

Bewertung von Modellen (Validierung)

Modelle sind durch Vereinfachung und Abstraktion entstandene Abbilder realer Sachkomplexe. Sie werden zu bestimmten Zwecken entwickelt. Daraus ergibt sich am Ende die Notwendigkeit, Modell und Original gegenüberzustellen. Welche Punkte dabei im Detail zu untersuchen sind, hängt vom jeweiligen Fall ab. Es lassen sich dennoch einige allgemeine Kriterien angeben, welche die *Validität (Gültigkeit)* eines Modells stützen können.

Vorab ist ein methodologischer Grundsatz herauszustellen: Es kann nicht darum gehen, die "Richtigkeit" eines Modells ein für alle Mal zu *beweisen*. Zum einen reicht das Beweisen, wie wir es aus der Mathematik kennen, gar nicht in die Erfahrungswelt hinein. Zum anderen verlangt der Anwendungskontext einer Modellbildung meistens weniger. Um ein Modell als valide zu bestätigen, sollte man es *belasten*. Eine Belastungsprüfung ist -- zugespitzt gesagt -- der Versuch, ein Modell in der einen oder anderen Hinsicht zum Scheitern zu bringen. Dies ist der Inhalt des sog. *Falsifikationsprinzips* (von K. Popper in seiner *Logik der Forschung* aufgestellt). Ein Modell, das eine Reihe von Falsifikationsversuchen überstanden hat, erlangt damit eine provisorische Gültigkeit.

Nach Bossel [*Modellbildung und Simulation*. Vieweg: Braunschweig/Wiesbaden 1992, S. 36] lassen sich vier zentrale Bereiche für die kritische Befragung eines Modells ausmachen:

Verhaltensgültigkeit

(falls das Verhalten eines Systems beschrieben wird): Stimmt das Verhalten des Modells mit dem Verhalten des modellierten Systems (unter gleichen oder analogen Bedingungen) qualitativ soweit überein, wie es der Modellzweck verlangt?

Strukturgültigkeit

Gibt das Modell die innere Struktur des Sachkomplexes, d.h. die funktionalen und kausalen Beziehungen seiner Bestandteile, im wesentlichen (im Rahmen des Modellzwecks angenähert) richtig wieder?

Empirische Gültigkeit

Stimmen die Ausgabedaten, insbesondere durch das Modell vorhergesagte Ergebnisse, im Bereich des Modellzwecks mit den entsprechenden Werten aus dem Erfahrungskontext überein? Für den Fall, dass Beobachtungen fehlen, sollten sie zumindest konsistent und plausibel sein.

Anwendungsgültigkeit

Lässt sich das Modell tatsächlich so anwenden, dass der ursprüngliche Zweck der Modellbildung erfüllt ist?