

Epistemologische Grenzen der klassischen Validierung bei komplexen Simulationsmodellen

Dr. Marko A. Hofmann
marko@informatik.unibw-muenchen.de
Institut für Technik Intelligenter Systeme (ITIS)
Universität der Bundeswehr München
Werner-Heisenberg-Weg 39
85577 Neubiberg

Kurzfassung

Trotz wachsender Komplexität vieler Simulationsmodelle sind die Anforderungen an die Qualität der Modelle in etwa gleich hoch geblieben. Während aber die Verifikation von Simulationsmodellen durch viele (formale) Verfahren verbessert wurde, ist deren Validierung oft schwierig und unbefriedigend. Dabei handelt es sich nicht nur um ein methodisches Problem, es verbirgt sich dahinter auch eine grundsätzliche epistemologische Grenze des klassischen, auf dem Vergleich des Modells mit empirischen Daten vom Referenzsystem beruhenden Validierungsbegriffes bei komplexen Simulationsmodellen. Gründe hierfür sind die zahlreichen als Arbeitshypothesen dienenden Modellierungsannahmen, die zur Vervollständigung des Wissens über die Wirklichkeit bzw. zur Komplexitätsreduktion dieser Wirklichkeit dienen, die weitgehende Skepsis der zeitgenössischen Wissenschaftstheorie gegenüber allzu positivistischen Positionen sowie die dürftige Prognosesicherheit vieler Modelle. Ihr zum Teil subjektiver Charakter sowie die Schwierigkeit sie vollständig durch empirische Daten zu belegen, verhindern eine Modellvalidierung im klassischen strengen Sinne. Für die explorative Nutzung von Modellen ist diese Strenge daher unzweckmäßig. Hier muss ein flexibler Validierungsbegriff zur Anwendung kommen, der explizit Modellierungsannahmen adressiert und mit einem schwächeren Wirklichkeitsbezug auskommt.

1 Einleitung

Die seit Jahrzehnten erfolgreiche Anwendung von Simulationsmodellen in vielfältigen Anwendungsformen (Analyse, Planung, Ausbildung, Entscheidungsunterstützung etc.) und Anwendungsdomänen hat die Modellierer weltweit dazu ermutigt, immer komplexere Systeme mit immer komplexeren Modellen zu simulieren. Die qualitativen Anforderungen an die Modelle (Korrektheit, Eignung, Glaubwürdigkeit, Performanz, Sicherheit, Zuverlässigkeit, Wirtschaftlichkeit, etc.) sind dabei aber grundsätzlich gleich hoch geblieben. Mit zunehmender Modellkomplexität sinken jedoch insbesondere deren intuitive Glaubwürdigkeit und unmittelbar erfassbare Aussagekraft. Aus diesem Grund sind Modellverifikation und

Modellvalidierung zu herausragenden Forschungsthemen geworden, was sich unter anderem daran ablesen lässt, dass „V&V“ (Verifikation und Validierung) auf vielen Simulationstagungen mit einem eignen Track vertreten ist. Verifikation hat dabei die Aufgabe, die Modellrealisierung (Implementierung des Simulationsmodells) mit der (im Idealfall) eindeutigen und vollständigen Spezifikation dessen, was der Anwender vom Modell erwartet, zu vergleichen und gegebenenfalls anzugleichen (oft vereinfacht zu: Wurde das Modell richtig erstellt?). Validierung bezieht sich dagegen auf die zweckgebundene Eignung und Realitätsnähe des Modells (Wurde das richtige Modell erstellt?). Während sich Verifikation durch viele Verfahren (z. B. formale Spezifikationstechniken, Model Checking, etc.) nachprüfbar verbessern lässt, bleibt Validierung bisher meist mühsame Arbeit, bisweilen gar Kunst.

Im Folgenden wird der Versuch gemacht zu erläutern, warum es sich dabei nicht nur um ein methodisches Problem handelt, sondern um eine grundsätzliche epistemologische Grenze des klassischen Validierungsverständnisses bei komplexen Simulationsmodellen.

2 Komplexitätsbeherrschung durch Modellierungsannahmen

Gemäß der in der Systemtheorie üblichen Definition zeichnen sich komplexe Systeme durch eine große Anzahl an Systemelementen und unterschiedlichen Relationen zwischen diesen Elementen aus (siehe z.B. [1]). Diese äußerst dehnbare Begriffsbestimmung lässt sich durch Angabe typischer (aber weder notwendiger noch hinreichender) Eigenschaften komplexer Systeme präzisieren. Hierzu gehören beispielsweise Nichtlinearität, Emergenz, Selbstorganisation, Heterogenität, Pfadabhängigkeit, Rückkopplung und die Existenz von Attraktoren. Diesen Eigenschaften gemeinsam ist die Schwierigkeit ihrer unmittelbaren Erfassbarkeit und anspruchsvollen Formalisierbarkeit. Verständlicherweise führt die Modellierung derart ausgezeichneter Systeme daher häufig (aber nicht zwingend) auch zu komplexen Modellen. Die tatsächliche Komplexität des Modells wird neben der Referenzsystemkomplexität durch die Zielsetzung der Modellierung bestimmt, die das zu lösende Problem widerspiegelt. Grundsätzlich wird mit jeder Modellierung eines komplexen Systems auch eine Komplexitätsreduktion verbunden sein, da die Wirkungszusammenhänge im System selbst zu „komplex“ für den menschlichen Betrachter sind. Neben der einfachen Aggregation von Systemelementen und der Abstraktion von ausgewählten Systemattributen muss bei der Modellierung komplexer Systeme auch versucht werden, die oben genannten Eigenschaften vereinfachend zu berücksichtigen. Bei all diesen Vereinfachungen kommen Modellierungsannahmen ins Spiel: Da die genaue Art der Nichtlinearität, Emergenz, Selbstorganisation, usw. in aller Regel ebenso wenig a priori bekannt ist wie die „richtige“ Aggregation und Abstraktion, wird der Modellierer meist akzeptierte, oft aber auch nur plausible und manchmal gar kontrafaktische Annahmen über diese Eigenschaften machen. Ob die Annahmen in Modellen realistisch sein müssen oder nicht, wird seit Milton Friedmans Aufsatz [2] zumindest in der Ökonomie intensiv diskutiert (siehe [3] und [4]).

Als Beispiele für derartige Annahmen seien genannt: Die Rationalität oder gar Homogenität der individuellen Käufer in Marktmodellen, die Vernachlässigung psychologischer

Faktoren in Warteschlangenmodellen (nur globale Ankunft- und Bedienzeiten), die Stabilität der Mortalität in Versicherungsmodellen, die Linearisierung verschiedenster Sachverhalte für die klassische Optimierung oder die Gleichförmigkeit der Niederhaltewirkung durch Feuer in militärischen Modellen.

3 Hypothesen und Annahmen

Unter einer „Modellierungsannahme“ wird im Folgenden eine bisher noch nicht oder nur unzureichend empirisch belegte Aussage A eines Subjekts S mit der Intention I (Zweck der Modellierung) bezüglich eines Originals O (Referenzsystem) für einen Zeitrahmen T verstanden. Modellierungsannahmen sind immer dann notwendig, wenn über Wirkungsbeziehungen in der Realität bisher nichts (unzureichendes) bekannt ist oder vom Modell benötigte Eingangsdaten nicht zu beschaffen sind. Die Formulierung von Modellierungsannahmen in ihrer klassischen Form als Hypothese ist wesentlicher Grundpfeiler aller naturwissenschaftlichen Erkenntnis. Das Besondere an der Hypothese ist, dass sie im Zuge der Modellverwendung im Zentrum empirischer Untersuchungen steht. Sie wird aufgrund des Aufbaus des empirischen Experiments widerlegt oder kann nicht abgelehnt werden, gilt mithin also als vorläufig bestätigt. Im Gegensatz dazu stehen die vielfältigen Annahmen, die nötig sind um komplexe Systeme simulieren zu können, nicht alle tatsächlich zur Diskussion und können häufig empirischer Überprüfung auch nicht unterzogen werden. Sie dienen als Arbeitshypothesen zur Vervollständigung des Wissens über die Wirklichkeit bzw. zur Komplexitätsreduzierung dieser Wirklichkeit an sich, ohne die das Modell nicht zu erstellen wäre. Der Unterschied zwischen naturwissenschaftlichen Hypothesen und zur Modellierung komplexer Systeme benutzter Annahmen besteht somit darin, dass aus der Nichtübereinstimmung von empirischen Daten und *Modellergebnissen* nicht eindeutig auf die Fehlerhaftigkeit *einer bestimmten* Annahme geschlossen werden kann. Umgekehrt kann die bloße Übereinstimmung von empirischen Daten mit Simulationsläufen nicht beweisen, dass die modellierten (und vielfach angenommen) Wirkungszusammenhänge auch in der Realität erfüllt sind. Simulationen dieser Art basieren auf Modellen, denen der Anspruch auf Validität im klassischen Sinne fehlt. Aus diesem Grund können solche Modelle auch nicht als „Abbilder“ der Realität verstanden werden. Eine fundamentale Kritik der Vorstellung, dass (ökonomische) Modelle Abbilder der Wirklichkeit seien, findet sich beispielsweise in [5]. Schmidt und Schor [6] nennen Modelle gar „realistische Fiktionen“. (Besonders drastisch zeigt sich die Problematik im Falle noch nicht existierender Systeme, bezüglich derer Prognosen erstellt werden sollen. Wie sollen sie qualitativ beurteilt werden?)

Auch in der militärischen Domäne ist der problematische Wirklichkeits- bzw. Problembezug vieler Modelle seit Langem bekannt [7]. Man spricht dort von sogenannten „explorativen“ Modellen, die dazu dienen, durch Experimente an hypothetischen „Welten“ etwas über die Wirklichkeit zu lernen. Der Wirklichkeitsbezug dieser Modelle reicht alleine in der Regel nicht aus um ihre Validität zu bestimmen. Als Lösung haben Davis und Bigelow [8] daher vorgeschlagen, für derartige Modelle den Begriff der Validität zu erweitern. Danach sind die wesentlichen Qualitätsmerkmale explorativer Modelle ihre Verständlichkeit, Er-

klärbarkeit und ihr Potential mit Unsicherheit auf transparente Weise umzugehen. Hierauf aufbauend definieren sie die Eigenschaften „valide für explorative Analysen“ und „valide, in Abhängigkeit von (folgenden) Modellierungsannahmen für explorative Analysen“. Das erste Prädikat kann beispielsweise dann Verwendung finden, wenn sich ein Modell in vergleichbaren Fällen als nützlich erwiesen hat (empirisch belegt wurde), im vorliegenden Fall jedoch Unsicherheit über die Qualität der verwendeten Eingabedaten besteht und Annahmen über diese Eingabedaten getroffen werden mussten. Das Modell liefert dann sicher keine präzisen Prognosen, kann aber beispielsweise zur Generierung und dem Vergleich verschiedener Handlungsalternativen oder auch nur zur Darstellung *möglicher* Entscheidungskonsequenzen verwendet werden. Das zweite Prädikat erscheint nicht ganz zu unrecht zirkulär im Sinne der Formulierung „Das Modell ist valide, wenn die Annahmen valide sind“. Trotzdem hat das Prädikat einen entscheidenden Vorteil: Der Modellierer macht seine Annahmen explizit und verknüpft dadurch seine Resultate mit Vorbedingungen, die von anderen akzeptiert oder abgelehnt werden können. Es stellt sich allerdings die Frage, wie der *Prozess der Validierung* bei Zielrichtung auf diese Prädikate aussehen soll? Sieht es doch ganz danach aus, als ob beide Prädikate „negativ“ bestimmt werden, also nur dann zum Einsatz kommen, wenn empirische Belege der Validität fehlen. Wie aber wird ein für „explorative Analysen valides“ Modell konstruiert? Hierzu scheint es (bisher?) kein Standardverfahren zu geben.

4 Validierung im Lichte verschiedener Epistemologien

Die grundsätzlichen Möglichkeiten und Grenzen der Validierung hängen unmittelbar mit der Epistemologie zusammen, an die der Modellierer *glaubt*. Die verschiedenen Positionen, die die Wissenschaftstheorie (als spezieller Zweig der Epistemologie) bezüglich der Validierung von Simulationen grundsätzlich bietet, wurden von Kleindorfer et. al. in [9] ausführlich beschrieben. Die eindrucksvollste Darstellung der unterschiedlichen Möglichkeiten ergibt sich aus dem Vergleich zweier Extrempositionen, beispielsweise dem logischen Positivismus Carnaps und dem radikalen Konstruktivismus von Glasersfelds. Während für den Positivisten ein Modell nur entweder wahr oder falsch (der Fall oder nicht der Fall) sein kann, sind dem Konstruktivisten alle Modelle nur nach dem Kriterium ihrer subjektiven Angemessenheit bewertbar. Somit ergeben sich fundamental voneinander unterschiedene Wirklichkeitsbezüge und in Folge dessen auch vollkommen unterschiedliche Qualitäten des Prädikats „valide“. Ungeachtet diesen Extrembeispiels wird die Unabhängigkeit der Modellvalidierung vom jeweiligen Modellierer/Validierer heute eher kritisch betrachtet (siehe [9] und [10]). Barlas und Carpenter [11] gehen sogar soweit zu behaupten, dass „Validierung eher eine Frage sozialer Konversation denn objektiver Konfrontation ist“. Auch wenn diese Einschätzung nicht auf alle Modelle zutrifft, so lässt sich doch schlussfolgern, dass ein ausschließlich auf objektiver Empirie begründeter Validitätsbegriff für Simulationen im Allgemeinen zu kurz greift.

5 Einsicht statt Prognosesicherheit

Als drittes Argument für den grundsätzlichen Charakter der Validierungsgrenzen bei komplexen Simulationsmodellen tritt deren geringer prognostischer Wert. So haben sich beispielsweise militärische Konflikte zwar sehr gut post festum analysieren lassen, korrekte Vorhersagen über den Verlauf dieser Auseinandersetzungen gelangen dagegen nur als Zufallstreffer (Beispielsweise lagen alle von den US-Streitkräften eingesetzten Modelle zur Abschätzung der (während des klassischen Gefechts) zu erwartenden eigenen Verluste in beiden Irakkriegen um mindestens eine Zehnerpotenz zu hoch [mündliche Mitteilung, RAND]). Im Rahmen des klassischen Validitätsbegriffes müssten solche Modelle als ungültig und mithin unnützlich bezeichnet werden. Tatsächlich haben diese Modelle jedoch erheblich zum besseren Verständnis („Insight“, „The purpose of computing is insight, not numbers“ [12], siehe auch [13]) vieler Informations- und Wirkprozesse geführt, die letztlich auch zur den tiefgreifenden Reformen geführt hat, die moderne westliche Armeen seit Ende der 90er Jahre in Gang gesetzt haben („Revolution in Military Affairs“, „Transformation“). Einsicht, so hat sich gezeigt, ist nicht an die empirische Validität der Modelle gebunden. Zentral ist vielmehr die Möglichkeit zur dynamischen Visualisierung jener Prozesse, die vom Modellierer aus der Wirklichkeit abstrahiert wurden. Erneut spielen dabei Transparenz und Nachvollziehbarkeit eine entscheidende Rolle bei der Modellbeurteilung. Die Modelle belegen ihre Eignung dadurch, dass sie die Komplexität und Unsicherheit der Systemanalyse reduzieren. Ob diese Reduktion gerechtfertigt ist, lässt sich erst a posteriori zeigen. A priori können die Modelle aber die Kommunikation zwischen all jenen, die am Systemverhalten interessiert sind, erheblich vereinfachen, da die Annahmen über Eingangswerte und postulierte Wirkungsbeziehungen explizit im Modell vorliegen und leicht variiert werden können. Im Wechselspiel von Simulation und Diskussion entsteht so im Idealfall jene Einsicht, die hilft, zumindest schwere Fehler zu vermeiden. Der dieser Betrachtungsweise zugrundeliegende hermeneutische Ansatz zur Validierung wurde bereits in [9] beschrieben und von Feinstein und Cannon ausführlich in [14] ausgearbeitet.

6 Schlussfolgerung

Die Schwierigkeiten der Validierung komplexer Simulationsmodelle allein auf der Grundlage empirischer Vergleichsdaten sind nicht allein methodologischer Art, sie ergeben sich auch aus epistemologischen Grenzen der Modellbildung. Es sind insbesondere die Subjektivität der Modellierung, die sich in unterschiedlichen Modellierungsannahmen manifestiert, und der geringe prognostische Wert im engeren Sinne, die eine Begriffserweiterung notwendig erscheinen lassen. Viele Simulationsmodelle komplexer Systeme dienen nicht zur harten Analyse von Fakten, sie erkunden vielmehr „explorativ“ mögliche Systemzusammenhänge auf der Basis angenommener Beziehungen und deren Implikationen. Sie sind dabei vielmehr Wirklichkeitskonstruktionen als Wirklichkeitsrepräsentationen. Auch wenn einige wünschenswerte Eigenschaften solcher Modelle bekannt sind (Transparenz, Verständlichkeit, Fähigkeit zur Repräsentation unsicheren Wissens) und sich ein neues Qua-

litätsprädikat „valide für explorative Analysen“ bei militärischen Modellen bewährt zu haben scheint, bleibt eine zentrale Frage derzeit weitgehend offen: Nach welchen Vorgehensweisen sollten Modelle, die ausdrücklich für explorative Zwecke bestimmt sind, entwickelt werden?

Literatur

- [1] R. L. Flood and E. R. Carson. *Dealing with complexity: An introduction to the theory and application of systems science*. Plenum Press, New York, 1993.
- [2] M. Friedman. *Essays in Positive Economics*. University of Chicago Press, Chicago, 1953.
- [3] U. Mäki. Kinds of assumptions and their truth: Shaking an untwisted f-twist. *Kyklos*, 53(3):317–336, 2000.
- [4] A. Musgrave. Unreal assumptions in economic theory. *Kyklos*, 34:377–387, 1981.
- [5] W. Bretzke. *Der Problembezug von Entscheidungsmodellen*. Mohr, 1980.
- [6] R. H. Schmidt and G. Schor. *Modelle in der Betriebswirtschaftslehre*. Gabler, 1987.
- [7] C. F. Hermann. Validation problems in games and simulations with special reference to models of international politics. *Behavioral Science*, 12:216–231, 1967.
- [8] P. K. Davis and J. H. Bigelow. *Implications for Model Validation of Multiresolution, Multiperspective Modeling (MRMPM) and Exploratory Analysis*. RAND, 2003.
- [9] G. B. Kleindorfer, L. O’Neill, and R. Ganeshan. Validation in simulation: Various positions in the philosophy of science. *Management Science*, 44(8):1087–1099, 1998.
- [10] S. R. Irvine, R. R. Levary, and M. S. McCoy. The impact of judgemental biases on the validation of simulation models. *Simulation & Gaming*, 29(2):152–164, June 1998.
- [11] Y. Barlas and S. Carpenter. Philosophical roots of model validation: Two paradigms. *Systems Dynamics Review*, 6:48–166, 1990.
- [12] R. W. Hamming. *Numerical Methods for Scientists and Engineers*. Dover Publications, 2nd (1st: 1962) edition, 1987.
- [13] G. Huntington, J. Weyant, and J. Sweeney. Modeling for insights, not numbers: The experience of the energy modeling forum. *Omega*, 10:449–462, 1982.
- [14] A. H. Feinstein and H. M. Cannon. A hermeneutical approach to external validation of simulation models. *Simulation and Gaming*, 34(2):186–197, 2003.