaus. Diese Berechnungen sind aufgrund der an das Unikat lediglich angenäherten Vorlagen ungenau. Sie spiegeln ein bereits gebautes Produkt wider. Bei dem Schiff, auf das die Templates angewendet werden, handelt es sich jedoch um ein Unikat. Somit muss der Planer berücksichtigen, dass auch die berechneten, vermeintlich exakten Werte weiterhin auf Unschärfen basieren.

Die Reduzierung von Schätzfehlern auf der Prozessseite kann durch den Einsatz der Simulation erfolgen, da im Rahmen der Simulationsmodelle bereits umfangreiche Prozesstabellen angelegt wurden. Die Simulation benötigt jedoch detaillierte Bauteil- und Baufolgedaten, um ihre volle Leistungsfähigkeit einbringen zu können. Der Einsatz der Simulation in der frühen Phase der Produktionsplanung bedarf daher einer Stückliste. Eine geschätzte Stückliste in ausreichender Genauigkeit kann daher helfen die Genauigkeit der Produktionsplanung zu erhöhen. Diesen Vorgang unterstützt der Einsatz des Datengenerators. Der Planer übergibt dem Datengenerator bauwerksspezifische Parameter und erzeugt eine Stahlstruktur, die er einem generischen Simulationsmodell zur Bestimmung der Durchführungszeit auf einem generischen Bauplatz übergibt. Die Simulationsergebnisse können im Nachhinein mit den realen Fertigungszeiten abgeglichen werden. Eine transparente Überprüfung der Simulationsergebnisse ist somit möglich.

Für die Simulation der Montageabläufe in den unterschiedlichen Stufen der schiffbaulichen Produktion werden die Simulationsbausteine des Simulation Toolkit Shipbuilding STS (Steinhauer 2011) sowie die damit einhergehenden Methoden zur Prozessdefinition verwendet. Über eine generische Vorgehensweise können die Prozesse so mit geringem Aufwand in verschiedenen Detaillierungsgraden definiert und verknüpft werden.

Für die generische Simulation wird ein Plant-Simulation-Modell mit Bauplätzen für unterschiedlich detaillierte Eingangsdaten verwendet. Daten können dem Modell auf Teile-, Sektions- oder Blockebene übergeben werden. Zielbauwerk des Modells können das Schiff, der Block oder die Sektion sein (Abbildung 2). Es kann durchgängig über verschiedene Detaillierungsgrade zum Beispiel vom Einzelteil zum Schiff oder nur auf wenigen Detaillierungsgraden, zum Beispiel von der Sektion zum Block erfolgen. Für die Aufwandsschätzung in der Blockmontage ist die Simulation vom Detaillierungsgrad Sektion zum Detaillierungsgrad Block interessant. Die Simulation liefert die Durchführungszeit für die Blockmontage auf dem Bauplatz nach einem transparenten Verfahren. Schätzfehler können so im Nachhinein bewertet und für zukünftige Schätzungen angepasst werden.

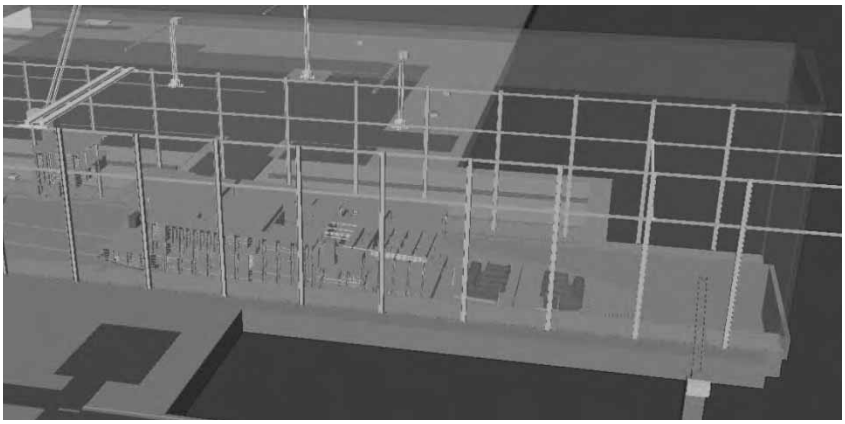


*Abbildung 2: Simulationsmodell der Schiffsmontage bei der FSG*

## 4 Ausblick

Im Rahmen des Forschungsvorhabens HEPP wird für die Zusammenführung, Bewertung und Generierung von Planungsdaten eine Plattform entwickelt, in welcher die konzipierten Werkzeuge vernetzt integriert sind. Dort können dann verschiedene Planungsdaten über die in HEPP erarbeitete Ontologie verknüpft und bzgl. der Unschärfen bewertet werden. Auch die Funktionalitäten zur Generierung von Details für die simulationsgestützte Belegungsplanung sind integraler Bestandteil der Webplattform. Die so zusammengeführten Daten können für die Simulationsstudien exportiert und die Ergebnisse wiederum importiert und visualisiert werden.

Die in HEPP entwickelten Vorgehensweisen werden derzeit bei den beteiligten Werften an Projekten sowohl aus der Vergangenheit als auch aus der Zukunft validiert.

## Literaturverzeichnis

Abebe, A. J.; Guinot, V.; Solomatine, D. P.: Fuzzy alpha-cut vs. Monte Carlo techniques in assessing uncertainty in model parameters. In: Proceedings 4-th International Conference on Hydroinformatics, Iowa City (USA), July 2000.

Gruß, R.: Schlanke Unikatfertigung - Zweistufiges Taktphasenmodell zur Steigerung der Prozesseffizienz in der Unikatfertigung auf Basis der Lean Production. Wiesbaden: Gabler Verlag 2010.

Hinrichs, H.: Datenqualitätsmanagement in Data Warehouse-Systemen. Dissertation Universität Oldenburg, 2002.

Kuhlmann, T.: Konzeption und Entwicklung eines Systems zur Koordinierung der Produktion komplexer Unikate. Dissertation Universität Bremen, 1994.

Steinhauer, D.; Hübler, M; Wagner, L.; Esins, E.; Pester, P.; Friedewald, A.; Wandt, R.: GeneSim - Generisches Daten- und Modellmanagement für die schiffbauliche Produktionssimulation. In: Forschungszentrum Jülich (Hrsg.): Statustagung Schifffahrt und Meerestechnik, Rostock, 2011, S. 7-23.

Steinhauer, D.: The Simulation Toolkit Shipbuilding – 10 Years of Cooperative Development and Interbranch Applications. In: Proceedings of the 10th Inter-

national Conference on Computer and IT Applications in the Maritime Industries (COMPIT), Berlin (Germany), 2011, S. 453-465.

Storch, Richard Lee: Ship Production. Centerville, Maryland: Cornell Maritime Press, 1995.

Werners, B.; Thorn, J.: Supply Chain Planning bei Nachfrageunsicherheit. In: PPS Management 7 (2002) 2, S. 51–54.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Picker Blocking in manuellen Kommissioniersystemen – Eine simulationsbasierte Analyse

## *A Simulation-based Analysis of the Effects of Picker Blocking in Manual Order Picking Systems*

Ralf Elbert, Torsten Franzke, Christoph H. Glock, Eric H. Grosse, TU Darmstadt, Darmstadt (Germany), elbert@log.tu-darmstadt.de, franzke@log.tu-darmstadt.de, glock@pscm.tu-darmstadt.de, grosse@pscm.tu-darmstadt.de

Abstract: This Paper analyses the effects of Picker Blocking (congestion) in a manual rectangular warehouse layout. Picker Blocking occurs if passing in an aisle is not possible or if several order pickers try to pick from the same shelf. Based on an agent-based simulation model developed in AnyLogic, the effects of Picker Blocking under different routing policies (return, s-shape, midpoint, largest gap, composite, and combined) and storage assignment rules (turnover based) are quantified. The shortest mean throughput times resulted for the composite strategy (up to 10 agents) and a vertical allocation of products. The midpoint strategy leads to the shortest mean throughput times for 15 order picker (horizontal allocation of products). The lowest mean numbers of blockings for each simulation run results for the s-shape policy and a horizontal allocation of products.

## 1 Einführung

Unter Kommissionierung wird der Prozess der auftragsgerechten Zusammenstellung von Artikeln oder Waren aus einem Lagerbereich verstanden (Arnold und Furmans 2005). Schätzungen zufolge wird diese in über 80% der Läger noch manuell, gemäß dem Prinzip Mann-zur-Ware, durchgeführt (de Koster et al. 2007; Napolitano 2012). Dies betrifft statische Lagerplätze, die der Kommissionierer entsprechend einer bestimmten Wegeführung abläuft und dort die für den Auftrag benötigten Artikel entnimmt.

Die manuelle Kommissionierung zählt üblicherweise zu den arbeits- und zeitintensivsten Tätigkeiten in der innerbetrieblichen Logistik (Frazelle 2000). Sie ist oftmals für über 50% der operativen Kosten von Warenlagern verantwortlich und dadurch häufig das Ziel von Outsourcing-Vorhaben (Tompkins et al. 2010). Trotz einer Vielzahl an vorhandenen technischen Hilfsmitteln (z. B. Förderfahrzeuge, Barcodescanner, Pick-by-light-Systeme) ist der Mensch in der Kommissionierung, mit sei-

nen individuellen kognitiven und motorischen Fähigkeiten und der daraus resultierenden hohen Flexibilität, nicht wirtschaftlich durch automatisierte Systeme ersetzbar (Arnold und Furmans 2005; Grosse et al. 2015). Darüber hinaus sind automatisierte Systeme meist mit hohen Investitionskosten verbunden und weisen eine geringere Flexibilität auf.

Zur Erreichung effizienter und stabiler Prozesse in der Kommissionierung existieren verschiedene Planungsansätze. Beim Vergleich der Zeitanteile, der für die Kommissionierung benötigten Tätigkeiten (Greifzeit, Basiszeit, Totzeit, Wegzeit), nimmt die Wegzeit mit ca. 55% den größten Anteil ein. Daher ist das Ziel der meisten Planungsansätze, die Wegzeit zu verkürzen (de Koster et al. 2007). Üblicherweise wird versucht, dieses Ziel insbesondere über die Lagerlayoutgestaltung, die Wegeführung durch das Lager (Routenpolitiken), die Lagerplatzvergabe sowie die Auftragsaggregation zu erreichen. Letzteres bezeichnet das Zusammenlegen oder Aufteilen von mehreren Kundenaufträgen zu Kommissionieraufträgen (Grosse et al. 2014). Bei der Layoutgestaltung wird die Anordnung, Größe und Aufteilung der Bereiche, Gänge und Regale im Lager bestimmt (Roodbergen und Vis 2006). Die Routenpolitiken geben dem Kommissionierer den Weg durch das Lager vor (Petersen und Aase 2004). Bei der Lagerplatzvergabe wird entschieden, wie die Produkte den Lagerfächern zugeordnet werden sollen (Glock und Grosse 2012).

Gerade im Bereich der manuellen Kommissionierung existiert eine Vielzahl an Arbeiten, die sich mit den genannten Planungsansätzen für eine effiziente Gestaltung des Kommissionierprozesses beschäftigen (de Koster et al. 2007). Die meisten Arbeiten nehmen jedoch an, dass lediglich ein Kommissionierer im Lagerbereich arbeitet und vernachlässigen somit die Auswirkungen, die durch die simultane Arbeit von mehreren Mitarbeitern im selben Kommissionierlager entstehen. Es ist leicht ersichtlich, dass in der Praxis hierbei Blockierungen entstehen können, wenn mehrere Kommissionierer im gleichen Bereich Artikel einsammeln (Pan & Wu 2012) oder wenn sie sich innerhalb der Lagergänge begegnen und es aufgrund der geringen Lagerbreite keine Möglichkeit gibt, aneinander vorbeizukommen (dies wird dann erschwert, wenn Handwägen als Hilfsmittel eingesetzt werden). Derartige Blockierungen werden in der Literatur mit dem Begriff „Picker Blocking“ bezeichnet (z. B. in Mowrey und Parikh 2014) und können zu zusätzlichen Wartezeiten, längeren Wegzeiten oder -strecken und zu einer verringerten Effizienz der Prozesse führen (Chen et al. 2013).

Daher befassen sich immer mehr Arbeiten mit der Thematik des Picker Blocking, um die dadurch entstehenden negativen Effekte zu verringern und die Effizienz der Kommissionierung zu erhöhen. So wurde bereits erarbeitet, wie sich der Unterschied zwischen breiten und schmalen Gängen (Parikh und Meller 2009), unterschiedliche Pickdichten (d. h. die Anzahl der Picks pro Gang im Lager) (Gue et al. 2006), mehrdimensionale Lager (Parikh und Meller 2010a), Batching-Heuristiken (Hong et al. 2012a; 2012b) oder verschiedene Routenpolitiken (Chen et al. 2013) auf das Picker Blocking auswirken.

Bei näherer Betrachtung der Literatur wird deutlich, dass die meisten Studien im Bereich des Picker Blocking jedoch sehr restriktive Annahmen in Bezug auf Lagerplatzvergabe und Routenführung treffen. So wird größtenteils die zufällige Lagerplatzvergabe angenommen, bei der Artikel zufällig auf freie Lagerplätze zugeordnet werden (z. B. Parikh und Meller 2010a; 2010b; Hong et al. 2010). Effekte, die durch

eine nachfrageorientierte Lagerplatzvergabe (z. B. ABC-Lagerplatzvergabe) entstehen, werden hingegen nur selten untersucht (z. B. Ruben und Jacobs 1999; Pan und Wu 2012). Dies trifft auch auf die Routenpolitiken zu, bei denen in der überwiegenden Anzahl an Arbeiten die praxisrelevante S-shape-Strategie (d. h. jeder Gang, der einen Bedarf enthält, wird vom Kommissionierer komplett durchlaufen) angewendet wird. Die Auswirkungen des Picker Blocking auf die Kommissionierzeit unter Berücksichtig der Interdependenzen unterschiedlicher Routenpolitiken (wie z. B. S-shape, Return, Midpoint, Largest gap, Composite, Combined) und unterschiedlicher Lagerplatzvergabestrategien (wie z. B. ABC-vertikal, -horizontal, -diagonal) wurden demnach noch nicht untersucht.

Das Ziel dieses Beitrags ist zu untersuchen, wie sich die Effekte des Picker Blocking in einem manuellen Kommissionierlager unter Berücksichtigung von unterschiedlichen Routenpolitiken sowie verschiedenen Lagerplatzvergabestrategien auf die Effizienz des Kommissioniervorgangs auswirken. Zu diesem Zweck wurde ein agentenbasiertes Simulationsmodell entwickelt, welches im folgenden Kapitel näher erläutert wird. Schließlich werden die Ergebnisse einer umfangreichen Simulationsstudie besprochen und zukünftige Forschungsmöglichkeiten sowie Implikationen für die Unternehmenspraxis aufgezeigt.

## 2 Simulationsstudie

Im Bereich der Simulationsmodelle werden in der Literatur verschiedene Kategorisierungen vorgenommen. Einerseits lassen sich Simulationsmodelle nach diskreten oder kontinuierlichen Veränderungen über die Zeit unterscheiden sowie nach den Ansätzen systemdynamisch, ereignisdiskret und agentenbasiert (Borshchev und Filippov 2004). Vor allem der agentenbasierte Ansatz zeigt sich als sehr gut geeignet, um individuelles Verhalten von Menschen (Agenten) zu simulieren (Borshchev 2013) und soll im Folgenden daher näher betrachtet werden.

### 2.1 Agentenbasierte Simulation

Agentenbasierte Simulationsmodelle bestehen im Wesentlichen aus drei Bestandteilen. Sie haben einerseits ein oder mehrere Agenten, welche in diesem Beitrag die Kommissionierer darstellen. Ferner besitzen die Agenten Beziehungen untereinander (diese sind hier die Möglichkeit, sich gegenseitig zu blockieren) und die Agenten befinden sich in einer Umwelt, mit der sie interagieren können. Letzteres wird durch das Lagerlayout definiert. Bisher gibt es noch keine einheitliche Definition der Agenten (Heath et al. 2013). Meist werden verschiedene Eigenschaften angegeben, die sie auszeichnen. Dazu zählen insbesondere die Fähigkeit, unabhängig handeln zu können, die Möglichkeit, zu interagieren sowie dabei autonom zu entscheiden. Ferner können Agenten ihr Verhalten an bestimmte Situationen anpassen (Klügl 2001; Macal und North 2009).

Bei der agentenbasierten Simulation wird ein bottom-up Ansatz verfolgt. Das bedeutet, dass individuelles Verhalten einzelner Agenten im Modell implementiert wird und dann durch Analysen des Outputs verfolgt werden kann, wie sich dies auf das Gesamtsystem auswirkt. Dadurch können emergente Phänomene, die durch das Zusammenwirken der einzelnen Einheiten des Systems erst deutlich werden, beobachtet und analysiert werden. Ferner kann so eine hohe Komplexität, wie sie bei

der Implementierung von menschlichem Verhalten vorhanden ist, abgebildet werden (Macal 2010). Als eine der ersten Arbeiten, die die agentenbasierte Simulation im Bereich des Picker Blocking nutzen, zählt die Arbeit von Heath et al. (2013), die deutlich aufzeigt, dass gerade im Bereich der Kommissionierung das Verhalten der Mitarbeiter durch den Simulationsansatz dargestellt werden kann und Auswirkungen untersucht werden können. Die Autoren legen ihren Fokus auf die Untersuchung von Auswirkungen menschlichen Verhaltens (z. B. proaktive und eigenständige Vermeidung von Blockierungen durch die Kommissionierer) auf die Kommissionierleistung. Auch analysieren Furmans et al. (2012) mit der Bedientheorie und einem agentenbasierten Simulationsmodell die Auswirkungen von Blockierungen auf die Produktivität in der manuellen Kommissionierung. Beide Arbeiten nutzen lediglich die S-shape-, Return- oder eigens definierte Routenpolitiken (bspw. zufälliges Einsammeln aller Artikel oder zufälliges Einsammeln der Artikel geordnet nach Gängen) sowie zufällige und ABC-Lagerplatzvergabe.

Im Bereich der Softwarelösungen gibt es nunmehr eine Vielzahl an Möglichkeiten, um agentenbasierte Simulationsmodelle zu erstellen (z. B. durch Software, wie MATLAB, MANA oder Swarm, sowie durch Programmiersprachen, wie JAVA oder C++). Trotz der seltenen Nennung der genutzten Software bei Veröffentlichungen im Bereich der agentenbasierten Simulationsmodelle wurde AnyLogic im Literaturüberblick von Heath, Hill und Ciarallo (2009) als mit am häufigsten genutzt identifiziert. AnyLogic ist für die Problemstellung der vorliegenden Arbeit sehr gut geeignet, da die Software einerseits agentenbasierte Simulationsmodelle unterstützt. Andererseits bietet AnyLogic bereits definierte Funktionsbibliotheken, eine intuitive Benutzeroberfläche und darüber hinaus die Möglichkeit, Simulationsmodelle flexibel durch zusätzliche Funktionen erweitern zu können (Borshchev und Filippov 2004).

## 2.2 Annahmen und Parameter

In der vorliegenden Arbeit wird ein rechtwinkliges Lagerlayout mit eindimensionaler Kommissionierung und Mann-zur-Ware-Prinzip untersucht, bei dem Agenten mehrere Artikel pro Kommissioniertour einsammeln. Es werden im Vorfeld 1.000 zufällige Auftragslisten pro Lagerplatzvergabestrategie generiert. Jede Auftragsliste enthält 20 einzusammelnde Artikel. Dabei ist die Lagerkapazität unbeschränkt, d. h. Artikel sind stets im Lagerfach verfügbar und der Wiederauffüllungsaufwand wird vernachlässigt. Die Geschwindigkeit der Kommissionierer ist mit 0,75 Meter pro Sekunde konstant. Auch die Zeit für einen Greifvorgang wird mit 20 Sekunden als fest vorgegeben angenommen.

Die Kommissionierer starten ihre Tour am Depot und sammeln gemäß der festgelegten Routenpolitik die 20 Produkte auf der Auftragsliste ein. Anschließend kehren sie zum Depot zurück. Die Anzahl der Kommissionierer im Lager wird dabei variiert und es werden drei Durchläufe jeweils mit 5, 10 und 15 Agenten durchgeführt. Für jeden Kommissionierer wird die Zeit, die er zur Erfüllung einer Auftragsliste benötigt, gemessen. Es wird angenommen, dass der nächste Agent startet, sobald der erste Agent den Eingang von Gang 1 erreicht hat (dafür werden 18,5 Sekunden festgelegt – die Zeit, die ein Agent für den Weg vom Depot zum ersten Gang benötigt). Bei Blockiersituationen bekommt der Agent mit dem kürzesten Weg zum nächsten Pick jeweils Priorität. Die anderen Agenten müssen dann warten, bis dieser den Gang verlässt oder den Artikel fertig kommissioniert hat.

Bei unterschiedlichen Lagerplatzvergabestrategien wurden die 20 Picks pro Auftrag jeweils anders angeordnet, da je nach Politik (vgl. Abbildung 1) die Verteilung der Artikel unterschiedlich ist (wenn z. B. eine Lagerplatzvergabestrategie gewählt wird, bei der Produkte, die häufig benötigt werden, nah an der Kommissionierbasis gelagert werden sollen). Es wurden drei Lagerplatzvergabestrategien implementiert, die das Lager jeweils in drei unterschiedliche Zonen unterteilen (A, B und C). Die Zonen unterscheiden sich entsprechend der erwarteten Verkaufshäufigkeit einzelner Produkte (Petersen et al. 2004). Es wird dabei eine Pareto-Verteilung der Produkte angenommen, d. h. dass 20% der Produkte für über 80% des Lagerumsatzes verantwortlich sind. Die Schnelldreher (höchste Umschlagshäufigkeit) werden demnach im ersten Bereich (A) angeordnet, welche 80% des Nachfragevolumens ausmachen und sich auf 20% (z. B. 2 Gänge bei vertikaler Anordnung) der Lagerfläche verteilen. Zone B enthält die Mitteldreher und nimmt ebenfalls 20% des Lagerplatzes in Anspruch. Der größte Bereich im Lager bildet der Bereich C (60% der Lagerplätze), wo die am seltensten verkauften Produkte (Langsamdreher) eingelagert sind, die einen Großteil des Artikelsortiments ausmachen, aber nur für einen geringen Teil Gesamtumsatzes verantwortlich sind (5% der Nachfragemenge). Ziel ist es, die Zone mit A-Produkten möglichst nahe an der Basis zu platzieren (Petersen und Schmenner 1999). Abbildung 1 zeigt die Verteilung der A-, B- und C-Bereiche im Lager. Bei der diagonalen Lagerplatzvergabe (rechts in der Grafik) musste jedoch eine andere Verteilung gewählt werden, da sonst in den einzelnen Gängen gegenüberliegenden Picks aus unterschiedlichen Zonen stammen. Daher wurde hier die Verteilung 20% (A), 30% (B), 50% (C) implementiert.

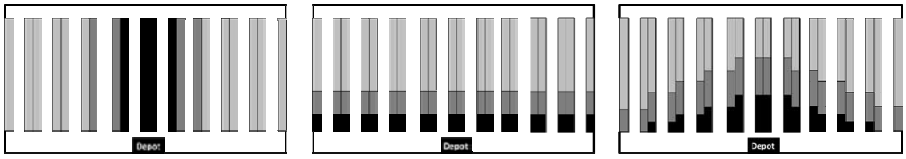



*Abbildung 1: Darstellung der ABC-Zonen (vertikal, horizontal und diagonal) im rechtwinkligen Lagerlayout mit dem Depot in der Mitte (A-Artikel: schwarz; B-Artikel: dunkelgrau; C-Artikel: hellgrau)*

Zur Analyse der Auswirkungen des Picker Blocking in manuellen Kommissionierlagern werden sechs Routenpolitiken und die drei beschriebenen Lagerplatzvergabepolitiken miteinander verglichen. Es werden dabei die im Bereich der manuellen Kommissionierung am häufigsten untersuchten Routenpolitiken Return, S-shape, Midpoint, Largest gap, Composite und Combined verwendet (Petersen und Schmenner 1999).

Bei der Return–Politik verlässt der Kommissionierer den Gang stets zu der Seite, in der er diesen betreten hat. Die S-shape-Politik verläuft S-förmig (meist von links nach rechts) durch das Lager, wobei Gänge mit Produkten stets komplett durchquert werden. Bei der Midpoint-Politik wird das Lager in eine obere und eine untere Hälfte unterteilt. Der Kommissionierer beginnt mit der zum Depot gewandten Seite und kommissioniert die Artikel. Dabei betritt er einzelne Gänge nur bis zur Mitte (mit Ausnahme der beiden äußeren Gänge, die komplett durchquert werden). Anschließend werden die Produkte in der zweiten Hälfte des Lagers eingesammelt. Die

Largest gap–Politik ist ähnlich der Midpoint-Politik. Hier betritt der Kommissionierer die jeweiligen Gänge bis zur größten Lücke zwischen zwei benachbarten Picks. Ebenfalls ähnlich sind die Politiken Composite und Combined, welche eine Kombination der Return- und der S-shape-Politik darstellen. Bei der Combined-Politik wird, im Unterschied zur Composite-Politik, der kürzeste Weg mithilfe dynamischer Programmierung bereits vor dem Loslaufen berechnet.

## 3 Ergebnisse

Nachfolgend werden die Ergebnisse der beschriebenen Simulationsdurchläufe besprochen. Abbildungen 2 bis 4 zeigen die durchschnittlichen Durchlaufzeiten für jeweils 1.000 Durchläufe und veränderlicher Zahl an Kommissionierern (5, 10 und 15) im Lager an. Es wird deutlich, dass je nach Lagerplatzvergabe unterschiedliche Routenpolitiken zu kürzesten durchschnittlichen Durchlaufzeiten führen. Bei den Lagerplatzvergabestrategien ABC-vertikal und -diagonal führt die Composite-Politik (bei 5, 10 und 15 Agenten) zu den kürzesten durchschnittlichen Durchlaufzeiten. Bei der Verteilung ABC-horizontal führt die Combined-Politik mit 5 Agenten zu den kürzesten Zeiten und für 10 bzw. 15 Agenten die Midpoint-Politik.

Auffällig ist hier, dass komplexere Heuristiken, wie Composite oder Combined, stets zu besseren Zeiten führen im Vergleich zu einfacheren Politiken, wie Return oder S-shape. Gerade die S-shape-Politik, welche in der Praxis aufgrund der einfachen Umsetzung häufig zu finden ist (de Koster et al. 2007), führt zu längeren durchschnittlichen Durchlaufzeiten. Dies ändert sich auch nicht, obwohl bei der S-shape-Politik mit ABC-horizontal-Verteilung die geringste durchschnittliche Anzahl an Blockierungen pro Durchlauf anfallen. In diesem Fall führte die Combined-/Midpoint-Politik trotz zusätzlicher Wartezeiten und Wegstrecken bei Blockierungen zu kürzeren Zeiten. Hier wirken sich die negativen Effekte auf die komplexeren Heuristiken nicht so stark aus, dass sie trotzdem zu besseren Ergebnissen führen.

Bei dem Vergleich der Politik, die zu kürzesten Durchlaufzeiten führt, mit den Politiken, die – unter den gleichen Parametern – jeweils zu den längsten Zeiten führen, fällt auf, dass bei der ABC-diagonal-Verteilung die Zeiten um ca. 8% (bei 5 Agenten) bis zu ca. 21% (bei 15 Agenten) höher sind. Bei der ABC-vertikal-Verteilung steigt der Unterschied noch weiter an. Hier ergeben sich bereits bei 5 Agenten Unterschiede von ca. 21% und bei 15 Agenten sind ca. 40% längere Zeiten zu erwarten. Sehr robust bezüglich der Routenpolitiken hat sich die ABC-horizontale-Verteilung gezeigt. Hier sind die Zeiten nur um ca. 6% (bei 5 Agenten) bis zu ca. 8% (bei 15 Agenten) länger.

Aufgrund der geringen Schwankungen bei den durchschnittlichen Durchlaufzeiten scheint es bei der ABC-horizontal-Verteilung weniger entscheidend, welche Routenpolitik verfolgt wird. Die komplexeren Heuristiken (Midpoint, Largest gap, Composite und Combined) führen über alle Durchläufe zu ähnlichen Zeiten, welche stets geringer sind als bei der S-shape- oder Return-Politik.

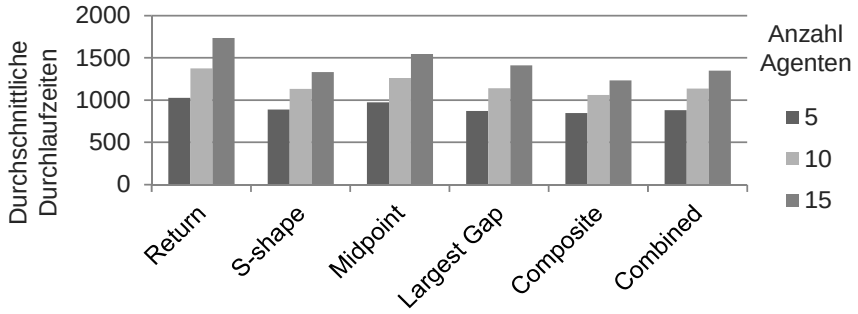



*Abbildung 2: Durchschnittliche Durchlaufzeiten bei vertikaler ABC-Lagerplatzvergabe*

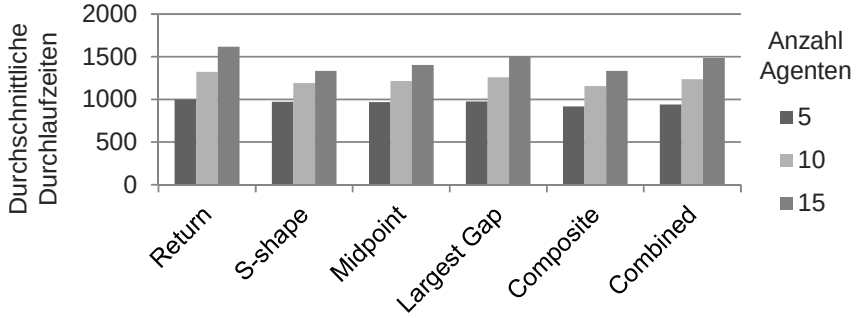



*Abbildung 3: Durchschnittliche Durchlaufzeiten bei diagonaler ABC-Lagerplatzvergabe*

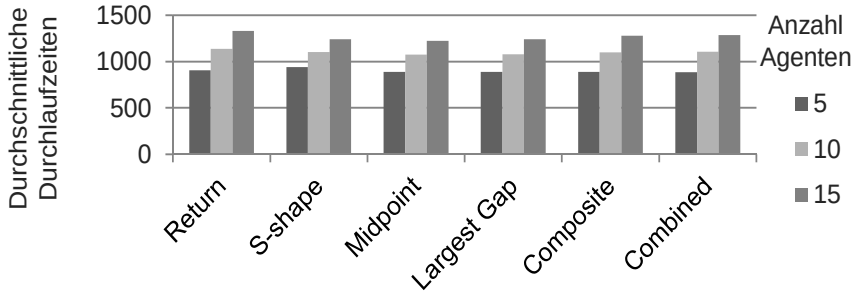



*Abbildung 4: Durchschnittliche Durchlaufzeiten bei horizontaler ABC-Lagerplatzvergabe*

Weiterhin wird deutlich, dass die durchschnittlichen Durchlaufzeiten bei der Composite-Routenpolitik und ABC-vertikal-Verteilung für 5 und 10 Agenten zu den

kürzesten durchschnittlichen Durchlaufzeiten über alle Lagerplatzstrategien führt. Bei 15 Agenten ist dies die Midpoint-Politik mit der ABC-horizontal-Verteilung.

## 4 Zusammenfassung und Fazit

Im Ergebnis wird deutlich, dass das Picker Blocking unter Berücksichtigung verschiedener Lagerplatzvergabestrategien und Routenpolitiken wechselnden Einfluss hat. Die ähnlichen Heuristiken Composite und Combined oder Midpoint führen zu kürzesten durchschnittlichen Durchlaufzeiten im Vergleich zu anderen Routenpolitiken (Return, S-shape, Midpoint, Largest gap). Vor allem bei der ABC-vertikal-Verteilung scheint es wichtig zu sein, welche Routenpolitik gewählt wird, da die durchschnittliche Durchlaufzeit bis zu 40% je nach Routenpolitik länger ausfällt. Gleiches gilt für die ABC-diagonal-Verteilung. Hier liegen die Unterschiede bei bis zu ca. 21% längeren Zeiten (bei 15 Agenten). Die ABC-horizontal-Verteilung zeigt sich hier als wesentlich robuster bei Veränderung der Routenpolitik sowie der Anzahl an Agenten (max. 8% längere Zeiten).

Als Implikationen für die Praxis zeigt sich, dass bei Berücksichtigung der Lagerplatzvergabe die Wahl der Routenpolitik wichtig ist. Zur Planung des Kommissionierprozesses in realen Kommissionierlagern sollten folgende Kombinationen gewählt werden: Bei der ABC-vertikal- und -diagonal-Verteilung zeigt sich die Composite-Politik für 5, 10 und 15 Agenten als am besten geeignet, bei der ABC-horizontal-Verteilung die Combined-Politik für 5 und 10 Kommissionierer und für 15 Kommissionierer die Midpoint-Politik. Hierbei ist die Anzahl der Agenten bei der ABC-vertikal- und -diagonal-Verteilung nicht relevant, da sowohl bei 5, 10 und 15 Agenten die gleiche Routenpolitik (Composite) zu den kürzesten durchschnittlichen Zeiten führt. Nur bei der ABC-horizontal-Verteilung ist die Midpoint-Politik ab 10 Agenten besser als die Combined-Politik, welche bei 5 Agenten zu den kürzesten durchschnittlichen Durchlaufzeiten führt. Ferner wird deutlich, dass der Einsatz komplexerer Heuristiken in der Praxis sinnvoll erscheint, auch wenn die Implementierung solcher Politiken schwieriger ist, da der Kommissionierer die Wegeführung nicht intuitiv wählen kann, wie es insbesondere bei der S-shape- oder Return-Politik der Fall ist.

Wenn das Ziel ist, eine möglichst geringe Durchlaufzeit zu erreichen, dann sollte bis zu 10 Kommissionierern die ABC-vertikal-Verteilung mit der Composite-Politik in rechtwinkligen Lagerlayouts mit eindimensionaler Kommissionierung und Mann-zur-Ware-Prinzip gewählt werden. Bei 15 Kommissionierern war die Midpoint-Politik mit der ABC-horizontal-Verteilung im Durchschnitt um lediglich 10 Sekunden schneller. Die geringste Anzahl an durchschnittlichen Blockierungen resultierte bei der ABC-horizontal-Verteilung und der S-shape-Politik.

Bei der vorliegenden Untersuchung wurde sich ausschließlich auf Kommissionierer beschränkt, die bei jedem Durchlauf den gleichen Regeln (Routenpolitiken, Ausweichen bei Blockierung) folgen. Verschiedene Aspekte des menschlichen Verhaltens, wie das Lernen, wurden nicht berücksichtigt. Auch wurde die Geschwindigkeit der Kommissionierer stets konstant gehalten. Ein Beschleunigen und Abbremsen, wie es in der Realität vorkommt, wurde ebenfalls nicht berücksichtigt. Dies gilt auch für die Kapazität der Lagerplätze, welche als unendlich angenommen wurde sowie für eine mögliche Beschädigung von Produkten, welche für die Untersuchung ebenfalls nicht relevant war.

In künftigen Studien können die eben genannten Aspekte in das Simulationsmodell implementiert und deren Auswirkungen untersucht werden. Auch die Auswirkung der Zeit, die ein Agent wartet, bis er seine Tour beginnt, kann in weiteren Arbeiten variiert werden. Ferner können weitere Routenpolitiken, wie die optimale Politik, oder zusätzliche Lagerplatzvergabestrategien, wie eine Paarverteilung, bei der häufig zusammengekaufte Artikel nebeneinander platziert werden, in die Analyse integriert werden.

## Literatur

Arnold, D.; Furmans K.: Materialfluss in Logistiksystemen. Berlin: Springer 2005.

Borshchev, A.; Filippov; A.: From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools. In: Proceedings of the 22nd International Conference of the System Dynamics Society No. 22, 2004 (ohne Seitenzahlen).

Borshchev, A.: The big book of simulation modeling: Multimethod modeling with AnyLogic 6. AnyLogic: North America 2013.

Chen, F.; Wang, H.; Qi, C.; Xie, Y.: An ant colony optimization routing algorithm for two order pickers with congestion consideration. Computers & Industrial Engineering 66 (2013) 1, S. 77-85.

de Koster, R.; Le-Duc, T.; Roodbergen, K. J.: Design and control of warehouse order picking: A literature review. European Journal of Operational Research 182 (2007) 2, S. 481-501.

Frazelle, E. H.: Trends in Transportation. Logistics Management Series, Georgia: The Logistics Institute at Georgia Tech 2000.

Furmans, K.; Huber, C.; Wisser, J.: Queueing models for manual order picking systems with blocking. Logistics Journal 1 (2009) 1, S. 1-16.

Glock, C. H.; Grosse, E. H.: Storage policies and order picking strategies in U-shaped order-picking systems with a movable base. International Journal of Production Research 50 (2012) 16, S. 4344-4357.

Grosse, E. H.; Glock, C. H.; Jaber, M. Y.; Neumann, W. P.: Incorporating human factors in order picking planning models: framework and research opportunities. International Journal of Production Research 53 (2015) 3, S. 695-717.

Grosse, E. H.; Glock, C. H.; Ballester-Ripoll, R.: A simulated annealing approach for the joint order batching and order picker routing problem with weight restrictions. International Journal of Operations and Quantitative Management 20 (2014) 2, S. 65-83.

Gue, K. R.; Meller, R. D.; Skufca, J. D.: The effects of pick density on order picking areas with narrow aisles. IIE Transactions 38 (2006) 10, S. 859–868.

Heath, B.; Hill, R.; Ciarallo, F.: A survey of agent-based modeling practices (January 1998 to July 2008). Journal of Artificial Societies and Social Simulation 12 (2009) 4, S. 9-44.

Heath, B.; Ciarallo, F.; Hill, R.: An agent-based modeling approach to analyze the impact of warehouse congestion on cost and performance. International Journal of Advanced Manufacturing Technology 67 (2013) 1, S. 563-574.

Hong, S.; Johnson, A. L.; Peters, B. A.: Analysis of picker blocking in narrow-aisle batch picking. In: Ellis, K.; Gue, K.; de Koster, R.; Meller, R.; Montreuil, B.; Oglep, M. (Hrsg.): Proceedings of 2010 International Material Handling

Research Colloquium (IMHRC), Charlotte: The Material Handling Institute 2010 (ohne Seitenzahlen).

Hong, S.; Johnson, A. L.; Peters, B. A.: Large-scale order batching in parallel-aisle picking systems. IIE Transactions 44 (2012a) 2, S. 88-106.

Hong, S.; Johnson, A. L.; Peters, B. A.: Batch picking in narrow-aisle order picking systems with consideration for picker blocking. European Journal of Operational Research 221 (2012b) 3, S. 557-570.

Klügl, F.: Multiagentensimulation: Konzepte, Werkzeuge, Anwendungen. München: Addison-Wesley 2001.

Law, A. M.: Simulation modeling and analysis. 5. Aufl., New York: McGraw-Hill Education 2007.

Macal, C. M.; North, M. J.: Agent-based modeling and simulation. In: Rossetti, M. D.; Hill, R. R.; Johansson, B.; Dunkin, A.; Ingalls, R. G. (Hrsg.): Proceedings of the 2009 Winter Simulation Conference 2009, S. 86-98.

Macal, C. M.: To agent-based simulation from system dynamics. In: Johansson, B.; Jain, S.; Montoya-Torres, J.; Hugan, J.; Yücesan, E. (Hrsg.): Proceedings of the 2010 Winter Simulation Conference 2010, S. 371-382.

Mowrey, C. H.; Parikh, P. J.: Mixed-width aisle configurations for order picking in distribution centers. European Journal of Operational Research 232 (2014) 1, S. 87-97.

Napolitano, M.: 2012 Warehouse/DC Operations Survey: Mixed signals. Modern Materials Handling 51 (2012) 11, S. 48–56.

Pan, J. C. H.; Wu, M. H.: Throughput analysis for order picking system with multiple pickers and aisle congestion considerations. Computers & Operations Research 39 (2012) 7, S. 1661-1672.

Parikh, P. J.; Meller, R. D.: Estimating picker blocking in wide-aisle order picking systems. IIE Transactions 41 (2009) 3, S. 232-246.

Parikh, P. J.; Meller, R. D.: A travel-time model for a person-onboard order picking system. European Journal of Operational Research 200 (2010a) 2, S. 385-394.

Parikh, P. J.; Meller, R. D.: A note on worker blocking in narrow-aisle order picking systems when pick time is non-deterministic. IIE Transactions 42 (2010b) 6, S. 392-404.

Petersen, C. G.; Schmenner, R. W.: An evaluation of routing and volume-based storage policies in an order picking operation. Decision Sciences 30 (1999) 2, S. 481-501.

Petersen, C. G.; Aase, G.: A comparison of picking, storage, and routing policies in manual order picking. International Journal of Production Economics 92 (2004) 1, S. 11-19.

Petersen, C. G.; Aase, G. R.; Heiser, D. R.: Improving order-picking performance through the implementation of class-based storage. International Journal of Physical Distribution and Logistics Management 34 (2004) 7, S. 534-544.

Roodbergen, K. J.; Vis, I. F.: A model for warehouse layout. IIE Transactions 38 (2006) 10, S. 799-811.

Ruben, R. A.; Jacobs, F. R.: Batch construction heuristics and storage assignment strategies for walk/ride and pick systems. Management Science 45 (1999) 4, S. 575-596.

Tompkins, J. A.; White, J. A.; Bozer, Y. A.; Tanchoco, J. A.: Facilities planning. 4. Aufl., New York: Wiley 2010.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Avoiding Equidistances when Routing by Shortest Paths

## *Vermeidung von Äquidistanzen beim Routing nach kürzesten Wegen*

Christian Hammel, Robert Schmaler, Thorsten Schmidt, Technische Universität Dresden, Dresden (Germany), christian.hammel@tu-dresden.de, robert.schmaler@tu-dresden.de, thorsten.schmidt@tu-dresden.de

Matthias Schöps, Globalfoundries Module One LLC & Co. KG, Dresden (Germany), matthias.schoeps@globalfoundries.com

Abstract: A considerable part of automated material handling systems (AMHS) use some sort of shortest path routing for defining the transport path. If the shortest paths are ambiguous (multiple shortest paths for one relation) there is often no explicit rule. Therefore, the exact behaviour is not completely predictable in simulation models. This reduces usability of simulation for analysing and optimising system behaviour, especially when examining behaviour on over-utilised tracks. To assess the potential impact of multiple shortest paths, a method is presented identifying all relevant node pairs. Also a method to eliminate these equidistances is discussed. Both methods are tested in AMHS simulations of an existing semiconductor fab. The case study also covers an impact analysis of multiple shortest paths in a real system. This shows outcome deviations of more than 10% in terms of throughput or delivery times – solely dependent on the exact behaviour in case of ambiguous paths.

## 1 Introduction

In areas like the semiconductor industry requirements to be fulfilled by installed material handling systems are steadily rising. This results in bigger and more complex systems. Current automated material handling systems (AMHS) in modern semiconductor fabs comprise a track system of several kilometres in length, more than 1,000 intersections and use several hundred vehicles to carry out some thousand transport jobs per hour (Gaxiola et al. 2012; Hammel et al. 2012).

The immense investment in the tool outfit in semiconductor industry makes it especially important to avoid any inefficiency in the production process. This includes the transportation of material between each two of more than 1,000 process steps. Because of the complexity simulation models are indispensable when checking and optimising the material handling system's functionality and performance. The routing strategy and its parameters are a key component of the transport system to be

reproduced in the model. To provide reasonable results in simulation runs this means the vehicles have to choose the same route in the model as they do in the real system. The exact reproduction may be challenging if the routing is based on shortest paths and these shortest paths are not unique. Differences in throughput and delivery times of more than 10 % have been revealed in simulation runs of a real world use case (see chapter 4). These differences just arose from differing routing schemes that resulted from exactly the same cost parameters, solely the selection from the equidistant paths varied. Hence, usability of simulation as prediction instrument for system performance in such cases suffers considerably.

This paper will introduce the routing problem in material handling systems in chapter 2. In chapter 3 the problem of equidistant shortest paths is described along with an algorithm to identify all origin-destination relations with multiple shortest paths and an approach to eliminate them. The method is demonstrated for an AMHS of a running semiconductor fab in chapter 4. The paper ends with concluding remarks in chapter 5.

## 2 Routing in Automated Material Handling Systems

In many existing automated material handling systems routing is based on some sort of static shortest path approach. This means all path segments are associated with a parameter based on e.g. length, travel time or some combination. This parameter is understood as costs for using this segment. Algorithms are then used to identify the path from origin to destination for which the total cost is minimal. Hence, ‘shortest’ path explicitly means the cheapest path regarding these cost parameters, not the one with physically shortest length.

Even though the system size would suggest more sophisticated (dynamic) approaches, necessary communication effort and the existing IT infrastructure mostly prohibit this. Sometimes simple functions are implemented to circumvent emerged traffic congestions, hence only acting reactively. More elaborate countermeasures require dynamic routing approaches (see e.g. Bartlett et al. 2014).

The mentioned shortest path routing is rather simple to implement in simulation models – as long as the shortest paths for all origin-destination relations are unique. If multiple different shortest paths (by cost parameters) exist additional information is needed to decide which of these paths to select. However, in many cases this additional information is not available. In real systems the existence of equidistant shortest paths often seems to be ignored, i.e. no explicit function covers this case. The decision of which of the shortest paths to choose is rather made implicitly. It depends on the specific implementation of the used search algorithm and the storage or search order of the path segments. As this is rarely documented to the last detail it is not possible to reproduce it in simulation models. This problem cannot be solved in the simulation alone. A possibility to avoid this problem right from the start is to eliminate equidistant paths in the real system. At the same time this would provide a higher predictability of system behaviour but it also means affecting the utilisation of path segments as some of the cost parameters have to be adjusted slightly. Especially if the utilisation of path segments is to be optimised by manipulating the cost parameters (Hammel et al. 2012) any such influence has to be examined carefully.

For analysing routing in general and shortest paths in particular a graph model of the transport network is an appropriate measure. Therefore, a graph $G(V, E)$ with vertices $V$ and edges $E$ is constructed by modelling all intersections as well as in- and outputs as vertices $v_i \in V$ and the interconnecting paths / pieces of track by edges $e_j \in E$. Each of the edges is furthermore associated with a cost parameter $c_j = c(e_j)$. A path $p_{odk}$ is a sequence of edges from origin vertex $v_o$ to destination vertex $v_d$ with path costs $c(p_{odk}) = \sum_{e_j \in p_{odk}} c_j$. The distance $d_{od}$ between two vertices $v_o$ and $v_d$ is defined as the path costs of the shortest path, i.e. the path with minimum path costs $d_{od} = \min_k(c(p_{odk}))$. Accordingly, the set of shortest paths from $v_o$ to $v_d$ is given by $MP_{od} = \{p_{odk} | c(p_{odk}) = d_{od}\}$. If $MP_{od}$ consists of at least two paths they are called equidistant paths for relation $(v_o, v_d)$.

# 3 Equidistant Shortest Paths

In material handling systems experience of the authors tells that the existence of multiple shortest paths for single origin-destination relations is mostly ignored. This may be (if the case is anticipated anyway) because occurrence of such equidistances is assumed scarcely and only little impact to global system behaviour and performance is expected. Furthermore, in the original idea of routing by shortest paths it does not make a difference which of the shortest paths to take, the distance is the same. However, this becomes an issue in case congestions impact travel times. E.g. uncertainty in terms of routes selected by vehicles may cause problems when traffic capacity of single pieces of track gets relevant. A real world application described in chapter 4 shows that the impact of equidistances may in fact be substantial. Additionally, algorithms often used for finding the shortest path (Dijkstra 1959; Floyd 1962; Warshall 1962) are designed to identify a shortest path; to be exactly i.e. one of the potentially multiple shortest paths. In their original form they are explicitly not checking for equidistances. Hence, computer implementations neither throw errors in such cases.

## 3.1 Issues for Simulation Studies

As mentioned above simulation studies, in particular discrete event simulations, are indispensable for some AMHSs. They are employed e.g. to predict maximum throughput, delivery times and impacts of different control strategies/functions. Even though simulation models as all models are abstractions of real world systems and the level of abstraction may vary, correct reproduction of routing decisions is a basic requirement for the reliability of such predictions.

In case of existing equidistances when routing happens according to shortest paths the accuracy of simulation studies to predict real system behaviour and subsequently the value of such studies is questionable. For assessment it is important to check for equidistances first. Even though this problem affects simulation and its value it is hard to solve without changes in the real system. Achieving this would require implementing the same procedures in the simulation model as occurring in the real system in case of more than one shortest path. While the identification of shortest paths in case they are unique is independent of which search algorithm is used, for multiple shortest paths the identified one might differ according to different algorithms used. Even for one and the same algorithm the respective implementation

and storage of the network structure might give different results. E.g. Dijkstra's algorithm chooses the first shortest path it finds. This means it depends among others on the order it searches through the edges. Consequently, reproducing real system routing in simulation is nearly infeasible in this case.

## 3.2 Identification

In order to describe the occurrence of multiple shortest paths in a network graph it is sufficient to list a minimum set $M$ of origin-destination relations in which for each node pair the following holds true: origin and destination nodes are the only nodes appearing in all shortest paths. Given that this minimum set of node pairs and their sets of shortest paths are known, all shortest paths for any other node pair can be generated from any one of the shortest paths by substituting respective parts. In particular this means only diverts appear as origins and merges as destinations in this minimum set. In Fig. 1 $M$ would consist of node pairs $(A, D)$, $(E, H)$, $(E, J)$ and $(G, J)$. Even though e.g. for $(A, J)$ the shortest path is not unique either, all shortest paths can be generated if the mentioned set is known. Moreover, an empty minimum set means that there are no equidistances in the network in general.

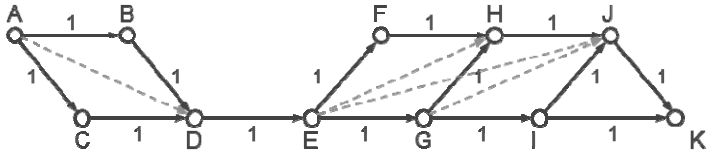



*Figure 1: Graph with uniform edge costs, the dashed lines depict those node pairs of the minimum set of node pairs with multiple shortest paths sufficient to describe all equidistances in the graph*

For finding this minimum set $M$ of node pairs a two-step method is proposed: The first step is a modification of the Floyd-Warshall algorithm commonly used to generate a matrix consisting of the next node on a shortest path for all node-to-node relations. Out of this a shortest path can be reconstructed. Storing not one but multiple next nodes in one position of the matrix in case of multiple shortest paths (if new distance is not less but equal to old distance) allows reconstruction of all shortest paths (Alg. 1).

*Algorithm 1: Modified Floyd-Warshall algorithm for finding all equidistant shortest paths*

```
matrix D with d_ij = minimum cost of edge from v_i to v_j,
                     ∞ if no edge between v_i to v_j
matrix N with n_ij = {v_j} if d_ij < ∞, NULL else
for i from 1 to |V|
     for j from 1 to |V|
          for k from 1 to |V|
               if d_ji + d_ik < d_jk then
                    n_jk = n_ji
                    d_jk = d_ji + d_ik
               else if d_ji + d_ik = d_jk then
                    n_jk = n_jk ∪ n_ji
```

The entries $n_{ij}$ of $N$ provide all next nodes on any shortest path from $v_i$ to $v_j$. Hence, in the second step only those node pairs for which $n_{ij}$ consists of multiple nodes have to be considered when searching for the minimum set $M$ of pairs describing all equidistances. Constructing all shortest paths from $v_i$ in the direction of $v_j$ by ascending length (similar to the approach used in the Dijkstra algorithm) and checking for the first node $v_k$ where all shortest paths combine again gives a node pair $(v_i, v_k)$ to include into $M$ (Alg. 2).

*Algorithm 2: Finding the elements of the minimum set M of all relevant node pairs for which the shortest path is not unique (run time optimisations omitted for better understanding)*

```
set M := {}
for i from 1 to |V|
     for j from 1 to |V|
          if |n_ij| > 1 then
               set LV := n_ij (ordered list)
               set LD := () (ordered list)
               for k from 1 to |LV|
                    v_s := LV_k
                    append d_is to LD
               while |LV| > 1
                    v_s := LV_k with LD_k = min(LD)
                    for r from 1 to |n_sj|
                         v_t := n_sj,r
                         if v_t ∉ LV
                              append v_t to LV
                              append d_it to LD
                    delete LV_k from LV
                    delete LD_k from LD
               v_s := LV_1
               if (v_i,v_s)∉ M
                    add (v_i,v_s) to M
```

## 3.3 Avoidance Approaches

Avoiding/eliminating equidistances means adjusting cost parameters in order to make all shortest paths unique ('shortest' regarding the path with minimum costs). Cost parameter settings can theoretically be constructed in a way that all shortest paths are unique. If e.g. the edges $e_j$ are ordered from $j = 1..n$ and their costs are defined by $c(e_j) = 2^j$ the cost of a path is determined by the edge with the highest index as $2^j > \sum_{i=1}^{j-1} 2^i$. However, in applications such an approach is infeasible because of limited ranges of used data types.

In real world systems the approach for eliminating equidistances from an existing shortest path routing scheme depends on the minimum steps when changing cost parameters. If this minimum step is (arbitrarily) small cost parameters may be adjusted to eliminate equidistant shortest paths without affecting the general routing. This can be achieved by increasing cost parameters for edges associated with all but one of the shortest path for each relevant node pair by an amount small enough not to raise the path costs to or above the second shortest paths (one example shown in Fig. 2). In case of a minimum step (e.g. integer values with a certain upper limit) it is more complicated to avoid impact to the general routing. Potentially any cost increase of one edge could change the shortest paths of all origin-destination pairs previously using this edge, i.e. all second-shortest paths may then have the same costs. This effect could be lessened by multiplying all cost parameters by an integer value first (e.g., by 10) thus generating gaps between the costs of paths and decreasing the probability of impact on the general routing. But the allowed range for path costs has to be considered before.

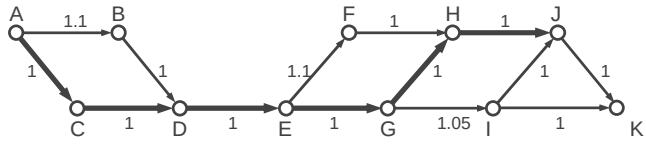



*Figure 2: Example for cost parameter adjustments to eliminate multiple shortest paths by small amounts with shortest path for $(A, J)$ in bold*

If impact on the general routing cannot be prevented completely, avoiding equidistances must be incorporated in the method generating the cost parameters in the first place. One possibility is the alternating application of the optimisation method and an iterative approach for eliminating equidistances until the requirements of both parts are fulfilled. The latter iteration would first generate the minimum set of relevant node pairs $M$ and then increase all but one outgoing edge of shortest paths starting at all origin nodes in $M$ (or all but edge of shortest paths related to the incoming edges of all destination nodes in $M$). For the graph in Fig. 1 for example cost increases in one of the outgoing edges of nodes $A$, $E$ and $G$ would be sufficient to eliminate all equidistances. These steps are repeated until the generation of $M$ provides an empty set. Fig. 3 shows the iteration for the example graph of Fig. 1. It also illustrates the potential of changing the original routing by the adjustments. The bottom part shows the new shortest path for $(G, K)$ which originally ran through node $I$ only.

As the identification of a solution cannot be guaranteed an appropriate exit condition should be implemented. For a better compatibility of the optimisation and equidistance elimination approaches the optimisation could e.g. increase cost parameters in multiples of the minimum increment, e.g. in even numbers only.

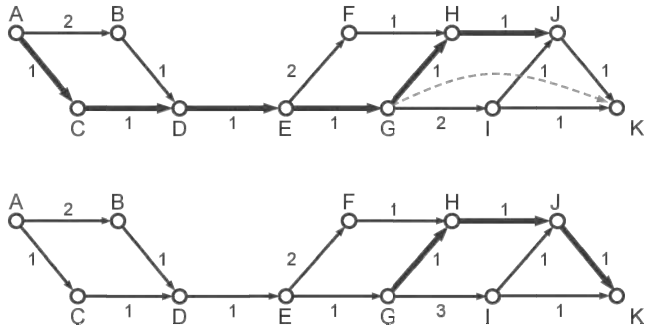



*Figure 3: Example for cost parameter adjustments to eliminate multiple shortest paths by minimum increments (integer values) generating new multiple shortest paths for $(G, K)$ in the first step (top, with shortest path for $(A, J)$ in bold) and changing the shortest path for $(G, K)$ (bold) in the second step (bottom)*

## 4 Case Study

The described facts and methods will be demonstrated on one of GLOBALFOUNDRIES' overhead hoisting transport systems in a fabrication plant in Dresden, Germany. It has been installed in 2004 and covers the whole production area of one building. It consists of 6,500 metres of track, nearly 1,000 intersections and 280 vehicles. Currently it performs around half a million transport jobs a week.

Vehicles in this system select the specific route to execute an assigned job by employing the Dijkstra algorithm. Cost parameters are therefore assigned to each piece of track. In the original setting these parameters roughly resembled the length of the respective piece of track with some manual adjustments, for computational reasons integer values were set analogue to full decimetres. This rounding is one of the factors raising the probability of equidistant paths occurring. Another factor is the mostly parallel arrangement of tracks in the same direction with several shunts to change from one to the other lane (Fig. 4).

The AMHS has been installed more than ten years ago and requirements especially in terms of throughput have risen (and still further rise) because of new or upgraded tools. Despite some technical upgrades also to the AMHS optimisation measures have to be taken in the control domain as well to be able to fully support production needs. Therefore, an optimisation of track utilisation by adjusting cost parameters as described by Hammel et. al (2012) should be executed. In the optimisation described in the paper equidistances have not been a big issue because for the highly utilised tracks cost parameters had been changed several times before so the occurrence was limited to areas where their impact was indeed negligible. In this new optimisation run cost parameters were to be reset to track lengths before adjustment. After this

reset and the subsequent optimisation the equidistance identification described in section 3.2 showed 162 intersection pairs with multiple shortest paths (Fig. 4). Especially for the typically highly utilised tracks connecting different areas (like in the far right in the figure) these equidistances may have unpredictable consequences as track utilisations are close to their capacity.

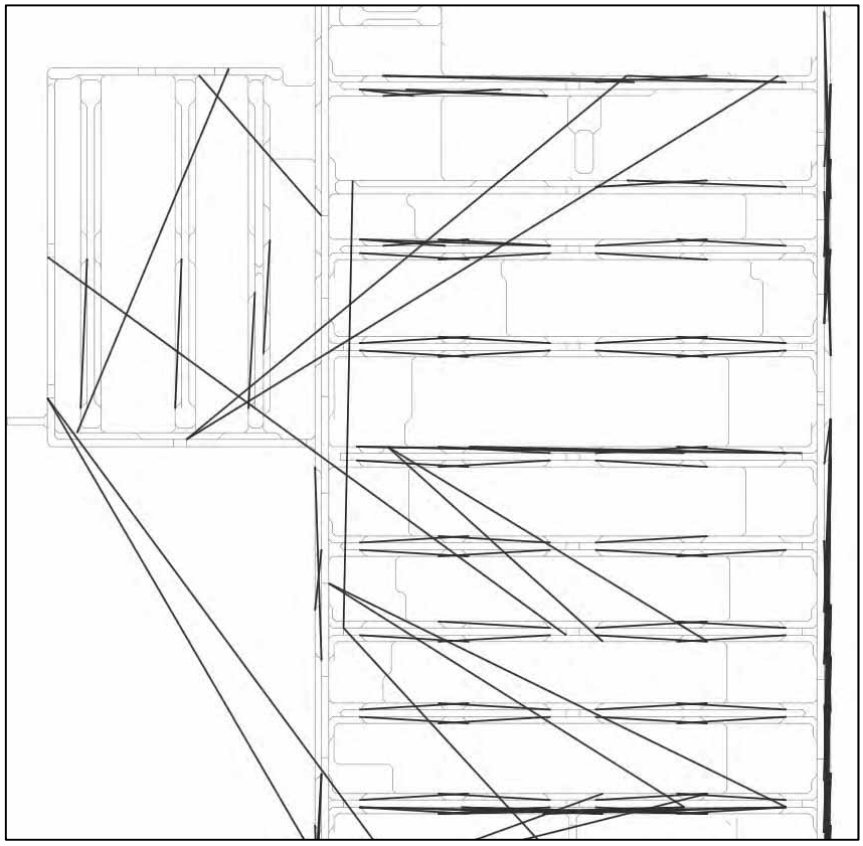


*Figure 4: Section of AMHS track layout of one of GLOBALFOUNDRIES' Dresden fabs (gray) with marked intersection pairs with multiple shortest paths (black)*

While in the real system cost parameters are limited to discrete values originally resembling length rounded to decimeters, in simulation models this limit does not apply. It is assumed that in reality as well as in simulation the path taken for one origin-destination relation is static, no splitting among equidistant paths takes place. To assess the potential impact of equidistances in this system cost parameters were adjusted in simulation runs slightly by small fractions of minimum differences of the real system (similar to Fig. 2). By this equidistances were eliminated, all shortest paths were made unique without changing the routing in general. Several of these cost parameter settings (each with other unique shortest paths) were generated, each showing one potential outcome of the actual real settings with multiple shortest paths; i.e. the ambiguous real cost parameters may lead to any of the unique

behaviors induced by these parameter settings adjusted for simulation, it is just not possible to predict which one is in fact the result in the real system. The simulation results (Tab. 1) therefore show a range of potential system behaviors and performance measures initiated by one and the same cost parameter setting. The sole difference is the selection of one of the shortest paths.

*Table 1: Result range of simulation runs with varying settings for multiple shortest paths in terms of delivery times (DT) normalised to values of best scenario*

| | | Medium load | High load |
|---|---|---|---|
| Best scenario | Average DT | 1.00 | 1.00 |
| | $95^{th}$ percentile DT | 1.00 | 1.00 |
| Worst scenario | Average DT | 1.01 | 1.05 |
| | $95^{th}$ percentile DT | 1.01 | 1.19 |

The results for medium load scenarios are quite similar: In case no traffic congestions occur the delivery time is barely affected as the track length is obviously nearly the same. However, this does not hold true for a high load scenario. The differences in average delivery time of 5 % and in the $95^{th}$ percentile of delivery time of 19 % show that there have been substantially more congestions in the worst case than in the best one. In fact, to achieve similar delivery times for the worst case would need to lower the load by 3 %.

To further test the potential impact of non-unique shortest paths the scenario with cost parameters reset to rounded decimetres of track length was simulated without subsequent optimisation. The throughput limit of this scenario was naturally considerably lower. However, executing several runs with only changing selection in case of multiple shortest paths showed an even bigger range for this throughput limit: The limit of the worst case was 16 % below the one of the best case. The optimisation process mentioned above obviously already reduced the number and impact of non-unique shortest paths as the difference for the maximum throughput was much smaller then.

In the real system the equidistances had to be eliminated differently as the limitation to discrete values resembling lengths rounded to decimetres applies. Therefore, the approach presented in section 3.3 had been employed in combination with the original optimisation algorithm. The two methods were executed in alternation. Thus, results of the respective simulation runs nearly identical to the ones of the best case in Tab. 1 could be achieved. In this scenario there are no ambiguities present in routing anymore. Considering and avoiding multiple shortest paths for origin-destination relations hence paid off in terms of predictability of system behaviour by simulation models as well as usability of simulation for system optimisation.

## 5 Conclusion

Even though shortest paths (regarding cost parameters) are often employed in some way as routing scheme for automated material handling systems, the case of non-

unique shortest paths is mostly ignored in both the real system and the simulation model. Explicit rules hence do rarely exist. Incorporating the same implicit rules in simulation models as taking effect in the real system is therefore hardly practicable.

For a real world case study it has been shown that the results of simulation studies in terms of delivery times and maximum throughput may vary in a range of up to more than 10 % relative to the best scenario. This makes simulation studies as prediction tool for the impact especially of different scenarios for cost settings affecting the routing unsuitable.

To assess this effect a method has been presented to identify a minimum set of origin-destination relations for which multiple shortest paths do exist. Out of this approaches have been described eliminating the occurrence of such equidistances with no or little impact to the general routing, depending on degree of freedom for cost parameters.

The importance of considering equidistances in a real system has been proven in a case study. Furthermore, the functionality of the described methods has been demonstrated providing a setting without any equidistances but achieving performance measures equal to the best case scenario without eliminating equidistances.

## References

Bartlett, K.; Lee, J.; Ahmed, S,; Nemhauser, G.; Sokol, J.; Na B.: Congestion-aware dynamic routing in automated material handling systems. Computers & Industrial Engineering 70 (2014), pp. 176-182.

Dijkstra, E.W.: A note on two problems in connexion with graphs. Numerische Mathematik 1 (1959) 1, pp. 269-271.

Floyd, R.W.: Algorithm 97 Shortest Path. Communications of the ACM 5 (1962) 6, p. 345.

Gaxiola, G.; Christensen, E.; Hammel, C.; Stachura, P.: Methodology to best extend AMHS for site expansion. In: Laroque, C.; Himmelspach, J; Pasupathy, R.; Rose, O,; Uhrmacher, A.M. (eds.): Proceedings of the 2012 Winter Simulation Conference, Berlin (Germany), 9-12 December 2012, pp. 1967-1976.

Hammel, C.; Schmidt, T.; Schöps, M.: Network optimization prior to dynamic simulation of AMHS. In: Laroque, C.; Himmelspach, J; Pasupathy, R.; Rose, O,; Uhrmacher, A.M. (eds.): Proceedings of the 2012 Winter Simulation Conference, Berlin (Germany), 9-12 December 2012, pp. 1956-1966.

Warshall, S.: A Theorem on Boolean Matrices. Journal of the ACM 9 (1962) 1, pp. 11–12.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Correlated Random Number Generation for Simulation Experiments

## *Generierung korrelierter Zufallszahlen für Simulationsexperimente*

Falko Bause, Jan Kriege, TU Dortmund, Dortmund (Germany),
falko.bause@tu-dortmund.de, jan.kriege@tu-dortmund.de

Abstract: The design of adequate input models is crucial for the validity of simulation experiments. Nowadays it is common to analyse trace data and to determine appropriate probability distributions for the generation of identically and independently distributed random numbers. Although it is well known that ignoring the correlation structure might give false results, modelling and generation of correlated random numbers is not supported by most simulation tools. In this paper we present an approach for generating correlated random numbers within the well-known simulator Arena. Fitting of input processes is done using the free available toolkit ProFiDo. Resultant process definitions are loaded into an Arena module with specific Visual Basic code for the generation of correlated random numbers. The effect of correlated versus uncorrelated input data is shown for the simulation results of an Arena assembly line model.

## 1 Introduction

As for any system analysis the GIGO (garbage in, garbage out) principle also holds for simulation experiments and often huge effort is invested to collect and analyse input data from real world. It is common practice to inspect trace data and to derive appropriate probability distributions (Law 2012). Nowadays this is supported by several tools like ExpertFit (Law and McComas 2003) or the Arena Input Analyser (Rathmell and Sturrock 2002) and resultant input models are used for the generation of identically and independently distributed random numbers. Even though it is well known that ignoring the correlation structure of input data might give false results (Livny et al. 1993) and modelling of correlation is considered as one of the most important issues in simulation (Biller and Nelson 2002), modelling and generation of correlated random numbers is not supported by most simulation tools. Since the work of Box and Jenkins (1970) well-known approaches to model correlation are time series models like AR (Auto Regressive), ARMA (Auto Regressive Moving Average) or ARIMA (Auto Regressive Integrated Moving Average) models or multivariate normal or Johnson distributions (Biller and Gunes 2010; Law 2014). Most of these input models are able to capture the correlation well, but have a

normal distribution, which is presumably one of the reasons why corresponding support is usually missing in simulation tools. An often applied workaround to capture correlation of input data is performing a trace driven simulation which has several disadvantages. E.g., traces may have certain irregularities or might be too short for simulation runs, so that theoretical input models should normally be preferred (Law 2014).

In recent years and decades research concentrated on the development of more general process definitions and appropriate fitting methods as e.g. for Markovian Arrival Processes (MAPs) (Neuts 1979) and Auto-Regressive-to-Anything (ARTA) processes (Cario and Nelson 1996). In most cases only prototype tools giving a proof of concept were developed, many with their own formats and user interfaces which makes comparisons and their use in practice difficult. Even if an appropriate process type and corresponding parameterisation has been found their use in simulation is not straightforward and often extra programming is required.

A multi-year project (DFG-project 2014) supported by the Deutsche Forschungsgemeinschaft (DFG) addressed these problems and resulted in the toolkit ProFiDo (Processes Fitting Toolkit Dortmund) (Bause et al. 2010; ProFiDo 2014), a free software, released under the terms of the GNU General Public License. ProFiDo provides a graphical user interface supporting the use of a variety of tools for the fitting and modelling of processes. Included are tools for fitting common distributions, phase-type distributions, MAPs, AR/ARIMA and ARTA processes. ProFiDo comes along with a well-defined interchange format such that its output can be easily used as input for simulators (Kriege and Buchholz 2011). In this paper we show how process descriptions resulting from ProFiDo's fitting tools can be used for the generation of correlated random numbers in the simulator Arena.

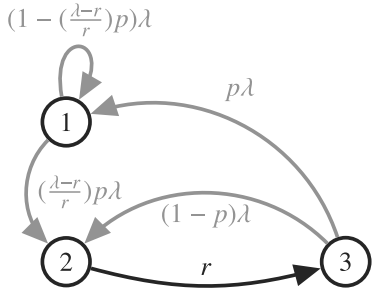
The structure of the paper is as follows. In Section 2 we present theoretical foundations and basics of the toolkit ProFiDo. Section 3 describes the integration of ProFiDo's output into Arena and the generation of correlated random numbers. Section 4 shows the effects of ignoring correlations for an assembly line model from Arena's standard example library. The paper ends with the conclusions in Section 5.

## 2 Modelling and Fitting of Input Data

### 2.1 Theoretical Background

This section concentrates on Markovian Arrival Processes (MAPs), since this class of processes is used in the example of Section 4. For details on other types of processes we refer to the mentioned literature.

A MAP is described by two matrices $D_0$ and $D_1$ such that $D_0 + D_1$ is the generator of an irreducible Markov process (see Fig. 1). $D_0$ is non-singular and contains non-negative values outside the diagonal and $D_1$ contains only non-negative values. $D_0$ describes internal transitions of the Markov process whereas transitions of $D_1$ specify arrivals. The arrival instances generated by a MAP can be described by a discrete time Markov chain (DTMC) with transition probability matrix $P = (-D_0)^{-1}D_1$. For a valid MAP $P$ is irreducible and the unique steady state vector $\pi$ given by $\pi P = \pi$, $\pi e = 1$ exists with $e$ being a column vector of ones of appropriate dimension. In steady state the probability density function (PDF) $f$ and the cumu-




$$\boldsymbol{D}_0 = \begin{bmatrix} -\lambda & 0 & 0 \\ 0 & -r & r \\ 0 & 0 & -\lambda \end{bmatrix}$$

$$\boldsymbol{D}_1 = \begin{bmatrix} (1-(\frac{\lambda-r}{r})p)\lambda & (\frac{\lambda-r}{r})p\lambda & 0 \\ 0 & 0 & 0 \\ p\lambda & (1-p)\lambda & 0 \end{bmatrix}$$

$$\boldsymbol{\pi} = \left(\frac{r}{\lambda}, \frac{\lambda-r}{\lambda}, 0\right)$$

*Figure 1: Example of a MAP ($p \in ]0,1[$, $\lambda > r > p\lambda/(1+p)$)*

lative distribution function (CDF) $F$ of the random variable $X$ specifying the interarrival times of a MAP are given by

$$f(x) = \pi \exp(D_0\, x)(-D_0)e \tag{1}$$

$$F(x) = 1 - \pi \exp(D_0\, x)e \tag{2}$$

resulting in the moments $\mu_i := E[X^i] = i!\pi(-D_0)^{-i}e$. Consecutive interarrival times $x_0$, $x_1$, … , $x_k$ ($k \in N$) are usually not independent and their joint distribution is given by

$$f(x_0, x_1, ..., x_k) = \pi \exp(D_0 x_0) D_1 \exp(D_0 x_1) D_1 \ldots \exp(D_0 x_k) D_1 e \tag{3}$$

giving the basis for the calculation of several characteristical values to express dependency of interarrival times. E.g., the lag-k autocorrelations $\rho_k$ are given by (Buchholz et al. 2014):

$$\rho_k = \frac{\mu_1^{-2}\pi(-D_0)^{-1} P^k (-D_0)^{-1} e - 1}{2\mu_1^{-2}\pi(-D_0)^{-1}(-D_0)^{-1} e - 1} \tag{4}$$

Figure 1 presents an example of a MAP with three parameters: $\lambda$ and p can be used to define different "degrees" of correlation and r specifies the rate of interarrivals. Interarrival times of this MAP are exponentially distributed with mean 1/r for all parameter values ($p \in ]0, 1[$, $\lambda > r > p\lambda/(1 + p)$) and the lag-1 autocorrelation is given by $\rho_1 = (p + 1)\, r/\lambda - p - (r/\lambda)^2$ and is non-zero if $p \neq r/\lambda$. For $p = r/\lambda$ the MAP coincides with an exponential distribution. Figure 2 shows the values of the autocorrelation function for the first lags for parameters $r = 5/3$, $p = 0.05$ and $\lambda = 3$.

The process of determining appropriate matrices $D_0$, $D_1$ for given statistical characteristics, as e.g. moments and autocorrelations derived from traces, is referred to as MAP fitting. In the last decades several fitting techniques have been developed with specific pros and cons (Buchholz et al. 2014), so that comparison of different fitting results is advisable. Unfortunately most fitting techniques come along, if at all, with their own prototype tools having their own formats and interfaces. This does not only makes comparisons difficult, but also impedes the integration of fitting results

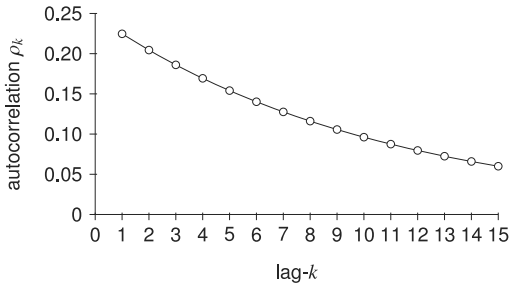



*Figure 2: Autocorrelation function of example MAP (r=5/3, p=0.05, λ=3)*

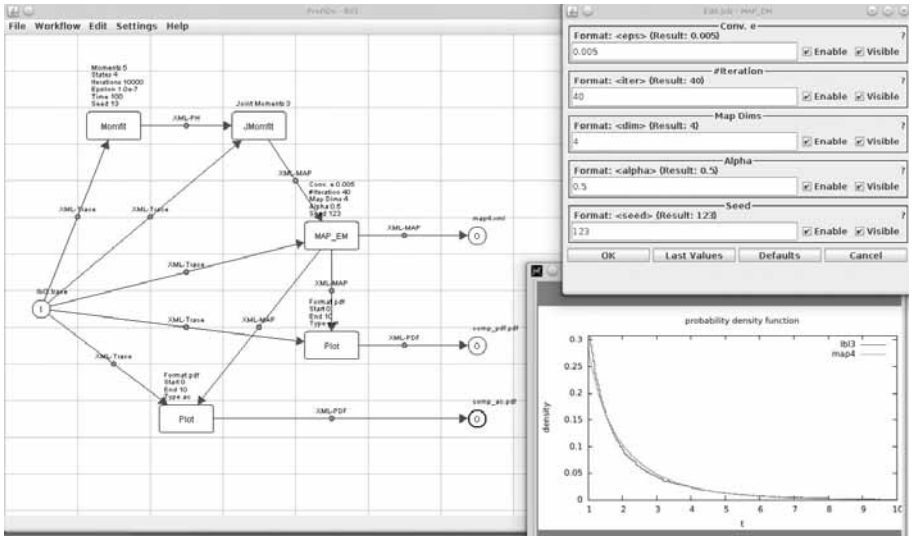



*Figure 3: Example workflow in ProFiDo*

into simulation tools. The freely available toolkit ProFiDo, described in the next subsection, intends to reduce these deficits.

## 2.2 ProFiDo - A Toolkit for Fitting Processes

ProFiDo (Bause et al. 2010; ProFiDo 2014) is a Java-based toolkit which integrates several tools for fitting the parameters of distributions or processes to be used as traffic sources or service processes in stochastic modelling. The basic idea of the toolkit is to provide an XML based interchange format for the interoperability of different fitting tools and a graphical interface to define workflows that specify the different steps of data preprocessing, parameter fitting and analysing the resulting processes. Figure 3 shows an example of such a workflow. The workflow starts with an input node named *I* which is associated with a trace file (*lbl3.trace*). Its content is used for three fitting tools named *Momfit*, *JMomfit* and *MAP_EM*. *Momfit* essentially fits the distribution of a MAP, i.e. matrix $D_0$. *JMomfit* uses the joint moments of a trace and the distribution of a MAP in order to determine a suitable matrix $D_1$. *MAP_EM* uses an EM algorithm to determine a MAP from a trace and usually profits from an initial guess which is determined by the tool *JMomfit* in this

workflow. ProFiDo offers additional tools which can also fit a complete MAP description in one step, but are not used in the example. In figure 3 the result of the tool *MAP_EM* is used for a special node *Plot* which allows the comparison of several statistical characteristics of fitted processes and original trace. Parameters of fitting tools can also be set within ProFiDo's graphical user interface to influence the quality of fitting results.

All fitting tools in ProFiDo are encapsulated by translator modules which transform output of tools into ProFiDo's XML based interchange format and data in interchange format into the tool specific input format. The interchange format allows for the description of a variety of probability distributions and stochastic processes and is the basis for the integration of fitting results into simulation tools. Kriege and Buchholz (2011) describe the use of such input models in the open source simulation framework OMNeT++. In the next section we present their integration into the simulator Arena.

## 3 Integrating Correlated Random Number Generation into Arena

The well-known simulator Arena (Kelton et al. 2015; Rathmell and Sturrock 2002) is widely used for modelling manufacturing and logistics systems. Models in Arena consist of modules that encapsulate a specific behaviour or perform specific tasks and are organised in templates. As mentioned, we have developed a template for Arena that allows for an easy integration of stochastic processes fitted with ProFiDo, and thus enables Arena to generate correlated random numbers from those processes.

In the following we will present one of the modules from the template in detail that can be used to sample random numbers from a MAP. The template also contains a module to generate random numbers from ARTA processes (Cario and Nelson 1996), but we will focus on the module for MAPs in this work. We first outline the module internals and some implementation details to give an overview how the module actually works. For a user who just wants to use the module in his model these details are of course hidden, but there are some parameters that can be used to influence the module's behaviour which will be explained at the end of this section.

Figure 4 shows the logic view of the MAPArrivals module in Arena. The module is composed of seven standard modules. The behaviour can be split into three parts. The first two modules in figure 4 are only executed once at the beginning of the simulation and are used to load a file with ProFiDo's XML description and initialise variables accordingly. Most of this work is done in the first *VBA* block in figure 4. *VBA* blocks can be used in Arena to integrate Visual Basic for Applications code into the model. The code is executed every time the process flow reaches the corresponding module. After everything is set up the main loop is entered that consists of three modules. *VBA* module 2 is used to generate the next interevent time from the MAP. The module contains code that simulates the underlying Markov chain of the MAP. The algorithm in figure 5 shows a pseudocode version of what is computed in this module. The *Delay* module is actually the only module that consumes model time in figure 4. Here the process flow is delayed until the next arrival is due. After that the process is split into two branches. The first branch will

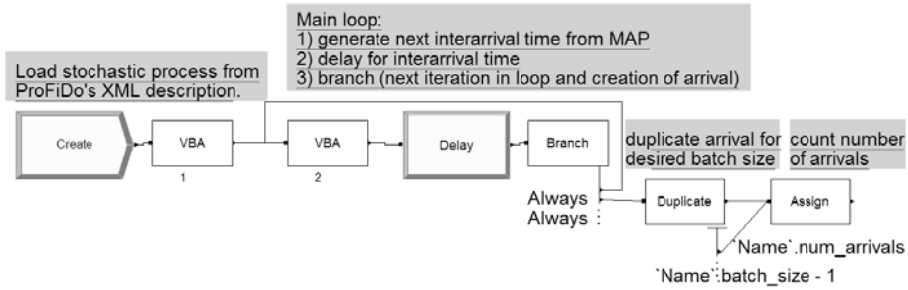



*Figure 4: Logic of the MAPArrivals Module*

**Input:** $\lambda_1, \cdots, \lambda_n, \boldsymbol{P}_0, \boldsymbol{p}_0, \boldsymbol{P}_1$, current state $i$;
**Output:** Time to the next event $x$ and new state $i$;
1: initialise $x = 0$ and $found = false$ ;
2: **repeat**
3: $x = x + exponential(\lambda_i)$;
4: $u = uniform(0, 1)$;
5: **if** $u \leq \boldsymbol{p}_0(i)$ **then**
6: $sum = 0$;
7: $j = 1$;
8: **repeat**
9: $sum = sum + \boldsymbol{P}_0(i, j)$;
10: $j = j + 1$ ;
11: **until** $sum > u$ ;
12: **else**
13: $sum = \boldsymbol{p}_0(i)$;
14: $j = 1$;
15: **repeat**
16: $sum = sum + \boldsymbol{P}_1(i, j)$;
17: $j = j + 1$ ;
18: **until** $sum > u$ ;
19: $found = true$ ;
20: **end if**
21: $i = j$ ;
22: **until** $found == true$ ;
23: return $x$ and $i$ ;

*Figure 5: Algorithm: Generate_Random_Number_from_MAP*

start the next iteration of the loop, while the second branch leaves the MAPArrivals module after a postprocessing step. The postprocessing step consists of two modules. In the first module the arrival is duplicated in case of batch arrivals and finally the number of generated arrivals is counted.

As already mentioned, figure 5 shows code to generate random numbers from a MAP (Buchholz et al. 2014) that is implemented in one of the VBA modules from figure 4. To simplify the notation of the algorithm we define

$$\lambda_i = -D_0(i,i) \qquad P_0(i,j) = \begin{cases} D_0(i,j)/\lambda_i & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases}$$
$$p_0 = P_0\, e \qquad P_1(i,j) = D_1(i,j)/\lambda_i \tag{5}$$

The algorithm draws random numbers from an exponential and a uniform distribution. The exponential random number is used to determine the time spent in the current state of the MAP (line 3). The uniform random number is used to determine the next state. A transition to another state can either occur according to $D_0$ (lines 5 to 11) or according to $D_1$ (lines 12 to 20). Transitions according to $D_0$ are internal transitions, while transitions according to $D_1$ trigger an event. Consequently, the

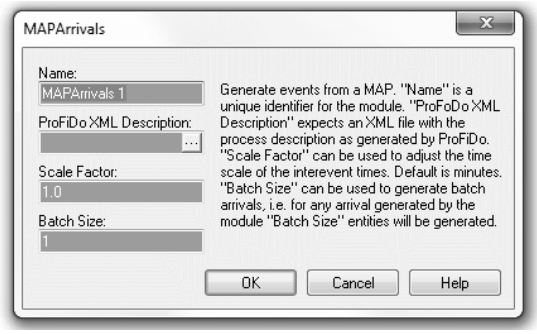



*Figure 6: Attribute window of the MAPArrivals module*

main loop of the algorithm is repeated until a transition according to $D_1$ has taken place. The interevent time is then returned as the sum of all times spent in the different states of the MAP until that transition occurred.

The details of the implementation presented so far are usually hidden from the user. If the module is used within a larger model it can be completely parameterised using the attribute window shown in figure 6. Apart from a name and a file with a MAP description in ProFiDo's XML format, one can specify a scale factor and the batch size. The MAPArrivals module assumes that the model time is given in minutes. If another time scale should be used this can be accomplished by setting the scale factor. The batch size determines how many arrivals should be generated at every event generated by the MAP.

## 4 Example

To demonstrate the effect of correlation on simulation results we use an assembly line model shown in figure 7 that is based on one of Arena's example models. The assembly line consists of 4 multi-server stations each associated with a buffer. The first buffer is infinite, while the remaining three buffers have limited capacity. Arriving parts are first processed at station 1, then passed to station 2 and so on. The details on the number of machines at each station, buffer sizes and deterministic processing times are summarized in table 1. Due to the limited buffer sizes of stations 2, 3 and 4, processed parts may not be able to leave a machine until space in the next buffer is available. This leads to the situation that a machine might be idle although parts are waiting to be served in the corresponding queue. This blocking phenomenon is known as transfer blocking (Tempelmeier and Kuhn 1993).

The model was simulated with arrivals generated by two different MAPs that are constructed according to figure 1 with mean arrival rates of r = 1.3 and r = 1.6 ($p=0.05$, $\lambda=3$). For comparison we simulated the model with arrivals according to exponential distributions with the same rates. Recall, that the interarrival time distribution of the MAPs is exponential. Thus, differences in the results are only caused by the autocorrelation of the MAPs. Table 2 shows the mean queue length for the buffers and the mean number of busy servers for the stations. If we compare the results of the MAP and the distribution with rate 1.3 we see that the main differences can be observed for the mean number of busy servers at station 1 and the mean queue length for the first two buffers. While the first buffer is almost empty for the

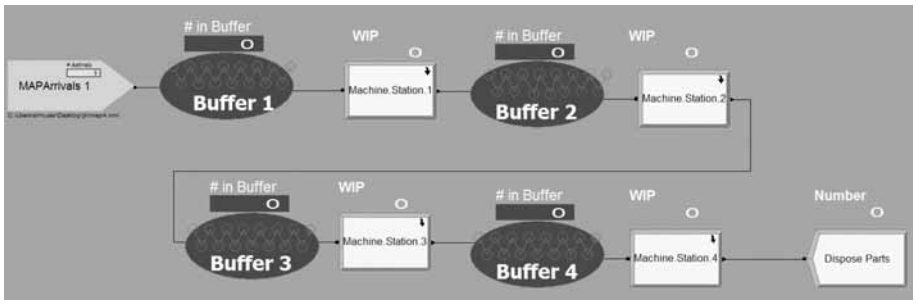



*Figure 7: Assembly line example model*

exponential distribution we have a mean queue length of 1.55 for the MAP. Of special interest are the values denoted as *Enter Buffer 2* in the table. These values indicate the mean number of parts that have been processed at station 1 but cannot proceed to buffer 2 because it is full, i.e. that are blocking a server at station 1. Note, that the values for the mean number of busy servers at station 1 include these values, implying that with correlated arrivals the servers are not really busy but are occupied due to transfer blocking for most of the time. With an increased arrival rate of 1.6 these effects become more significant. The mean number of parts that are blocked from entering buffer 2 increases to 1.67, implying that more than 50% of the server capacity at station 1 is lost due to transfer blocking.

*Table 1: Servers and buffers for assembly line model*

| Station | Buffer size | Number of servers | Service time |
|---|---|---|---|
| Station 1 | $\infty$ | 3 | 3 |
| Station 2 | 5 | 9 | 5 |
| Station 3 | 6 | 15 | 8 |
| Station 4 | 1 | 11 | 6 |

For the previous example we have chosen a very simple distribution to limit the effects on the difference between correlated and independent arrivals.

For our second example we used a trace with real data from Paxson and Floyd (1995) (LBL-TCP-3). The trace contains packet arrivals in a network and therefore, the data is of course not fully realistic for arrivals in an assembly line model. However, the data contains a complicated empirical distribution and correlation and serves well in demonstrating what data can be fitted if we do not limit ourselves to the standard distributions that are provided with simulation software. We adjusted the trace by a factor of approx. 150 in order to match the time scale of the model and used ProFiDo to generate a MAP with 4 states according to the workflow shown in figure 3. The best result obtained by the Arena Input Analyzer was $24 \cdot \text{beta}(0.475, 17.6)$, i.e. a Beta distribution scaled to the interval [0, 24]. Again, we simulated the assembly line model with the MAP and the distribution. Some results are summarised in Table 3.

*Table 2: Simulation results (mean population in queues and servers with 95% confidence intervals)*

| | MAP (rate 1.3) | Expo. (rate 1.3) | MAP (rate 1.6) | Expo. (rate 1.6) |
|---|---|---|---|---|
| Buffer 1 | 1.55 (±0.21) | 0.005 (±0.002) | 16.38 (±4.57) | 0.42 (±0.06) |
| Buffer 2 | 1.52 (±0.07) | 0.37 (±0.02) | 3.30 (±0.17) | 1.80 (±0.06) |
| Buffer 3 | 0.14 (±0.01) | 0.03 (±0.002) | 0.27 (±0.02) | 0.13 (±0.004) |
| Buffer 4 | 0.11 (±0.01) | 0.04 (±0.002) | 0.21 (±0.01) | 0.15 (±0.004) |
| Enter Buffer 2 | 0.59 (±0.04) | 0.02 (±0.004) | 1.67 (±0.12) | 0.42 (±0.03) |
| Station 1 | 0.66 (±0.04) | 0.09 (±0.004) | 1.75 (±0.12) | 0.50 (±0.03) |
| Station 2 | 6.50 (±0.07) | 6.48 (±0.03) | 8.02 (±0.11) | 7.99 (±0.03) |
| Station 3 | 10.42 (±0.12) | 10.38 (±0.04) | 12.85 (±0.17) | 12.80 (±0.05) |
| Station 4 | 7.80 (±0.09) | 7.78 (±0.03) | 9.62 (±0.13) | 9.68 (±0.04) |

*Table 3: Simulation results for fitted MAP and distribution (mean population in queues and servers with 95% confidence intervals)*

| | MAP(4) | Beta distribution |
|---|---|---|
| Buffer 1 | 24.63 (±5.40) | 2.98 (±0.55) |
| Station 1 | 1.96 (±0.10) | 1.12 (±0.06) |
| Enter Buffer 2 | 1.88 (±0.10) | 1.04 (±0.06) |
| Buffer 2 | 3.58 (±0.14) | 2.73 (±0.09) |

From the experiments we have seen, that autocorrelation might have a serious impact on the simulation results. Apparently, transfer blocking, which is common in assembly line models, even increases this effect. With correlated arrivals, sequences of short interarrival times tend to fall together, resulting in some periods with higher load. Due to the transfer blocking the servers cannot use their full capacity and therefore need considerable time to finish the load that has piled up, resulting in the effects shown in our examples.

## 5 Conclusion

The example of Section 4 demonstrates that correlations of arrival events might have a significant impact on simulation results. Modelling and fitting of a variety of corresponding input models came up in recent years, but their use has not found its way into most simulation tools. At present simulation experts often lack appropriate support which is partly caused by the diversity of input processes and associated tools and formats. The toolkit ProFiDo intends to reduce such deficits and offers fitting tools for several types of input models being unified under a precisely defined interface providing the basis for their use in simulation.

As mentioned, results of ProFiDo's fitting tools can be directly used in the open source simulation framework OMNeT++ (Kriege and Buchholz 2011). In this paper

we showed their direct use in the common simulator Arena. The Arena Arrival Module described in Section 3 is freely available (ProFiDo 2014) and might serve as a basis for the extension of other simulators.

## References

Bause, F.; Buchholz, P.; Kriege, J.: ProFiDo - the processes fitting toolkit Dortmund. In: Proceedings of the 7th International Conference on Quantitative Evaluation of SysTems (QEST), Williamsburg (VA) 2010, pp. 87–96.

Biller, B.; Gunes, C.: Introduction to simulation input modeling. In: Proceedings of the Winter Simulation Conference, Baltimore (MD) 2010, pp. 49–58.

Biller, B.; Nelson, B.: Answers to the top ten input modeling questions. In: Proceedings of the Winter Simulation Conference, San Diego (CA) 2002, pp. 35-40.

Box, G.; Jenkins, G.: Time series analysis - forecasting and control. San Francisco: Holden-Day 1970.

Buchholz, P.; Kriege, J.; Felko, I.: Input modeling with phase-type distributions and Markov models - theory and applications. Cham et al.: Springer 2014.

Cario, M.; Nelson, B.: Autoregressive to anything: time-series input processes for simulation. Operations Research Letters 19 (1996) 2, pp. 51–58.

DFG-project: Markovian and rational arrival and service processes for performance and reliability analysis. Website http://ls4-www.cs.tu-dortmund.de/cms/de/forschung/projekte/dfg-maps/index.html 2014.

Kelton, W.; Sadowski, R.; Zupick, N.: Simulation with Arena, 6th edition. Bosten et al.: McGraw-Hill Higher Education 2015.

Kriege, J.; Buchholz, P.: Simulating stochastic processes with OMNeT++. In: Proceedings of the 4th International Workshop on OMNeT++, Barcelona 2011, pp. 367-374.

Law, A.: A tutorial on how to select simulation input probability distributions. In: Proceedings of the Winter Simulation Conference, Berlin 2012, pp. 1-15.

Law, A.: Simulation modeling and analysis. 5th edition. New York: McGraw Hill 2014.

Law, A.; McComas, M.: How the ExpertFit distribution-fitting software can make your simulation models more valid. In: Proceedings of the Winter Simulation Conference, San Diego (CA) 2002, pp. 199-204.

Livny, M.; Melamed, B.; Tsiolis, A.: The impact of autocorrelation on queuing systems. Management Science 39 (1993) 3, pp. 322-339.

Neuts, M.: A versatile Markovian point process. Journal of Applied Probability 16 (1979) 4, pp. 764-779.

Paxson, V.; Floyd, S.: Wide-area traffic: the failure of Poisson modeling. IEEE/ACM Transactions on Networking 3 (1995) 3, pp. 226-244.

ProFiDo: Website http://ls4-www.cs.tu-dortmund.de/profido 2014.

Rathmell, J.; Sturrock, D.: Arena: The Arena product family: enterprise modeling solutions. In: Proceedings of the Winter Simulation Conference, San Diego (CA) 2002, pp. 165-172.

Tempelmeier, H.; Kuhn, H.: Flexible manufacturing systems: decision support for design and operation. New York: Wiley 1993.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Implementation of Dynamic Biological Process Models into a Reference Net Simulation Environment

## *Implementierung von dynamischen biologischen Prozessmodellen in eine Referenznetz-Simulationsumgebung*

Liam Pettigrew, Stefan Hubert, Frauke Groß, Antonio Delgado, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen (Germany),
liam.pettigrew@fau.de, stefan.hubert@fau.de, frauke.gross@fau.de,
antonio.delgado@fau.de

Abstract: This work explores a different approach to the standard hybrid Petri net. A subclass of object-orientated Petri nets, known as reference nets, are used to simulate the discrete states of a system. The objects represented by tokens in the net are then able to refer to continuous processes as their own objects. The reference net itself acts as a discrete system, but the continuous processes are able to be modelled for a given period and to also return state events that influence the system as a whole. This leads to a hybrid modelling environment that is able to model complex biotechnological manufacturing systems. This modelling environment was tested using two different biological processes commonly found in industrial production systems, the fermentation of beer and anaerobic digestion of wastewater.

## 1 Introduction

In the industrial use of biological processes to develop products (biotechnology), system performance depends on the behaviour of many different continuous processes. However, a large number of events in an industrial batch production system that implements biological processes can be discrete. Therefore, such a system is neither completely discrete nor completely continuous and requires a simulation environment that can combine aspects of both discrete event and continuous simulation, otherwise known as a hybrid simulation. This simulation model will need to know when a discrete event has affected a continuous variable or when the continuous variable reaches a state that causes a discrete event, a so called state event (Law 2014).

Accurate modelling of the discrete and continuous processes occurring within larger industrial production chains is important for later optimisation of scheduling and control strategies (Busch et al. 2007; Nahor et al. 2005; Park and Moon 2007;

Pettigrew et al. 2015; Torres Zúñiga and Vande Wouwer 2014; Tsafarakis et al. 2013). When more detailed information is needed from industrial production systems, in a networked sensor system for example, it is important to understand the behaviour of the continuous processes (Gudiño-Mendoza and López-Mellado 2013).

Many software tools already exist for simulation of industrial processes, including Aspen HYSYS®, gPROMS®, OpenModelica and MATLAB®. These tools are used to model systems that do not yet exist, retrofit existing complex facilities, determine effects of process changes on the environment and cost, and find the optimal solution. These simulators often operate as either a process flowsheet, a mathematical programming language or both and can model discrete and continuous (hybrid) processes (Lam et al. 2011).

Petri nets are another modelling system, where the marking on a place can represent the number of machines and the current state of the machines. However, when a system contains a large number of obtainable states the practical limitation of a discrete Petri net is reached. In a continuous Petri net the markings of places are real numbers and the firing of transitions is a continuous process. A hybrid Petri net contains discrete and continuous parts (David and Alla 2010).

This work investigates the simulation modelling and analysis of combined discrete and continuous systems using a high-level Petri net formalism in combination with Java inscriptions. The Petri net graph is used to layout and describe discrete events, while continuous processes are simulated using the Java programming language.

# 2 Reference nets and the Java programming language

## 2.1 Reference nets overview

Petri nets are a mathematical tool that have been used extensively to model and analyse manufacturing and production process chains. Petri nets are suitable for modelling events or defined states and implement a graphical environment, which allows for formal analysis of a system while providing a form of communication to the user that is easy to interpret (David and Alla 2010; Petri 1962; Zurawski and Zhou 1994).

High level Petri nets are an extended version of the standard Petri net formalism that introduces data attributes, timing and hierarchy. The reference net formalism is a subclass of object-oriented Petri nets that introduces synchronous channels and net instances. Reference nets allow the organisation, visualisation and running of complex, simultaneous processes through the dynamic creation of new net instances (Kummer 2002). The adding of temporal and attribute elements to the Petri net increases the modelling and analysis capabilities of the simulation (van der Aalst 1994).The reference net formalism can be described as an Elementary Object System (EOS) defined as a tuple:

$$EOS = (SN, \widehat{ON}, Rho, type, \widehat{M})$$

Where $SN = (P, T, W)$ is the system net defined by a finite set of places $P$, $T$ a finite set of transitions (disjoint from $P$) and $W$ a flow relation $W \subseteq (P \times T) \cup (T \times P)$. $\widehat{ON} = \{ON_1, ON_2, \dots, ON_n\}$ $(n \geq 1)$ is a finite set of elementary nets called object nets respectively systems of EOS with $ON_i = (B_i, E_i, F_i, m_{0i})$ where $B$ is a

finite set of places, $E$ a finite set of transitions, $F$ a flow relation $F \subseteq (B \times E) \cup (E \times B)$ and $m_{0i}$ the initial marking. $Rho = (\rho, \sigma)$ is the interaction relation, consisting of a system/object interaction relation $\rho \subseteq T \times \boldsymbol{E}$ where $\boldsymbol{E} \coloneqq \cup\{E_i | 1 \leq i \leq n\}$ and $\sigma$ a symmetric object/object interaction relation $\sigma \subseteq (\boldsymbol{E} \times \boldsymbol{E}) \backslash id_E$. The arc type function $type : W \longrightarrow 2^{\{1,\ldots,n\}} \cup \mathbb{N}$ and $\widehat{M}$ is a marking (Miyamoto and Kumagai 2005; Valk 1998).

The Reference Net Workshop (Renew) is a free and open-source Java-based multi-formalism editor and simulator developed by the Department of Informatics at the University of Hamburg, Germany. It provides a flexible modelling approach based on, but not limited to, reference nets. Renew utilises a plugin system allowing any Java class to be incorporated into the program, where transitions and arcs are then able to call functions from them. It is therefore possible to implement existing java libraries into a net graph to extend its functionality beyond that of a typical reference net. Figure 1 represents how a model stored in a java class object is referenced as a token in the reference net, the inputs and event handlers are separate tokens that are fed into the model object before solving. The solved model is output at the pre-defined or event driven time (t).

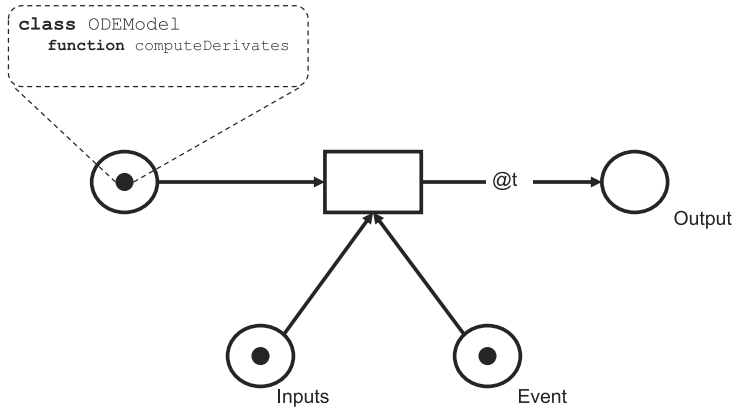



*Figure 1: A representation of a reference net containing a Java class reference from a token. The inputs and event handler are added separately at the transition where the system is solved. The solved system is sent on to the output place at the pre-defined or event driven time (t).*

## 2.2 Solving equations in Java

The Java programming language is a general-purpose, concurrent, class-based, object-oriented language (Gosling 2000). The Java virtual machine is a technology that allows any computer system to host a Java application without rewriting or recompiling (Lindholm and Yellin 1999). Java is one of the most popular programming languages, with mobile application development on the Android operating system increasing its popularity over the last few years (Barua et al. 2014; Destefanis 2014).

The Apache Commons Mathematics Library contains lightweight, self-contained mathematics and statistics components. These components are designed to address common problems that are not specifically addressed in the Java programming language (Andersen et al. 2011). The ode package provides classes and functions to solve Ordinary Differential Equation (ODE) problems. The package comes with a number of pre-defined fixed step and adaptive step-size integrators that provide dense output for obtaining the state and its derivative between time steps.

Open source modelling platforms implementing Java ODE solvers have been found to be more stable, accurate and even faster than more traditional implementations (Hemmerling et al. 2013; Ward et al. 2014). The solver can be run on a separate thread within reference net simulations so as to allow independent transitions to fire while the ODE is being solved, reducing simulation time on large systems. The solved model is passed through the final transition and exists in the reference net as a Java object, where outputs from the model can be retrieved. A continuous output model can also be saved in the Java object for visualisation and graphing of data over time.

Discrete event detection is handled using switching functions. Integrators can implement multiple discrete event handlers that when triggered stop the integration, change the state vector or allow the integration simply to continue. These event handlers must be specified before beginning the simulation.

## 3 Biological processes in industrial systems

This study focuses on using reference nets in combination with the Java programming language for the simulation modelling and analysis of two different biological processes commonly found in industrial production systems. One describing the anaerobic digestion of waste activated sludge in a wastewater treatment facility and a second regarding the fermentation of beer in the food and beverage industry.

### 3.1 Anaerobic digestion model no. 1 (ADM1)

The ADM1 was proposed by the IWA Task Group for Mathematical Modelling of Anaerobic Digestion Processes in 2001. The goal of the model is to increase utilisation of anaerobic process technologies for treatment of waste and wastewater, while concurrently reducing energy demand and greenhouse gas emissions (Batstone et al. 2002). ADM1 was a response to the need for a generic model of anaerobic digestion that could be built upon and easily validated. It was developed due to the multitude of models focusing on different aspects of the process in very specific settings. The model has been implemented in a number of popular aquatic modelling packages and languages including BioWin, EFOR, STOAT, JASS, WEST, GPS-X, MATLAB/Simulink, SIMBA, AquaSim, and FORTRAN (Batstone et al. 2006). The benchmark simulation model (BSM2) is a plant-wide modelling tool that implements the ADM1 into MATLAB/Simulink and considers computational aspects in regard to dynamic simulations of integrated systems. This tool was developed as a standardised simulation benchmark allowing for testing of different control strategies under realistic conditions (Rosen et al. 2006).

Anaerobic digestion involves a number of interconnected biological reactions, where organic matter is transformed into biomass, carbon dioxide and methane in an oxygen-free environment. The ADM1 model describes 24 species involved in 19 bioconversion processes, meaning it requires many kinetic parameters and is very complicated (Donoso-Bravo et al. 2011). Detailed substrate information, not readily available from widely applied measurements, is also required for the model to accurately simulate a real-world system (Kleerebezem and van Loosdrecht 2006). The ADM1 includes three overall biochemical (cellular) steps, acidogenesis, acetogenesis and methanogenesis, as well as extracellular disintegration and hydrolysis steps (Batstone 2006).

## 3.2 Fermentation model

The fermentation of wort during the production of beer is usually a batch production process. However, the ethanol production over time is an important continuous process that takes place in the reactor. A kinetic model for the fermentation process during the production of beer is based on the work of de Andrés-Toro et al. (1998).

The model describes five response values, including biomass, substrate, ethanol, diacetyl and ethyl acetate. It is based on a set of ordinary differential and algebraic equations. The process scheme considers the biomass in three different states, namely in lag-phase, active and dead (Fig. 2). In order to ferment the wort, yeast cells therefore need to be in an active state. Because there is no mechanical stirring applied, sedimentation of the yeast cells is also being considered. As the only variable to be manipulated within the process of fermentation is the temperature, special focus was laid on modelling the temperature effects on the process dynamics. Optimisation of the temperature profile has been achieved with this simulation model using dynamic programming and genetic algorithms. The model itself was validated using experimental data.

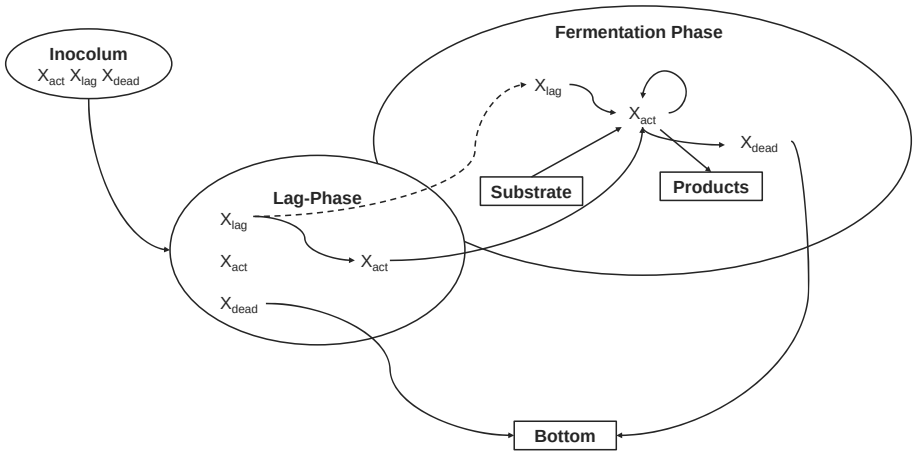



*Figure 2: Process scheme considered in the kinetic model for fermentation of beer wort (adapted from de Andrés-Toro et al. 1998).*

# 4 Simulation modelling and analysis approach

## 4.1 ADM1 implementation in Java and reference nets

The ADM1 model implemented in Java is based on that produced for the BSM2. This means the model can handle dynamic inputs as well as time-discrete and event-driven control actions. It also means that the stiffness problem inherent in the original ADM1 model is addressed. The standard ADM1 model is considered very stiff due to time constants from different processes within the system ranging from fractions of a second to months. The BSM2 implementation omits the fastest states by reformulating some of the system equations. The fast states are considered to be instantaneous in this implementation and are described by algebraic equations. Therefore, the implementation becomes a differential algebraic equation (DAE) system.

A Dormand-Prince 8(5,3) integrator is instantiated inside a simulation function of a model class with pre-defined maximum and minimum step sizes and allowed error values. A step handler is attached to the integrator object if the values over time are being recorded. Multiple event handlers are also attached to the integrator before beginning simulation.

The integrator calls a separate class containing the ODEs for solving giving starting values and start and end times. Separate classes are also used to store digester output and parameter values. A reference net is able to easily solve the given ODEs using the separate model class. Figure 3 shows a simple implementation of the ADM1 model in a reference net.

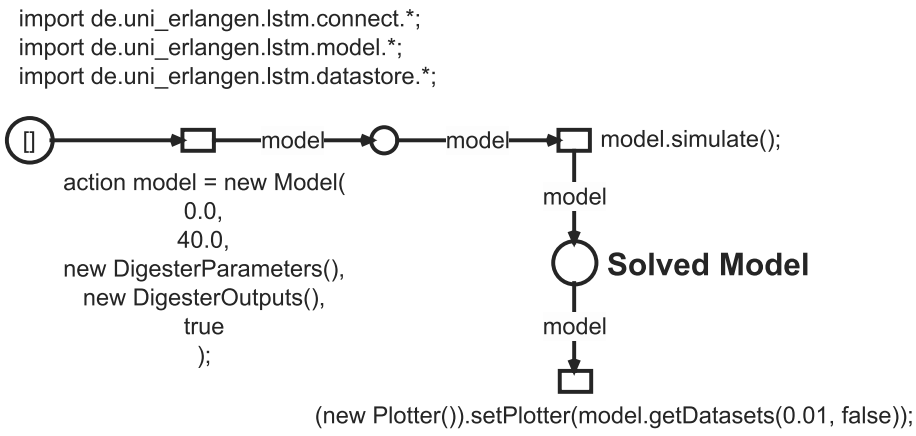



*Figure 3: A simple implementation of the ADM1 model into a reference net.*

## 4.2 Fermentation model implementation in Java and reference nets

A complex hybrid model was built dealing with the fermentation of wort and enhancing it with additional calculations. The aim was to completely portray the process, including discrete events such as the selection of fermentation, pumps and storage tanks. Events are modelled using reference nets, while the wort fermentation

and maturation are both modelled within a custom Java class and included within the net as an object.

The recipe model starts with the selection of the fermentation tank from a list of possible tanks and with the selection of the pumping system. The wort, given as a set of input parameters, is then pumped into the tank while simultaneously adding the yeast. An object of the appropriate recipe class is initiated and a method called to begin simulation of the fermentation process. The batch gets pumped into a storage tank and the maturation (or conditioning) period is then simulated.

## 5 Results

The computer system used for running all simulations is an Intel® Core™ i7 870 CPU running at 2.93 Ghz with 16.0 GB RAM. The computer is running the Windows 7 Enterprise 64-bit operating system.

A 200 day simulation was run in steady state to test the ADM1 outputs. The MATLAB implementation required 14 minutes and 11 seconds using the ode45 solver. The Java implementation of the DAE required 2 minutes and 43 seconds and gave exactly the same results. Both implementations used a non-stiff solver for testing due to the hybrid nature of the future system and the later inclusion of noise for control strategy investigations.

The Java implementation of the fermentation model was solved using a simple fixed step size Euler method. Figure 4 shows the course of biomass yeast (X in g/l) in different states (total, active, lag and dead) over time during an example wort fermentation. The 170 hour simulation had a runtime of approximately 100 ms and was successfully implemented into a whole plant brewery reference net model.

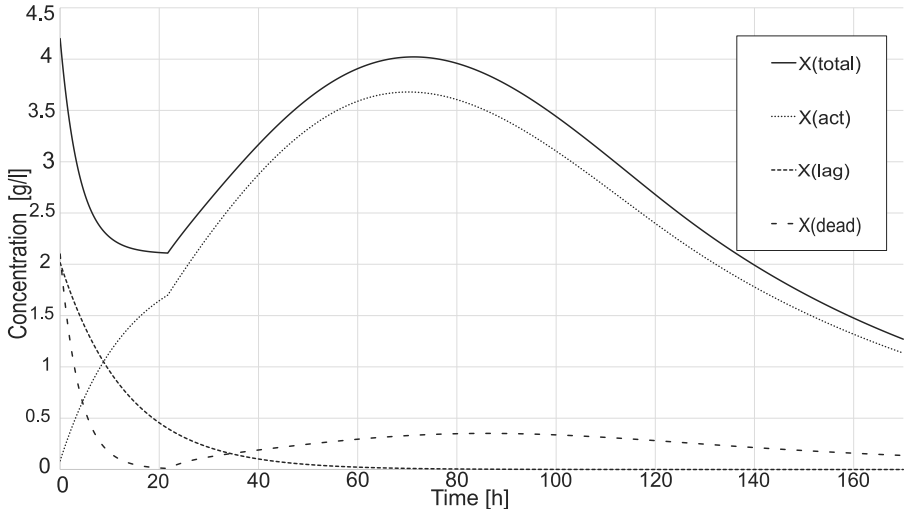



*Figure 4: A 170 hour simulation using the Java implementation of the fermentation model. The course over time for total, active, lag and dead Biomass in g/l is plotted. Simulation runtime was approximately 100 ms.*

## 6 Conclusion

Two examples of biotechnological industrial processes (anaerobic digestion of wastewater sludge and fermentation of wort) were successfully modelled and simulated using a combined reference net and Java modelling system. The implementation of Java plugins in the Renew reference net editor and simulator allows for advanced Java implemented functions to run in a familiar Petri net modelling environment. ODE and DAE implementations produced acceptable results and simulation times when compared with other existing models, while using a completely free and open source software package.

The incredible popularity of the Java programming environment and its large user base means solutions to problems and libraries for functions are quickly and easily available. Optimisations to the Java virtual machine and better hardware support means programs written in Java are becoming increasingly faster and more widely available. Java frameworks and libraries that already exist can be referenced directly from a Reference net with little to no changes or ported using a wrapping class for ease-of-use. The problem however remains, that any advanced models and functions must be initially programmed in Java before being implemented into the reference net environment if they do not already exist. A level of programming skill is required to properly define and construct these more complex models. Future research will focus on increasing accessibility to existing models and libraries while refining and focusing the task of constructing new models.

It is an opinion that mathematical modelling of biotechnological systems will primarily focus on the biological aspect in the future. So that widespread use occurs, this new generation of models will require a tool that can layer the complexity of microbial diversity and activity under simplified discrete batch process systems that describe their effects on an industrial scale.

## References

Andersen, M.M.; Barker, B.; Chou, A.D.; Diggory, M.; Donkin, R.B.; O'Brien, T.; Maisonobe, L.; Pietschmann, J.; Pourbaix, D.; Steitz, P.; others: Commons Math: The Apache Commons Mathematics Library. URL http://commons.apache.org/math/, online (2011).

Andrés-Toro, B. de; Girón-Sierra, J.M.; López-Orozco, J.A.; Fernández-Conde, C.; Peinado, J.M.; García-Ochoa, F.: A kinetic model for beer production under industrial operational conditions. Mathematics and Computers in Simulation 48 (1998) 1, pp. 65–74.

Barua, A.; Thomas, S.W.; Hassan, A.E.: What are developers talking about? An analysis of topics and trends in stack overflow. Empirical Software Engineering 19 (2014) 3, pp. 619–654.

Batstone, D.J.; Keller, J.; Angelidaki, I.; Kalyuzhnyi, S.V.; Pavlostathis, S.G.; Rozzi, A.; Sanders, W. T. M.; Siegrist, H.; Vavilin, V.A.: The IWA Anaerobic Digestion Model No 1(ADM 1). Water Science & Technology 45 (2002) 10, pp. 65–73.

Batstone, D.J.: Mathematical modelling of anaerobic reactors treating domestic wastewater: Rational criteria for model use. Reviews in Environmental Science and Bio/Technology 5 (2006) 1, pp. 57–71.

Batstone, D.J.; Keller, J.; Steyer, J.P.: A review of ADM 1 extensions, applications, and analysis: 2002-2005. Water science and technology 54 (2006) 4, pp. 1–10.

Busch, J.; Oldenburg, J.; Santos, M.; Cruse, A.; Marquardt, W.: Dynamic predictive scheduling of operational strategies for continuous processes using mixed-logic dynamic optimization. Computers & Chemical Engineering 31 (2007) 5–6, pp. 574–587.

David, R.; Alla, H.: Discrete, Continuous, and Hybrid Petri Nets. Berlin, Heidelberg: Springer 2010.

Destefanis, G.: Which programming language should a company use? A Twitter-based analysis. Technical Report Montréal, CRIM 2014.

Donoso-Bravo, A.; Mailier, J.; Martin, C.; Rodríguez, J.; Aceves-Lara, C.A.; Wouwer, A.V.: Model selection, identification and validation in anaerobic digestion: A review. Water Research 45 (2011) 17, pp. 5347–5364.

Gosling, J.: The Java Language Specification. Boston: Addison-Wesley Professional 2000.

Gudiño-Mendoza, B.; López-Mellado, E.: Modelling networked agents' behaviour using Timed Hybrid Petri Nets. Procedia Technology 7 (2013), pp. 289–296.

Hemmerling, R.; Evers, J.B.; Smoleňová, K.; Buck-Sorlin, G.; Kurth, W.: Extension of the GroIMP modelling platform to allow easy specification of differential equations describing biological processes within plant models. Computers and Electronics in Agriculture 92 (2013), pp. 1–8.

Kleerebezem, R.; van Loosdrecht, M. C. M.: Critical analysis of some concepts proposed in ADM1. Water Science and Technology 54 (2006) 4, pp. 51–57.

Kummer, O.: Referenznetze. Berlin: Logos 2002.

Lam, H.L.; Klemeš, J.J.; Kravanja, Z.; Varbanov, P.S.: Software tools overview: process integration, modelling and optimisation for energy saving and pollution reduction. Asia-Pacific Journal of Chemical Engineering 6 (2011) 5, pp. 696–712.

Law, A.: Simulation Modeling and Analysis. New York: McGraw-Hill Education 2014.

Lindholm, T.; Yellin, F.: Java Virtual Machine Specification. Boston: Addison-Wesley Longman 1999.

Miyamoto, T.; Kumagai, S.: A survey of object-oriented Petri nets and analysis methods. IEICE transactions on fundamentals of electronics, communications and computer sciences 88 (2005) 11, pp. 2964–2971.

Nahor, H.B.; Scheerlinck, N.; Verboven, P.; van Impe, J.; Nicolaï, B.M.: A continuous/discrete simulation of controlled atmosphere (CA) cool storage systems: Validation using industrial CA cool storage. International Journal of Refrigeration 28 (2005) 4, pp. 461–470.

Park, W.; Moon, I.: A discrete multi states model for the biological production of hydrogen by phototrophic microalga. Biochemical Engineering Journal 36 (2007) 1, pp. 19–27.

Petri, C.A.: Kommunikation mit Automaten. Dissertation, Universität Hamburg 1962.

Pettigrew, L.; Blomenhofer, V.; Hubert, S.; Groß, F.; Delgado, A.: Optimisation of water usage in a brewery clean-in-place system using reference nets. Journal of Cleaner Production 87 (2015), pp. 583–593.

Rosen, C.; Vrecko, D.; Gernaey, K.V.; Pons, M.-N.; Jeppsson, U.: Implementing ADM 1 for plant-wide benchmark simulations in Matlab/Simulink. Water Science & Technology 54 (2006) 4, pp. 11–19.

Torres Zúñiga, I.; Vande Wouwer, A.: Optimization of VARICOL SMB processes using hybrid modeling and nonlinear programming. Computers & Chemical Engineering 71 (2014), pp. 1–10.

Tsafarakis, S.; Saridakis, C.; Baltas, G.; Matsatsinis, N.: Hybrid particle swarm optimization with mutation for optimizing industrial product lines: An application to a mixed solution space considering both discrete and continuous design variables. Industrial Marketing Management 42 (2013) 4, pp. 496–506.

Valk, R.: Petri nets as token objects. In: Application and Theory of Petri Nets 1998. Berlin, Heidelberg, New York: Springer 1998, pp. 1–24.

van der Aalst, W.M.P.: Putting high-level Petri nets to work in industry. Computers in Industry 25 (1994) 1, pp. 45–54.

Ward, E.M.; Warner, J.G.; Maisonobe, L.: Do open source tools rival heritage systems? A comparison of tide models in OCEAN and Orekit. In: AIAA/AAS Astrodynamics Specialist Conference, San Diego (CA), 4-7 August 2014.

Zurawski, R.; Zhou, M.: Petri nets and industrial applications: A tutorial. Transactions on Industrial Electronics 41 (1994) 6, pp. 567–583.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# A Natural-language-based Simulation Modelling Approach

## *Mit natürlicher Sprache zum Simulationsmodell*

Thomas Mayer, Falk Stefan Pappert, Oliver Rose, Universität der Bundeswehr München, Neubiberg (Germany), thomas.mayer@unibw.de, falk.pappert@unibw.de, oliver.rose@unibw.de

Abstract: Formal methods to describe complex production and logistic systems for the purpose of simulation are increasing in importance. Looking at modern software development methodologies, i.e., Behaviour Driven Development (BDD), we see promising approaches to requirements analysis. BDDs principles are based on the idea of automated execution of requirement specifications which are formulated in semi-formal, executable natural language constructs, thereby improving the communication between stakeholders and development team. In this paper, we introduce natural language constructs for modelling systems from the domain of production and logistics. Their usage is showcased with the help of an example production system and their usability and suitability are evaluated with respect to the modelling process. Furthermore, challenges in the use of natural language as a new formalism to support simulation modelling are discussed.

## 1 Introduction

One of the main tasks in almost every simulation project is the development of a valid simulation model, To this end, conceptual modelling is one of the major aspects (Robinson et al. 2011). The common modelling and validation approaches emphasise the importance of communication between experts and stakeholders (Sargent 1996; Rabe et al. 2008). Pace (2000) also highlights the problem of understanding of what solution to implement. A starting point to support the communication between experts and stakeholders is a suitable representation of the problem. Robinson (2006; 2013) identifies the need of research in developing methods for designing conceptual models. Huang et al. (2007) approached this topic by using SysML to visualize processes and structures of a real world system. Another graphical approach was introduced by Ahmed at al. (2014) into the domain of Discrete Event Simulation using the Structured Analysis and Design Technique (SADT). SADT is an established concept in software engineering, specifically during the requirements analysis. In general, there is a similar communication challenge in the process of developing software. There, the software developer or archi-

tect has to understand the requirements for the product and the stakeholder need to be able to validate the delivered software based upon the agreed-on requirements. Beside graphical approaches it is common best practice to formalize requirements in natural language to increase understanding (Soeken et al. 2012). Furthermore, formalized executable requirements are used to automate the validation process. A well-known natural-language-based requirement formalization approach in software engineering is implemented within the software testing tool Cucumber. Cucumber automates acceptance testing based on requirements formalized in the language Gherkin that aims to use natural language constructs (Wynne and Hellesøy 2012). A broader approach was introduced as "Lean Modelling" (Seifert et al. 2013).

We investigate whether it is feasible to adopt this non-graphical requirements engineering approach based on natural languages to describe systems for modelling and simulation. Furthermore, it is evaluated whether a non-graphical approach is a suitable representation of a system. Therefore, language constructs are developed, which are based on the ideas behind Gherkin. A transformation into a simulation model is implemented that can be executed in our own simulation environment.

From our point of view, there are four mandatory components in every simulation environment (fig. 1). The editor component is responsible for user interaction, nowadays usually a graphical editor for model development which is providing support functionality to improve usability. With the help of the editor the conceptual model is implemented into a custom model, which represents the real world model within a tool-specific formalization and therefore the metamodel of the editor. The custom model is transformed into a simulator model that is based on the simulation tool's metamodel. This metamodel is the inherent model specification description of a simulator.

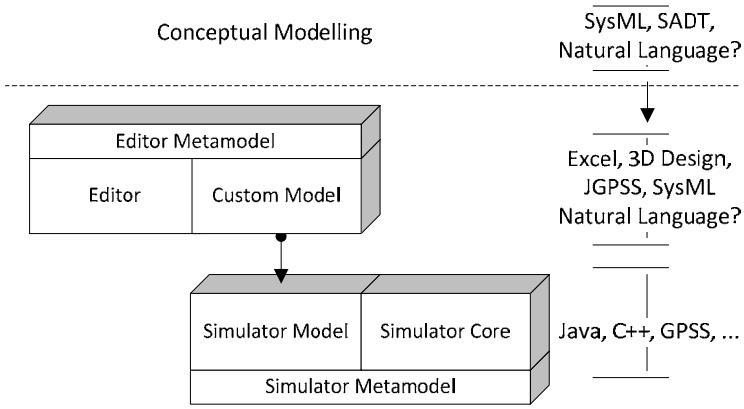



*Figure 1: Component model*

The presented natural language modelling approach focusses first on the conceptual modelling phase and second on the components editor and custom model. The editor component is a simple text editor for natural language constructs to create the custom model. After transformation into the simulator model, it is executed by the simulation core. Metamodel and simulation core are implemented in Java.

In the following, we will first give an insight into Gherkin and the ideas behind it. A small production system is shown which is used as a first test case. In Section 4 our language constructs are introduced. The transformation process of our language constructs into our metamodel is addressed in Section 5. Conclusions are drawn in Section 6.

## 2 The Gherkin Language

In Behaviour Driven Development (BDD), a software development approach building upon the foundation of Test Driven Development (TDD), there is a focus on involving every member of the development team with the goal of the software (Wynne and Hellesøy 2012). This focus is to ensure that only valid features are implemented and to have everyone from the stakeholder to the developer to talk on a level playing field. This should be done in the same language, thereby avoiding common problems of misunderstanding and miscommunication. Although certain experts will still focus the main part of their work in their area of expertise everyone in the team is supposed to be able to read and understand all documents of the project. To ensure that this approach is possible, the language in which specification and documentation of the system are written in needs to be ubiquitous. A further step to clarify specifications is to write them in an executable way. Thereby, specifications can be used to directly test and evaluate the state of the software system development without the need for further interpretation and translation into a testable format by a test professional. A well-known technology in the area of software development to implement this approach is Cucumber with the underlying language Gherkin. It is used to communicate, discuss, document and later automate specifications of the software in development.

Features are described with the help of use cases in the language Gherkin that is very close to natural language. Algorithm 1 shows a small Gherkin example of a calculator.

*Algorithm 1: Gherkin example*

```
Feature addition of two numbers.
Scenario basic addition.
      Given the calculator is open
      When I enter the number 3
      And I press the plus button
      And I enter the number 4
      Then I should see the number 7 as result.
```

There are only very few meta symbols and key words used. They are the first words in every line. The majority of the specification is actually close to real English sentences. This clear and simple way of describing features is, therefore, easy to understand to all project members. In contrast to other specification approaches, the major attention in Gherkin is set on having a simple and general language with only a few key words and an intelligent mapping of semantics. The interested reader can find a discussion on the benefits of using BDD and Gherkin as well as thoughts on the completeness of this language construct in (Martin 2008). As mentioned before,

the focus is not only on having a simple natural-language-based specification. Rather there is significant emphasis on being able to execute these specifications automatically as tests when the features are implemented by the software developers in the team. This automation is done by mapping sentences of the generated specification to methods and functions of the source code.

The language Gherkin is a very useful tool to tackle descriptions of use cases. Although, it is not possible to describe a manufacturing system only with use cases and expect an algorithm to generate a valid simulation model of the described system. We therefore see the necessity in a language to support more than simply describing use cases. There is in fact the need to describe structure and general behaviour of a system to be able to generate a simulation model.

## 3 Forklift Example

It is a common practise to evaluate the capabilities of a modelling approach starting with a small model from the intended domain. We have chosen a well-known model from the domain of production and logistics from Law and Kelton (2000). It is illustrated in Figure 2. The system modelled is a small production system utilizing forklift trucks for transport. It consists of five processing areas (work station) and one area for receiving and shipping. Within each production area there are a number of equivalent machines, which share a single queue. Three different types of jobs move through the system, differing mainly in processing times and order of their tasks.

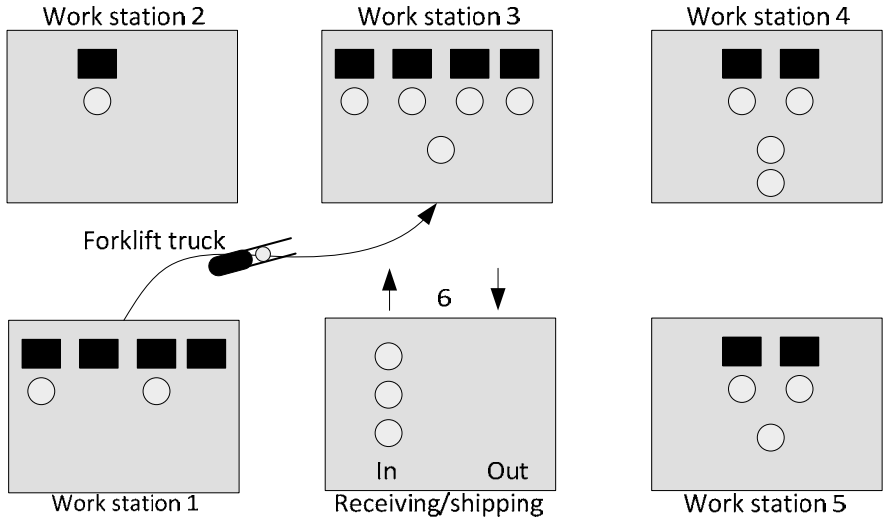



*Figure 2: Small production system with forklift trucks (Law and Kelton 2000, p. 685)*

# 4 Our New Natural-language-based Modelling Approach

In the following chapter, our basic language constructs are introduced with the help of the small example production system presented in Section 3. We start with describing the structure and then move on to behaviour.

## 4.1 Representation of Structure

Looking at the simple sentence “There is a tree in the forest, it has green leafs.”, real world objects are intuitively addressed with the phrase “There is a” followed by the name or type of the object. Hierarchical relations between objects can be described with “… in …” and properties are formulated with “… has a …”. Following this simple sentence we derive our language construct for modelling structure. Algorithm 2 shows an example of the structural constructs. It demonstrates the three structural elements: `Model`, `Workstation` and `Queue`. The `Workstation` with the name `WS1` is a sub-component of `Model M1.` It has two properties with the name `capacity` and `content`. These properties have initial values of 3 and 0 respectively. In the “has a …” statement the object is not directly addressed again, it is derived from context. Every “There is a …” phrase creates a context which can be used by subsequent phrases. Type and object names are used as identifiers, which an object is referred to globally. All structural elements can be modelled with these two phrases.

*Algorithm 2: Structure elements with properties*

```
There is a Model M1.
There is a Workstation WS1 in M1.
     has a capacity of 3.
     has a content of 0.
There is a Queue WS1_Q1 in WS1.
     has a capacity of 100.
     has a content of 0.
```

Algorithm 3 illustrates the representation of distances between the workstations from the small production system presented in chapter 3. Each distance is modelled as a property in the context of the structural element `TransportMatrix`. In common practise distances are not described in this way but in our feasibility study we use this approach as shortcut to avoid increased implementation effort.

*Algorithm 3: Structure element TransportMatrix*

```
There is a TransportMatrix TM1 in M1.
          has a WS1_WS2 of 150
          has a WS1_WS3 of 213
```

The flow through the production system is represented by additional phrases, which are describing job types. They are modelled from the viewpoint of the flow item. In

our sample production system a flow item can use machines and forklifts for a specified time and change their properties, e.g. a forklifts location. These resource allocations and state changes are modelled with the phrase "uses … for … from … to …". The variable elements of the phrase represent the allocated object, duration, and locations.

*Algorithm 4: Structure element JobType*

```
There is a JobType JT1 in M1.
      is created in RS1_Source every 5
      uses RS1_Q1
      uses FL1 for TM1.(FL1.location)_RS1/FL1.speed
           from FL1.location to RS1
      uses FL1 for TM1.RS1_WS3/FL1.speed from RS1 to WS3
      uses WS3_Q1
      uses WS3 for 0.25
```

Algorithm 4 illustrates the modelling of the structural element `JobType` with the name `JT1` as sub-component of `Model M1`. A flow item is created every 5 time units in the structural element that has the name `RS1_Source`. Every "uses …" phrase refers to one structural element to be allocated. Properties of other structural elements can be referenced by the element name and its properties name concatenated with a dot. This way of referring to properties seems more common for programming languages than for natural language, but is chosen for this demonstration as it is easier to handle for our current tool chain. With an improved tool chain it is possible to reference properties in a more natural way. Gelhausen (2010) is offering ideas on interpreting natural language which we are going to consider in future research.

## 4.2 Representation of Behaviour

While moving through the system a flow item is using different structural elements, which can change properties. Furthermore, certain conditions to use objects must be considered. In the sample production system a single machine can only process one flow item at a time.

Figure 3 illustrates the concept of our behaviour model. Only if the `enter condition step 1` is met, the `usage state step 1` is entered. As long as the succeeding structural enter condition of the element is not met, the flow item is kept in a loop between the states `after usage step 1` and `undo after usage step 1`. This loop is a simple way to get a full picture of the system state after leaving the current step. We therefore roll out all after usage changes to the system. In case the conditions to proceed are not met, the changes are simply rolled back to keep a consistent system state. This check loop is triggered again when relevant system changes have happened. The flow item can only leave the loop, if the enter condition for the following element is met, only by passing the `after usage state`.

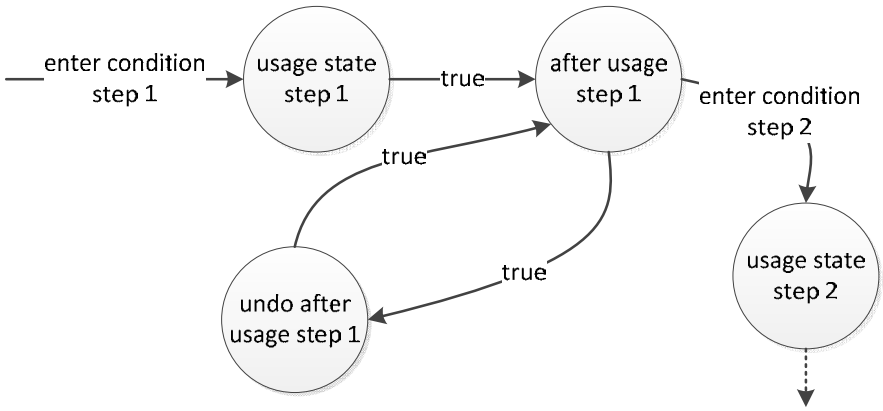



*Figure 3: Step processing state chart*

Algorithm 5 showcases the modelling of a workstation's behaviour. It is represented as `Businessrule` with the name `BR_W`. The same approach is used to access properties as in modelling structural elements. Additionally, mathematical operations and expressions to model conditions and state changes are supported. The delay of a flow item is modelled with the phrase "During use delay process for param1" and is dependent on the flow item. Values, given in the job type definition, can be accessed with the keywords "param0", "param1", "param2", and "param3". As an example, the JobType description in Algorithm 4 is considered again. The last step of the JobType is: "uses WS3 for 0.25". When a step is executed, businessrules can access its parameters. In this case "param0" is "WS3" the resource to be used and param1 refers to the duration of 0.25.

*Algorithm 5: Definition of behaviour*

```
There is a Businessrule BR_W for Workstation
      Before use check if content<capacity
      During use set content++
      During use delay process for param1
      After use set content--
      Undo after use set content++
```

## 5 Model Transformation

This section gives an overview on the technologies used to transform the model into executable code or into the model description of an arbitrary simulator. As mentioned before, there is a significant advantage in an automated transformation of a specification into a target model or program. In the case of Gherkin this transformation can be done with a number of tools. One of the most prominent ones is Cucumber. Since our natural language constructs differ from Gherkin not only in syntax but also in grammar it is not possible to do this mapping with the help of Cucumber. In fact a tool with a much broader scope is necessary. One such tool is

NatSpec from DevBoost GmbH. NatSpec is an Eclipse plugin, which is supporting the mapping of natural language phrases to Java methods (Seifert et al. 2013). The general idea behind these tools is to map clauses and sentences of a specification to methods of source code. During this mapping values of parameters are identified and handed to the target method. Then the method is executed.

In the following, this mapping is demonstrated for the “There is ...” clauses. Given our phrases NatSpec generates method headers to specify the mapping. The java method header for the “There is …” clause is shown in Algorithm 6.

*Algorithm 6: Generated method header for “There is …” clause*

```
@TextSyntax("There is a #1 #2 in #3.")
public SimulationEntity thereIsA_In_
     (String clazz, String name, String contextId,
          SimulationEntity context) { … }
```

The annotation, `@TextSyntax("There is a #1 #2 in #3.")`, is the basis of the mapping, as it contains the fixed parts of a clause and the position of all parameters of the method. Sentences like `There is a Workstation WS1 in M1.` are now recognizable and are interpreted as a method call with their specific parameters. The values written in the parameter positions are mapped to the corresponding ones of the method. The actual method (not shown here) contains functionality to create the referred object, in our example a workstation object, within the model and relates it to `M1`, as defined in the third parameter. This third parameter and the return value of the method represent a very common feature in natural languages, the semantic context. Usually a sentence in natural language does not stand for its own but has some links and references to previous and following sentences to avoid repetition and improve the flow of a text. A method can take into account the previous context by utilizing the context parameter and can hand its own context to subsequent sentences. An example of sentences using context can be seen in Algorithms 2 the “There is …” clause creates the context of a new workstation. The following “has…” statements build on this context and define parameters without additionally mentioning the workstation object.

Although this might seem complicated, this is a quite simple process. The implementation of nine methods has been sufficient to transform all possible models that can be represented in our natural language constructs.

Another important part of the model is the actual behaviour within the business rules. Besides clauses, which are interpreted based on the previously discussed mapping approach; there are conditions and state changes to be evaluated during the simulation run. These statements are transformed into strings of Java code and executed during runtime using a framework called Javassist. Javassist is a framework that allows manipulating java bytecode during runtime. It can be used to generate and use new Java classes during the execution of a Java program. We consider Javassist to be a very interesting and powerful technology but, with respect to the scope of this paper no further technical detail is given here (see Chiba and Nishizawa 2003 for more information on Javassist).

In summary, every sentence is mapped to a parameterized method call which creates parts of the model during its initialization. Behaviour is generated during model execution using Javassist.

# 6 Conclusion and Outlook

Based on the ideas of requirements engineering and BDD, we introduce a modelling approach based on natural language. The main idea is to utilize sentences in natural language to improve the communication between stakeholders and simulation experts during the early stages of simulation projects by formalizing specifications in a way that comes natural to everyone involved. We pointed out the benefits of writing executable specifications, which in an ideal case should be automatically translated into a simulation model of an arbitrary simulation system. The main benefit is the reduced need for interpretation of requirements by project members who are not domain experts. In addition, we see benefits for model validation by reducing miscommunication and helping to improve understanding of the real world system and its model. As this is only a first test case there are several open issues to be addressed in future research.

## 6.1 Extend the Language Construct with more Production-related Information

Currently, our language constructs only capture some major features of production and logistics systems. Although, they are able to show case the possibilities of the idea behind the modelling approach, we do feel it is necessary to include more production details to be able to model real systems.

## 6.2 Moving the Language Closer to a Natural Language

Looking at the way we are currently describing behaviour, there surely is some work to do before it is close to proper natural language. As mentioned before those statements which are still similar to source code can be easily replaced by natural language within the constructs but their interpretation and automated execution demand a tool chain with more capabilities than our current one. We therefore see the need to improve our toolchain and the tools within it to reach a language quality which is no longer cryptic to non-IT professionals. Due to the trend of language-based assistants in modern electronic gadgets and smartphones there is a broad interest in technologies that process and interpret natural language. We aim to utilize this new research to further improve the quality with which models can be developed and communicated.

## 6.3 Development of Editor Support

There is a need for a good development environment for modelling. Looking at modern software development and its integrated development environments we see enormous benefits in adapting common practises like auto-completion and automated error reporting to improve the usability of an environment and to help users develop their models.

## References

Ahmed, F.; Robinson, S.; Tako, A.A.: Using the structured analysis and design technique (SADT) in simulation conceptual modeling. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): Proceedings of the 2014 Winter Simulation Conference (WSC), Savannah (USA), 7.-10. December 2014, S. 1038-1049.

Chiba, S.; Nishizawa M.: An easy-to-use toolkit for efficient Java bytecode translators. In: Czarnecki, K; Pfenning F.; Smaragdakis, Y. (Hrsg.): Proceedings of the 2nd international conference on Generative programming and component engineering, Erfurt (Germany), 22.-25. September 2003, S. 364-376.

Gelhausen, T.: Modellextraktion aus natürlichen Sprachen. Karlsruhe: KIT Scientific Publishing 2010.

Huang, E.; Ramamurthy, R.; McGinnis, L.F.: System and simulation modeling using SysML. In: Henderson, S.G.; Billder, B.; Hsieh, M.-H.; Shortle, J.; Tew, J.D.; Barton, R.R. (Hrsg): Proceedings of the 2007 Winter Simulation Conference (WSC), Washington, D.C. (USA), 7.-10. December 2007, S. 796-803.

Law, A.M.; Kelton, W.D.: Simulation Modeling & Analysis. New York: McGraw-Hill, Inc. 2000.

Martin, R.C., 2008: The Truth about BDD. Accessed March 31. 2015. https://sites.google.com/site/unclebobconsultingllc/the-truth-about-bdd.

Pace, D.K.: Ideas about simulation conceptual model development. Johns Hopkins APL Technical Digest 21 (2000) 3, S. 327-336.

Rabe, M.; Spieckermann, S.; Wenzel, S.: Verifikation und Validierung für die Simulation in Produktion und Logistik. Berlin, Heidelberg: Springer 2008.

Robinson, S.: Conceptual modeling for simulation: Issues and research requirements. In: Perrone, L.F.; Wieland, F.P.; Liu, J.; Lawson, B.G.; Nicol, D.M. Fujimoto, R.M. (Hrsg.): Proceedings of the 2006 Winter Simulation Conference (WSC), Monterey (USA), 3-6 December 2006, S. 792-800.

Robinson, S.: Conceptual modeling for simulation. In: Pasupathy, R.; Kim, S.-H.; Tolk, A.; Hill R.; Kuhl M.E. (Hrsg.): Proceedings of the 2013 Winter Simulation Conference (WSC), Washington, D.C. (USA), 8-11 December 2013, S.377-388.

Robinson, S.; Brooks, R.; Kotiadis, K.; Van Der Zee, D.-J.: Conceptual Modeling for Discrete-Event Simulation. Boca Raton (FL): CRC Press 2011, S. 3-30.

Sargent, R.G.: Verifying and validating simulation models. In: Charnes, J.M.; Morrice, D.M.; Brunner, D.T.; Swain, J.J. (Hrsg.): Proceedings of the 1996 Winter Simulation Conference (WSC), Coronado (USA), 8.-11. December 1996, S. 55-64.

Seifert M.; Wende C.; Heidenreich F.; Nestler T.: Lean Modelling: Mit natürlicher Sprache zum Modell. Java Magazin 12.13 (2013).

Soeken, M.; Wille R.; Drechsler R.: Assisted behavior driven development using natural language processing. In: Furia, C.A.; Nanz, S. (Hrsg.): Objects, Models, Components, Patterns. Berlin: Springer 2012, S. 269-287.

Wynne, M.; Hellesøy, A.: The Cucumber Book: Behavior-driven development for testers and developers. USA: The Pragmatic Programmers, LLC 2012.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Farming for Mining – Entscheidungsunterstützung mittels Simulation im Supply Chain Management

## *Farming for Mining – Decision Support by Simulation in Supply Chain Management*

Markus Rabe, Anne Antonia Scheidler, TU Dortmund, Dortmund (Germany),
markus.rabe@tu-dortmund.de, anne-antonia.scheidler@tu-dortmund.de

Abstract: This paper constitutes an approach for combining Simulation and Knowledge Discovery techniques in the field of Supply Chain support and demonstrates relationships among them. We discuss how simulation and Knowledge Discovery will be applied as a part of data farming. In view of the above, we give an insight into first experiment setups using a prototypical implementation. In the experiment, the prototype for the Plant Simulation model is introduced and first results of the Supply Chain simulation scenarios are presented. We demonstrate how to transform the farmed model output into a transactional database which serves as a basis for the Knowledge Discovery process.

## 1 Einleitung

Die vielfältigen Aufgaben in der Supply Chain (SC) werden heutzutage oftmals mit Unterstützung des Knowledge Discovery in Databases (KDD) gelöst (Gabriel et al. 2009). Zu den Aufgaben zählen insbesondere auch Kernkompetenzen des Supply Chain Managements (SCMs), wie beispielsweise der Decision Support. Typische Fragestellungen des SCMs sind beispielsweise wie eine fristgerechte Lieferung sichergestellt werden kann, welche Transportmittel unter welchen Bedingungen am effizientesten sind oder wo und wann Komponenten in der SC benötigt werden. Betrachtet man diese Fragestellungen des SCMs, so sind insbesondere Fragen nach Zusammenhängen von großer Bedeutung. In diesem Umfeld können Techniken der Wissensentdeckung sinnvoll eingesetzt werden, um die Daten der SC-Systemlandschaft zu untersuchen. Da Techniken der Wissensentdeckung spezifische Anforderungen an die Datenbasis stellen (z. B. Stichprobenumfang, vorliegende Datenformate und eingehende Datenqualität) und auch prozessbegleitende Verifikation und Validierung nicht Teil der Wissensentdeckungsmodelle sind (Düsing 2010), genügt dieser Ansatz den heutigen Anforderungen der Industrie nicht. Um die vorhandenen Verfahrensschwächen aufzulösen, werden in diesem Beitrag Techniken der Wissensentdeckung mit Simulation kombiniert.

## 2 Supply Chain, Wissensentdeckung und Simulation - aktuelle Problemfelder

Nachfrageschwankungen sind ein hochaktuelles Thema in der SC und zählen zu den größten Herausforderungen im SCM (Capgemini 2011). In Zeiten globaler SCs sind Nachfrageschwankungen selten isolierte Effekte; vielmehr gibt es Korrelationen zwischen unterschiedlichen Effekten. So kann eine erhöhte Nachfrage eines Artikels x mit erhöhten Nachfragen der Artikel y und z korrelieren, obwohl die Artikelgruppen von x, y und z nicht in betriebswirtschaftlichen Beziehungen wie Produktgruppen zueinander stehen. Erhöhte Nachfragen können in der SC zu unterschiedlichen Problemen führen, beispielhaft wären das zunächst erhöhte Transportaufkommen und der darauf folgende Lieferengpass.

Die zugrundeliegenden Daten, in denen Korrelationen zu entdecken sind, sind schwer zu kategorisieren. Dies liegt zum einen an der Vielzahl der beteiligten Systeme und somit an den redundanten Dateninformationen, stark aggregierten Daten sowie heterogenen Datenformaten. Erschwerend wirkt sich zudem aus, dass die betriebswirtschaftliche Sicht auf die Daten und somit auch eine Vielzahl der zugehörigen Systemdokumentation nicht die Datenstrukturen und die möglichen Operationen auf den Datenstrukturen wiedergeben. Die Daten der Basissysteme, wie beispielsweise Datenbanken, dienen jedoch als Eingabedaten für Wissensentdeckungstechniken. Ein Großteil der SC-Daten sind Transaktionsdaten, auch als Bewegungsdaten bezeichnet, die Zustandsänderungen mittels Mengenangaben und Zeitstempeln beschreiben. Diese bilden die Grundlagen für die Wissensentdeckungen, beispielsweise von Wissen über Korrelationen, da die Transaktionen für die Daten- und Managementprozesse von grundlegender Bedeutung sind (Anane et al. 2002).

Die beschriebenen Korrelationen in der SC lassen sich mit Techniken des KDDs, einem Prozessmodell zur Wissensentdeckung (Fayyad et al. 1996), aufdecken. Hierbei ist der zentrale Prozessschritt - das Data Mining - so bedeutend, dass heutzutage die Begriffe KDD und Data Mining oftmals synonym verwendet werden (Adriaans und Zantinge 1996). Data Mining deckt Aufgabenstellungen aus den Bereichen der Segmentierung, Klassifizierung, Prognose oder Abhängigkeitsanalyse ab (Petersohn 2005). In der Logistik wurden Problemstellungen aus der Kundensegmentierung, Prognosen im Bereich Controlling, Materialbedarfsplanung und große Bereiche der Finanzdienstleistungen bereits mittels des Einsatzes von Data Mining gelöst (Gabriel et al. 2009). Auch Problemstellungen der SC, wie die bereits erwähnten korrelierten Nachfrageschwankungen, aber auch Fragestellungen nach Gruppierungen von Artikeln oder Transportfahrzeugen, profitieren von der Unterstützung des Data Minings (Rabe und Scheidler 2014). So lässt sich eine korrelierte Nachfrageschwankung mit Regellernern aufdecken, und verschiedene Clusteralgorithmen gruppieren Artikel, Fahrzeuge oder auch Kunden und Lieferanten (Düsing 2010).

Der wesentliche Unterschied von KDD und traditionell-logistischen Verfahren, wie der Clusteranalyse zur Set-Bildung bei Artikeln, liegt in der Mächtigkeit der Verfahren. Während logistische Verfahren im konkreten Beispiel reine Artikeleigenschaften zur Gruppenbildung heranziehen, kann sich das KDD aller vorhandenen Kennzahlen bedienen, die im Zeitalter von Big Data vollautomatisch von den logistischen Systemen protokolliert werden (Mayer-Schönberger und Cukier 2013).

Aus dem entdeckten Wissen des KDDs kann das SCM Handlungsanweisungen, beispielsweise für das SC-Bestellwesen, ableiten. Das Bestellwesen kann im konkreten Fall der Artikelkorrelation eine Clusterung der betroffenen Artikel bilden. Eine Alternative wäre eine Clusterung der artikelproduzierenden Unternehmen, um ein günstiges Transport-und Lagernetzwerk zu entwickeln. Diese Cluster können jedoch im Regelfall nicht ohne Veränderungen in die bestehende SC-Landschaft integriert werden.

Veränderungen können in einer SC aufgrund ihrer emergenten Eigenschaften und inhärenten Wechselwirkungen in der Regel nicht partiell durchgeführt werden. Um die Auswirkungen der unterschiedlichen Wissensintegrationsmöglichkeiten auf Produktion, Lagerhaltung und Distribution zu testen, bedient sich die Logistik der ereignisdiskreten Simulation (Discrete Event Simulation, DES), da es sich bei ihr um eine sehr effiziente Methode zur Handhabung zeitvarianter Systeme handelt (Hellström und Johnsson 2002). Bei dem Aufbau des Simulationsmodells spielt die Frage nach dessen Zweck eine herausragende Rolle (Robinson, 1994). Denn nur unter Berücksichtigung des „Warum?“ erhält man geeignete Simulationsergebnisse, die den Einsatz von Simulation rechtfertigen (Rabe et al. 2008).

Am Anfang ist es sinnvoll, das KDD auf Echtdaten durchzuführen, um eine konkrete Hypothese für die Simulation zu entwickeln. Jedoch können Echtdaten häufig nicht in ausreichender Menge oder angemessener Qualität zur Verfügung gestellt werden. Gründe hierfür sind mangelnde Datenqualität oder Sicherheitsbestimmungen in Unternehmen (Hildebrand 2011). Zusätzlich erschwert eine komplexe Systemlandschaft heterogener Systeme mit redundanter Datenhaltung eine zweckgebundene Datenextraktion. Unter diesen Bedingungen muss die Ausgangshypothese durch Expertenwissen oder andere Quellen gebildet werden, da ohne adäquate Datenbasis kein Anwendungsfall für das KDD gegeben ist. In beiden Fällen ist die Bildung der Starthypothese nur ein Spezialfall des nachfolgenden Prozesses aus DES und KDD, der im nachfolgenden Kapitel explizit beschrieben wird.

## 3 Verbindung DES und KDD

Die Veränderungen der SC werden in der Praxis des SCMs als „What-if-Szenarien“ simuliert (Konrad und Stache 2005). Ein „What-if-Szenario“ der Simulation produziert als Simulationsergebnis große Datenmengen. Der Unterschied dieser Datenmenge zu den Echtdaten ist, dass hier die Ausgangssituation (Simulationsaufbau und Parameter) bekannt und kontrollierbar ist. Veränderungen in der Ausgangssituation können gezielt vorgenommen werden, um Veränderungen in der Datenbasis zu produzieren. Dies ist ein Vorgang, der in der Literatur auch im Kontext des Data Farmings bekannt ist (Horne 1997). Der Begriff stammt von der Idee des bewussten „Kultivierens“ von Daten mittels Simulationsszenarien und dem nachfolgenden „Ernten“ der Daten aus Simulationsläufen, oder wie Sanchez (2014) sagt: “They grow data from their models“. Das Ergebnis des Data Farmings ist eine Datenlandschaft, die mittels verschiedener Techniken analysiert werden kann. Horne (2005) nennt als mögliche Techniken Data Mining, Data Analysis und Data Visualization. Diese Techniken sind jedoch nicht ohne Zwischenschritte oder Transformationen auf ein Simulationsergebnis anwendbar. Rabe und Scheidler (2014) haben gezeigt, dass insbesondere die Vorverarbeitung der Daten, die ein essentieller Teil der Wissensentdeckung ist, auch in der Verbindung von Simulation und KDD

von grundlegender Bedeutung ist. Daher wird in diesem Beitrag die Verbindung der Simulation zum KDD, also dem Prozessmodell der Wissensentdeckung (Fayyad et al. 1996), gesetzt. Die erzeugte Datenlandschaft ist durch zwei Dimensionen gekennzeichnet. Zum einen die Dimension Volumen, die sich aus der Anzahl der einzelnen Replikationen ergibt. Zum anderen die Dimension Komplexität, die sich im Kontext von Datenmengen auch als Varietät in der Literatur finden lässt (Che et al. 2013). Die Komplexität der Datenlandschaft wird durch den Aufbau der Simulation bestimmt. Die erzeugte Datenlandschaft kann mittels des KDDs analysiert werden, um zu prüfen, ob durch die Umstrukturierung der SC-Landschaft neue, eventuell ungewollte, Korrelationen entstanden sind (Farming for Mining).

Im konkreten Fall der Clusterung des Materials kann ein Simulationsexperiment die unterschiedlichen Neustrukturierungen von Lagerzonen und Transportmodalitäten simulieren. Auf den so gewonnenen Ergebnisdaten kann mittels der Assoziationsanalyse nach neuen Korrelationen, beispielsweise mit weiteren Artikelgruppen, gesucht werden. Eine Bewertung der neuen Korrelation oder eine Ursachenanalyse ist jedoch durch eine Assoziationsanalyse nicht möglich. Die aus dem KDD gewonnen Erkenntnisse können über neue Hypothesen in einem iterativen Prozess in die Simulation zurückgeführt werden. Hier führt die Hypothese zu einer Veränderung des Simulationsaufbaus und der Simulationsparameter. Dadurch erfolgt durch mehrere Simulationsläufe aus der ursprünglichen Ist-SC ein Vorschlag für die umgestaltete Soll-SC (Abb. 1).

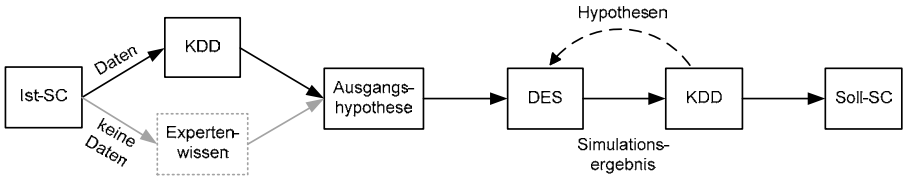



*Abbildung 1: KDD und Simulation*

## 4 Das Simulationsmodell für Transaktionsdaten

Das vorliegende Simulationsmodell basiert auf der Arbeit von Arndt (2014) und ist am Fachgebiet „IT in Produktion und Logistik“ entwickelt worden. Das Modell ist in Plant Simulation implementiert worden. Plant Simulation wird von Siemens Industry Software GmbH & Co. KG vertrieben. Die Simulationssoftware, die Einsatz im Umfeld der digitalen Fabrik findet, gestattet dem Benutzer Prozesse aus den unterschiedlichen Bereichen zu modellieren und zu simulieren. Das Leistungsspektrum umfasst die Abbildung der Materialflussteuerung, Visualisierungen in 2D und 3D, Bausteine und Klassenbibliotheken sowie Analysetools (Bangsow 2010).

Abbildung 2 zeigt einen einfachen Ausschnitt des Plant Simulation SC-Prototyps mit Lieferanten und Kunden. Hierbei liegt der fachliche Fokus auf der unternehmensübergreifenden Modellierung einer globalen SC. Das Simulationsmodel dient zu Generation von Transaktionsdaten einer globalen Automobilwertschöpfungskette. Die Eingabegrößen des Modells sind auf der linken Seite dargestellt und umfassen Warenausgang, Lieferdauer und Lieferabweichung. Die Ausgabegrößen sind Lieferdatum (LD), Soll-Wareneingangsdatum (Soll-WE). Ist-Wareneingangsdatum

(Ist-WE) sowie die Anzahl der Wareneingänge, die momentan auf einen fixen Wert eingestellt sind. Die Simulation startet mit dem Warenausgang eines Lieferanten, der die bewegten Elemente (BE) in erster Annäherung mit normalverteiltem Abstand erzeugt. Ein erzeugtes BE wird an die nachfolgenden Stationen verteilt. Die Ankunft eines BEs an einer nachfolgenden Station symbolisiert das Eingangsdatum einer Lieferung. In Stationen können BEs eine bestimmte Zeit warten, was die Lieferdauer einer Sendung nachstellt. Die Lieferdauer ist in das Modell als diskrete, empirische Verteilung eingebunden. Zuletzt werden Sendungen nach diversen Bearbeitungszeiten, Zwischenlagern und Lieferungen zur Senke transportiert. Die Senke stellt den tatsächlichen Wareneingang beim Original Equipment Manufacturer (OEM) dar.

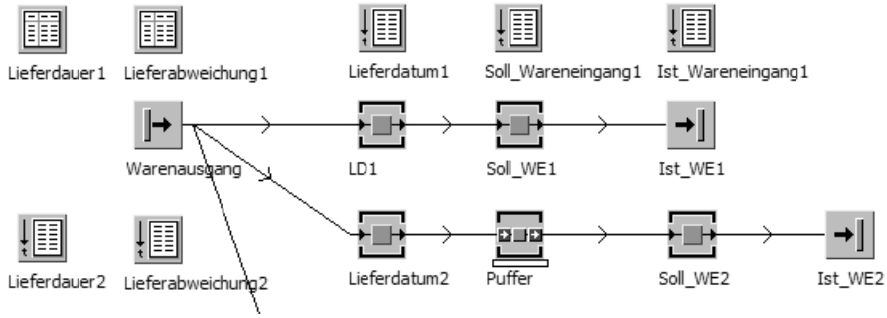



*Abbildung 2: Ausschnitt des SC-Modells (Plant Simulation)*

Technisch gesehen wird nach Eintreffen des BEs in der Senke die BE vernichtet und der Liefervorgang ist abgeschlossen. Die Dokumentation der Simulationsausgabedaten erfolgt über Zeitleisten in Tabellenform die aus Werten (1 für Ankunft, 0 für Ausgang) und Zeitstempeln bestehen. Die Zeitstempel sind, wie auch der Zeitanteil von Soll-WE und Ist-WE, vom Datentyp „DateTime“. In der Tabelle wird für jedes ankommende BE in den entsprechenden Stationen ein Eintrag erstellt und das Ergebnis einer Replikation ist für jede relevante Station eine Tabelle von Ausgabedaten. Tabelle 1 zeigt einen Ausschnitt einer aggregierten Sicht aus Warenausgang, Soll-WE und Ist-WE eines Lieferpfades. Die Tabelle wurde um eine alphanumerische ID angereichert. Der alphabetische Teil der ID stellt die Replikation dar, den numerische Teil einer BE-Aktion in der jeweiligen Replikation.

Tabelle 1 zeigt vier Transaktionsdatensätze (A1 bis A4) und ihre Zeitstempel. Diese Tabelle kann, mit leichter Modifikation der ID zu Surrogatschlüsseln, in eine relationale Transaktionsdatenbank überführt werden. Die Transaktionsdatenbank dient als Quelle für KDD-Verfahren. Je nach eingesetztem Verfahren ist zu prüfen, ob die Daten im Anschluss an die Simulation für die Weiterverarbeitung durch das KDD in andere Darstellungsformen transformiert werden müssen. Alternativ besteht die Möglichkeit, die benötigten Datenformate für das KDD direkt in der Simulation festzulegen. Beide Vorgehensweisen haben Vor-und Nachteile, die wir abschließend an dem konkreten Praxisbeispiel des vorliegenden Simulationsmodells diskutieren wollen.

*Tabelle 1: Transaktionsdaten Plant Simulation (Auszug)*

| | Warenausgang | | Soll-WE | | Ist-WE | |
|---|---|---|---|---|---|---|
| ID | Zeit | Wert | Zeit | Wert | Zeit | Wert |
| A1 | 12.02.2015 00:00 | 1 | 12.02.2015 00:00 | 1 | 13.02.2015 00:00 | 1 |
| A1 | 12.02.2015 00:00 | 0 | 13.02.2015 00:00 | 0 | 13.02.2015 00:00 | 0 |
| A2 | 26.03.2015 16:42 | 1 | 27.03.2015 16:00 | 1 | 27.03.2015 16:44 | 1 |
| A2 | 27.03.2015 16:42 | 0 | 27.03.2015 16:42 | 0 | 27.03.2015 16:44 | 0 |
| A3 | 25.04.2015 15:56 | 1 | 25.04.2015 15:56 | 1 | 02.05.2015 15:56 | 1 |
| A3 | 25.04.2015 15:56 | 0 | 02.05.2015 15:56 | 0 | 02.05.2015 16:26 | 0 |
| A4 | 14.05.2015 23:25 | 1 | 17.05.2015 23:25 | 1 | 18.05.2015 23:25 | 1 |
| A4 | 17.05.2015 23:25 | 0 | 18.05.2015 23:55 | 0 | 18.05.2015 23:25 | 0 |

Als Beispiel soll die Erweiterung der Transaktionen im Simulationsmodell um Mengenattribute diskutiert werden. Tabelle 2 zeigt vier Transaktionen mit den Mengenangaben der Lieferkette.

*Tabelle 2: Alternative A - Transaktionsdaten mit Mengenattributen (Auszug)*

| | Warenausgang | | Soll-WE | | Ist-WE | |
|---|---|---|---|---|---|---|
| ID | Zeit | Menge | Zeit | Menge | Zeit | Menge |
| A1 | 12.02.2015 00:00 | 100 | 12.02.2015 00:00 | 100 | 13.02.2015 00:00 | 100 |
| A2 | 12.02.2015 00:00 | 105 | 13.02.2015 00:00 | 105 | 13.02.2015 00:00 | 105 |
| A3 | 26.03.2015 16:42 | 150 | 27.03.2015 16:00 | 160 | 27.03.2015 16:44 | 150 |
| A4 | 27.03.2015 16:42 | 2000 | 27.03.2015 16:42 | 2000 | 27.03.2015 16:44 | 2000 |

Eine Vielzahl von KDD-Verfahren kann die numerischen Mengenangaben nicht verarbeiten. Alternative A besteht in einer Modifikation des Simulationsmodells. Die Mengenangaben können in der Simulation durch Container (z.B. M1 – M4) in geeigneter Abstraktionsstufe ersetzt werden (Tab. 3).

*Tabelle 3: Alternative B - Transaktionsdaten mit Containern (Auszug)*

| | Warenausgang | | Soll-WE | | Ist-WE | |
|---|---|---|---|---|---|---|
| ID | Zeit | Menge | Zeit | Menge | Zeit | Menge |
| A1 | 12.02.2015 00:00 | M1 | 12.02.2015 00:00 | M1 | 13.02.2015 00:00 | M1 |
| A2 | 12.02.2015 00:00 | M1 | 13.02.2015 00:00 | M1 | 13.02.2015 00:00 | M1 |
| A3 | 26.03.2015 16:42 | M1 | 27.03.2015 16:00 | M2 | 27.03.2015 16:44 | M1 |
| A4 | 27.03.2015 16:42 | M4 | 27.03.2015 16:42 | M4 | 27.03.2015 16:44 | M4 |

Die Alternative A besteht in einer nachträglichen Diskretisierung der numerischen Attribute in eine benutzerspezifische Anzahl von Containern mit geeigneter Zuordnungsfunktion (z.B. nach Häufigkeitsverteilung der Werte). Alternative B erspart mögliche Transformationsschritte vor den KDD-Verfahren, was sich positiv auf die Gesamtdauer des Verfahrens auswirkt. Allerdings ist die erzeugte Datenlandschaft spezifisch auf das Verfahren angepasst und kleine Verfahrensänderungen erfordern mitunter größere Anpassungen im Simulationsmodell. Alternative A ist im Durchschnitt zeitaufwendiger als Alternative B, da die nachträglichen Transformationen zusätzliche Zeit in der Prozesskette benötigen. Der Vorteil bei Alternative A besteht jedoch in der Flexibilität, denn bei Veränderungen im KDD-Bereich müssen nur die Transformationsschritte adaptiert werden. Dies ist zumeist schneller und kostengünstiger zu bewerkstelligen als Anpassungen im Simulationsmodell.

## 5 Zusammenfassung und Ausblick

Der vorliegende Beitrag diskutiert die Grundlagen der Entscheidungsunterstützung mittels Simulation im Umfeld des SCMs. Hierzu wurden die aktuellen wissenschaftlichen Grundlagen und Problemstellungen beleuchtet und die beteiligten Themengebiete miteinander verknüpft. Es wurde gezeigt, wie Daten generell mittels eines Data-Farming-Simulationsmodells in Plant Simulation erzeugt werden können. Im Anschluss wurde demonstriert, wie diese Daten aus dem Simulationsmodell (Simulationsergebnisse) exportiert werden können. Die damit verbundenen grundlegenden fachlichen Überlegungen wurden an zwei Alternativen für den Import in die KDD-Umgebung diskutiert. Als Ergebnis der Untersuchungen, wurde festgestellt, dass bei dem Experimentaufbau (vgl. Abb. 1) eine Entscheidung zwischen Prozessgeschwindigkeit (Alternative B) und Prozessflexibilität (Alternative A) zu treffen ist. Beide Varianten zeigen, wie die prinzipielle Überführung der Simulationsergebnisse in das KDD durchzuführen ist. Ebenso wird dargelegt, welche Konfigurationsmöglichkeiten dem Prozessmodellierer zur Verfügung stehen.

Die kommenden Forschungsarbeiten umfassen zum einen weitere Untersuchungen an der Schnittstelle von Simulation und KDD im SC-Kontext, um die automatische Datenverarbeitung zu unterstützen. Hierzu wird in den nächsten Schritten der Import der KDD-Ergebnisse (Hypothesen) in das Simulationsmodell diskutiert. Gleichzeitig

wird das Simulationsmodell kontinuierlich erweitert, um die Komplexität der resultierenden Datenbasis zu erhöhen und in Folge den Informationsgehalt der Simulationsergebnisse zu steigern. Die Erweiterungen der nächsten Phase umfassen Mengenattribute und kurzfristige Änderungen innerhalb der Bestellprozesse sowie die Generierung zusätzlicher Transaktionsattribute.

## Literatur

Anane, R.; Younas, M.; Tsai, C.-F.; Chao, K.-M.: Agent-based transactional framework for the supply chain. In: International Conference on Machine Learning and Cybernetics, Peking (China), 2002-11-04, 2002, S. 1956–1961.

Arndt, V.: Ereignisdiskrete Simulation einer Supply Chain zur Generierung von Transaktionsdaten. Dortmund, Technische Universität Dortmund, Fachgebiet IT in Produktion und Logistik, Masterarbeit, 2014. Online verfügbar unter http://www.itpl.mb.tu-dortmund.de/rabe/?page=theses&lang=en

Adriaans, P.; Zantinge, D.: Data mining. Harlow (England), Reading (MA): Addison-Wesley 1996.

Bangsow, S.: Manufacturing simulation with Plant Simulation and SimTalk. Berlin: Springer 2010.

Capgemini: Nachfrageschwankungen sind größte Herausforderung im Supply Chain Management. Pressemitteilung, online verfügbar unter http://www.de.capgemini.com/news/herausforderung-supply-chain-management, geprüft am 11.02.2015.

Che, D.; Safran, M.; Peng, Z.: From Big Data to Big Data Mining: Challenges, issues, and opportunities. In: Hutchison, D.; Kanade, T.; Kittler, J.; Kleinberg, J.M.; Mattern, F.; Mitchell, J.C. et al. (Hrsg.): Database Systems for Advanced Applications. Berlin, Heidelberg: Springer 2013, S. 1–15.

Düsing, Roland (2010): Knowledge Discovery in Databases. In: Chamoni, P.; Gluchowski, P. (Hrsg.): Analytische Informationssysteme. Business-Intelligence-Technologien und -Anwendungen, 4. Aufl. Berlin [u.a.]: Springer, S. 282–295.

Fayyad, U.M.; Piatetsky-Shapiro, G.; Smyth, P.: From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence 17 (1996) 3, S. 37–54.

Gabriel, R.; Gluchowski, P.; Pastwa, A.: Data warehouse & data mining. Herdecke, Witten: W3L 2009.

Hellström, D.; Johnsson, M.: Using discrete event simulation in supply chain planning. In: Solem, O. (Hrsg.): Proceedings of the 14th Annual Conference for Nordic Researchers in Logistics, Trondheim (Norway), June 13-14, 2002, S. 13–28.

Hildebrand, K. (Hg.): Daten- und Informationsqualität. Wiesbaden: Vieweg + Teubner 2011.

Horne, G.E.; Brandstein, A.G., 1997: Data Farming: A meta-technique for research in the 21st Century. Naval War CollegeNewport, RI, November 1997.

Horne, G.E.; Meyer, T.E.: Data Farming: Discovering surprise. In: Kuhl, M.E.; Steiger, N.M.; Armstrong, F.B.; Joines, J.A. (Hrsg.): Proceedings of the 2005 Winter Simulation Conference, Piscataway: IEEE Press 2005, S. 1082–1087.

Konrad, G.; Stache, U.: Theorie, Anwendbarkeit und strategische Potenziale des Supply Chain Management. Wiesbaden: Deutscher Universitätsverlag/GWV Fachverlage 2005.

Mayer-Schönberger, V.; Cukier, K.: Big Data. Die Revolution, die unser Leben verändern wird. München: REDLINE 2013.

Petersohn, H.: Data Mining. Verfahren, Prozesse, Anwendungsarchitektur. München, Wien: Oldenbourg 2005.

Rabe, M.; Wenzel, S.; Spieckermann, S.: Verifikation und Validierung für die Simulation in Produktion und Logistik. Berlin, Heidelberg: Springer 2008.

Rabe, M.; Scheidler, A.A.: An approach for increasing the level of accuracy in Supply Chain Simulation by using patterns on input data. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): Proceedings of the 2014 Winter Simulation Conference, Piscataway: IEEE Press 2014, S. 1897-1906.

Robinson, S.: Successful simulation. A practical approach to simulation projects. London, New York: McGraw-Hill 1994.

Sanchez, S.M.: Simulation experiments: Better data, not just big data. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): Proceedings of the 2014 Winter Simulation Conference, Piscataway: IEEE Press 2014, S. 805–816.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Domänenspezifische Sprache für ein simulationsunterstütztes Supply Chain Design

## *Domain-specific Language for Simulation Application in Supply Chain Design*

Matthias Parlings, Marco Motta, Fraunhofer-Institut für Materialfluss und Logistik, Dortmund (Germany), matthias.parlings@iml.fraunhofer.de, marco.motta@iml.fraunhofer.de

Philipp Sprenger, Graduate School of Logistics, Dortmund (Germany), Philipp.Sprenger@iml-stipendiat.fraunhofer.de

Abstract: Supply Chain Design (SCD) as a strategic task of supply chain management is carried out based on aggregated forecasted information. Therefore, a dynamic discrete event simulation is not used regularly in SCD. In contrast, the method of simulation, especially the model-based approach, offers various advantages in supply chain design. In this paper, results from the research project "Supply Chain Design", founded by the German Federal Ministry of Education and Research, are presented. The paper focuses on the idea of a domain-specific modelling language. This language translates the modelling elements of the application (application language) into the technical modelling elements of the tool (simulation tool language) and vice versa on different levels of aggregation. As a result, the complexity of configuring a simulation model is minimised. Within the project, a service-based software prototype was developed and applied in a practical case of an automotive first tier supplier.

## 1 Einleitung – Motivation und Problemstellung

Das Supply Chain Design als strategische und langfristige Aufgabe des Supply Chain Managements unterliegt einer Verstetigung. Die Ursachen hierfür sind volatile Umfeldbedingungen und die immer schnellere Einführung neuer Produkte in neue Märkte. In der Folge wird das Planungsumfeld auf Grund dieser rasanten Entwicklungen dynamischer, ist mit diversen Unsicherheiten, wie z. B. Nachfrageschwankungen, behaftet und erfordert eine erhöhte Anpassungsfähigkeit der Supply Chain. Die Gestaltung von Strukturen und Prozessen wird in den Fokus gerückt, da die hohen Anpassungsbedarfe alleine aufgrund von taktischen und operativen Maßnahmen kaum zu bewältigen sind.

Die Maßnahmen zur Gestaltung werden im Supply Chain Design (SCD) nach Strategie, Struktur- und Prozessgestaltung untergliedert (Parlings et al. 2013). Die so entstehenden vielfältigen Planungs- und Gestaltungsaufgaben des SCD werden in der Praxis durch unterschiedliche Teams bearbeitet. Die zum Teil in der Praxis nur rudimentäre Integration von angewendeten Planungsansätzen führt zur Bildung von Planungsinseln, d. h. abgegrenzte Bereiche oder Teams, die an einem Problem arbeiten ohne Wechselwirkungen zu benachbarten Bereichen zu beachten. Zudem sind die Planungsprozesse zeitintensiv und nicht aufeinander abgestimmt und werden zum Teil sogar auf unterschiedlichen Datenbasen bzw. Planungsständen ausgeführt. Strategische Entscheidungen bestimmen jedoch die Infrastruktur ganzer Wertschöpfungsnetzwerke und determinieren die Wettbewerbsfähigkeit von SC-Partnern über mehrere Jahre. Mit den gegenwärtigen in der Praxis gelebten Planungsprozessen und -strukturen können die steigenden Herausforderungen daher nicht mehr bewältigt werden.

Eine Methode, die detaillierte Bewertungen komplexer Systeme und die Bewertung von Handlungsalternativen ermöglicht, ist die Simulation. Ein wesentliches Hemmnis zur Nutzung der Simulation ist jedoch, dass der Modellierungsaufwand sehr hoch ist und daher i.d.R. ein großes Expertenwissen erfordert. Zudem ist die Modellierung spezifisch an die Nutzung des Simulationswerkzeugs gekoppelt, was die Wieder- und Weiterverwendbarkeit des Modells erschwert (Kuhn et al. 2010).

Gegenstand dieses Beitrags ist die Vorstellung der Methode der Simulation zur Vernetzung der heutigen Planungsinseln. Dies geschieht anhand der Darlegung eines Modellierungsdienstes für das SCD, welcher der Anforderung eines erhöhten Anpassungsbedarfs der Supply Chains und Ihrer Elemente gerecht wird. Die Komplexität des Netzwerks wird für die Planungsaufgaben der Netzwerkgestaltung beherrschbar, indem die Modellierung auf einem für die Fragestellungen des SCD angemessenen Abstraktionsniveaus durchgeführt wird. Umgesetzt wird die Modellierung durch eine bausteinbasierte Domain-specific Modelling Language (DSL).

Der Beitrag ist wie folgt aufgebaut: Im Anschluss an die Einleitung werden bestehende Planungsansätze und Bausteinkonzepte für die Modellierung im SCM vorgestellt. Dadurch erfolgt eine Einordnung des Beitrags in das Forschungsthema. Anschließend wird das Konzept für eine bausteinbasierte DSL im SCD beschrieben, die den Modellierungsaufwand auf das benötigte Maß des SCD reduziert und für eine Anwendung ohne spezifisches Simulationswissen nutzbar macht. Anschließend wird die Integration der DSL in das Konzept eines szenariofähigen Modellierungsdienstes dargestellt und am Beispiel eines Anwendungsfalls in der Automobilbranche durch Konfiguration unterschiedlicher Szenarioalternativen demonstriert.

# 2 Stand der Forschung

## 2.1 Planungsansätze und Werkzeugunterstützung im Supply Chain Design

In der Forschung werden unterschiedliche Vorgehensmodelle und Planungsmethoden für die langfristigen Gestaltungsaufgaben beschrieben, welche entweder sehr generell sind oder extrem spezifisch auf bestimmte Gestaltungsaufgaben, wie beispielsweise die Standortentscheidung, angepasst sind (Sprenger et al. 2014). Zu-

gleich werden Bewertungs- und Analysemethoden aber auch die Optimierung oftmals isoliert und nicht integriert genutzt, sodass eine Bewertung komplexer Supply Chains ganzheitlich kaum möglich ist.

Eine Analyse von Planungsansätzen für das Supply Chain Design hat offenbart, dass nahezu alle Ansätze einem ähnlich strukturierten Problemlösungsprozess folgen. Die Prozesse können dabei in die Hauptschritte Strategie- und Zieldefinition, Problemdefinition, Datensammlung, Modellbildung, Szenarioanalyse, Auswertung und Reporting unterteilt werden (Parlings et al. 2015).

Die meisten existierenden Planungsansätze verwenden keine durchgängige Werkzeug-Unterstützung zur Lösung der Aufgabenstellungen. In der Regel werden in den einzelnen Lösungsschritten spezifische softwarebasierte Analysemethoden eingesetzt (Parlings et al. 2015). Grundsätzlich können die eingesetzten Werkzeuge jedoch in die beiden Klassen der analytischen Methoden und der Simulation eingeteilt werden (Seidel 2009).

Um einen umfassenden Überblick über den Stand der Technik im Bereich der softwarebasierten Werkzeugunterstützung im SCD zu erhalten, wurde im Rahmen des Forschungsprojektes SCD eine interne Analyse (unter den Anwendungspartnern) und eine externe Analyse (Anbieter-Recherche) durchgeführt (Parlings et al. 2015). Die Recherche-Ergebnisse zeigen, das nur wenige Werkzeuge flexibel genug sind, um unterschiedliche Aufgaben des SCD, wie z. B. Standortentscheidungen, Transportnetzgestaltung, Kapazitätsdimensionierung zu lösen. Zudem wird nur spezielle Planungsfunktionalität geboten, die eine Verbindung der einzelnen Planungsprozessschritte nicht unterstützen. Ein in der Praxis häufig eingesetztes Werkzeug ist Microsoft Excel.

Um einen Beitrag zu leisten, die Forschungslücke der durchgängigen Werkzeugunterstützung im SCD zu schließen, wurde im BMBF-geförderten Forschungsprojekt "Supply Chain Design" im EffizienzCluster LogistikRuhr ein generalisierter Planungsworkflow zur Vernetzung der Planungsinseln erarbeitet. Eine ganzheitliche Entscheidungsfindung wird durch servicebasierte on-demand Technologien unterstützt, die in dem Planungsworkflow adressiert werden. Kernelement des Lösungsansatzes sind drei interoperable szenariofähige Dienste (Modellierungs-, Simulations- und Reportingdienst), welche den Einsatz der Methode der Simulation zur Entscheidungsfindung im SCD begünstigen sollen. Da der Gegenstand dieses Beitrags der Modellierungsdienst ist, wird im folgenden Abschnitt explizit auf den Forschungsstand auf dem Gebiet der Modellierung von Supply Chains eingegangen.

## 2.2 Bausteinkonzepte für die Modellierung von Supply Chains

Vor dem Hintergrund steigender Komplexität, Unsicherheit und Dynamik und im Hinblick auf eine angestrebte „Vernetzung“ der Planungsinseln bietet die Simulation als Werkzeug zur ganzheitlichen Betrachtung des Netzwerks große Potentiale. Es stehen seit Jahren Simulationswerkzeuge wie OTD-NET (Wagenitz 2007) oder Supply Net Simulator (Stäblein et al. 2007) zur Verfügung, welche die Komplexität und Dynamik der Logistiknetzwerke durchgängig abbilden können (vgl. Stäblein et al. 2007; Reeker et al. 2011; Wagenitz 2007). Speziell für Simulations-Tools kann aber festgestellt werden, dass die Komplexität der Bedienung in der Regel einen Simulationsexperten erfordert. Trotz hoher Potentiale wird daher die Simulation zu selten zur Unterstützung des SCD eingesetzt (vgl. Kuhn et al. 2010).

Zur Reduzierung der Modellierungskomplexität existieren für Ausschnitte von Supply Chains Bausteinkonzepte. In diesem Zusammenhang sind die Arbeiten von Hermes und Kemper zu nennen, die im Rahmen der Entwicklung von simulationsgestützten Konzepten für die Programm-und Reihenfolgeplanung in der Automobilindustrie Bausteine zur Darstellung von Beschaffungs- und Distributionsprozesse und den jeweiligen Wirkzusammenhängen verwendet haben. Dabei haben sich die Autoren an den jeweiligen Logistikprozessen orientiert und Modellierungsbausteine für die Überführung in ein Simulationsmodell konzipiert (Hermes 2011; Kemper 2012). Den strukturellen Rahmen für die konzipierten Modellierungsbausteine bildet in beiden Ansätzen das Dortmunder Prozessketteninstrumentarium (Kuhn 1995). Die Ansätze sind jedoch nur für Teilausschnitte der Supply Chain verwendbar und die Granularität der Bausteine ist für die Anwendung im Supply Chain Design zu hoch. Dennoch bieten die Arbeiten eine gute Basis für die Entwicklung eine DSL für das Supply Chain Design.

Einen intralogistischen Bausteinentwicklungsansatz liefert Lütjen, welcher zur Beschreibung intralogistischer Prozesse in der Fertigung für die Automobilindustrie verwendet werden kann. Lütjen benutzt die eigene domänenspezifische Modellierungssprache GRAMOSA (Graphische Modellierung und simulationstechnische Analyse), speziell für eine Modellierung von Fertigungs- und Logistikprozessen. Das Metamodell GRAMOSA umfasst 12 Objektklassen, wie Auftragsauslöser, Fertigungselement, Materialflussobjekt, Organisationseinheit, Prozess, Prozessauftrag, Prozessfolge, Puffer, Stückliste, Zeitgeber, Vorgang und Fertigungsfehler (Lütjen 2014). Für die Aufgaben des Supply Chain Design ist der Ansatz von Lütjen nicht direkt zu verwenden, da der Fokus auf intralogistischer Modellierung liegt. Für zukünftige Entwicklungen ist zu prüfen, inwiefern eine Kopplung der SCD-spezifischen DSL auf Makroebene mit der Modellierungssprache GRAMOSA ein Potenzial zur durchgängigen Modellierung bis auf die Werkstattebene bietet.

# 3 Bausteinbasierte Domain-specific Language für das Supply Chain Design

## 3.1 Grundsätze einer Domain-specific Language im SCD

Die grundsätzliche Anforderung an die Modellierung im Supply Chain Design besteht darin, die Komplexität des Netzwerks für die Planungsaufgaben der Netzwerkgestaltung beherrschbar zu machen, indem die Modellierung auf einem für die Fragestellungen des SCD angemessenen Abstraktionsniveau durchgeführt wird. Umgesetzt wird die Modellierung durch eine bausteinbasierte Domain-specific Modelling Language. Die DSL übersetzt zwischen den Anwendungen der realen Umgebung mit dem immanenten fachlichen Logistikwissen in die technischen Elemente eines Simulationsmodels. Zugleich werden die wesentlichen Informationen und Parameter der logistischen Elemente einer Supply Chain auf einer aggregierten Ebene in Modellierungsbausteinen zusammengefasst. Die Informationen und Parameter beschreiben die Strukturen, die interne Kontrolllogik und die Zusammenhänge der wesentlichen Bestandteile einer realen Supply Chain, wie beispielsweise Werke, Transportrelationen, Lager und Cross-Docks (Hegmanns et al. 2014). Durch die Entwicklung vorkonfigurierter Modellierungsbausteine wird der Bedarf an benötigten Informationen auf ein Minimum reduziert. Zugleich wird die

Modellierung ohne spezifisches Hintergrundwissen über die technischen Ausführungen der Simulation ermöglicht, indem die relevanten logistischen Spezifikationen abgefragt werden und die technischen Details im Hintergrund hinterlegt sind. Für die technische Übersetzung wird das standardisierte XML-Format genutzt, das zugleich eine weitergehende Nutzung des Modells für unterschiedliche Simulationswerkzeuge ermöglicht. Im folgenden Abschnitt wird das Konzept der im Forschungsprojekt „Supply Chain Design" im EffizienzCluster LogistikRuhr entwickelten DSL innerhalb des Modellierungsdienstes erläutert.

## 3.2 Ebenen der DSL für die durchgehende Aufgabenbearbeitung im Supply Chain Design

Die entwickelte DSL umfasst drei Ebenen, die fachliche, die disaggregierende und die aggregierende Ebene. Die Bedeutung der drei Ebenen wird im Workflow der Einbettung der Simulation in die Planungsaufgaben deutlich.

Grundlegend wurde als durchgängige Datenbasis das Konzept einer Supply Chain-Database (SCDB) umgesetzt, welche objektorientiert die Parameter und Attribute der in der fachlichen DSL definierten „Makro-Objekte“ der Supply Chain aufnimmt und als durchgängige Datenbasis funktioniert.

Wie in der folgenden Abbildung 1 dargestellt, werden zunächst mittels eines Szenariodienstes ein leeres Szenario angelegt oder ein neues Szenario aus einem bestehenden Modell abgeleitet. Mittels des Modellierungsdienstes kann der manuelle Aufbau der Supply Chain Struktur erfolgen. Auch die Parametersetzung wird hier auf einem dem SCD angepassten Aggregationsniveau ermöglicht. Alternativ ist die Anbindung an andere externe Datenquellen wie bspw. bestehende ERP-Systeme möglich. Als Austauschformat wurde ein generisches XML-Format definiert.

Die Supply Chain Database (SCDB) beinhaltet die zu untersuchenden Szenarien in der Makro-Bausteinwelt des SCD. Diese hochaggregierte Bausteinwelt beinhaltet bspw. ein gesamtes Crossdock mit seiner Routingfunktionalität als gekapselten Baustein.

Da es nicht das Ziel des Projektes war, ein vollständig neues spezifisches Simulationswerkzeug zu schaffen, wurde die disaggregierende Ebene der DSL entwickelt. „Dissaggregierend“ bedeutet in diesem Zusammenhang, dass im Rahmen der DSL-Entwicklung generische Referenzmodelle für die Struktur und das dynamische Verhalten der Elemente der Macro-Bausteinwelt definiert wurden. Die DSL löst nun diese Struktur- und Verhaltensmuster in die für Simulationsmodelle typischen Bausteine mit Warteschlangen- und Pufferfunktionalität auf. Speziell für das Simulationswerkzeug OTD-NET wird hier der fachliche „Makro-Baustein“ in die erforderlichen „Micro-Bausteine“ des Simulationswerkzeugs und deren Verknüpfung aufgelöst. Parallel ist eine Ausleitung der Beschreibung der „Micro-Ebene“ in einem generischen XML-Format möglich, so dass auch andere Simulationswerkzeuge wie bspw. PlantSimulation mit entsprechenden vormodellierten Netzwerken zur Abbildung der Micro-Bausteine angebunden werden können.

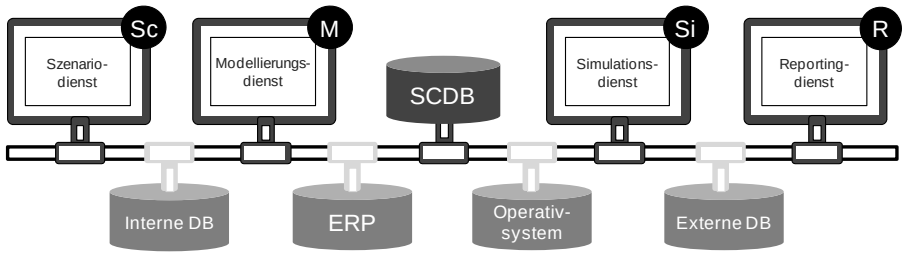



*Abbildung 1: Planungsworkflow und Informationsquellen*

Mittels des folgenden Simulationsdienstes kann der Simulationskern angebunden werden und es können konzeptionell bspw. vordefinierte Experimentserien simuliert werden. Die Ergebnisse werden in Datenbanken mit einem szenariofähigen Core-Schema gespeichert und können durch den Reporting-Dienst ausgewertet werden. An dieser Stelle ist die aggregierende DSL relevant, da aufgrund der Übersetzungslogik die Einzelergebnisse auf „mikroskopischer" Ebene wieder auf das SCD-Aggregationsniveau übertragen werden. Dies bedeutet, dass der Anwender Kennzahlen auf der Ebene der „Macro-Bausteine" auswertet und interpretiert und die Analyse eben nicht bis auf die werkzeugspezifische Ebene der „Micro-Bausteine" auszuführen ist. Durch die Übersetzung der Referenzmodelle der Macro-Bausteine in Form der DSL kann diese „Interpretationsleistung" in der Aggregation weitestgehend automatisiert werden. Ebenso ist für die Interpretation der Ergebnisse und die Rückschlüsse auf die zu gestaltende Supply Chain die fachliche DSL zu berücksichtigen. Der beschriebene Ablauf ist in Abbildung 2 noch einmal als Kreislauf dargestellt.

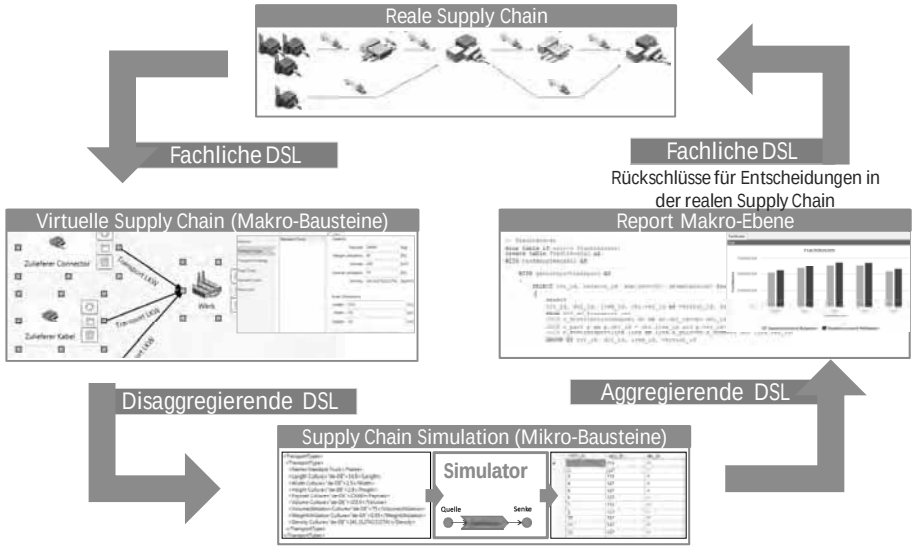



*Abbildung 2: Ebenen der Domain-specific Language (DSL)*

# 4 Anwendung am Beispiel eines internationalen Automobilzulieferers

Der entwickelte Lösungsansatz wird am Beispiel eines international agierenden 1st-Tier Lieferanten der Automobilindustrie erprobt. Dabei kann insbesondere die Bedeutung des modellbasierten Simulationsansatzes zur Vernetzung von Planungsinseln herausgestellt werden. Im folgenden Anwendungsfall steht die Gestaltung von Kabelsatz Supply Chains, genauer die Standortentscheidung für Kabelsatzproduktionsstandorte in interkontinentalen Wertschöpfungsnetzwerken im Fokus. Die Standortplanung als Problemkomplex des SCD unterliegt in der Automobilindustrie diversen auftrags- und standortspezifischen Planungsunsicherheiten. Um diesen Unsicherheiten und den sich schnell ändernden Anforderungen an Standorte und Wertschöpfungsnetzwerke gerecht zu werden, bedarf es sowohl eines permanenten Vorgehens zur Generierung potentieller Standortszenarien als auch Methoden zur Modellierung und Bewertung komplexer Supply Chains im dynamischen Umfeld.

Ausgehend von dem Planungsansatz für robuste Standortentscheidungen (siehe Sprenger et al. 2014) kann mit Hilfe verschiedener Voruntersuchungen die Anzahl möglicher Standortszenarien auf ein handhabbares Level potentieller Standortalternativen reduziert werden. Diese Standortalternativen können mit Hilfe des Modellierungsdienstes spezifiziert, parametrisiert und beschreiben werden und so die unterschiedlichen Supply Chain Designs modelliert werden. Die Überführung der realen Supply Chains des Automobilzulieferers in ein Simulationsmodell erfolgt mit dem oben beschriebenen DSL-Konzept und dem Modellierungsdienst. In Abbildung 3 ist ein typisches Kabelsatz Supply Chain Netzwerk abgebildet.

Das dargestellte Supply Chain Netzwerk kann mit Hilfe der Makro-Bausteine „Supplier (BTO/BTS)“, „Transport“, „Consolidation Center“, „Crossdock“, „Production“ und „Sink“ in eine virtuelle Supply Chain überführt werden.

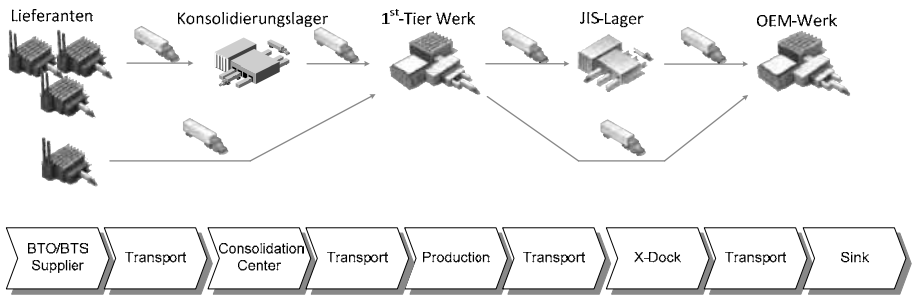



*Abbildung 3: Beispiel für eine Kabelsatz-Supply-Chain*

Die fachliche DSL wurde so entwickelt, dass notwendige Parameter manuell oder durch Datenimporte aus verknüpften Datenquellen gesetzt werden können. Am Beispiel des Makro-Bausteins „Production“ können z. B. Stücklisten hinterlegt, Produktionsprogramme und Kapazitäten definiert sowie Lieferteile zu unterschiedlichen Lieferanten inkl. Bestellvorlaufzeiten zugeordnet werden. Im Folgenden werden zwei alternative fiktive Standortszenarien betrachtet. Das notwendige Simulationsmodell wurde mit dem Modellierungsdienst erstellt (Abb. 4).

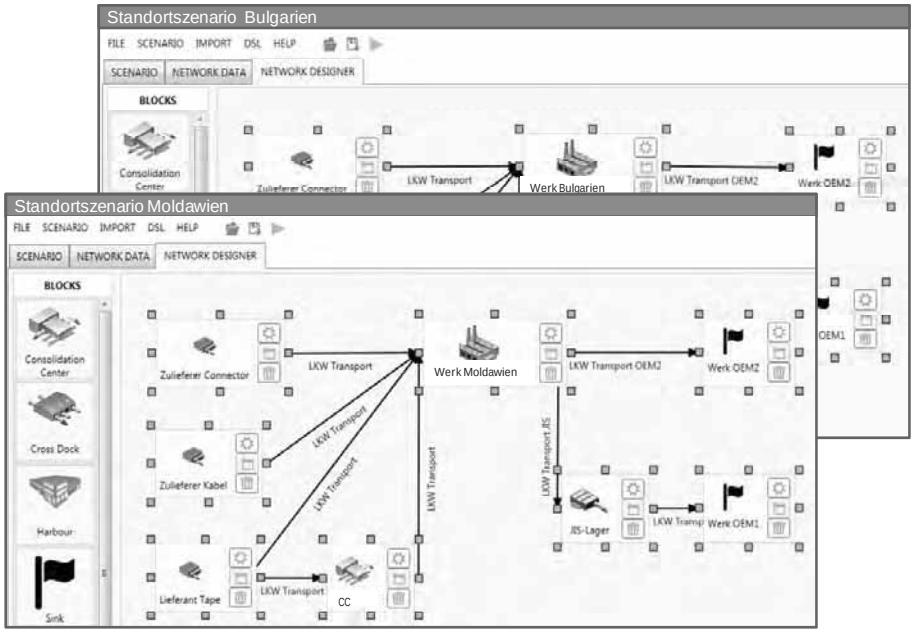



*Abbildung 4: Modellierte Supply-Chain-Netzwerke für fiktive Standortszenarien*

Mit Hilfe der disaggregierenden DSL wird das Modell im Simulationsdienst in simulatorverständliche Bausteine übersetzt und ermöglicht so die Durchführung unterschiedlicher Simulationsläufe. Mittels aggregierender DSL können die einzelnen Simulationsläufe im Report auf der Makro-Ebene hinsichtlich zuvor definierter Zielgrößen ausgewertet werden. Der Reportingdienst ermöglicht einen Szenariovergleich und stellt je nach Anwenderfunktion z. B. Reports für Entscheidungsträger des Top-Managements bzw. für die Rolle eines Supply Chain Planers bereit. Durch die Darstellung von Unsicherheiten wie z. B. Transportzeiten mittels stochastischer Verteilungsfunktionen oder die Berücksichtigung von ereignisdiskreten Zusammenhängen wie z. B. Bestelltermine, Fahr- oder Werkspläne in Verbindung mit der Simulation als Bewertungsmethode können Sensitivitätsanalysen durchgeführt werden, die Aussagen bezüglich der Robustheit unterschiedlicher Standortszenarien ermöglichen. In Bezug auf das Beispiel der fiktiven Standortszenarien kann der Einfluss erfolgskritischer Einflussgrößen wie z. B. schwankende Frachtraten und Wechselkurse oder Volumenvariationen bezogen auf die Fracht- oder auch Gesamtkosten ausgewertet werden, wie in Abbildung 5 dargestellt ist.

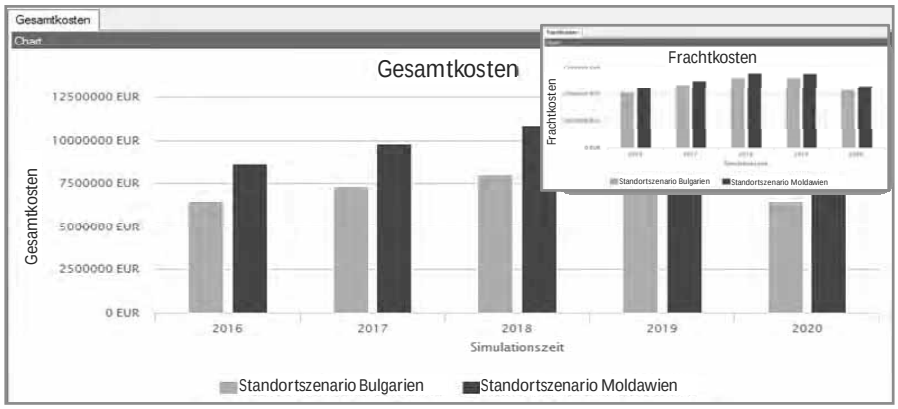



*Abbildung 5: Szenarienvergleich im Reportingdienst*

Durch die Gestaltung verknüpfter Dienste und die Integration von Datenquellen kann die Ergebnisqualität entscheidend gesteigert und die beanspruchte Planungszeit erheblich reduziert werden.

## 5 Fazit und Ausblick

Die Methode der Simulation kann trotz der Verwendung unsicherer und aggregierter Eingangsdaten einen Beitrag zur Steigerung der Qualität der Planungsunterstützung im SCD leisten. Essentielle Voraussetzung hierzu ist es, die Einstiegshürde für die Anwendung der Simulation zu minimieren. Mit den vorliegenden Ergebnissen des Forschungsprojekts konnte gezeigt werden, dass über die Ebenen einer fachlichen, disaggregierenden und aggregierenden Domain-specific Language der Modellierungsaufwand gegenüber der eingesetzten Basissimulationssoftware (hier OTD-NET und PlantSimulation) drastisch reduziert werden konnte. Insbesondere die disaggregierende Ebene der DSL, die zu einer Modellübersetzung in die existierenden, „mikroskopischen" Bausteine des eingesetzten Simulationsinstrumentariums führt, stellt eine hohe Anschlussfähigkeit dar. Die im SCD entstandenen Modelle sind damit kompatibel zu den im Rahmen des Supply Chain Planning bereits erfolgreich eingesetzten Logistischen Assistenzsystemen (vgl. Henke und Motta 2014), die zur Entscheidungsunterstützung den betriebsbegleitenden Einsatz von Simulationsmodellen ermöglichen.

## Danksagung

Die Autoren danken dem Bundesministerium für Bildung und Forschung für die Förderung des EffizienzCluster LogistikRuhr-Forschungsprojekt Supply Chain Design (Förderkennzeichen 01IC12L03B). Unser Dank richtet sich ebenfalls an die Projektpartner des Forschungsprojektes für Ihre anwendungsorientierte Unterstützung und die Graduate School of Logistics der TU Dortmund.

## Literatur

Hegmanns, T.; Schwede, C.; Motta, M.; Henke, M.: Model-based technologies for tool support in supply chain risk management. In: BVL International: The Global Supply Chain Network: Logistics in the Networked Industry: Conference Papers, Köln 2014, S. 135 - 152.

Henke, M.; Motta, M.: IT im Supply Chain Management: Simulationsgestützte logistische Assistenzsysteme als Ansatz zur Steigerung der Supply Chain Agilität. In: Kille, C. (Hrsg.): Navigation durch die komplexe Welt der Logistik: Texte aus Wissenschaft und Praxis zum Schaffenswerk von Wolf-Rüdiger Bretzke. Wiesbaden: Springer Fachmedien 2014, S. 153-169.

Hermes, A. M.: Modellbasierte Bewertung von Potentialen einer distributionsorientierten Programm- und Reihenfolgeplanung in der Automobilindustrie. Dortmund: Praxiswissen 2011.

Kemper, J.: Modellgestützte Bewertung der Potentiale einer beschaffungslogistikorientierten Programmplanung in der Automobilindustrie. Dortmund: Praxiswissen 2012.

Kuhn, A.: Prozessketten in der Logistik: Entwicklungstrends und Umsetzungsstrategien. Dortmund: Praxiswissen 1995.

Kuhn, A.; Wagenitz, A.; Klingebiel, K.: Praxis Materialflusssimulation - Antworten oft zu spät. In: Wolf-Kluthausen, H. (Hrsg.): Jahrbuch der Logistik. Korschenbroich: Free Beratung 2010, S. 206-221.

Lütjen, M.: Modellierungskonzept zur integrierten Planung und Simulation von Produktionsszenarien entwickelt am Beispiel der CFK-Serienfertigung. Bremen, Universität Bremen, Dissertation 2014.

Parlings, M.; Klingebiel, K.; Cirullies, J.: A literature-based state of the art review on the identification and classification of supply chain design. In: Proceedings of the 17th Cambridge International Manufacturing Symposium, Cambridge 2013, Publikation ohne Seitenzahlen.

Parlings, M.; Sprenger, P.; Hegmanns, T.; Koßmann, D.: Modular IT-support for integrated supply chain design. In: Klumpp, M. (Hrsg.): Springer LNL Book – Bridging the Gap. Zur Veröffentlichung angenommen, 2015.

Reeker, C.; Hellingrath, B.; Wagenitz, A.: Bewertungsansatz zur ökologischen Gestaltung automobiler Logistiksysteme im Werksumfeld von Produktionsstandorten. In: Bogaschewsky, R.; Eßig, M.; Lasch, R.; Stölzle, W. (Hrsg.): Supply Management Research. Wiesbaden: Gabler 2011, S. 139-168.

Seidel, T.: Ein Vorgehensmodell des softwareunterstützten Supply Chain Design. Dortmund, Technische Universität Dortmund, Dissertation. Dortmund: Praxiswissen 2009.

Sprenger, P.; Parlings, M.; Hegmanns, T.: Planning Approach for Robust Manufacturing Footprint Decisions. In: Kersten, W.; Blecker, T.; Ringle, C. M.: Next Generation Supply Chains. Berlin: epubli 2014.

Stäblein, T.; Baumgärtel, H.; Wilke, J.: The Supply Net Simulator SNS: An artificial intelligence approach for highly efficient supply network simulation. In: Günther, H.-O.; Mattfeld, D. C.; Suhl, L. (Hrsg.): Management logistischer Netzwerke. Entscheidungsunterstützung, Informationssysteme und OR-Tools. Heidelberg: Physica 2007, S. 85–110.

Wagenitz, A.: Modellierungsmethode zur Auftragsabwicklung in der Automobilindustrie. Dortmund, Technische Universität Dortmund, Dissertation, 2007.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Entwicklung einer menschenähnlichen Puppe als Eingabegerät für die Menschsimulation

## *Design of a Humanlike Puppet as an Input Device for Human Simulation*

David Wiegmann, Holger Brüggemann, Ostfalia Hochschule für angewandte Wissenschaften, Wolfenbüttel (Germany), david.wiegmann@ostfalia.de, holger.brueggemann@ostfalia.de

Abstract: Ergonomic simulation offers high potential to improve ergonomics and productivity of future manual production processes. Nevertheless ergonomic simulation is used quite rarely in production planning processes, due to the high effort in time which is required to create human movements in the simulation. This paper presents a new approach to minimise the time required to create human simulation. A puppet is developed to be used as a tangible input device for digital humans in ergonomic simulation. The paper describes the state of the art of human motion creation in ergonomic simulation and the using of tangible input devices for other purposes. Furthermore the design, the properties and the functioning of the new input device is shown in this paper.

## 1 Einleitung

Innerhalb der arbeitswissenschaftlichen Menschsimulation wird der Mensch in ergonomischer Hinsicht in der digitalen Fabrik abgebildet. Sie dient dem Zweck, Arbeitssysteme unter ergonomischen, arbeitsmethodischen und zeitwirtschaftlichen Gesichtspunkten zu analysieren und zu bewerten (VDI 4499 Blatt 4). Die frühzeitige Durchführung von arbeitswissenschaftlichen Menschsimulationen im Produktionsplanungs- und Gestaltungsprozess bietet die Möglichkeit, frühzeitig Einfluss auf zukünftige Produkte, Betriebsmittel und Produktionsprozesse zu nehmen. Dies stellt ein großes Potenzial zur Verbesserung der Ergonomie und der Produktivität von manuellen Arbeitsplätzen und deren Planungsprozessen dar. Im Gegensatz zum Einsatz in der Produktentwicklung, z. B. im Fahrzeug- oder Flugzeugbau, hat die Bewertung von Arbeitsplätzen mithilfe der Menschsimulation in der industriellen Anwendung bisher weniger an Bedeutung erlangt (VDI 4499 Blatt 4).

Häufig erfolgt die Absicherung manueller Arbeitsplätze durch Prototypen, aus der Erfahrung von Experten oder virtuell ohne den Einsatz eines Menschmodells. Kommt dennoch die Menschsimulation zum Einsatz, dann meist bei gezielten,

kritischen Vorgängen oder neuen Konzepten (Rodríguez 2010). Der gravierendste Nachteil und somit das größte Hindernis eines stärkeren Einsatzes von arbeitswissenschaftlicher Menschsimulation ist der hohe zeitliche Aufwand für die Erstellung der Simulation (Mühlstedt 2012; Rodríguez 2010; Mun und Rim 2011).

Das BMBF-geförderte Forschungsprojekt Human Engineering Computer Interface (HECI), das am Institut für Produktionstechnik der Ostfalia durchgeführt wird, verfolgt einen neuen Ansatz zur Verkürzung der Erstellungszeit für eine arbeitswissenschaftliche Menschsimulation. Mit einem Eingabegerät in Form einer menschenähnlichen Puppe soll der Aufwand für die Generierung von Bewegungen des Menschmodells erheblich verringert werden. Das Eingabegerät soll direkt an das Simulationsprogramm Process Simulate von Siemens angebunden werden und dort das Menschmodell *Jack* steuern.

## 2 Stand der Wissenschaft und Technik

Der verhältnismäßig größte Zeitbedarf bei der Erstellung einer Menschsimulation fällt bei der Generierung von Bewegungen des digitalen Menschmodells an. Hierzu bieten die am Markt bestehenden Simulationstools unterschiedliche Methoden an. Diese und die Anwendung gestaltartiger Eingabegeräte soll im Folgenden aufgezeigt werden.

### 2.1 Bewegungsgenerierung in der Menschsimulation

Eine häufig angewandte Möglichkeit zur Bewegungsgenerierung ist das so genannte Keyframing, das ursprünglich in der Filmindustrie angewandt wurde. Hierbei werden charakteristische Körperhaltungen bzw. Posen (Keyframes) des Menschmodells generiert. Eine diskretisierte Bewegung wird durch die meist lineare Interpolation zwischen den Posen erzeugt (Abb. 1). Diese Art der Bewegungsgenerierung erzeugt aber nur dann realistische Bewegungen, wenn sich Start- und Endpose nicht zu stark voneinander unterscheiden. Dementsprechend viele Keyframes werden für eine realistische Bewegung benötigt (Rodríguez 2010). Die einzelnen Posen lassen sich grundsätzlich mit den Methoden der Vorwärtskinematik oder der inversen Kinematik generieren. Bei der Vorwärtskinematik muss jedes einzelne Gelenk ausgewählt und eingestellt werden. Das Ermitteln und Einstellen geeigneter Gelenkwinkel gestaltet sich hierbei als schwierig und zeitaufwändig. Bei der inversen Kinematik werden die Zielpositionen und -orientierungen der Endeffektoren des Menschmodells angegeben. Die daraus resultierenden Gelenkwinkel der gesamten kinematischen Kette des Menschmodells werden dann errechnet. Allerdings besitzen die digitalen Menschmodelle in der Regel mehr als sechs Freiheitsgrade, was zu Mehrdeutigkeiten führt (Rieseler 1992; Weber 2013). Als Abhilfe wird bisher versucht, Randbedingungen zu formulieren wie etwa die Minimierung der potentiellen Energie oder die minimale Quadratsumme der Winkelinkremente, damit das Menschmodell eine natürliche Haltung einnimmt (Arlt 1999). Bisher wurde allerdings noch keine korrekte Randbedingung gefunden, sodass hier Vereinfachungen getroffen werden (VDI 4499 Blatt 4).

Eine weitere Möglichkeit der Bewegungsgenerierung stellt das Motion Capturing dar. Hierbei führen Probanden die Bewegungsvorgänge der zukünftigen Arbeitsvorgänge durch. Dabei werden die Bewegungsdaten durch Verfolgung bestimmter

Schlüssel-Punkte am menschlichen Körper räumlich und zeitlich aufgezeichnet. Anschließend werden die Bewegungsdaten auf das digitale Menschmodell übertragen (Mapping) (Menache 2010). Das Motion Capturing führt zwar aus Prinzip zu realistischen Bewegungen (Arlt 1999). Jedoch ist der finanzielle Aufwand je nach System mit Hardwarekosten bis zu € 250.000 nicht unerheblich. Weiterhin ist die Durchführung der Bewegungsaufzeichnung mit hohem Aufwand verbunden, da häufig ein Prototyp von Produkt und/oder Betriebsmittel notwendig ist. Zusätzlich ist in der Regel eine extensive Datenaufbereitung notwendig (Menache 2010). Die resultierenden Bewegungen sind zudem unabhängig von Einflussparametern, d. h. Änderungen (z. B. durch eine neue Schrauberposition) sind zunächst nicht direkt in der Simulationsumgebung durchführbar (Arlt 1999).

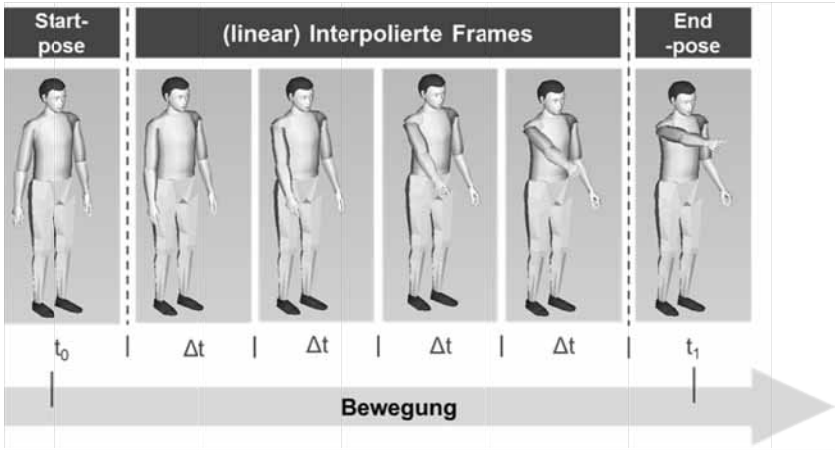



*Abbildung 1: Bewegungsgenerierung durch Keyframing*

Ein weiterer Ansatz zur Generierung von Bewegungen ist die parametrische Bewegungsgenerierung. Die Bewegung des Menschmodells erfolgt hierbei auf Basis von typischen, von Parametern abhängigen, idealen Bewegungsmustern, die in Laborstudien abgeleitet wurden (Schönherr 2014; Gläser et al. 2014). Bei der Erstellung der Simulation werden die Parameter zum einen vom Benutzer gesetzt (z. B. Art der Verrichtung) und zum anderen vom Programm aus der Simulationsumgebung ermittelt (z. B. Position des Betriebsmittels). Bei dieser Art der Bewegungsgenerierung besteht die Problematik, dass die Ableitung von Bewegungen aller möglichen Verrichtungen mit allen möglichen Parametern einen immensen Aufwand darstellt. Zusätzlich ist es bei komplexen Bewegungen (etwa zur Kollisionsvermeidung) bisher noch nicht möglich ideale Bewegungsmuster abzuleiten (Gläser et al. 2014).

## 2.2 Gestaltartige Eingabegeräte

Als gestaltartige Eingabegeräte sollen in diesem Beitrag solche bezeichnet werden, die in ihrer physikalischen Form ein Abbild eines zu manipulierenden digitalen Modells darstellen. In anderen Bereichen wurden gestaltartige Eingabegeräte bereits erfolgreich eingesetzt und zeigten insbesondere, dass diese einfach und intuitiv zu bedienen sind. An dieser Stelle sollen ausgewählte Beispiele dargestellt werden.

Über den Einsatz gestaltartiger Eingabegeräte in der Filmindustrie berichten bereits Knep et al. (1995) mit ihrem *Dinosaur Input Device (DID)*. Im Film Jurassic Park wurden mit Eingabegeräten in Form von Dinosauriern virtuelle Dinosaurier in computeranimierten Szenen beim Keyframing manipuliert. Eine Weiterentwicklung dieser Technik wurde im Film Starship Troopers angewandt, wo für computeranimierte Szenen sog. *Bugs* mit einem entsprechenden Eingabegerät kontrolliert wurden (Hayes 1998). Den Einsatz eines menschenähnlichen Eingabegeräts beschreiben erstmals Esposito et al. (1995). Ihr Eingabegerät *Monkey* stellt ein elektromechanisches Skelett dar, dessen Maße vom 50. Perzentil der nordamerikanischen Bevölkerung abgeleitet sind und im Maßstab 1:3 verkleinert sind. Es besaß 35 Freiheitsgrade und wurde für Menschsimulationen im Rahmen des Produktentwicklungsprozess bei Flugzeugen eingesetzt. Abbildung 2 zeigt das Eingabegerät *Monkey* und das *DID*.

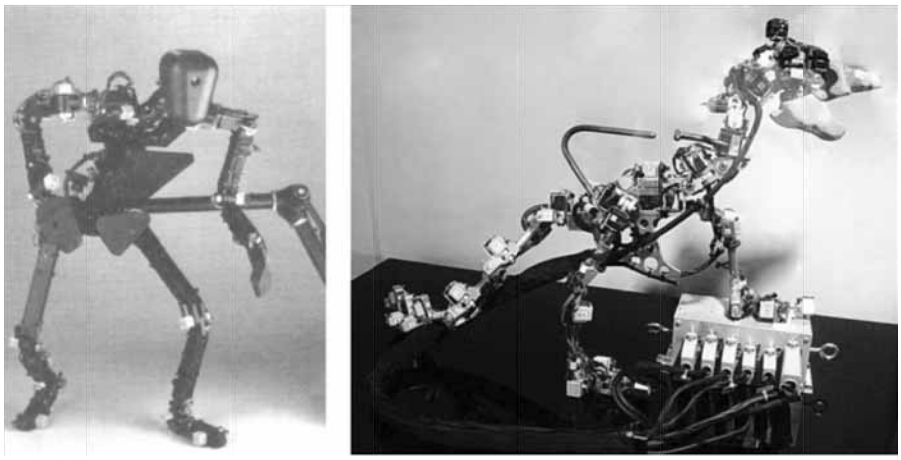


*Abbildung 2: Monkey (links, Hayes 1998) und DID (rechts, Knep 2014)*

Yoshikazi et al. (2011) entwickelten ein Eingabegerät in Form einer menschenähnlichen Puppe, deren Gelenke mit Antrieben ausgestattet sind. Sie verglichen in einer Nutzerstudie ihr Eingabegerät mit aktivierten Antrieben, mit passiven Antrieben und mit der Maus beim Posing. Im Rahmen der Studie konnten die Probanden mit der Puppe das Menschmodell schneller und präziser in eine Zielpose überführen. Bei aktiven Antrieben waren die Probanden zudem schneller als bei passiven Antrieben. Dennoch wird erwähnt, dass die Aussagekraft dieses Ergebnisses begrenzt ist, da keine auf das Menschmodell zugeschnittene passive Puppe zur Verfügung stand. Eine weitere Nutzerstudie zeigen Jacobson et al. (2014) auf. Sie verglichen ihr modulares, konfigurierbares System, das Gestalten in Form unterschiedlicher Skelette annehmen kann, mit der Maus. Ihre Studie ergab unter anderem, dass man mit ihrem System sich schneller an Zielposen annähert als mit der Maus.

Da die bisherigen Nutzerstudien zur Anwendung gestaltartiger Eingabegeräte keinen Bezug zur arbeitswissenschaftlichen Menschsimulation nehmen, ist es fragwürdig, inwieweit sich die Ergebnisse auf die Simulation menschlicher Arbeitsprozesse übertragen lassen.

# 3 Entwicklung eines gestaltartigen Eingabegeräts für die Menschsimulation

Im Folgenden soll der Aufbau und die Nutzung des auf die arbeitswissenschaftliche Menschsimulation zugeschnittenen Eingabegeräts vorgestellt werden, das im Rahmen des HECI-Projektes entwickelt wurde. Mit dem Eingabegerät in Form einer kleinen menschenähnlichen Puppe sollen Bewegungen des Menschmodells nach dem Keyframing-Ansatz erzeugt werden, indem die Gelenkwinkel der Puppe auf das digitale Modell übertragen werden.

## 3.1 Sensorik

Die Gelenkwinkel des Eingabegerätes werden durch kleine optische inkrementelle Drehgeber erfasst. Zusammen mit den Körperteilen der Puppe bilden sie ein elektromechanisches Trackingsystem. Die Drehgeber haben einen Gehäusedurchmesser von 7,2 mm bei einer Länge von 13,5 mm und einen Wellendurchmesser von 1,5 mm. Durch die Sensoren werden Gelenkdrehungen in Schritten von 0,225° erfasst. Die Drehbewegungen werden durch binäre Signale über einen Open-Collector-Ausgang kodiert ausgegeben. Durch das binäre Signal ist die Übertragung der Signale sehr robust. Im Gegensatz zu anderen Sensoren, die bei elektromechanischen Trackingsystemen zum Einsatz kommen, wie etwa Potentiometern, stellt ein Rauschen des Signals hierbei kein nennenswertes Problem dar. Auch die Erfassung des Winkels an sich ist sehr unempfindlich gegenüber Störungen im Vergleich zu anderen Trackingverfahren. Bei magnetischen Trackingverfahren kann es zu Störungen durch elektromagnetische Quellen oder metallische Komponenten kommen, bei optischen Verfahren können äußere Lichteinflüsse stören und bei Trackingverfahren auf Basis von Inertialsensoren kann der mit der Zeit ansteigende Winkelfehler (Drift) zu Problemen führen.

Da die Drehgeber nur einen Freiheitsgrad besitzen, müssen Gelenke mit mehreren Freiheitsgraden, durch eine entsprechende Anzahl seriell angeordneter Gelenke ersetzt werden. Die statischen Gelenkstellungen der mehrdimensionalen Gelenke, die für die Bewegungserzeugung nach dem Keyframing-Ansatz notwendig sind, können hierdurch weiterhin durch das Eingabegerät abgebildet werden. Weiterhin ist es notwendig die Gelenke vor der Verwendung des Eingabegerätes einmalig, durch eine Drehbewegung bis zu einer bestimmten Stellung zu referenzieren, da inkrementelle Drehgeber verwendet werden. In Abbildung 3 ist der Drehgeber sowie eine Gelenkeinheit mit einem und mit drei Freiheitsgraden dargestellt.

## 3.2 Mechanik

Das Eingabegerät basiert auf dem Menschmodell *Jack* aus Process Simulate und entspricht dem 50. Perzentil männlich nach DIN 33402 im Maßstab 1:5. Das Menschmodell *Jack* verfügt über insgesamt 68 Gelenke mit 135 Freiheitsgraden (Spanner-Ulmer und Mühlstedt 2009). Hieraus wurde für die Puppe ein reduziertes kinematisches Skelett mit insgesamt 22 Freiheitsgraden abgeleitet. Gelenkgruppen des Jack-Modells wie etwa der Wirbelsäule oder des Schlüsselbein-Schulter-Verbundes wurden hierbei durch je ein Gelenk im Skelett der Puppe ersetzt. Die Winkelstellung dieser einzelnen Gelenke dient als Stellglied, um die gesamte Gelenkgruppe des Jack-Modells zu manipulieren. Trotz dieser Vereinfachung gibt es eine hohe Übereinstimmung zwischen der Körperhaltung des Eingabegeräts und der des Jack-

Modells. Das digitale Modell kann weiterhin in vollem Umfang mit der Puppe manipuliert werden. Das Skelett sieht keine Manipulation der Hand- und Fußgelenke vor. Abbildung 4 stellt das Skelett des digitalen Modells *Jack* und das abgeleitete Skelett der Puppe gegenüber.

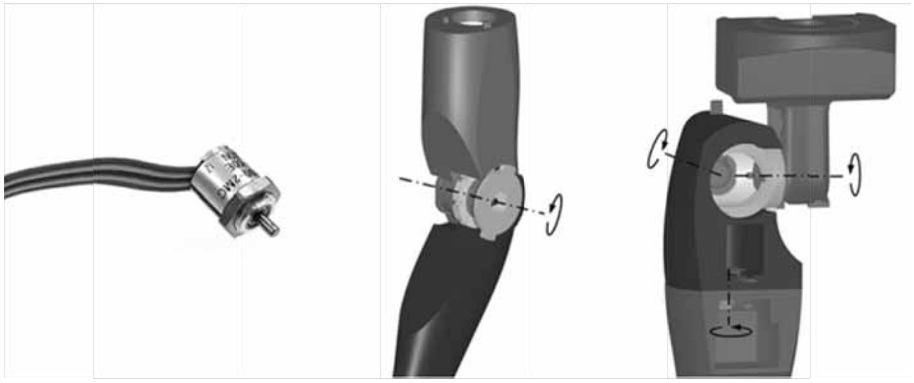


*Abbildung 3: Drehgeber (links), Gelenkeinheit mit einem Freiheitsgrad (Mitte) und mit drei Freiheitsgraden (rechts)*

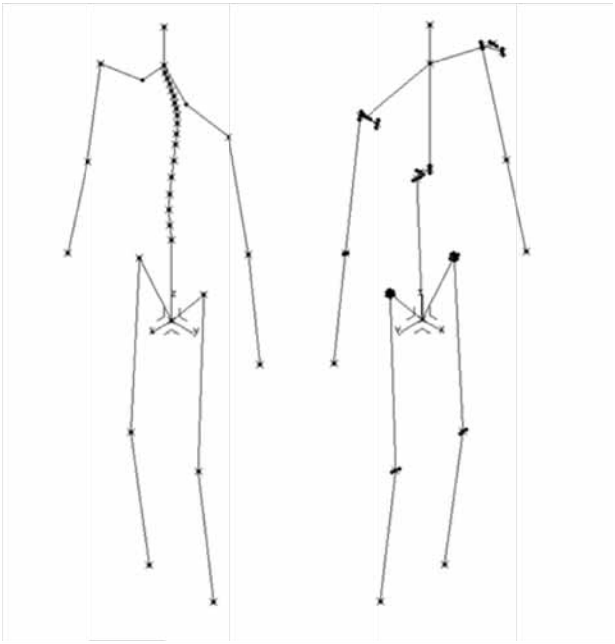


*Abbildung 4: Skelett des digitalen Modells (links) und abgeleitetes Skelett (rechts)*

Das Eingabegerät wurde anschließend auf Basis des Skeletts und der verwendeten Drehgeber konstruiert. Hierbei wurde ebenfalls versucht sich an der äußeren Gestalt des digitalen Menschmodells anzunähern. Während sich die äußere Haut des digitalen Modells selbst durchdringen kann und sich zudem auch verformen lässt, ist dies bei der Puppe nicht möglich. Damit trotzdem die Gelenke der Puppe die gleichen Winkelbereiche wie die des digitalen Modells abdecken können, wurden die Körperteile der Puppe an entsprechenden Stellen ausgespart. Die einzelnen Glieder der Gelenke werden durch Reibung in ihrer Stellung gehalten, sodass diese nach Betätigung in ihrer Position bleiben. Abbildung 5 stellt die Konstruktion des Eingabegeräts und das Menschmodell *Jack* gegenüber.

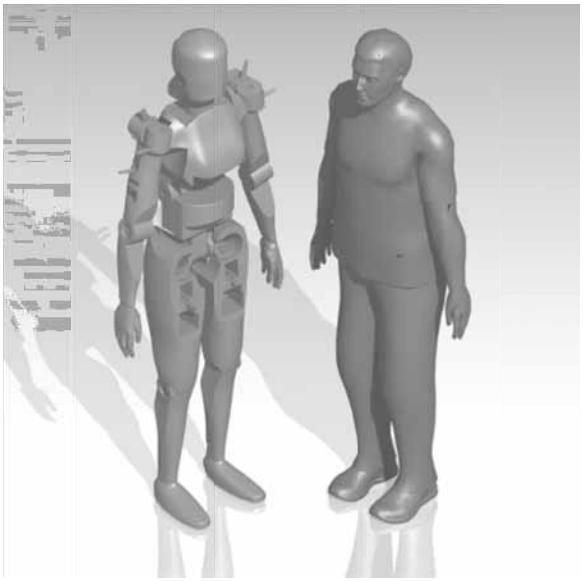


*Abbildung 5: Konstruktion des Eingabegeräts im Vergleich zum Menschmodell Jack*

## 3.3 Schnittstelle

Für das Eingabegerät wurde eine Schnittstelle zu Process Simulate geschaffen, sodass das Eingabegerät direkt in einer Simulationsumgebung genutzt werden kann. Hierfür wurde eine Steuereinheit entwickelt, die die Winkelstellungen des Eingabegeräts auf das digitale Menschmodell übertragen.

In der Steuereinheit werden die kodierten binären Signale der Sensoren zunächst von Mikrokontrollern erfasst. Diese dekodieren die Signale und wandeln sie in einen Winkel um. Die Mikrokontroller sind als Slave-Teilnehmer an einen Bus angeschlossen und senden den Winkel an einen Master-Teilnehmer, der über die USB-Schnittstelle an den Computer angeschlossen wird, auf dem die Simulation durchgeführt wird. Über die Programmierschnittstelle von Process Simulate wurde eine Anwendung programmiert, die die Winkelstellungen des Eingabegeräts aus der Steuereinheit liest. Anschließend erfolgt ein Mapping der eingelesenen Winkelstellungen auf das digitale Menschmodell in der Simulationsumgebung, sodass die

Körperhaltung der Puppe mit der des digitalen Menschmodells annähernd übereinstimmt.

## 3.4 Bedienung und erster Test des Eingabegeräts

Gemäß beschriebener Konstruktion wurde ein Prototyp angefertigt, der in Abbildung 6 dargestellt ist. Dabei wurden funktionsgebende Teile durch klassische Fertigungsverfahren und komplexe formgebende Bauteile durch additive Fertigungsverfahren hergestellt. Zur besseren Handhabung wurde die Puppe an einem Mast befestigt, sodass der Anwender mit beiden Händen die Gelenke der Puppe einstellen kann. Das Eingabegerät wird direkt am Simulationsarbeitsplatz in der Nähe des Bildschirms positioniert, auf dem die Simulationsumgebung dargestellt wird. Indem die Gelenke der Puppe eingestellt werden und dadurch diese eine gewünschte Körperhaltung einnimmt, wird direkt das digitale Menschmodell manipuliert. Der Anwender kann dann Mithilfe des Eingabegerätes innerhalb der Anwendung Keyframes erstellen, aus denen dann die gewünschte Bewegung erzeugt wird.

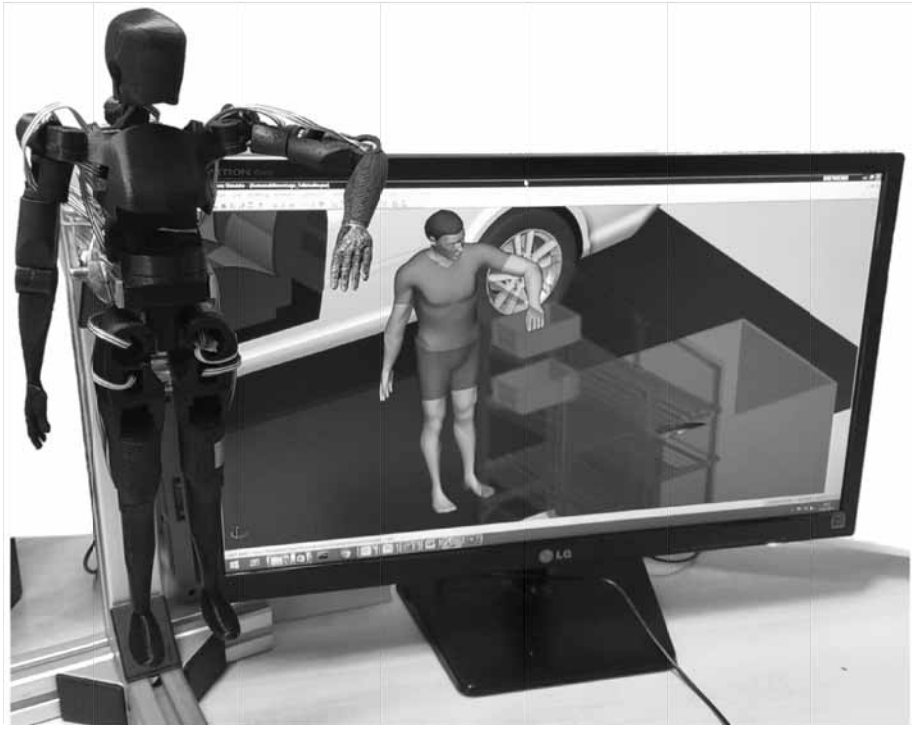


*Abbildung 6: Prototyp des Eingabegeräts*

In einem ersten Test, in denen Probanden das Menschmodell sowohl mit dem Eingabegerät als auch mit der Maus in bestimmte Körperhaltungen brachten, erhielt die Puppe ein positives Feedback. Die Probanden gaben an, dass sie im Vergleich zur Maus zufriedener mit dem Eingabegerät waren. Weiterhin nahmen sie die Bedienung als komfortabler wahr und konnten zudem Keyframes in kürzerer Zeit erzeu-

gen. Für verlässliche quantitative Aussagen hierzu werden in Zukunft weitere Versuche durchgeführt.

Im weiteren Ablauf der Entwicklung soll das Eingabegerät weiter optimiert werden, mit dem Ziel die Eingabezeiten während der Simulation zu verkürzen. Ein Aspekt ist hierbei die Übereinstimmung der Körperhaltung zwischen der Puppe und dem digitalem Menschmodell zu verbessern, damit der Benutzer während des Posings weniger Rückmeldung aus der Simulationsumgebung benötigt.

## 4 Zusammenfassung und Ausblick

Durch frühzeitige Einflussnahme auf den Produktionsplanungs- und Gestaltungsprozess bietet die arbeitswissenschaftliche Menschsimulation großes Potenzial die Ergonomie und die Produktivität manueller Produktionsprozesse zu verbessern. Dennoch wird die arbeitswissenschaftliche Menschsimulation bisher verhältnismäßig selten angewandt, was hauptsächlich auf den hohen zeitlichen Aufwand für die Generierung von Bewegung des digitalen Menschmodells zurückzuführen ist. Die bisherigen Methoden zur Bewegungsgenerierung bieten bisher noch keine zufriedenstellende, allgemeine Lösung, mit der man natürliche Bewegungen in kurzer Zeit erzeugen kann. In anderen Bereichen haben gestaltartige Eingabegeräte gezeigt, dass man mit diesen einfach und intuitiv die Körperhaltung digitaler kinematischer Modelle einstellen kann. Ziel des Forschungsprojektes HECI ist es mit einem für die arbeitswissenschaftliche Menschsimulation zugeschnittenen gestaltartigen Eingabegerät den Zeitbedarf für die Bewegungsgenerierung erheblich zu verkürzen. Auf Basis eines Prototyps konnten erste informelle Tests zur Anwendung dieses Eingabegerätes vorgenommen werden. In Zukunft soll die Entwicklung des Eingabegerätes abgeschlossen werden und in umfassenden Nutzerstudien die Gebrauchstauglichkeit des Eingabesystems evaluiert werden.

## Literatur

Arlt, F.: Untersuchung zielgerichteter Bewegungen zur Simulation mit einem CAD-Menschmodell. München: Herbert Utz 1999.

Esposito, C.; et al.: Of mice and monkeys: A specialized input device for virtual body animation. In: Symposium on Interactive 3D Graphics (I3D ’95), Monterey (USA), 9-12 April 1995, S. 109ff.

Gläser, D.; et al.: Auf den Spuren der Humanmotorik. In: Gesellschaft für Arbeitswissenschaft e.V. (Hrsg.): Gestaltung der Arbeitswelt der Zukunft – Bericht zum 60. Kongress der Gesellschaft für Arbeitswissenschaft. Dortmund: GfA-Press 2014, S. 363-365.

Hayes, C.: Starship troopers. In: The 25th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH ’98), Orlando (USA), 19-24 Juli, S. 311.

Jacobson, A.; Panozzo, D.; Glauser, O.; Pradalier, C.; Hilliges, O.; Sorkine-Hornung, O.: Tangible and modular input device for character articulation. ACM Transactions on Graphics 33 (2014) 4, S.82:1-82:12.

Knep, B.: Brian Knep.: Dinosaur input device. http://www.blep.com/rd/special-effects/dinosaur-input-device/, Zugriff am 16.07.2014.

Knep, B.; Hayes, C.; Sayre, R.; Williams, T.: Dinosaur input device. In: Katz, I.; Mack, R.; Marks, L.; Rosson, M.B.; Nielsen, J. (Hrsg.): Proceedings of the Conference on Human Factors in Computing Systems (CHI ’95), Denver (CO), 7-11 Mai 1995, S. 304-309.

Menache, A.: Motion capture primer. In: Computer Animation Complete. All-in-One: Learn Motion Capture, Characteristic, Point-Cased, and Maya Winning Techniques. Amsterdam et al.: Morgan Kaufmann Publishers 2010, S. 71–104.

Mühlstedt, J.: Entwicklung eines Modells dynamisch-muskulärer Arbeitsbeanspruchung auf Basis digitaler Menschmodelle. Chemnitz: Universitätsverlag Chemnitz 2012.

Mun, J. H.; Rim Y.H.: Human body modeling. In: Canetta, L. (Hrsg.): Digital Factory for Human-oriented Production Systems. London: Springer 2011.

Rieseler, H.: Roboterkinematik – Grundlagen, Invertierung und Symbolische Berechnung. Braunschweig/Wiesbaden: Vieweg 1992.

Rodríguez Flick, D.: Virtuelle Absicherung manueller Fahrzeugmontagevorgänge mittels digitalem 3-D-Menschmodell – Optimierung der Mensch-Computer-Interaktion. München: TU München 2010.

Schönherr, R.: Simulationsbasierte Absicherung der Ergonomie mit Hilfe digital beschriebener Bewegungen. Chemnitz: TU Chemnitz 2014.

Spanner-Ulmer, B.; Mühlstedt, J.: Virtuelle Ergonomie mittels digitaler Menschmodelle und anderer Softwarewerkzeuge. In: Schenk, M. (Hrsg.): Digital Engineering – Herausforderung für die Arbeits- und Betriebsorganisation. Berlin: GITO 2009, S.153 -170.

VDI 4499, Blatt 4: Digitale Fabrik – Ergonomische Abbildung des Menschen in der Digitalen Fabrik. Berlin: Beuth 2015.

Weber, W.: Industrieroboter - Methoden der Steuerung und Regelung, 2.Aufl. München: Carl Hanser 2013.

Yoshikazi, W.; et al.: An actuated physical puppet as an input device for controlling a digital manikin. In: Conference on Human Factors in Computing System (CHI ’11), Vancouver (Kanada), 7-12 Mai 2011, S. 637-646.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Agentenbasierte Simulation zur Abbildung von Intralogistikprozessen in der Grobblechindustrie zur Kampagnenoptimierung

## *Agent-based Simulation for Modelling Intra-logistics Processes in the Heavy Plate Industry*

Alexander Hübl, FH Oberösterreich, Logistikum, Steyr (Austria),
alexander.huebl@fh-steyr.at

Georg Strasser, voestalpine Grobblech GmbH, Linz (Austria),
georg.strasser@voestalpine.com

Abstract: In heavy plate steel industry heat treatment furnaces are used to achieve desired material properties of the manufactured plates. In this work such a heat treatment furnace is considered, which is constructed as continuous type and allows the sequential processing of materials in a typical flow direction. Additionally, the observed furnace is considered as a bottleneck in the higher-level production system. Materials being processed are scheduled in advance upon a given plate inventory, whereas technical constraints like the required temperature but also the sequential order of processed plates affect the possible throughput. In this work an agent based simulation model is presented, whereby the logic used of the agents is described.

## 1 Einleitung

In der Grobblech-Stahlindustrie werden häufig Wärmebehandlungsöfen zur Einstellung von gewünschten Materialeigenschaften der erzeugten Bleche eingesetzt. Im Rahmen dieses Beitrages wird ein Wärmebehandlungsofen betrachtet, welcher als Durchlaufofen ausgeführt ist und die sequentielle Abarbeitung von Materialien in einer typischen Materialflussrichtung ermöglicht. Der betrachtete Ofen stellt ferner einen Engpass im übergeordneten Produktionssystem dar und die zu bearbeitenden Materialien werden vorab aus einem Blechbestand eingeplant (die Blechherstellung ist nicht Bestandteil der Betrachtung), wobei neben technischen Kriterien wie der Blechabmessungen und der erforderlichen Temperatur die Reihenfolge der bearbeiteten Materialien den möglichen Durchsatz beeinflusst. Die Materialien müssen eingangsseitig fristgerecht aus dem Bestand angeliefert werden, was aufgrund der Lagerung in strukturierten Blechstapeln und dadurch oft erforderlicher Umstapelvorgänge, sowie limitierter Transportkapazitäten teils nur eingeschränkt möglich ist. Verspätungen bei der Anlieferung führen zum Verlust wertvoller Engpass-Kapazität.

Im nachfolgend präsentierten agentenbasierten Ansatz werden die Anlieferungen in der gewünschten Reihenfolge und die hierfür ausgelösten Transportvorgänge simuliert, um die Machbarkeit aus logistischer Sicht zu beurteilen.

## 2 Umsetzung

In diesem Artikel wird für das oben aufgezeigte Umfeld mit ihren spezifischen Charakteristiken ein logistisches Simulationsmodell aufgebaut mit dem der Ist-Status des Unternehmens abgebildet wird ohne die vorgelagerten Prozesse der Stahlherstellung zu berücksichtigen. Dazu wurde ein weitgehend flexibel konfigurierbares agentenbasiertes Modell mittels des Simulationswerkzeugs „Anylogic“ erstellt, welches die Abbildung des Produktionssystems mit spezifischen Layout- Elementen, Fördermitteln und individuellen Materialien ermöglicht. Agentenbasierte Simulation erweist sich aufgrund der vielfachen Anwendung in der Logistik und Produktionsplanung als geeignetes Werkzeug (Negahban und Smith 2014; Nikolopoulou und Ierapetritou 2012; Holmgren et al. 2012; Weng et al. 2008; Mula et al. 2006). Einen Schwerpunkt stellen autonom arbeitende Brückenkräne dar, welche im gegebenen Umfeld als hauptsächliches unstetiges Fördermittel eingesetzt werden. Wie in der Literatur üblich (Hübl et al. 2011) wird das erstellte Modell extern über Stamm- und Bewegungsdaten konfiguriert um sogenannte „Throw-away“ Lösungen zu vermeiden (Thompson 1994; Randell und Bolmsjo 2001).

Der vorgelagerte Prozess der Stahlerzeugung wird nicht detailliert abgebildet, sondern als stochastischer Ankunftsprozess in das Modell implementiert. Die Produktionsaufträge treffen häufig mit Abweichungen gegenüber dem eingeplanten Fertigstellungstermin im abgebildeten Bereich ein. Für eine herausgelöste Betrachtung eignen sich daher klassische Kennzahlen wie Lieferzeit, Verspätung und Liefertreue nicht. Für diese Simulationsaufgabe werden stattdessen folgende Kennzahlen definiert:

Anzahl der abgearbeiteten Bleche im Betrachtungszeitraum als Performance-Indikator der geplanten Reihenfolge, Belegung des Engpassaggregates Ofen, kumulierter Vorgabetemperaturgradient (summiert die nötigen Änderungen der Vorgabetemperatur als Maß für die durchlaufenen Rüstvorgänge auf) und die Summe der Transportoperationen zur Abschätzung der Inanspruchnahme der vorhandenen Transportkapazitäten.

## 3 Simulationsmodell

Eine Kernfunktionalität des erstellten Modells stellt die Abbildung des Materialflusses um den betrachteten Engpass dar. Es wird von einem definierten Anfangsbestand sowie einem Vorgriff-Bestand an Blechen ausgegangen, welche im Laufe des Simulationsexperiments abgearbeitet werden und letztendlich im Versand-Bereich eintreffen. Vom Engpass-Durchlaufofen werden Materialien in einer definierten Reihenfolge angefordert. Je Anforderung wird im System ein Fahrauftrag für Brückenkräne mit Referenzen auf das zu transportierende Material und das gewünschte Ziel platziert (Schema siehe Abbildung 1). Die bestehenden Fahraufträge werden für die Brückenkräne bewertet. Hierzu werden einerseits die Erreichbarkeit von Quell- und Zielposition, die aktuelle Anfahrdistanz zum Transport sowie die Priorität des Fahrauftrags berücksichtigt und resultieren in einer Bewertungs-

kennzahl. Der am besten für den Transport bewertete Brückenkran fährt zur aktuellen Position des jeweils angeforderten Materials und positioniert die Magnettraverse auf diesem. Nach Aufnahme des Blechs wird die Traverse angehoben und der Kran fährt nach Freistellung vom Stapel zum Einlaufrollgang des Ofens. Nach fertiger Positionierung über der Soll-Position am Einlaufrollgang wird das Blech auf diesem abgelegt.

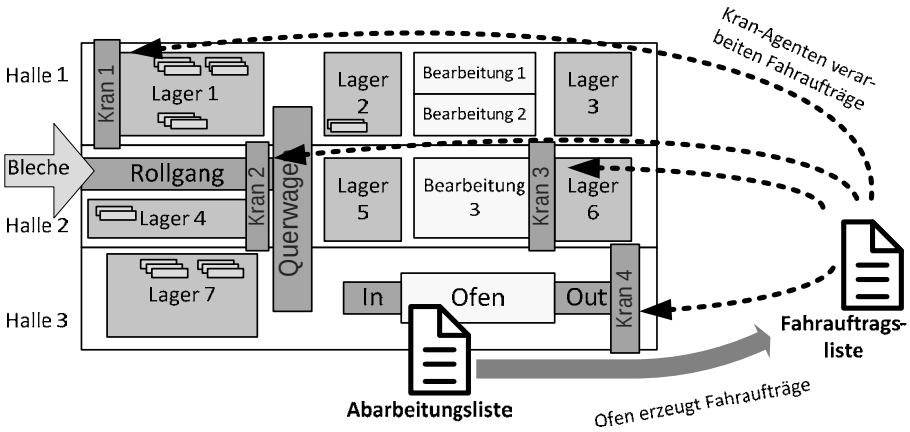



*Abbildung 1: Schema des Systems*

Als für die resultierende Anlieferungszeit eines gewünschten Blechs maßgeblich kann dessen Lage im Produktionssystem betrachtet werden. Falls das anzuliefernde Blech tiefer in einem Stapel liegt, müssen die darüber liegenden Materialien zuerst umgelagert werden, um das gewünschte Blech zu erreichen. Der Brückenkran nimmt hierzu die darüber liegenden Bleche sukzessive auf und legt diese an eine Ersatz- Position. Als Ersatzposition werden primär existierende Lagerstapel gesucht, welche in den Stapelbildungskriterien des Systems definierte Bedingungen erfüllen. Falls hierbei keine passenden existenten Stapel gefunden werden, wird ein neuer Stapel innerhalb der Lagerflächen gebildet. Für das abzulegende Blech wird vorrangig versucht, dieses am Stapel neben schmäleren oder kürzeren Blechen zu positionieren, um lagestabile Stapel zu gewährleisten. Falls keine derartige verschachtelte Ablage möglich ist, wird das Blech am Stapel zentriert abgelegt. Falls das anzuliefernde Blech in einem anderen Hallenschiff als die Anlage liegt, kann dieses nicht direkt per Brückenkran transportiert werden. Hierzu ist ein mehrstufiger Transport erforderlich, wobei das Blech durch einen Brückenkran des aktuellen Hallenschiffs an einen Gleiswagen zum Quertransport angeliefert wird. Nach der Positionierung des Quertransports zur Zielhalle, wird das Blech von einem Brückenkran des dortigen Hallenschiffs entladen und entweder zur Anlage angeliefert oder gegebenenfalls zum nächsten erforderlichen Quertransport transportiert.

Falls mehrere Bleche verschachtelt durch den Ofen gefahren werden, wird auf die vollständige Anlieferung aller Bleche am Einlauf gewartet. Nach fertiger Anlieferung des Blechs oder der Blechgruppe erfolgt die Positionierung an den Anfang der Ofenkammer. Die Ofenkammer wird bis zur gewünschten Soll-Temperatur angepasst und nach Erreichen dieser für das Blech freigegeben. Nach dem Eintritt in

die Ofenkammer wird das Blech oder die Blechgruppe je nach technischer Anforderung an das Produkt bis zum Erreichen der Blech-Solltemperatur und Haltezeit erwärmt, wobei der Durchlauf entweder kontinuierlich oder im Pendelbetrieb erfolgt und eine spezifische Erwärmungszeit beansprucht. Nach Fertigstellung der Erwärmung wird das Blech entweder direkt in den Auslaufbereich zur Luftabkühlung oder zuvor zur beschleunigten Abkühlung in eine Einspann-Vorrichtung (Quette) gefahren und dort mittels Wasser abgeschreckt. Es erfolgt ferner eine Rückmeldung der Jobausführung und die Aktivierung des nächsten Jobs im Arbeitsplan. Am Auslaufrollgang liegende Bleche werden nach Erreichen der für einen Magnettransport zulässigen Temperatur vom Brückenkran abtransportiert und für die folgende Bearbeitung eingelagert.

Die Kapazität des Ofenraums wird über dessen Länge determiniert. In der Ofenkammer werden in der Regel mehrere Bleche bzw. Blechgruppen gleichzeitig bearbeitet, wobei am Ofenrollgang liegende Bleche nachkommende Bleche blockieren. Da eine Kollision in der Ofenkammer infolge eines Auflaufens zu einer unzulässigen Erwärmung von Blechen führen würde, dürfen nur Bleche in die Ofenkammer gefahren werden, wenn deren Bearbeitungszeit nach der Bearbeitungszeit des jeweiligen Vorgängerblechs endet und im Falle eines Pendeldurchlaufs der notwendige Platz im Ofenraum nicht durch ein Vorgängerblech belegt ist. Um auslaufseitig eine Aufstauung von Blechen aus dem Ofen und der Quette zu vermeiden, erfolgt der Abtransport von fertigen Blechen mit hoher Priorität. Falls eine Blechgruppe bearbeitet wurde, müssen sämtliche Bleche der Gruppe abtransportiert werden.

## 3.1 Modellierungsmuster für Bewegungsabläufe

Für die gegebene Simulationsaufgabe sollen Materialien und Fördermittel im System Bewegungen ausführen, welche mit begrenzten Beschleunigungen und Verzögerungen erfolgen (diese werden zur Vereinfachung als konstant angenommen). Es wird zur weiteren Vereinfachung ferner davon ausgegangen, dass Geschwindigkeiten, Beschleunigungen und Verzögerungen von Fördermitteln unabhängig von der Beladung sind. Im Modell wurde ein rein ereignisgesteuerter Ansatz mittels Statecharts zur Abbildung von Bewegungen gewählt, wobei nach Eintreffen eines Fahrbefehls mit den Zielkoordinaten die Positionierung mit dem sukzessiven Durchlaufen einzelner States für Beschleunigen, Fahren und Abbremsen bis zum Stillstand erfolgt.

## 3.2 Modellobjekte

Innerhalb der Simulationsumgebung wurden Klassen für Layout, Ressourcen, Bewegungsdaten und rein logistische Steuerungsaufgaben erstellt. Das Modell wurde in einer kontinuierlichen 3D-Umgebung als Umwelt aufgebaut.

### *3.2.1 Halle und Bereich*

Mittels der Objekte Halle und Bereich wird das Layout des Produktionssystems definiert. Zur Vereinfachung wird eine ausschließlich rechteckige Form von Hallen und Bereichen vorausgesetzt.

Die Abbildung der Produktionshallen stellt die unterste Ebene im Modell dar. Attribute umfassen die Abmessungen und die Position der Halle sowie die Hallenbezeichnung. Um die Hallenflächen in zweckgebundene Bereiche zu unterteilen,

wurde ein Objekt „Bereich“ definiert. Attribute umfassen den Bereichsname, dessen Abmessungen und Position. Ein Bereich kann etwa als Lagerplatz oder als Bearbeitungszone für definierbare Bearbeitungen konfiguriert werden. Bei der Suche nach Lagerplätzen für Bleche wird bevorzugt ein Ablageort nahe eines Bereichs ermittelt, welcher für die anstehende Bearbeitung konfiguriert wurde.

### 3.2.2 Blech

Ein Blech wird im System als Exemplar der Blech-Klasse abgebildet, welche eine Reihe von Basis-Parametern, wie den Blechabmessungen, der Masse, dem Liefertermin, einem Auftragsbezug sowie einer Prioritätskennzahl enthält. Die Visualisierung des Blech-Objekts erfolgt durch einen Quader mit den proportionalen Blechabmessungen. Jedes Blech-Exemplar kann zudem individuell Referenzen zu anderen Modellobjekten aufweisen. Zur Abbildung des Fertigungspfades im System kann zudem individuell je Blech ein Arbeitsplan hinterlegt werden. Jedes Blech-Exemplar verfügt über ein Statechart, welches beim Transport und bei einer Bearbeitung durchlaufen wird und etwa nach Abschluss einer Bearbeitung die nächste anstehende Bearbeitung auf dem Arbeitsplan aktiv schaltet oder für mehrstufige Transporte (d.h. falls ein Blech mittels Quertransporten über mehrere Hallenschiffe transportiert werden muss) nach Abschluss eines Transports den nächsten Transportauftrag erzeugt.

### 3.2.3 Stapel

Ein Stapel-Objekt wird zur logischen Gruppierung von strukturiert abgestapelten Blechen verwendet. Im erstellten Modell wird davon ausgegangen, dass gelagerte Bleche stets einem eindeutigen Stapel zugeordnet sind. Jedes Exemplar der Stapel-Klasse verfügt über eine Reihe von einfachen Basis-Attributen wie der Stapel-Bezeichnung, den geometrischen Abmessungen und der erlaubten Stapelhöhe. Darüber hinaus können jedem Stapel-Exemplar eine Reihe von Bearbeitungen zugeordnet werden, welche bei der Stapelplatzsuche im Zuge der Einlagerung von Blechen herangezogen werden. Zweck hierbei ist die Vermeidung einer Durchmischung von Blechen mit verschiedenen anstehenden Bearbeitungen. Jeder Stapel kann zudem mittels eines Flags gesperrt werden, welches beispielsweise im Zuge einer gewünschten Anlieferung eines spezifischen im Stapel enthaltenen Blechs gesetzt wird und den Stapel als Ablageplatz für möglicherweise zwischenzeitlich stattfindende Umlagerungen sperrt.

Im gegebenen System werden Fördermittel wie Gleiswagen und Rollgänge eingesetzt, welche den Transport von gestapelten Blechen ermöglichen. Die Funktionalität dieser Fördermittel kann hierbei auf eine einfache Positionieraufgabe in Längs- oder Querrichtung reduziert werden. Für das Modell wurde diese Charakteristik durch das Objekt „Fördermittel“ umgesetzt, welche das Stapel-Objekt über eine Referenzierung optional erweitert. Im Modell können die Fördermittel hierdurch mittels normaler Stapel adressiert sowie auf deren Eigenschaften zugegriffen werden. Attribute des Fördermittels umfassen die maximalen Verfahrwege in Längs- und Querrichtung sowie die maximalen Geschwindigkeiten, Beschleunigungen und Verzögerungen. Die Steuerung des Fördermittels erfolgt durch den Empfang von Fahraufträgen. Das Fördermittel verfügt über ein Statechart zur Abbildung des aktuellen Zustands, wobei zwischen dem Fördermittel Gleiswagen und Rollgang unterschieden wird. Beim Gleiswagen wird der gesamte referenzierte Stapel ver-

schoben und seine Lage im System geändert, beim Rollgang ist der Stapel hingegen ortsfest und es werden lediglich die auf den Rollen aufliegenden Materialien bewegt. Bewegungsabläufe des gesamten Stapels werden entsprechend dem zuvor beschriebenen Modellierungsmuster mit Beschleunigungs-, Fahr- und Abbremszustand im Statechart ausgeführt, bei Rollgangstransporten wird hingegen ein temporärer Cluster der zu transportierenden Bleche erzeugt und zum Zielpunkt geschickt (das Cluster-Objekt wird im nachfolgenden Abschnitt beschrieben).

### 3.2.4 Cluster

Das Cluster-Objekt stellt im Modell, ähnlich wie das Stapel-Objekt, ein Hilfsobjekt zur Gruppierung von Blechen dar. Cluster-Objekte werden im erstellten Modell zur eher kurzfristigen Verwendung vorwiegend dynamisch erzeugt und aufgelöst und speichern Offset-Koordinaten der enthaltenen Bleche. Im Modell werden Cluster beim verschachtelten Transport mehrerer Bleche mittels Kran oder Rollgang erzeugt. Für den Rollgangstransport von nebeneinanderliegenden Blechen besitzen Cluster ein Statechart, in dem der Bewegungsablauf entsprechend dem zuvor beschriebenen Muster durchlaufen wird. Zusätzlich wurde eine einfache Kollisionslogik implementiert, um Kollisionen von Clustern am gleichen Rollgang zu vermeiden.

### 3.2.5 Fahrauftrag

Das Objekt „Fahrauftrag“ dient im Modell allgemein zur Spezifikation einer Transportaufgabe und enthält Referenzen auf das zu transportierende Blech und einen möglichen Zielstapel. Zusätzlich stehen Attribute zur Vorgabe einer definierten Zielposition, einer Orientierung zur Ablage am Stapel, einer Fahrauftrags-Priorität sowie einer Referenz auf einen Anfragekran zur Verfügung. Jeder Fahrauftrag verfügt ferner über einen Timestamp, im dem der Entstehungszeitpunkt des Fahrauftrags abgelegt wird und als Sortierkriterium herangezogen werden kann. Mittels eines Steuer-Flags wird zudem definiert, ob ein transportiertes Blech an der Zielposition auf einen aktiven Blechcluster abgelegt werden darf. Hierdurch wird beispielsweise die Blechablage auf gerade am Rollgang bewegten Blechen gesperrt.

Im Modell werden Exemplare der Klasse Fahrauftrag dynamisch erzeugt und stellen temporäre Objekte dar, welche nach Durchführung einer Transportaufgabe wieder aufgelöst werden. Fahraufträge werden nach unterschiedlichen Typen eingeteilt und können primär entweder eine Anlieferung von Blechen an einen bestimmten Stapel oder eine Umstapelung, etwa zur Freilegung unterhalb liegender Bleche bezwecken. Es wurden ferner eigene Fahrauftragstypen für die Aushandlung von Streckenfreigaben zwischen Kränen implementiert.

### 3.2.6 Kran

Das Kran-Objekt dient zur Abbildung von autonom arbeitenden Brückenkränen im Modell, welche Bewegungen in drei Achsen ausführen. Zur Realisierung dieser Charakteristik wurden mehrere Objekte verschachtelt. Der Kran-Agent wurde hierfür aus einem Haupt-Objekt „Kran“ zur Ausführung von Bewegung entlang der X-Achse, einem untergeordneten Objekt „Laufkatze“ für Bewegungen entlang der Y-Achse sowie einem wiederum untergeordneten Objekt „Hubwerk“ zur Durchführung von Hubbewegungen entlang der Z-Achse aufgebaut. Erfolgt eine Positionierung von übergeordneten Objekten, werden sämtliche untergeordneten Objekte

mit bewegt. Das Kran-Objekt enthält eine Reihe von Basis-Parametern, wie der bautechnisch vorgebebenen Abmessungen, eine Bezeichnung sowie die im System möglichen und optimalen Längspositionen entlang der Hallenschiffe. Für die Charakterisierung der Bewegungseigenschaften stehen Parameter für Geschwindigkeit, Beschleunigung und Verzögerung zur Verfügung. Jeder Kran verfügt ferner über konstruktiv bedingte technische Grenzen für maximales Beladungsgewicht und der Eindringtiefe der Magnetanlage für verschachtelte Transporte von übereinander liegenden Blechen. Jedes Kran-Exemplar verfügt weiters über verschiedene Referenzen, etwa zu den übernommenen Fahraufträgen oder zum jeweils direkt vorgelagerten bzw. nachfolgenden Kran im gleichen Hallenschiff. Im Haupt-Objekt und den untergeordneten Objekten werden die Achsbewegungen mittels eines Statecharts nach dem bereits vorgestellten Muster ausgeführt. Der Kran-Agent verfügt ferner über eine Statechart-Logik zur Prüfung und Bewertung von im System vorhandenen Fahraufträgen, welche bei passender Eignung (Kriterien sind etwa die generelle Erreichbarkeit im Anfahrraum und die Nähe zum Quell-Punkt) ausgeführt werden. Bei Übernahme eines Fahrauftrags muss der Kran zuerst zur Quellposition des zu transportierenden Blechs bewegt werden, nimmt dieses auf, fährt im Anschluss zu einer Zielposition und legt das Blech auf dieser ab. Der Haupt-Agent Kran führt hierbei ausschließlich Bewegungen in X-Richtung aus. Zur Positionierung in Y-Richtung wird vom Haupt-Agenten ein Fahrbefehl an den untergeordneten Agenten Laufkatze gesendet, welcher die Positionierung eigenständig mittels der bereits bekannten Statechart-Logik ausführt. Bei Erreichen des Ziels wird vom Haupt-Agenten Kran ein Absenkbefehl an den untergeordneten Hubwerk-Agenten gesendet, welcher ebenfalls autonom durch die Abarbeitung eines Statecharts die Absenkung in Z-Richtung ausführt, das Blech aufnimmt oder ablegt und im Anschluss wieder in Z-Richtung anhebt. Mittels Referenzen auf direkt vorgelagerte bzw. nachfolgende Kräne werden im Haupt-Kran-Agenten mögliche Kollisionen erkannt und gegebenenfalls rechtzeitig abgebremst. Der blockierende Kran erhält in diesem Fall während des Stillstands zyklisch Anfrage-Fahraufträge zur Freigabe der Strecke, welche dieser vor einer Weiterfahrt des blockierten Krans erst bestätigen und ausführen muss. Im Kran-Agenten wurden ferner Funktionalitäten implementiert um für Einlagerungs-Vorgänge (etwa nach Abschluss einer Bearbeitung) eigenständig einen passenden Stapelplatz zu suchen und Bleche zur verbesserten Lagerraumnutzung nach Möglichkeit auch verschachtelt (nebeneinander oder hintereinander) zu bereits im Stapel liegenden Blechen abzulegen.

### *3.2.7 Ofen*

Das Ofen-Objekt charakterisiert den in dieser Arbeit untersuchten Engpass-Durchlaufofen und verfügt einerseits über Basisattribute wie den bautechnischen Abmessungen des Ofens und andererseits über unterschiedliche Steuerungsparameter wie der Soll-Temperatur und einer Liste der zu bearbeitenden Bleche mit deren Bearbeitungsvorgaben. Der Ofen setzt sich aus dem Einlaufrollgang, der Ofenkammer zur Erwärmung, der Quette zur optionalen beschleunigten Abkühlung und dem Auslaufrollgang zusammen. Die verwendeten Rollgänge werden mittels der bereits vorgestellten Objekte Stapel und Fördermittel implementiert.

Die erwartete Ofenzeit wird nicht wie in der Realität möglichst exakt durch ein physikalisches Ofenmodell berechnet, sondern basiert im ausgeführten Modell auf einer einfachen, für den Ofen spezifischen Approximationsrechnung. Die tatsäch-

liche Istzeit wird im erstellten Modell mittels einer stochastischen, aus Realdaten abgeleiteten Funktion ermittelt. Diese liefert eine zu erwartende prozentuelle Abweichung der Istzeit von der theoretischen Ofenzeit.

Die primäre Aufgabe des Ofens besteht in der Abarbeitung der in der Abarbeitungsliste geplanten Einträge und der Veranlassung von An- und Abtransporten, was mittels eines mehrstufigen Statecharts ausgeführt wurde. Vom Ofen werden hierfür durch den Aufruf einer entsprechenden Funktion am angeforderten Blech die nötigen Fahraufträge zur Anlieferung an den Einlaufrollgang erzeugt. Liegt ein angefordertes Blech tiefer in einem Stapel werden auch für die darüber liegenden Bleche Fahraufträge zur Umstapelung erzeugt, um dieses freizulegen. In weiterer Folge werden die generierten Fahraufträge von den Fördermitteln im Modell (primär den Brückenkranen) sukzessive abgearbeitet, bis das angeforderte Blech letztendlich angeliefert wurde. Nachdem der Einlauf-Rollgang auf diese Weise vollständig beladen wurde, wird das darauf befindliche Blech oder die Blechgruppe an den Beginn der Ofenkammer positioniert und gegebenenfalls bis zum Erreichen der gewünschten Ofenraumtemperatur gewartet. Falls in der Ofenkammer bereits Vorgängerbleche vorhanden sind, wird ferner sichergestellt, dass keine Kollision stattfinden kann und das Blech zusätzlich zurückgehalten. Nach dem Start der Bearbeitung wird das Blech oder die Blechgruppe entsprechend der gewünschten Charakteristik durch den Ofenraum bewegt und wird nach vollständigem Durchlauf vom Einlauf- auf den Auslaufrollgang übernommen. Falls in der Abarbeitungsliste für das Blech eine Schnellkühlung in der Quette vorgesehen wurde, wird dieses mittig in der Quette positioniert, indem ein entsprechendes Positionier-Kommando an den Auslaufrollgang übermittelt wird. Ist keine beschleunigte Abkühlung erforderlich, wird das Blech direkt an die Endposition am Auslaufrollgang gesendet. Nach Erreichen einer berechneten Abkühlzeit werden vom Ofen für die am Auslaufbereich liegenden Bleche Fahraufträge zur Umstapelung erzeugt. Da der Abtransport zur Vermeidung von Material-Stauungen mit hoher Priorität erfolgen soll, wird diesem Fahrauftrag eine hohe Prioritätskennzahl zugewiesen.

## 4 Ergebnisse

Das Modell bildet neben dem betrachteten Engpassaggregat einerseits Layout-Elemente des Produktionssystems, Förderelemente wie Brückenkräne, Gleiswagen und Rollgänge sowie individuelle Materialien und deren Fertigungspfade ab. Es wurden ferner die speziellen Charakteristiken des Systems, wie die Verschachtelung bei Lagerung, Transport und Abarbeitung abgebildet (hierbei werden mehrere neben- oder übereinanderliegende Bleche gleichzeitig manipuliert). In Abbildung 2 ist eine exemplarische Visualisierung der verschiedenen Agenten in zwei Hallenschiffen dargestellt.

Mittels des Modells können Simulationsexperimente im untersuchten Produktionssystem ausgeführt werden, um basierend auf einem Anfangsbestand und einem Zulaufbestand unterschiedliche Abarbeitungsreihenfolgen am Ofen oder auch unterschiedliche Modell-Regeln auf deren logistische Auswirkungen zu untersuchen. Als Ergebniskennzahlen werden über die Simulationsdauer die Anzahl der bearbeiteten Bleche, die hierfür erforderlichen Transportvorgänge und die Belegung des Ofenraums (U) ermittelt. Die kumulierten Änderungen der Vorgabetemperatur stellen überdies ein Maß für die erforderlichen Rüstvorgänge dar, wobei ein hoher Wert

zum stärkeren Verschleiß der Ofenanlage führt. In Abbildung 3 wurden exemplarische Ergebnis-Charts eines Simulationslaufs zusammengefasst.

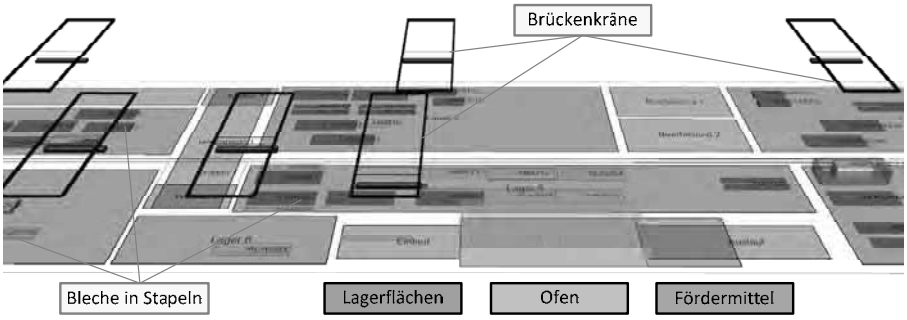
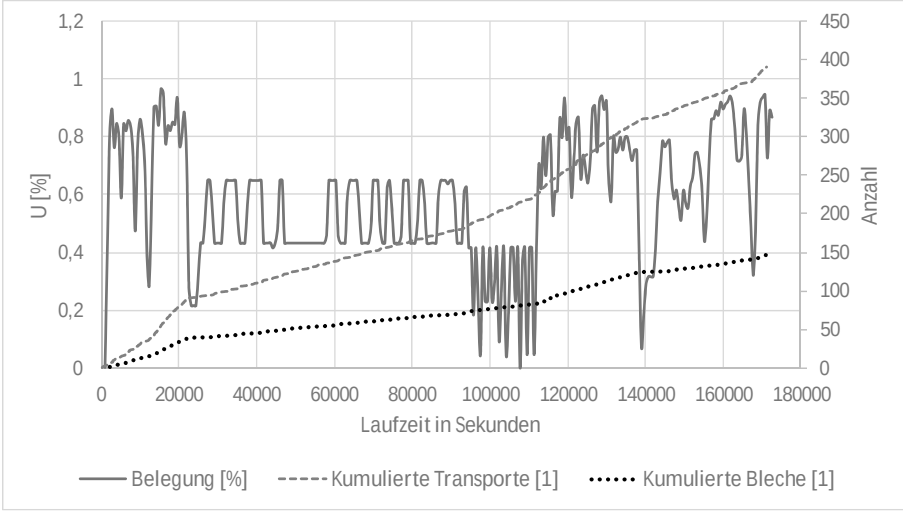



*Abbildung 2: Visualisierung eines Beispiel-Systems mit Modellelementen*



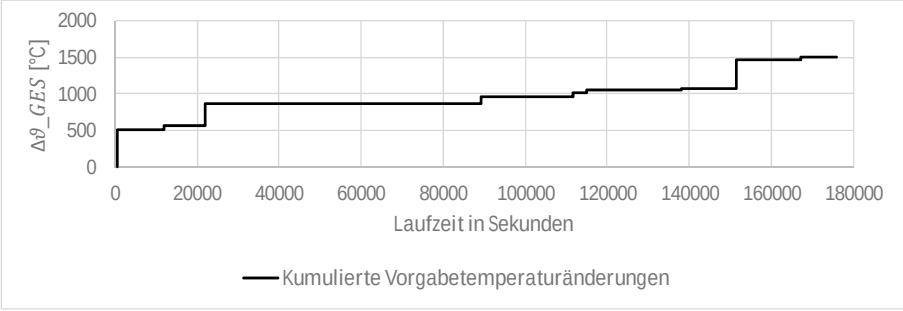




*Abbildung 3: Ergebnisse eines Simulationslaufs*

Das flexible Simulationsmodell ist mit realen Daten eines Unternehmens konfiguriert und getestet worden, wobei noch instabile Verhaltensweisen aufgetreten sind, z. B. nicht Erreichbarkeit von Materialien, Blockierung von Kränen oder eine erfolglose Stapelplatzsuche. Im realen System werden derartige Situationen durch die individuellen Fehlerlösungskompetenzen der beteiligten Mitarbeiter schnell behoben, im Modell konnten hingegen teils lediglich Workaround-Lösungen zur Stabilisierung implementiert werden. Diese Lösungen im Simulationsmodell bilden reale Sachverhalte erwartungsgemäß nicht akkurat ab und führen daher auch bei den Simulationsergebnissen zu Abweichungen gegenüber realen Werten. Die Abbildung der Blechtemperaturen im Bereich des Ofens stimmte bei individueller Betrachtung der Bleche hierdurch ebenfalls nicht mit der Realität überein, da zur Berechnung im Modell anstatt des eigentlich erforderlichen, physikalischen Modells (dieses ist bedingt durch die Ofenbauart und verwendete Steuerungsmechanismen für jede Anlage individuell) lediglich einfache Basis-Zusammenhänge herangezogen wurden. Im Modell erfolgten etwa Erwärmung und Abkühlung hiermit meist zu langsam.

## Danksagung

Dieser Artikel wurde im Rahmen des FFG (Forschungsförderungsgesellschaft) COMET Projektes „Heuristische Optimierung in Produktion und Logistik“ (HOPL), #843532 und dem FH OÖ Forschungsprojekt NODE verfasst.

## Literatur

Holmgren, J.; Davidsson, P.; Persson, J.A.; Ramstedt, L.: TAPAS: A multi-agent-based model for simulation of transport chains. Simulation Modelling Practice and Theory 23 (2012), S. 1–18.

Hübl, A.; Altendorfer, K.; Jodlbauer, H.; Gansterer, M.; Hartl, R.F.: Flexible model for analyzing production system with discrete event simulation. Proceedings of the Winter Simulation Conference, Phoenix (AZ) 2011, S. 1554–1565.

Mula, J.; Poler, R.; García-Sabater, J.P.; Lario, F.C.: Models for production planning under uncertainty: A review. International Journal of Production Economics 103 (2006) 1, S. 271–285.

Negahban, A.; Smith, J.S.: Simulation for manufacturing system design and operation: Literature review and analysis. Journal of Manufacturing Systems 33 (2014) 2, S. 241–261.

Nikolopoulou, A.; Ierapetritou, M.G.: Hybrid simulation based optimization approach for supply chain management. Computers and Chemical Engineering 47 (2012), S. 183–193.

Randell, L.; Bolmsjo, G.: Database driven factory simulation: A proof-of-concept demonstrator. Proceedings of the Winter Simulation Conference, Arlington (VA) 2001, S. 977–983.

Thompson, M.B.: Expanding simulation beyond planning and design. Industrial Engineering 26 (1994) 10, S. 64.

Weng, M.X.; Wu, Z.; Qi, G.; Zheng, L.: Multi-agent-based workload control for make-to-order manufacturing. International Journal of Production Research 46 (2008) 8, S. 2197–2213.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Agentenbasierte Planung und simulationsunterstützte Bewertung präferenzorientierter Einsatzzeiten des Personals

## *Agent-based Planning and Simulation-based Assessment of Preference-oriented Working Time Configuration of the Workforce*

Thilo Gamber, thilo@gamber.de, Gert Zülch, Karlsruher Institut für Technologie, Karlsruhe (Germany), gert.zuelch@gefora-beratung.de

Abstract: Operational focussed working time systems may lead to conflicts between work and private life of the employees. To reduce these conflicts, this paper deals with the organisation of socially acceptable working times. In order to improve an initial solution, an agent-based approach for an appropriate working time system has been developed. To proof the efficiency of this approach, a simulation procedure is used which allows for an adequate modelling of the working time system already in its planning phase and for evaluating an improved solution in a multi-criterial way. The approach is illustrated with the help of a case study in an in-house logistics department.

## 1 Kombination von Verbesserungsverfahren und Simulation bei der Einsatzzeitgestaltung

Zur Einsatzzeit- oder Dienstplanung geht man in der betrieblichen Praxis vielfach in der Weise vor, dass die Mitarbeiter im Anschluss an die Festlegung der individuellen Einsatzzeiten durch einen Personalverantwortlichen die Möglichkeit haben, einzelne Schichten oder Teile davon untereinander zu tauschen. Dabei besteht allerdings die Gefahr, dass sich besonders durchsetzungsfähige Mitarbeiter mehr Vorteile sichern als andere, die sich hinsichtlich ihrer Einsatzzeitpräferenzen eher zurückhalten. Dies kann dann nicht nur zu einer gewissen Unzufriedenheit innerhalb der Belegschaft, sondern auch zu vermehrten zeitlichen Konflikten zwischen Berufs- und Privatleben führen.

Einer solchen Vorgehensweise wird hier ein agentenbasiertes Planungsverfahren gegenüber gestellt, bei dem die Mitarbeiter innerhalb eines begrenzten Arbeitssystems und bei einem bereits vorliegenden Einsatzzeitplan durch ein Softwareverfahren repräsentiert werden, in dem dann mögliche Tausche virtuell durchgeführt werden. Voraussetzung hierfür ist, dass die Mitarbeiter gleichartige Qualifikationen

aufweisen und ihre Einsatzzeitpräferenzen für den betreffenden Planungszeitraum benennen.

Hierfür wird ein agentenbasiertes Planungsverfahren entwickelt, dessen Ergebnis allerdings als statisch zu betrachten ist, da die Arbeitsabläufe nicht explizit berücksichtigt werden. Daher wird das Ergebnis der erzielten Verbesserungslösung anschließend an ein Simulationsverfahren übergeben, in dem das Arbeitssystem mit den im Planungszeitraum durchzuführenden Aufgaben dynamisch bewertet werden kann. Hierdurch entsteht eine Kombination aus agentenbasiertem Planungsverfahren und ereignisdiskreter Simulation. Damit lässt sich dann überprüfen, welche Nutzenverbesserungen für die einzelnen Mitarbeiter im Arbeitssystem erreicht werden können und wie sich die Fairness innerhalb der Belegschaft gestaltet. Unter Fairness ist dabei zu verstehen, dass die individuelle Nutzenverbesserung durch die Umplanung gerechter und gleichmäßiger innerhalb der Belegschaft verteilt ist.

## 2 Wissenschaftliche Abgrenzung zu bisherigen Arbeiten

Bislang existiert kein agentenbasiertes Planungsverfahren, das auf präferenzgeleiteten Einsatzzeittauschen basiert und auch Fairnessaspekte berücksichtigt. Außerdem werden in den bisherigen Verfahren, die mit Agenten realisiert sind und Einsatzzeitpräferenzen benutzen (z.B. Chiaramonte und Chiaramonte 2008; Wang und Wang 2009), weder das deutsche Arbeitszeitrecht noch arbeitswissenschaftliche Empfehlungen berücksichtigt.

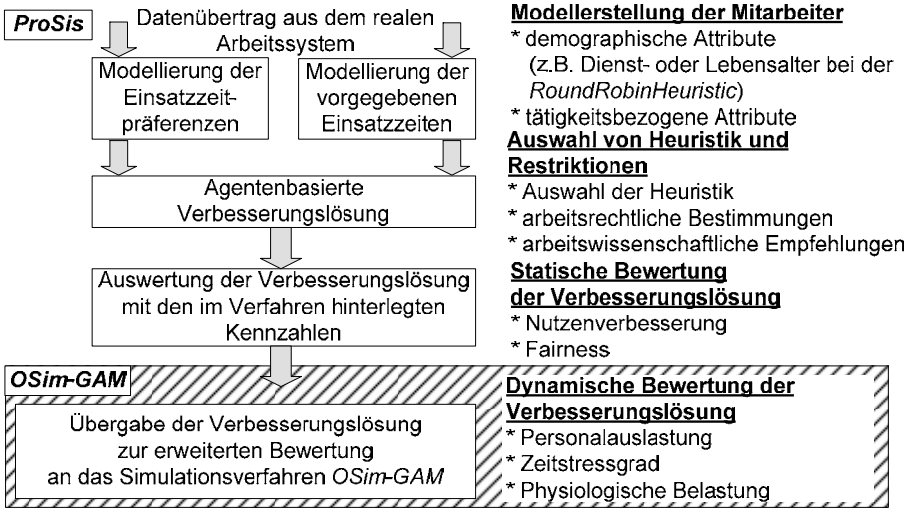



*Abbildung 1: Aufbau des Planungsverfahrens ProSis und Kopplung an das Simulationsverfahren OSim-GAM (Gamber 2015, S. 103; modifiziert)*

Weiterhin existiert kein Verfahren, das die Modellierung der Arbeitsabläufe und den Einsatz der Simulation in Kombination mit einem agentenbasierten Verbesserungsverfahren vorsieht (vgl. Cheang et al. 2003; Ernst et al. 2004; Haspeslagh und De

Causemaecker 2007). Abbildung 1 stellt den Zusammenhang zwischen beiden Verfahren dar. Das Planungsverfahren *ProSis* (*Pr*äferenz*o*rientierte Planung und *Si*mulation von Arbeits*s*chichten; Gamber 2015) und das Simulationsverfahren *OSim-GAM* (*O*bjekt*sim*ulator für die *G*estaltung vor *A*rbeitszeit*m*odellen; Bogus 2002) wurden beide am Institut für Arbeitswissenschaft und Betriebsorganisation des Karlsruher Instituts für Technologie entwickelt. Die hier beschriebene Verfahrenskombination kann auf eine Vielzahl von Fällen angewendet werden, bei denen Mitarbeiter gleicher Qualifikation betrachtet werden.

## 3 Kennzahlen für die Bewertung von Verbesserungslösungen

Für die Bewertung von Verbesserungslösungen wurden im Verfahren *ProSis* neue Kennzahlen definiert: „Anzahl Tausche“, „Mittlere Tauschanzahl“ und „Mittlerer Zufriedenheitsgrad“ innerhalb der Belegschaft. Im Simulationsverfahren *OSim-GAM* können die bereits implementierten Kennzahlen „Personalauslastung“, „Physiologische Belastung“ und „Zeitstressgrad“ genutzt werden (Abb. 2). Mittels einer lexikographischen Zielfunktion lassen sich diese dann zu einem Gesamtergebnis aggregieren. Auf diese Zielfunktion wird in Kapitel 5 näher eingegangen.

*Abbildung 2: Kennzahlen für die Bewertung von Einsatzzeiten (nach Zülch, Gamber, Freytag 2011, S. 294; übersetzt und modifiziert)*

Zur Überprüfung einer Verbesserungslösung muss diese auf ihre Konformität mit arbeitsrechtlichen Bestimmungen und arbeitswissenschaftlichen Empfehlungen überprüft werden. Die hierzu hinterlegten Gestaltungsregeln können vom Benutzer priorisiert werden, um mögliche Widersprüche zwischen ihnen zu vermeiden.

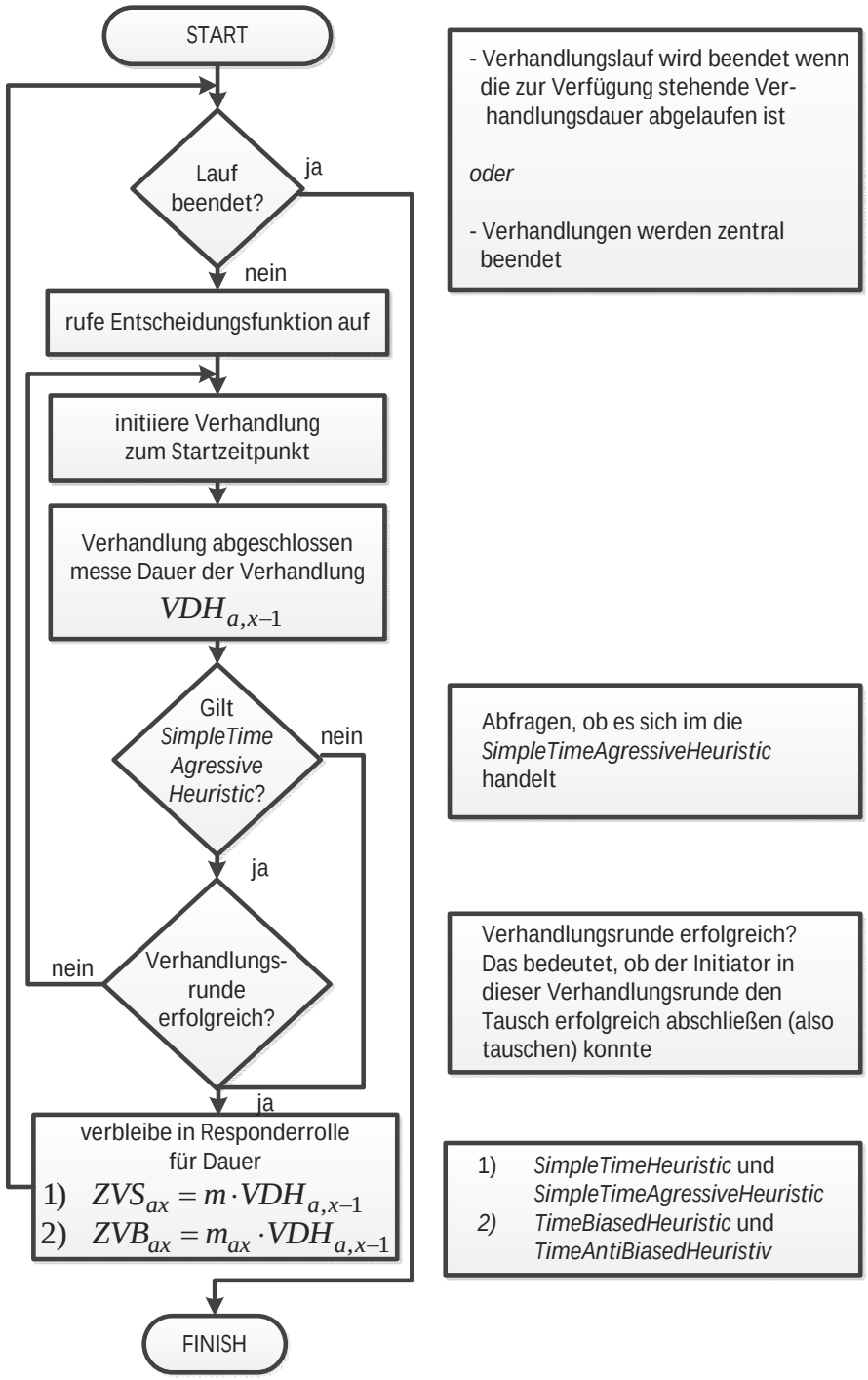



*Abbildung 3: Flussdiagramm für zeitbasierte Heuristiken (Gamber 2015, S. 134; modifiziert)*

# 4 Heuristiken zur Ableitung von Verbesserungslösungen

Agentenbasierte Heuristiken laufen nach den Prinzipien der agentenbasierten Paradigmen (Wooldridge und Jennings 1995, S. 118 ff.) und Kommunikation ab. Dies bedeutet vor allem, dass die Verhandlungen der autonom agierenden Agenten über ein Kommunikationsprotokoll (Kontraktnetzprotokolls) ablaufen und die Entscheidungen vom jeweiligen Agenten selbst getroffen werden (FIPA 2002, S. 1 ff.; vgl. Freytag 2011, S. 39). In der Kommunikation existieren zwei Rollen, die ein Agent annehmen kann. Zum einen ist dies die Rolle des Initiators, zum anderen die Rolle des Responders. Der Initiator startet die Kommunikation und schlägt den anderen Agenten im System einen gewünschten Tausch vor. Diese prüfen dann den Vorschlag und teilen dem initiierenden Agenten mit, ob sie seinen Vorschlag akzeptieren oder nicht.

Bezüglich der eigenen Nutzenentwicklung hat der Initiator grundsätzlich einen Vorteil gegenüber dem Responder, da er über die Möglichkeit verfügt, allen anderen Agenten seinen Tauschwunsch vorzuschlagen. Daher muss eine die Verhandlung koordinierende Heuristik gewährleisten, dass diese Rolle, die über mehrere Verhandlungsrunden abläuft, wechseln kann. Dies entspricht dem Fairnessanspruch des Verfahrens, dass sich möglichst alle Agenten verbessern können. Dazu lässt sich nach jedem Tausch *x-1* ein Zufriedenheitsgrad $ROS_{a,x-1}$ für jeden Agenten *a* bestimmen, mit dem dessen individuelle Präferenzsituation für einen neuen Tausch *x* bewertet wird.

Um Erkenntnisse darüber zu gewinnen, welche Heuristiken sich für virtuelle Einsatzzeittausche am besten eignen, wurde eine Auswahl davon implementiert. Die verwendeten Heuristiken wurden teilweise aus der bestehenden Literatur übernommen (z. B. Chiaramonte und Chiaramonte 2008) und teilweise selber entwickelt (Gamber 2015, S. 127 ff.). Mittels eines Testdatensatzes wurden die Heuristiken miteinander verglichen und die gefundenen Verbesserungslösungen mit den in *ProSis* und *OSim-GAM* hinterlegten Kennzahlen bewertet. Die untersuchten Heuristiken lassen sich in die Kategorien zeitbasiert, zufallsbasiert und triggerbasiert einteilen.

Abbildung 3 zeigt beispielhaft das Flussdiagramm für zeitbasierte Heuristiken. Im Folgenden werden die entwickelten zeitbasierten Heuristiken exemplarisch näher beschrieben. Bei den zeitbasierten Heuristiken werden die Zeitdauern $ZVS_{ax}$ oder $ZVB_{ax}$ bestimmt, die aussagen, wie lange ein Agent *a*, der gerade einen Tausch *x-1* initiiert hat, erneut warten muss, bis er wieder initiieren darf. Dafür kann der Agent *a* die dafür benötigte Verhandlungsdauer $VDH_{a,x-1}$ für die von ihm initiierte, vorhergehende Verhandlung messen (Abb. 4).

Für die Heuristiken *SimpleTime* und *SimpleTimeAgressive* ist zusätzlich ein Faktor *m* zu definieren, der im Verfahren konstant bleibt. Die Zeitdauer $ZVS_{ax}$ bis zur nächsten Initiierung des Agenten *a* nach dem Tausch *x-1* berechnet sich aus diesem Faktor *m* multipliziert mit der gemessenen Dauer $VDH_{a,x-1}$. Der Faktor *m* wird in der Heuristik einmalig festgelegt, wobei zum Start des Verfahrens eine initiale Zeit $VDH_0$ vorliegt. Je kleiner *m* ist, desto schneller initiiert der Agent *a* wieder, je größer er ist, desto langsamer initiiert er. Die Heuristik *SimpleTimeAgressive* initiiert solange Verhandlungen, bis ein Tausch durchgeführt wird, und verharrt dann wie bei der Heuristik *SimpleTime*.

Für die Ermittlung der Zeitdauer $ZVB_{ax}$ bis zur nächsten Initiierung des Agenten *a* nach Tausch *x-1* wird analog zu $ZVS_{ax}$ die Verhandlungsdauer $VDH_{a,x-1}$ mit einem festgelegten Faktor *m* multipliziert. Die Zeitdauer $ZVS_{ax}$ bzw. $ZVB_{ax}$ ist also weder bei der Heuristik *SimpleTime* noch bei der Heuristik *SimpleTimeAgressive* konstant.

| Zeitbasierte Heuristiken | |
|---|---|
| Richte den Zeitpunkt für eine Entscheidung nach der letzten Verhandlungsdauer:<br>• Messe die Verhandlungsdauer $VDH_{a,x-1}$ für jeden Agenten *a* und Tausch *x-1*.<br>• Lege den Zeitpunkt fest, zu dem erneut ein Tausch *x* initiiert werden soll.<br>Faktor *m bzw.* $m_{ax}$ bestimmt die Aggressivität der Verhandlung. | |
| Ausführungsvarianten | Verbale Beschreibung |
| *SimpleTime* | *m* ist konstant |
| *SimpleTimeAgressive* | *m* ist konstant; verhandle, bis Tausch erfolgt |
| *TimeBiased* | $m_{ax}$ wächst mit Zufriedenheitsgrad $ROS_{ax-1}$ |
| *TimeAntiBiased* | $m_{ax}$ sinkt mit Zufriedenheitsgrad $ROS_{ax-1}$ |

*Abbildung 4: Verhandlungsstrategien mit zeitbasierten Heuristiken (nach Gamber 2015, S. 133)*

Für die Heuristiken *TimeBiased* und die *AntiTimeBiased* wird durch den Faktor $m_{ax}$ zusätzlich der aktuelle Zufriedenheitsgrad $ROS_{a,x-1}$ des Agenten *a* nach dem Tausch *x-1* berücksichtigt. Der Faktor $m_{ax}$ variiert somit im Verlauf des Verfahrens.

Die Heuristik *TimeBiased* erhöht den zeitlichen Abstand zwischen den Verhandlungsinitiierungen abhängig vom erreichten Zufriedenheitsgrad $ROS_{a,x-1}$. Bei dieser Heuristik ist der Agent *a* also umso passiver, je mehr seiner Präferenzen bislang erfüllt wurden. Damit soll erreicht werden, dass die Agenten bezüglich ihrer Möglichkeiten zur Nutzenverbesserung untereinander in fairer Weise berücksichtigt werden. Hingegen verringert die Heuristik *TimeAntiBiased* den zeitlichen Abstand zwischen den Initiierungen eines Agenten.

Formalisiert ist dies durch den Multiplikationsfaktor $m_{ax}$, der von $ROS_{a,x-1}$ abhängig ist. Wenn im Fall dieser Heuristik $ROS_{a,x-1}$ steigt, wird $m_{ax}$ größer, wodurch Agent *a* weniger häufig initiiert. Die Heuristik *TimeAntiBiased* führt also dazu, dass der Agent umso aggressiver handelt, je höher sein Zufriedenheitsgrad ist bzw. die Verbesserung voranschreitet. Es ist zu erwarten, dass diese Heuristik vergleichsweise unfairere Ergebnisse liefert, da die Anzahl der Initiierungen für einen Teil der Agenten steigt und für andere sinkt.

## 5 Exemplarische Anwendung in einer innerbetrieblichen Logistikabteilung

Als Anwendungsfall wird die innerbetriebliche Logistikabteilung eines Unternehmens der Prozessautomatisierung herangezogen. Die von den Mitarbeitern indivi-

duell bevorzugten Einsatzzeiten wurden dort durch mündliche Befragungen ermittelt. Als Grundlage für die durchzuführenden Aufgaben wurden Daten aus einer früheren Untersuchung von Graichen, Schmidt und Zülch (2011) verwendet.

Im Arbeitssystem sind zehn Mitarbeiter in Vollzeit beschäftigt und dabei unterschiedlichen Funktionsbereichen zugeordnet. Ihre Qualifikation ist teilweise generalisiert und teilweise spezialisiert auf einzelne Funktionen. Es liegt eine Fünf-Tage-Woche und ein Schichtsystem mit Früh- und Spätschicht vor. Jeder Mitarbeiter wird im wöchentlichen Schichtwechsel acht Stunden pro Tag eingesetzt.

Die Mitarbeiter in der untersuchten Abteilung haben die klassischen Aufgaben einer innerbetrieblichen Logistik zu erfüllen (Abb. 5). Sie sind zunächst für die Materialbereitstellung und Entsorgung im Fertigungsbereich zuständig. Auch der Transport von Fertigerzeugnissen von der Fertigung zum Logistikbereich wird von diesen Mitarbeitern erledigt. Im betrachteten Arbeitssystem gibt es zwei verschiedene Kommissionierbereiche. Diese unterscheiden sich nach den Erzeugnissen, z.B. abhängig von Gewicht und Leistungsklasse der Geräte, sowie nach den Kunden, z.B. Versand nach Europa oder Asien gemäß Zollbestimmungen oder Anforderungen an die Transportsicherheit. Die Verpackung und die Zusammenstellung von Lieferungen zum Versand gehören ebenfalls zu den Aufgaben der Logistikmitarbeiter. In der Simulation werden die Arbeitsabläufe als netzwerkartige Durchlaufpläne modelliert (Gamber 2015, S. 179).

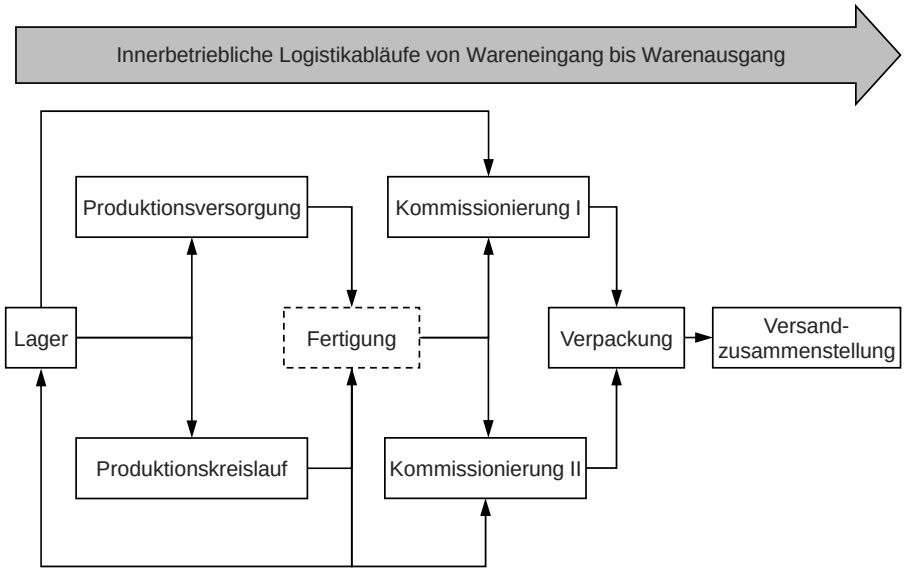



*Abbildung 5: Arbeitsabläufe in der innerbetrieblichen Logistikabteilung (Quelle: Andris 2011, S. 28; übersetzt)*

Im agentenbasierten Verbesserungsverfahren sollen Zeitscheiben im Umfang von mindestens zwei Stunden getauscht werden können. Wenn die Möglichkeit besteht, sollen dazu vorzugsweise Zeitscheiben an den Randzeiten der Schichten getauscht werden, um den Mitarbeitern möglichst flexible Einsatzzeiten zu ermöglichen.

Durch Parametereinstellung können Tauschrestriktionen aufgrund arbeitsrechtlicher Bestimmungen und arbeitswissenschaftlicher Empfehlungen aktiviert werden. Im vorliegenden Anwendungsfall wird auf einige zentrale Bestimmungen des Arbeitszeitgesetzes (ArbZG 1994) und auf arbeitswissenschaftliche Empfehlungen (z. B. Hornberger und Knauth 2000; Knauth 2002) zurückgegriffen. Im Einzelnen sind dies:

- die maximale tägliche Arbeitszeit von 10 Stunden,
- die durchschnittliche wöchentliche Arbeitszeit von maximal 48 Stunden sowie
- die kürzeste Dauer einer Schicht von 4 Stunden.

Speziell werden im Anwendungsfall folgende Heuristiken beispielhaft untersucht (Gamber 2015, S. 131 ff.):

- die zeitbasierte Heuristik *TimeBiased (h=1),*
- die triggerbasierte Heuristik *RoundRobinAgressive (h=2)* sowie
- die ebenfalls triggerbasierte Heuristik *BiasedTrigger (h=3).*

Zunächst wird das agentenbasierte Verbesserungsverfahren *ProSis* und anschließend eine dynamische Analyse der gefundenen Lösung mit dem Simulationsverfahren *OSim-GAM* durchgeführt. Dabei werden die Auslöser der Durchlaufpläne, also die Initiierung der im Planungszeitraum von 26 Wochen anfallenden Arbeitsaufgaben, über die gesamte Simulationsdauer hinweg stochastisch verteilt.

| | **Heuristik** | **Kennzahl aus** *ProSis* | | | |
|---|---|---|---|---|---|
| *h* | | Anzahl Tausche *rng* ($AZT_h$) | Mittlere Tauschanzahl *rng* ($MTA_h$) | Durchschnittlicher Zufriedenheitsgrad *rng* ($DSR_h$) | |
| 1 | *TimeBiased* | 3 (24) | 3 (4,8) | 3 (0,96) | |
| 2 | *BiasedTrigger* | 1 (37) | 1 (7,4) | 1,5 (0,98) | |
| 3 | *RoundRobinAggressive* | 2 (36) | 2 (7,2) | 1,5 (0,98) | |
| | | **Differenz der Zielerreichungsgrade aus** *OSim-GAM* | | | **Gesamtergebnis** |
| | | Differenz Auslastung *rng* ($DAU_h$) | Differenz Zeitstressgrad *rng* ($DZG_h$) | Differenz Belastung *rng* ($DBE_h$) | $GTR_h$ |
| 1 | *TimeBiased* | 1 (-0,15 %) | 1 (-0,63 %) | 1,5 (0,09 %) | 2 |
| 2 | *BiasedTrigger* | 2 (-0,25 %) | 2 (-1,59 %) | 1,5 (0,09 %) | 1 |
| 3 | *RoundRobinAggressive* | 3 (-0,35 % ) | 3 (-1,89 %) | 3 (0,01 %) | 3 |

*Abbildung 6: Bewertung von Heuristiken anhand des Anwendungsfalls "Innerbetriebliche Logistik" (Daten teilweise nach Huck 2011, S. 47 ff. und S. 73 ff.)*

Die Ergebnisse aus den Verfahren *ProSis* und *OSim-GAM* sind in Abbildung 6 dargestellt. Für die verwendeten Heuristiken *h* aus *ProSis* ist dort die Anzahl der durchgeführten Tausche $AZT_h$, die durchschnittliche Tauschanzahl $MTA_h$ und der durch-

schnittliche Zufriedenheitsgrad $DSR_h$ angegeben. Aus dem Verfahren *OSim-GAM* werden die Differenzen der Zielerreichungsgrade Personalauslastung $ZAU_h$, Physiologische Belastung $ZBE_h$ und Zeitstressgrad $ZZG_h$ berechnet. Die Differenzen der Zielerreichungsgrade aus der Verbesserungslösung und der Ausgangslösung sind in Abbildung 6 bezüglich der Kennzahlen Personalauslastung $DAU_h$, Zeitstressgrad $DZG_h$ und Physiologische Belastung $DBE_h$ wiedergegeben.

Das Ergebnis wird gemäß einer lexikographischen Präferenzordnung beurteilt. Diese Präferenzordnung weist dem besten Wert eines jeden Bewertungskriteriums den Rang 1 zu, nachfolgende Ränge werden entsprechend höher bewertet. Abschließend wird eine lexikographische Entscheidungsregel angewendet, bei der das Kriterium mit der größten Häufigkeit des besseren Kriteriums den besten Rang erhält (Schwarze 2014, S. 96). Entsteht keine einfache Mehrheit bezüglich der Häufigkeit des besten Kriteriums, entscheidet die einfache Mehrheit der Häufigkeiten des zweitbesten Kriteriums usw. Das Gesamtergebnis wird als Güte der Heuristik $GTR_h$ in Abbildung 6 dargestellt.

## 6 Schlussfolgerung und Ausblick

Für die vorliegende Studie zeigt sich, dass die Heuristik *BiasedTrigger* die zielführendste ist. Festzuhalten ist jedoch, dass sich diese Aussage nicht verallgemeinern lässt. In einem anderen Anwendungsfall im Krankenhausbereich stellte sich die Heuristik *RoundRobinAggressive* als für diesen Fall geeignetste dar (Gamber 2015, S. 175 f.). Bezüglich der Wirkungsweise der Heuristiken ergibt sich demnach noch Forschungsbedarf. Weiterer Forschungsbedarf besteht darin, dass Personal mit sich überlappenden Qualifikationen in das Verbesserungsverfahren einbezogen wird.

Darüber hinaus kann der vorgestellte Ansatz auch auf andere Arbeitssysteme angewandt werden, so z.B. auf die Einsatzzeitplanung in ambulanten Sozialdiensten, die Schichtplangestaltung im Einzelhandel oder die Einsatzzeitplanung von Fluglotsen. Weiterhin könnte das Verfahren dahingehend erweitert werden, dass sich Wünsche in der Zusammenarbeit bei der Teambildung berücksichtigen lassen. Auch könnten Fragestellungen hinsichtlich Themen wie altersbezogene Einsatzzeitmodelle und Lebensarbeitszeitmodelle mit dem Verfahren untersucht werden.

## Literatur

ArbZG: Arbeitszeitgesetz. Vom 6. Juni 1994, zuletzt geändert am 20. April 2013. http://www.gesetze-im-internet.de/bundesrecht/arbzg/gesamt.pdf, Stand: 10.05.2015.

Andris, T.: Arbeitszeitgestaltung unter Berücksichtigung individueller Präferenzen der Mitarbeiter durch agentenbasierte Optimierung und Bewertung durch Simulation. Studienarbeit Universität Karlsruhe, Institut für Arbeitswissenschaft und Betriebsorganisation 2011.

Bogus, T.: Simulationsbasierte Gestaltung von Arbeitszeitmodellen in Dienstleistungsbetrieben mit kundenfrequenzabhängigem Arbeitszeitbedarf, Aachen: Shaker Verlag, 2002.

Cheang, B.; Li, H.; Lim, A. et al.: Nurse rostering problems – a bibliographic survey. European Journal of Operational Research 151 (2003) 3, S. 447-460.

Chiaramonte, M. V.; Chiaramonte, L. M.: An agent-based nurse rostering system under minimal staffing conditions. International Journal of Production Economics 114 (2008) 2, S. 679-713.

Ernst, A.T.; Jiang, H.; Krishnamoorthy, M.; Sier, D.: Staff scheduling and rostering: A review of applications, methods and models. European Journal of Operational Research 153 (2004) 1, S. 3-27.

FIPA – Foundation for Intelligent Physical Agents (Hrsg.): Contract Net Interaction Protocol Specification. Genf: FIPA, 2002. (Document Number SC00029H) http://www.fipa.org/specs/fipa00029/SC00029H.pdf, Stand: 18.04.2012.

Freytag, R.: Koordinationsmechanismen bei der Verhandlung in Agentensystemen. Bachelorarbeit, Karlsruher Institut für Technologie, Institut für Arbeitswissenschaft und Betriebsorganisation 2011.

Gamber, T.: Entwicklung eines Verfahrens zur agentenbasierten Einsatzzeitgestaltung. Aachen: Shaker 2015.

Graichen, S.; Schmidt, D.; Zülch, G.: Load-oriented work organisation due to an ageing workforce. In: Spath, D.; Ilg, R.; Krause, T. (Hrsg): Innovation in Product and Production. 21st International Conference on Production Research (ICPR 21), Stuttgart, 31.07.-04.08.2011, S. 427-432.

Haspeslagh, S.; De Causemaecker, P.: Distributed decision making in hospital wide nurse rostering problems. In: Baptiste, P.; Kendall, G.; Munier-Kordon, A.; Sourd, F. (Hrsg.): Proceedings of the 3rd Multidisciplinary International Conference on Scheduling Theory and Applications (MISTA 2007). Paris: MISTA 2007, S. 192-199.

Hornberger, S.; Knauth, P.: Innovative Flexibilisierung der Arbeit. In: Knauth, P.; Zülch, G. (Hrsg.): Innovatives Arbeitszeitmanagement, Aachen: Shaker 2000, S. 23-49.

Huck, S.: Gestaltung von Arbeitszeiten unter Berücksichtigung individueller Präferenzen der Mitarbeiter durch agentenbasierte Optimierung und Bewertung durch Simulation. Diplomarbeit Karlsruher Institut für Technologie, Institut für Arbeitswissenschaft und Betriebsorganisation 2011.

Knauth, P.: Arbeitszeitflexibilisierung aus arbeitswissenschaftlicher Sicht. In: Zülch, G.; Stock, P.; Bogus, T. (Hrsg.): Arbeitszeitflexibilisierung im Dienstleistungsbereich, Aachen: Shaker 2002, S. 51-74.

Schwarze, J.: Grundlagen der Statistik, Band 1: Beschreibende Verfahren, 12. Aufl. Herne: NWB-Verlag 2014.

Wang, Z.; Wang, C.: Automating nurse self-rostering: A multiagent systems model. In: Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio (TX) 2009, S. 4422-4425.

Wooldridge, M.; Jennings, N.: Intelligent agents: Theory and practice. In: Knowledge Engineering Review 10 (1995) 2, S. 115-152.

Zülch, G.; Gamber, T.; Freytag, R.: Improving the work life balance of hospital staff through an agent-based rostering procedure. In: Albolino, S.; Bagnara, S.; Bellandi, T.; Llaneza, J.; Rosal, G.; Tartaglia, R. (Hrsg.): Healthcare Systems Ergonomics and Patient Safety 2011, London: Taylor & Francis Group, 2011, S. 291-295.

*Simulation in Production*
*and Logistics 2015*
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

# Ressourcenbeschränkte Terminplanung mit einem System kollaborativer Agenten

## *Solving the Multi-mode Resource-constrained Project Scheduling Problem with Collaborative Agents*

Florian Wenzler, Willibald A. Günthner, Technische Universität München, München (Germany), wenzler@fml.mw.tum.de, guenthner@fml.mw.tum.de

Abstract: For the realisation of a construction project the quality of the created project plan is essential. Since there are a large number of constraints to be considered, this is a big challenge for the planners. The problem that has to solved here is known as the multi-mode resource-constrained project scheduling problem (MRCPSP). This paper presents a multi-agent approach, in which resources and processes are represented as collaborative agents. The autonomous process agents register themselves on a central blackboard where the negotiation about resource allocation to the activities takes place. The model is implemented in Discrete-Event-Simulation and evaluated with standardised project plans from the field of Operations Research.

## 1 Einführung

Die Erstellung von Terminplänen im Bauwesen wird meist manuell und nicht sehr detailliert durchgeführt. Daher kann es leicht zu einem erhöhten Koordinationsaufwand und geringerer Produktivität während der Bauausführung oder zum Verzug der Fertigstellung führen. Dies hat als Ursache die hohe Anzahl an Abhängigkeiten und Randbedingungen in einem derartigen Projekt, aber auch die hohe Rate an Neuplanungen aufgrund unvorhersehbarer externer Einflüsse wie dem Wetter, von denen insbesondere Baustellen betroffen sind. Aus diesen Gründen wird angestrebt, eine Methodik für ein Planungswerkzeug bereitzustellen, welche eine schnellere, detailliertere und zugleich objektivere Planung ermöglicht. Zudem soll eine eventuell notwendige Neuplanung aufwandsarm durchführbar sein.

In der Wissenschaft wird die zu Grunde liegende Problemstellung als Resource-Constrained Project Scheduling Problem (RCPSP) bezeichnet. Die Erweiterung, bei der die einzelnen Prozessschritte zusätzlich noch in unterschiedlicher Weise durchgeführt werden können, ist im Operations Research als Multi-Mode Resource-Constrained Project Scheduling Problem (MRCPSP) bekannt.

Im Anschluss wird zunächst auf die theoretische Beschreibung des MRCPSP eingegangen mit den für den restlichen Teil getroffenen Vereinfachungen. Daraufhin werden verschiedene Möglichkeiten zur Lösung der Problemstellung vorgestellt und zum Lösungsansatz, der in diesem Beitrag verfolgt wird, abgegrenzt. Dies ist auch die Überleitung zur Beschreibung des Multi-Agenten-Ansatzes und dessen Art der Ressourcenzuweisung zu den Prozessen. In einem weiteren Abschnitt werden die bisherigen Simulationsergebnisse vorgestellt. Den Abschluss bildet eine Zusammenfassung dieses Beitrags mit einem Ausblick auf die weiterführenden Forschungsarbeiten.

## 2 Theoretische Problemstellung des (M)RCPSP

Das zuvor erwähnte RCPSP wird durch einige Randbedingungen definiert. Die einzelnen Prozesse sind in ihrer Ausführungsreihenfolge durch Vorgänger-Nochfolger-Beziehungen eingeschränkt. Ein Prozess darf erst dann gestartet werden, wenn alle Vorgänger abgeschlossen wurden. Zudem ist für die Durchführung eine bestimmte Anzahl an Ressourcen notwendig, was zu weiteren Beschränkungen führt, da für das gesamte Projekt nur eine definierte Anzahl an Ressourcen zur Verfügung steht, die nicht überschritten werden darf. Innerhalb der Ressourcen wird zusätzlich noch zwischen zwei grundlegenden Ressourcentypen unterschieden. Erneuerbare Ressourcen stehen nach der Verwendung durch einen Prozess wieder für die nachfolgenden Prozesse zur Verfügung, können also wiederverwendet werden. Dazu zählen Maschinen oder Personen. Im Gegensatz dazu werden die nicht erneuerbaren Ressourcen während der Projektlaufzeit verbraucht. Beispiele hierfür sind Material, Treibstoff oder Geld.

Wie durch Blazewicz (1983) nachgewiesen, ist dies den NP-schweren Problemen zuzuordnen. Somit ist das allgemeinere MRCPSP ebenfalls ein NP-schweres Problem. Das MRCPSP kann folgendermaßen beschrieben werden:

- J: Anzahl an Prozessen mit J = {1, … , j}
- M: Anzahl Modi je Aktivität J
- $d_{jm}$: Dauer von Prozess j in Modus m
- $S_j$: Nachfolger von Prozess j
- $P_j$: Vorgänger von Prozess j
- R: Menge erneuerbarer Ressourcentypen
- N: Menge nicht erneuerbarer Ressourcentypen
- $r_{jmk}$: Bedarf des Prozesses j im Modus m an erneuerbaren Ressourcen vom Typ k ∈ R
- $n_{jml}$: Bedarf des Prozesses j im Modus m an nicht erneuerbaren Ressourcen vom Typ l ∈ N

Typischerweise ist das Optimierungsziel die Minimierung der Projektdauer bei Einhaltung aller gegebenen Randbedingungen.

Für das weitere Vorgehen können noch einige Vereinfachungen getroffen werden, ohne die das Optimierungsproblem deutlich komplexer wäre:

- Keine Unterbrechungen von gestarteten Prozessen
- Der gewählte Modus eines Prozesses kann nicht mehr verändert werden
- Es werden nicht mehrere Projektpläne gleichzeitig (an unterschiedlichen Orten) ausgeführt, die sich die Ressourcen teilen

Diese Punkte stellen jedoch im Vergleich zur Realität auf der Baustelle keine drastischen Einschränkungen dar und können deshalb akzeptiert werden.

# 3 Lösungsansätze für das (M)RCPSP

Für die Lösungsfindung existieren verschiedene Ansätze. Dazu zählen Branch-and-Bound und lower bounds als Verfahren, die die optimale Lösung sicher finden. Allerdings steigt die Rechenzeit bei großen Problemen stark an, da der zu durchsuchende Lösungsraum ebenfalls deutlich zunimmt. Des Weiteren sind diverse andere Ansätze bekannt, die bereits umfangreich untersucht wurden. Dazu zählen Simulated Annealing (Józefowska et al. 2001), Partikelschwarmoptimierung (Jarboui et al. 2008), genetische Algorithmen (van Peteghem und Vanhoucke 2010) und Ameisenalgorithmen (Li und Zhang 2013). Eine davon abweichende Herangehensweise ist die Formulierung eines Multi-Agenten-Systems (MAS). Die erhofften Vorteile bestehen darin, das Gesamtproblem in kleinere, leichter zu lösende Aufgabenstellungen zu unterteilen. Außerdem ist ein funktionierendes Agentensystem robust gegenüber äußeren Störeinflüssen und kann leicht um weitere Ressourcen oder Prozesse erweitert werden (Maes 1995). Knotts (2000) definiert in diesem Zusammenhang alle Prozesse als Agenten, die die benötigten Ressourcen selbstständig anwerben müssen. Ein weiteres MAS wurde von Horenburg (2014) entwickelt. Dabei sind neben den Prozessen auch alle Ressourcen als Agenten angelegt. Dadurch wird erreicht, dass auch innerhalb einer Ressourcengruppe die einzelnen Ressourcen mit individuellen Eigenschaften ausgestattet werden können. Auf diesem System, das für die Single-Mode Variante des RCPSP ausgelegt ist, basiert dieser Beitrag. Allerdings wurde die Beschränkung auf nur eine Ausführungsart je Prozess fallen gelassen und somit ein flexibleres Werkzeug erstellt.

## 3.1 Aufbau des Multi-Agenten-Systems

Agenten werden als eigenständige Einheiten verstanden, die ohne direkte Entscheidungshilfe versuchen, ihre Ziele zu erreichen (Woolridge 2009). Nwana (1996) unterteilt des Weiteren die Eigenschaften, die „smart agents“ erfüllen müssen, nach den folgenden 3 Kriterien:

- Autonom: Die Agenten handeln eigenständig und versuchen proaktiv ihre Ziele zu erreichen.
- Kooperativ: Die Agenten können miteinander kommunizieren und interagieren.
- Lernfähig: Intelligente Agenten können aus der Interaktion mit der Umwelt oder anderen Agenten lernen, um auf ähnliche Situationen besser reagieren zu können.

Im aktuellen Projektstand erfüllen die Agenten die ersten beiden Punkte. Nwana klassifiziert diese als kollaborative Agenten.

Für die Umsetzung des Multi-Agenten-Systems wurde die ereignisdiskrete Simulationssoftware Plant Simulation verwendet. Das System besteht aus verschiedenen Agententypen, die innerhalb einer Umwelt interagieren (Abb. 1).

In Abbildung 1 ist außerdem ersichtlich, welche Informationen jeder Agent besitzt (Wissen) und welche Aktionen er durchführen kann (Handlungsfähigkeit).

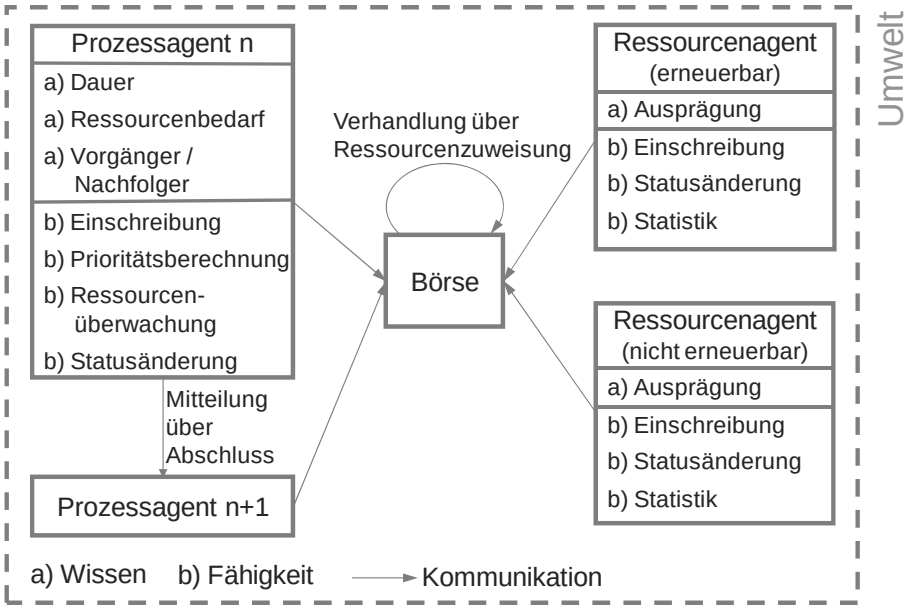



*Abbildung 1: Aufbau des Agentensystems mit der Kommunikationsstruktur*

Jeder Prozess des Projektplans wird durch einen Prozessagenten repräsentiert. Dieser verfügt über alle Informationen, die ihn selbst betreffen, wie seine Vorgänger und Nachfolger, die benötigten Ressourcen je Modus und die dazugehörige Bearbeitungszeit. Die Prozessagenten kommunizieren direkt mit anderen Prozessagenten über ihren derzeitigen Status. Zudem besteht ein Informationsaustausch zwischen Prozessagent und Börse.

Die Ressourcen können in zwei unterschiedliche Typen unterteilt werden: Erneuerbare und nicht erneuerbare Ressourcen. Jede Gruppe wird durch individuelle Agenten abgebildet, die den speziellen Eigenschaften angepasst sind.

Die Börse ist das zentrale Element dieses Systems. Hier melden sich alle Prozess- und Ressourcenagenten für die Zuordnung an. Alle Entscheidungen werden anschließend an dieser Stelle getroffen und auch die Informationen über den derzeitigen Systemstatus werden hier protokolliert.

## 3.2 Funktionsweise

Zu jedem Zeitpunkt, an dem ein Prozess beendet wird, wird ein neuer Bietvorgang angestoßen. Jeder Prozess, dessen Vorgänger abgeschlossen sind, schreibt sich mit jedem seiner Modi an der Börse ein, um an der Ressourcenvergabe teilzunehmen. Dabei ist der mit übergebene Prioritätswert (vgl. Kapitel 4) von zentraler Bedeutung. Nach dieser Größe wird entschieden, für welchen Prozess als erstes überprüft wird, ob ausreichend Ressourcen vorhanden sind. Im Gegensatz zu anderen Herangehensweisen (z. B. Boctor 1993), bei denen zunächst ein Prozess ausgewählt und anschließend ein ausführbarer Modus gesucht wird, sind bei dieser Vorgehensweise alle Modi der aktuell zugelassenen Prozesse gleichwertig. Ist beispielsweise Modus 1 von Prozess 1 nicht ausführbar obwohl priorisiert, könnte als nächstes ein Modus

von Prozess 2 getestet werden. Bei einer vorgeschalteten Festlegung auf einen Prozess würde dagegen zunächst Modus 2 von Prozess 1 überprüft.

Auf Seiten der Ressourcen ist das Vorgehen ähnlich. Alle Ressourcen, die zum aktuellen Bietzeitpunkt verfügbar sind, schreiben sich ebenfalls an der Börse ein. Hier kann die Reihenfolge, in der sie den Prozessen zugewiesen werden, verändert werden. Somit lässt sich beispielsweise durch Priorisierung von bisher wenig eingesetzten Ressourcen die Auslastung nivellieren.

Nachdem einem Prozess die Ressourcen zugeteilt wurden, werden die alternativen Modi gelöscht und es wird zum nächsten Prozess in der Liste übergegangen. Dies geschieht, bis die gesamte Liste abgearbeitet wurde. Anschließend werden alle neuen Prozesse gestartet und bis zum nächsten Bietzeitpunkt gewartet.

# 4 Prioritätsregeln

Die Entscheidung, welcher Prozess als erstes Ressourcen zugewiesen bekommt, wird über heuristische Prioritätsregeln bestimmt. Sie legen fest, auf welche Weise jeder Prozess den Wert berechnet, mit dem er sich an der Börse einschreibt. Der Startzeitpunkt und der Ausführungsmodus eines Prozesses sind die Freiheitsgrade der Optimierung, die mit dem Prioritätswert beeinflusst werden können.

## 4.1 Regeln für das RCPSP

In den Arbeiten von Kolisch (1996a), Yang (1998), Alvarez-Valdés (1989) und Alastuey Gonzalez (2012) haben sich die folgenden Prioritätsregeln für das Single-Mode-Problem als die besten Heuristiken herausgestellt:

- GNS (Greatest Number of Successors): Für die Priorisierung zählt die Anzahl aller Nachfolger des jeweiligen Prozesses.
- GRPW (Greatest Rank Positional Weight): Vorrang vor anderen Prozessen hat derjenige, dessen Dauer zusammen mit der Dauer all seiner Nachfolger am größten ist.
- LFT (Latest Finish Time): Entscheidend für die Priorisierung ist der letztmögliche Zeitpunkt, zu dem der Prozess ausgeführt werden muss, ohne dass dadurch Verzögerungen im Gesamtprojekt auftreten.
- LPF (Longest Path Following): Innerhalb der Nachfolger des aktuell betrachteten Prozesses lässt sich ein kritischer Pfad bis zum Projektende bestimmen. Nach dessen Länge werden die Prozesse priorisiert. Der kritische Pfad wird dabei ohne Beachtung der Ressourcenbeschränkung bestimmt.
- MSLK (Minimum Slack): Diese Regel bevorzugt diejenigen Prozesse, die den geringsten Puffer besitzen. Der Puffer beschreibt die Zeit, die die Prozesse nach hinten verschoben werden können, ohne dass es eine Auswirkung auf die Gesamtprojektlänge hat. Prozesse auf dem kritischen Pfad verfügen demnach über keinen Puffer.
- WCS (Worst Case Slack): Für die Priorität wird bei dieser Regel bestimmt, wie weit der fragliche Prozess verschoben werden muss, falls stattdessen ein anderer ausführbarer Prozess sofort den Zuschlag für die Ressourcen erhalten würde. Dieser Vergleich wird paarweise mit allen derzeit eingeschriebenen Prozessen durchgeführt und der größte Wert für die Priorisierung verwendet.

- ACS (Average Case Slack): Im Gegensatz zu WCS gehen hier alle Vergleichswerte in die Berechnung mit ein. Priorisiert werden die Prozesse, die im Durchschnitt bei Zurückstellung zugunsten der anderen Prozesse den größten Verzug aufweisen würden.

Bei einigen Regeln bleibt der Wert während der Simulationsausführung konstant. Eine einmalige Bestimmung der Prioritätswerte ist damit ausreichend. Im Gegensatz dazu kann sich die Berechnungsgrundlage bei den Methoden WCS und ACS bei jedem Bietzeitpunkt ändern.

Des Weiteren wurde die zufällige Bestimmung des Prioritätswerts mit aufgenommen. Dies dient zum einen zur Validierung des Modells und zum anderen als Referenz für die restlichen Prioritätsregeln.

## 4.2 Anpassungen für das MRCPSP

Die Berechnungsvorschriften für das Single-Mode-Problem können nicht direkt für das MRCPSP übernommen werden, sofern Informationen bezüglich der Prozessdauer und des Ressourcenbedarfs benötigt werden. Jeder Prozess müsste abhängig von den Modi seiner Nachfolger unterschiedliche Prioritätswerte erhalten. Allerdings sind die tatsächlichen Modi der Nachfolger zum aktuellen Bietzeitpunkt noch unbekannt. Aus diesem Grund wurden für die Berechnung zwei Alternativen eingeführt:

1. Für alle relevanten Größen wird immer der maximale Wert verwendet. Damit wird immer vom schlechtesten Fall ausgegangen, der noch eintreten kann. Die Prioritätsregeln werden durch ein angehängtes MAX gekennzeichnet (z. B. LPF MAX).
2. Als Alternative dazu wurde die Verwendung der Mittelwerte der jeweiligen Prozessdaten implementiert. Dies ist dadurch begründet, dass im Mittel davon auszugehen ist, dass bei verschiedenen Prozessen nicht immer der längste oder kürzeste (bzw. der mit dem kleinsten/größten Ressourcenbedarf) gewählt wird, sondern dass sich ein Durchschnittswert einstellt. In der Beschriftung ist diese Methode durch ein zusätzliches AVG ersichtlich.

# 5 Ergebnisse bei standardisierten Plänen

Für die bisherigen Untersuchungen wurden standardisierte Projektpläne der „project scheduling problem library – PSPLIB“ verwendet (Kolisch und Sprecher 1996b). Diese Bibliothek stellt eine Vielzahl unterschiedlicher Projektpläne zur Verfügung. Die Pläne sind in Gruppen eingeteilt, die sich durch Variation einzelner Parameter (Anzahl Prozesse, Anzahl Modi je Prozess, Anzahl Ressourcentypen, etc.) unterscheiden und so die Bestimmung des Einflusses dieser Merkmale auf die Qualität der Optimierungsstrategie möglich machen. Zudem haben sie den Vorteil, dass die jeweiligen Optima bekannt sind. In Tabelle 1 ist die Zusammenstellung der im weiteren Verlauf untersuchten Pläne aufgeführt.

*Tabelle 1: Aufbau ausgewählter Projektpläne*

| Bezeichnung | Anzahl Prozesse | Anzahl Modi je Prozess | Anzahl nicht erneuerbare Ressourcen | Anzahl Pläne |
|---|---|---|---|---|
| n0 | 10-20 | 3 | 0 | 470 |
| j30 | 30 | 3 | 2 | 552 |

In Abbildung 2 sind beispielhaft die Ergebnisse für die Prioritätsregel LPF AVG für Pläne ohne erneuerbare Ressourcen dargestellt.

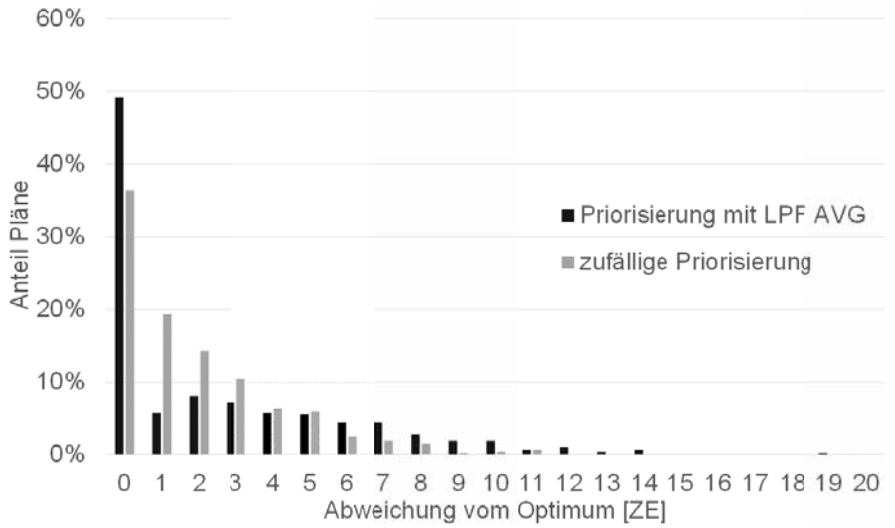



*Abbildung 2: Abweichung vom Optimum für Projektpläne ohne nicht erneuerbare Ressourcen*

In ca. 49% der vorhandenen Pläne wird mit dieser Regel bei einem einzigen Durchlauf des Plans das bekannte Optimum erreicht. In diesen Fällen wurde also zu jedem Entscheidungszeitpunkt der richtige Prozess im richtigen Modus priorisiert. Als Vergleichswert dient hierfür auch die zufällige Priorisierung der Prozesse (zweite Datenreihe in Abb. 2). Dafür wurden je Plan 10 000 Simulationsläufe durchgeführt und nur der Beste davon für die Auswertung verwendet. Trotzdem konnte hier nur in ca. 36% der Projekte das Optimum erreicht werden.

Im Durchschnitt wurde die optimale Projektdauer mit der Prioritätsregel LPF AVG um 10,9% überschritten (Tabelle 2), wobei bei einer durchschnittlichen minimalen Projektdauer von 22 Zeiteinheiten die Abweichung von einer Zeiteinheit (der kleinstmöglichen Differenz) bereits einer prozentualen Abweichung von 4,5% entspricht. Die Abweichung bei der zufälligen Festlegung des Prioritätswerts ist mit 8,4% etwas geringer, was aber mit der Verwendung nur des besten Ergebnisses jedes Plans erklärbar ist.

Weitere Ergebnisse mit ausgewählten Prioritätsregeln und Terminplänen sind in Tabelle 2 aufgelistet. Der Unterschied in der Ergebnisqualität zwischen MAX und AVG ist mit Ausnahme der Regel WCS kaum vorhanden. Deshalb ist in der nachfolgenden Tabelle nur AVG (und WCS MAX) eingetragen.

*Tabelle 2: Vergleich verschiedener Prioritätsregeln für das MRCPSP*

| Plan-Typ | Regel | Durchschnittliche Abweichung [%] | Anzahl Optimum | Anzahl ungelöst |
|---|---|---|---|---|
| n0 | GNS | 12,3 | 218 (46,4%) | 0 |
| | LPF AVG | 10,9 | 231 (49,1%) | 0 |
| | GRPW AVG | 10,9 | 222 (47,2%) | 0 |
| | LFT AVG | 11,8 | 222 (47,2%) | 0 |
| | MSLK AVG | 12,1 | 224 (47,7%) | 0 |
| | WCS AVG | 74,9 | 0 | 0 |
| | ACS AVG | 97,3 | 0 | 0 |
| | WCS MAX | 122,0 | 0 | 0 |
| j30 | GNS | 12,3 | 112 (20,3%) | 302 (54,7%) |
| | LPF AVG | 11,7 | 116 (21,0%) | 300 (54,3%) |
| | GRPW AVG | 11,8 | 113 (20,5%) | 300 (54,3%) |
| | LFT AVG | 12,3 | 114 (20,7%) | 301 (54,5%) |
| | MSLK AVG | 11,7 | 119 (21,6%) | 304 (55,1%) |
| | WCS AVG | 72,7 | 0 | 114 (20,7%) |
| | ACS AVG | 95,5 | 0 | 81 (14,5%) |
| | WCS MAX | 123,1 | 0 | 40 (7,2%) |

In den Ergebnissen (v. a. bei der Anzahl optimaler Lösungen) spiegelt sich die Komplexität der bearbeiteten Projektpläne wider. Ohne nicht erneuerbare Ressourcen (Typ n0) können alle Pläne gelöst werden. Bei j30 konnten diverse Projektpläne aufgrund der Überschreitung der Bedarfe an nicht erneuerbaren Ressourcen (was bei den Plänen der PSPLIB möglich ist) nicht korrekt erstellt werden. In allen Fällen liefern die Prioritätsregeln ACS und WCS bei der Projektlänge deutlich schlechtere Ergebnisse. Allerdings haben sie den Vorteil, dass die Wahrscheinlichkeit, einen gültigen Plan zu erhalten, deutlich größer ist.

Im Vergleich mit anderen Lösungsansätzen schneidet dieses Agentensystem zum aktuellen Zeitpunkt erwartungsgemäß schlechter ab (Tabelle 3). Bei einem einzigen Durchlauf des Plans (mit Prioritätsregel LPF AVG) ohne Optimierungsschleifen sind die Abweichungen vom Optimum höher als bei anderen Methoden, die viele Wiederholungen je Projektplan durchführen und durch Zufallseinflüsse komplett neue Lösungsalternativen generieren können.

Allerdings bleibt die Abweichung bei steigender Projektgröße nahezu konstant, während diese bei anderen Ansätzen teils stark zunimmt. In Tabelle 3 ist außerdem die durchschnittliche Rechenzeit angegeben, die für einen Plan der jeweiligen Größe benötigt wurde. Hier ist anzumerken, dass Plant Simulation u. a. aufgrund der Visualisierung und anderer Hintergrundprozesse nicht als Benchmark für die Rechenzeit dienen kann.

*Tabelle 3: Vergleich mit anderen Lösungsansätzen für das MRCPSP – durchschnittliche Abweichung vom Optimum in %*

| | j10 | j12 | j14 | j16 | j18 | j20 |
|---|---|---|---|---|---|---|
| LPF AVG | 10,51 | 10,17 | 10,52 | 10,27 | 10,26 | 10,79 |
| | 0,043s | 0,048s | 0,052s | 0,057s | 0,063s | 0,069s |
| Van Peteghem und Vanhoucke (2010) | 0,01 | 0,09 | 0,22 | 0,32 | 0,42 | 0,57 |
| Jarboui et al. (2008) | 0,03 | 0,09 | 0,36 | 0,44 | 0,89 | 1,1 |
| Ranjbar et al. (2009) | 0,18 | 0,65 | 0,89 | 0,95 | 1,21 | 1,64 |
| Józefowska et al. (2001) | 1,16 | 1,73 | 2,6 | 4,07 | 5,52 | 6,74 |

## 6 Zusammenfassung und Ausblick

Der Terminplanung im Bauwesen liegt das sogenannte Multi-Mode Resource-Constrained Project Scheduling Problem (MRCPSP) zu Grunde. Die freie Wahl der Ausführungsreihenfolge der Bauprozesse wird eingeschränkt durch Vorgänger-Nachfolger-Beziehungen sowie die begrenzte Anzahl verfügbarer Ressourcen. Für dieses Optimierungsproblem wurde eine agentenbasierte Methodik entwickelt, die anschließend in einer ereignisdiskreten Simulationsumgebung implementiert wurde. Mit dem MAS soll eine Alternative zu den gängigen Lösungsansätzen aufgezeigt und untersucht werden, ob diese Herangehensweise für derartige Probleme sinnvoll ist. Die ersten Versuchsläufe mit den verwendeten heuristischen Prioritätsregeln zu Ressourcenzuweisung und Priorisierung der Prozesse ergaben vielversprechende Resultate.

Die bisherigen Ergebnisse sind ein Zwischenschritt auf dem Weg zu einem lernfähigen Agentensystem. Einerseits sollen die aktuellen Prioritätsregeln verbessert werden. Dies betrifft vor allem die Berücksichtigung der nicht erneuerbaren Ressourcen, die bereits zu einem frühen Zeitpunkt zu ungültigen Plänen führen können. Andererseits sollen die Agenten noch um die fehlende Lernfähigkeit erweitert werden. Damit können Optimierungspotenziale innerhalb des Plans eigenständig erkannt und dort gezielt Änderungen für erneute Simulationsläufe durchgeführt werden. Nach der Validierung an den bisher verwendeten Projektplänen soll die Methodik an Daten von realen Bauvorhaben getestet werden. Bei einer Skalierung der Projektgröße ist die Anzahl der Restriktionen für die Funktionsweise irrelevant. Die Prozessagenten nehmen erst dann am Bietprozess teil, wenn alle Bedingungen erfüllt sind.

## Literatur

Alvarez-Valdez, R.; Tamarit, JM.: Heuristic algorithms for resource-constrained project scheduling: a review and an empirical analysis. In: Slowinski, R.; Weglarz, J. (Hrsg.): Advances in project scheduling. Amsterdam: Elsevier 1989, S. 113-134.

Alastuey Gonzalez, F.: Evaluation of different priority rules for the resource allocation in project scheduling. Diplomarbeit Technische Universität München, Lehrstuhl für Fördertechnik Materialfluss Logistik, 2012.

Blazewicz, J.; Lenstra, J.; Rinnooy Kann, A. H. G.: Scheduling subject to resource constraints: Classification an complexity. Discrete Applied Mathematics 5 (1983) 1, S. 11-24.

Boctor, F. F.: Heuristics for scheduling projects with resource restrictions and several resource-duration modes. International Journal of Production Research 31 (1993) 11, S. 2547-2558.

Horenburg, T.: Simulationsgestützte Ablaufplanung unter Berücksichtigung aktueller Baufortschrittsinformationen. Dissertation Technische Universität München, Lehrstuhl für Fördertechnik Materialfluss Logistik, 2014.

Jarboui, B.; Damak, N.; Siarry, P.; Rebai, A.: A combinatorial particle swarm optimization for solving multi-mode resource-constrained project scheduling problems. Applied Mathematics and Computation 195 (2008), S. 299-308.

Józefowska, J.; Mika, M.; Rózycki, R.; Waligóra, G.; Weglarz, J.: Simulated annealing for the multi-mode resource-constrained project scheduling. Annals of Operations Research 102 (2001), S. 137-155.

Knotts, G.; Dror, M.; Hartmann, B. C.: Agent-based project scheduling. IIE Transactions 32 (2000) 5, S. 387-401.

Kolisch, R. (1996a) Efficient priority rules for the resource constrained project scheduling problem. Journal of Operations Management 14 (1996) 3, S. 179-192.

Kolisch, R.; Sprecher, A. (1996b) PSPLIB – A project scheduling problem library. European Journal of Operational Research 96 (1996) 1, S. 205-216.

Li, H.; Zhang, H.: Ant colony optimization-based multi-mode scheduling under renewable and nonrenewable resource constraints. Automation in Construction 35 (2013), S. 431-438.

Maes, P.: Modeling adaptive autonomous agents. In: Langton, C. G. (Hrsg): Artificial life. Cambridge, MA: MIT Press, 1995, S. 135–162.

Nwana, H. S.: Software Agents: an overview. The Knowledge Engineering Review 11 (1996) 3, S. 205-244.

Ranjbar, M.; De Reyck, B.; Kianfar, F.: A hybrid scatter-search for the discrete time/resource trade-off problem in project scheduling. European Journal of Operational Research 193 (2009), S. 35-48.

Van Peteghem, V.; Vanhoucke, M.: A genetic algorithm for the preemptive and non-preemptive multi-mode resource-constrained project scheduling problem. European Journal of Operational Research,201 (2010), S. 409-418.

Woolridge, M. J.: An introduction to multiagent systems. 2. Auflage. Chichester, U.K.: John Wiley & Sons 2009.

Yang, K.-K.: A comparison of dispatching rules for executing a resource-constrained project witch estimated activity durations. Omega 26 (1998) 6, S. 729-738.

# Author Index


Abourizk, Simaan ..... 491
Aichele, Fabian ..... 249
Altendorfer, Klaus ..... 367, 377
Angelidis, Evangelos ..... 1
Angelidis, Evangelos ..... 89
Aurich, Paul ..... 29
Bargstädt, Hans-Joachim ..... 501
Baudach, Jens ..... 279
Bause, Falko ..... 641
Beißert, Ulrike ..... 147
Bendavid, Illana ..... 337
Bergmann, Sören ..... 481
Bernhard, Jochen ..... 239
Blauermel, Gregor ..... 319
Bohn, Daniel ..... 89
Bolton, Carlos Contreras ..... 49
Böttcher, Lars-Boris ..... 239
Brambring, Felix ..... 459
Brandau, Annegret ..... 289
Brüggemann, Holger ..... 691
Canyakmaz, Caner ..... 347
Claus, Thorsten ..... 79, 299, 575
Clausen, Uwe ..... 279, 565
Conrads, Alena ..... 523
Daniluk, Damian ..... 259
Deininger, Maik ..... 387
Delgado, Antonio ..... 651
Diekmann, Daniel ..... 279
Dietz, Thomas ..... 239
Doleschal, Dirk ..... 397
Dross, Felix ..... 167
Dunke, Fabian ..... 69
Eckstein, Bernd ..... 249

Eilers, Kevin ..... 269
Elbert, Ralf ..... 621
Ertl, Willi ..... 299
Fabig, Christian ..... 101
Fanghänel, Christin ..... 545
Fay, Alexander ..... 407
Feine, Immo ..... 501
Feist, Markus ..... 329
Felberbauer, Thomas ..... 367, 377
Feldkamp, Niclas ..... 481
Feller, Nico ..... 469
Fernández Alcalde, Ana Maria ..... 329
Fikar, Christian ..... 49
Fischer, Jörg W. ..... 417
Fowler, John ..... 447
Franzke, Torsten ..... 621
Freitag, Michael ..... 309
Friedewald, Axel ..... 613
Gamber, Thilo ..... 711
Glock, Christoph H. ..... 621
Golabchi, Alireza ..... 491
Groß, Frauke ..... 651
Grosse, Eric H. ..... 621
Gruler, Aljoscha ..... 49
Grundmeier, Nico ..... 555
Güller, Mustafa ..... 207
Günthner, Willibald A. ..... 721
Gust, Peter ..... 469
Habenicht, Ilka ..... 513
Hahn, Axel ..... 555
Hammel, Christian ..... 219, 229, 631
Han, SangUk ..... 491
Hanschke, Thomas ..... 111
Haux, Moritz Alexander ..... 613
Heck, Aron ..... 513
Heckmann, Iris ..... 177

Heger, Jens ..... 11
Hegmanns, Tobias ..... 137, 207
Heinrich, Marek ..... 119
Hellingrath, Bernd ..... 157
Herrer, Yale ..... 337
Herrmann, Frank ..... 79, 575
Hildebrandt, Torsten ..... 11, 309
Hinze, Ulrich ..... 481
Hirsch, Patrick ..... 49
Horstkemper, Dennis ..... 157
Hubert, Stefan ..... 651
Hübl, Alexander ..... 701
Iakovou, Eleftherios ..... 357
Ihle, Norman ..... 555
Jahn, Carlos ..... 187, 585
Jami, Neil ..... 147
Janssen, Larissa ..... 79
Jørgensen, Bent Aksel ..... 427
Juan, Angel A. ..... 49, 59
Kaffka, Jan ..... 565
Kamphues, Josef ..... 137
Karaesmen, Fikri ..... 347
Karakaya, Elif ..... 207
Keramydas, Christos ..... 357
Klaas, Alexander ..... 129
Klibi, Kamel ..... 129
König, Franz-Joseph ..... 119
König, Markus ..... 523
Kriege, Jan ..... 641
Kühn, Mathias ..... 89
Künne, Bernd ..... 595
Lämkull, Dan ..... 437
Lange, Ann-Kathrin ..... 187
Lange, Hans Rüdiger ..... 545
Liotta, Giacomo ..... 427
Lödding, Hermann ..... 613

Mallidis, Ioannis ..... 357
März, Lothar ..... 1
Mayer, Gottfried ..... 605
Mayer, Thomas ..... 661
Mieschner, Marielouise ..... 605
Miodrag, Zoran ..... 565
Møller, Charles ..... 427
Mönch, Lars ..... 447
Motta, Marco ..... 147, 681
Müller, Christiane ..... 459
Müller, Ulf ..... 469
Münsterberg, Torsten ..... 585
Mutzke, Harald ..... 319
Nahhas, Abdulrahman ..... 29
Ng, H.C. Amos ..... 437
Nickel, Stefan ..... 69, 177
Ohmann, Lars ..... 1
Ozekici, Suleyman ..... 347
Pappert, Falk Stefan ..... 661
Parlings, Matthias ..... 681
Peirleitner, Andreas Josef ..... 377
Peter, Tim ..... 535
Pettigrew, Liam ..... 651
Pfeilsticker, Lena ..... 59
Pirovano, Giovanni ..... 187
Pitsch, Holger ..... 565
Pöting, Moritz ..... 279
Puntel-Schmidt, Philipp ..... 407
Putz, Matthias ..... 545
Rabe, Markus ..... 59, 167, 387, 671
Rank, Sebastian ..... 219
Reggelin, Tobias ..... 29
Reuter, Christina ..... 459
Roidl, Moritz ..... 197
Rose, Oliver ..... 21, 39, 101, 661
Rossi, Tommaso ..... 187

Roßmann, Jürgen ..... 269
Roth, Stefan ..... 119
Scheffer, Markus ..... 523
Scheidler, Anne Antonia ..... 671
Schenke, Björn ..... 249
Schieweck, Steffen ..... 197
Schiffmann, Michael ..... 469
Schlegel, Andreas ..... 545
Schmaler, Robert ..... 229, 631
Schmidt, Thorsten ..... 219, 229, 631
Schneider, Germar ..... 219
Schnirzer, Rene Kurt ..... 367
Schönherr, Oliver ..... 101
Schöps, Matthias ..... 229, 631
Schröder, Michael ..... 147
Selmair, Maximilian ..... 575
Shufang, Xie ..... 21
Sikorra, Jan Niklas ..... 613
Sinha, Katrin ..... 595
Skoogh, Anders ..... 437
Sprenger, Philipp ..... 681
Steinhauer, Dirk ..... 613
Stoldt, Johannes ..... 545
Straßburger, Steffen ..... 481
Strasser, Georg ..... 701
Stumvoll, Ulrike ..... 299
Teich, Enrico ..... 575
ten Hompel, Michael ..... 197
ten Hompel, Michael ..... 259
Terlunen, Sebastian ..... 157
Thewes, Markus ..... 523
Tolujew, Juri ..... 29, 289
Uhlig, Tobias ..... 39
Vennemann, Andrea ..... 167
Vlachos, Dimitrios ..... 357
Völker, Michael ..... 89, 101

Völker, Sven ..... 417
Weigert, David ..... 289
Weigert, Gerald ..... 397
Wenzel, Sigrid ..... 535
Wenzler, Florian ..... 721
Wiegmann, David ..... 691
Winter, Elias ..... 101
Witthaut, Markus ..... 137
Wölck, Martin ..... 157
Woldt, Thomas ..... 545
Yucesan, Enver ..... 337
Zhang, Tao ..... 21
Zhou, Zhugen ..... 89
Zisgen, Horst ..... 111
Zülch, Gert ..... 711