hier sind Berechnungsaufgaben relativ zu den Eingaben zu sehen. Hier allerdings nur zu diesen und nicht auch zu irgendwelchen implizierten Modellannahmen. Damit kann das Bilanzverfahren als ideale Simulationsmaschine gesehen werden, was für die Beantwortung von Planungsfragen sehr wichtig ist: Die Erzeugung korrekter Informationen zuhanden der zu Beratenden ist zwar nicht hinreichende, wohl aber notwendige Voraussetzung für glaubwürdige Planung. Das Bilanzverfahren erlaubt einerseits Schätzungen künftiger Ereignisse aufgrund unscharfen Wissens darüber. Andrerseits kann aufgrund normativer Vorgaben über künftige Ereignisse geschätzt werden, was sich wie entwickeln müsste, damit die Vorgabe erreichbar wäre. Hiermit werden u.E. zwei wesentliche Informationsbedürfnisse der Planung in korrekter Weise abgedeckt.

Zürich, 7.10.1981/MS/el

Beat SCHMID

Institut für Orts-, Regional- und Landesplanung der Eidgenössischen Technischen Hochschule Zürich

PROBLEME MIT PROGNOSEMODELLEN - EIN LOESUNGSBEITRAG

Inhalt

Gegenstand der nachfolgenden Ausführungen sind quantitative Modelle, aufgefasst als Werkzeuge oder Hilfsmittel der Planung im weitesten Sinne. Das Subjekt solcher Planung werde mit dem Begriff 'Organisation' bezeichnet. Dieser Rahmen ergibt bestimmte Forderungen an Modelle. Im ersten Teil sollen diese, der verwendete Modellbegriff und die Probleme mit Modellen sozioökonomischer Systeme (z.B. Regionen) gestreift werden, um die verwendete Optik erkennbar werden zu lassen. Im zweiten Teil wird ein eigener Modellansatz vorgestellt: Bilanz- oder Datenmodelle. Dabei wurden Konzepte, die teilweise oder gesamthaft jedem quantitativen Modell zugrunde liegen, systematischer als üblich ausgewertet. Es wird sich zeigen, dass diese Basiskonzepte allein schon sehr viel leisten. Damit kann oft auf das Hinzunehmen höherer abgeleiteter und damit problemreicherer Konzepte wie normologische Hypothesen ("Modellgesetze") u.ä. verzichtet werden.

1. Handeln und Modelle

Modelle werden in verschiedenen Umfeldern konstruiert und eingesetzt. Ein wichtiges Feld ist etwa das Bemühen, mit Modellbildungen zur Erklärung von beobachtbaren Phänomenen beizutragen. Wir verfolgen hier nicht dieses Anliegen, sondern gehen vom Handeln aus, das Antworten benötigt auf Fragen der Art: Was ist, wenn...? Was muss sein, damit...? Auch wenn zwischen diesen beiden Umfeldern der Modellentwicklung und des Modelleinsatzes tiefe Zusammenhänge bestehen, so dürfen doch die Unterschiede nicht übersehen werden. Ein wichtiger

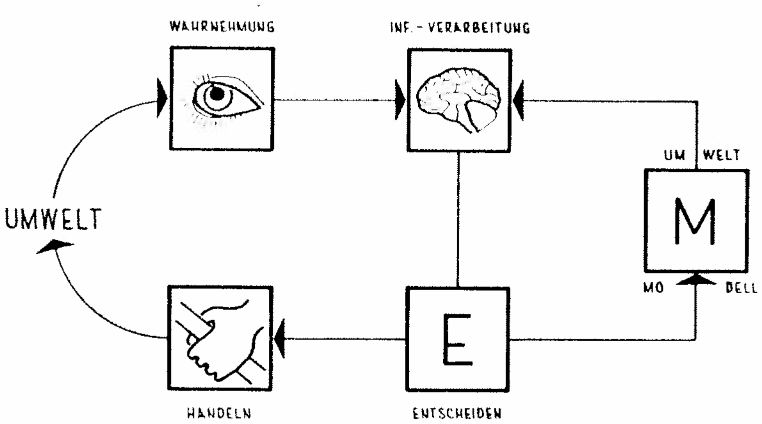
betrifft die vom Modell verwendete Information: Im Handlungskontext darf nur gesicherte oder vom Handlungssubjekt akzeptierte Information verarbeitet werden - das Modell fungiert hier als Informationstransformator. In der Laborsituation des Bemühens um Erklärungen dagegen ist es erlaubt, auch mit zunächst wilden Hypothesen zu experimentieren: Hier richten sie keinen Schaden an, sie können vielmehr oft zu nutzbringenden Erkenntnissen herangezüchtet werden, die dann - und erst dann - auch dem Handelnden dienen.

Modelle für Organisationen

Nach dem Gesagten interessieren hier Modelle im Rahmen planvollen Handelns. Zur Planung gehört eine Trennung der Welt in planendes Subjekt und zu planende Umwelt. Verbunden sind beide durch mögliche Aktionen des Subjektes in der Umwelt und durch eine bestimmte Wahrnehmung der Umwelt durch das Subjekt. Das planende Subjekt ist dabei ein Organismus oder eine Organisation, wobei im Zusammenhang mit der Planung mindestens folgende Funktionen unterschieden werden sollten:

- die informationsverarbeitende Funktion
- eine entscheidende Instanz und
- die handelnden, auf die Umwelt einwirkenden "Organe".

Das nachfolgende Schema veranschaulicht diesen Sachverhalt mit Symbolen, die biologischen Organismen entnommen sind und ihn gut bezeichnen.



Diese Funktionen sind auch bei niederen Organismen anzutreffen. Bei ihnen ist der Entscheidungsmechanismus im wesentlichen ein Reiz-Reaktions-Muster. Man kann dort deshalb nicht von Planung sprechen. Höhere Organismen wie der Mensch verfügen über die Möglichkeit, Erwägungen anzustellen: Sie können sich das Ausführen einer Handlung vorstellen, also die Wirkungen der Handlung extrapolieren und sie so mit denen anderer Handlungen vergleichen. Diese Fähigkeit setzt die Existenz eines Organs (in der biologischen Bedeutung des Wortes) voraus, das die Simulation der Wirkungen einer Handlung auf die Welt erlaubt. Man benötigt also eine Art Modell der Umwelt, neben der realen Umwelt, analog zum Sandkasten mit den Bleisoldaten etc. des Generals. Wir nennen dieses Organ daher kurz Umweltmodell. Gemeint ist die 'Infrastruktur' zur Funktion 'Simulation', egal, wie sie im Einzelfall realisiert wird.

Der Mensch besitzt in seinem Hirn einen solchen Apparat. Er ist sehr leistungsfähig. Die Evolution hat ihn mit jenen Fähigkeiten ausgestattet, die in den letzten 100 000en von Jahren wichtig waren. Sie hat ihn jedoch nicht für Situationen und Ereignisse ausgerüstet, die diesen Rahmen wesentlich sprengen. Genauer: er wird mit den Aufgaben fertig, die die Daseinsform des Menschen bis vor wenigen 1000 Jahren ausmachten, nämlich der Stamm oder die Sippe. Grössere Verbände wie Dorf- oder Stadtgemeinschaften bedurften bereits beachtlicher 'Ueberbeuleistungen'. Diese Grossgemeinschaften wurden durch die Existenz der Sprache möglich und mit entsprechenden Erweiterungen der Sprache beherrschbar, die ihren Niederschlag in Konstruktionen wie Recht und Verwaltungen fanden.

In der jüngsten Vergangenheit sind die Organisationen jedoch immer umfangreicher und komplexer geworden. Die Planung, die wir hier meinen, bezeichnet das Lenken der Handlungen solcher komplexer und sehr komplexer Organisationen, wie es die modernen Städte, Regionen, Wirtschaftssysteme etc. mit ihren Verwaltungen sind. Die uns angeborenen oder im "normalen" sprachlichen Ueberbau vorhandenen Modelle der Umwelt reichen zur planenden Lenkung solch komplexer Systeme jedoch nicht mehr aus.

Quantitative Modelle

Aus der eben beschriebenen Aufgabe eines Modelles im Entscheidungsprozess ergibt sich: Ein Modell ist eine Repräsentation einiger interessierender Aspekte der Umwelt. Dieser Modellbegriff setzt folgendes voraus:

- Neben der abzubildenden Umwelt wird ein weiterer Bereich von Objekten als Modellbereich benötigt.
- Beide Bereiche - Umwelt und Modellbereich - müssen in einem gemeinsamen Medium repräsentiert sein. In unserem Falle ist dieses Medium die Sprache.
- Die beiden Bereiche gemeinsamen Strukturmerkmale müssen bezeichnet werden. Es muss eine Abbildung α (Vorschrift, Rezept) angegeben werden, wie der erste Bereich B_1 hinsichtlich dieser Strukturmerkmale in den Modellbereich B_2 abzubilden ist:

$$B_1 \xrightarrow[\alpha]{} B_2$$

Ein Bereich B_2 kann also als Modell für einen anderen Bereich B_1 dienen, wenn zwischen B_1 und B_2 eine <u>Abbildung</u> existiert, die gewisse Strukturmerkmale von B_1 auf entsprechende Strukturmerkmale in B_2 abbildet. (Beispiel: B_1 = Stadt Zürich, B_2 = Stadtplan von Zürich.) Die durch die Abbildung erfassten <u>Strukturmerkmale</u> sind jeweils explizite zu nennen. Ihre Kenntnis erst macht es möglich, ein Modell zu erstellen (- im Beispiel: eine Karte von Zürich zu zeichnen).

Um dem Modell praktische Relevanz zu verleihen, muss die Modellabbildung mindestens teilweise <u>umkehrbar</u> sein. (Im Beispiel: Man sollte die Karte "lesen" können, um sich so in Zürich zurechtzufinden.)

Man sieht unmittelbar ein, dass dem gleichen Bereich B_1 verschiedene, ev. nicht vergleichbare Modelle zugeordnet werden können. Weiter: Der Modellbereich B_2 kann selbst als Ausgangsbereich für eine Modellabbildung in einen dritten Bereich B_3 dienen:

$$B_1 \xrightarrow[\alpha_1]{} B_2 \xrightarrow[\alpha_2]{} B_3$$

Falls die Abbildungen α_1 und α_2 zusammensetzbar sind (- auf die Bedingungen dafür soll hier nicht eingegangen werden -), kann B_3 auch als Modell von B_1 aufgefasst werden. Um zu quantitativen Modellen zu gelangen, wird meist eine solche mehrstufige Folge von Modellabbildungen durchlaufen. Von <u>quantitativen Modellen</u> spricht man dann, wenn der (so erreichte) Bildbereich der Modellabbildung eine mathematische Struktur ist. So ist es etwa seit Descartes möglich, geometrischen Strukturen, wie im Beispiel durch die Karte gegeben, eine mathematische (analytische) Repräsentation zuzuordnen. Diese Abbildung ist in den Naturwissenschaften bekanntlich sehr erfolgreich. Kein Wunder daher, dass sie im Gebiet der quantitativen Modelle sozioökonomische Systeme oft zu kopieren versucht wird. Dabei werden folgende zwei Umstände dieser Abbildung häufig verwendet:

- Die analytische Darstellung der räumlichen Aspekte eines Systems - z.B. einer Region - ermuntert zur Verwendung naturwissenschaftlicher Konzepte wie (Gravitations-) Wechselwirkung, Ströme etc.
- Die analytische Darstellung geometrischer Objekte wie Kurven oder Flächen besitzt die Gestalt

von mathematischen Gleichungen. Diese werden oft als Modellgesetze bezeichnet und als Kern des Modells genommen. Man sucht dann bei Modellentwicklungen nach solchen "Gesetzen" - oft zu voreilig: der eigentliche Kern eines Modelles ist das Konzept der strukturerhaltenden (hormomorphen) Abbildung!

Modelle sozioökonomischer Systeme

Es gibt inzwischen zahlreiche quantitative Modelle sozioökonomischer Systeme. In ihrer Mehrzahl sind sie den naturwissenschaftlichen Modellen nachgebildet und damit nomologisch orientiert, d.h. sie enthalten Modellgesetze. Gleichzeitig sind sie im allgemeinen wenig erfolgreich. Statt eine gerechtere und differenziertere Darstellung zu geben, sei an ihren Zweck der Politikberatung in (komplexeren) Organisationen erinnert. Dort findet man folgendes:

- Die zu organisierende Umwelt ist zunächst sprachlich repräsentiert.
- Quantitatives ist in dieser Sprache in der Form von gegebenen oder gesuchten Daten vorhanden. Zu deren Mitteilung existieren neben dem Code (- meist in Form von Zahlen -) Indikatorbegriffe (Beispiele: monetäre Grössen wie Bilanzposten, Mengenangaben für Güter etc.). Diese Begriffe sollten in dem Sinne operationalisiert sein, dass es klar ist, wie sie auf ein Objekt oder einen Objektbereich anwendbar sind und wie sie zu "messen" sind.
- Jedes quantitative Modell benötigt als Eingabe (Input) solche Daten.

- Die Ergebnisse einer Modellanwendung (Output) müssen in diese Form übersetzbar sein, um für den Entscheidungsträger verwertbar zu sein.

Wir bezeichnen im folgenden solche vorhandenen oder gesuchten Daten, zusammen mit den zugehörigen Indikatorbegriffen oder Variablen als (Teil-)Informationen.

Ein Modell als Planungsinstrument ist also - logisch betrachtet - eine 'Maschine', die bestimmte Outputinformationen O auf der Basis gewisser (Input-)Informationen I herleitet, indem ein bestimmtes Deduktionsverfahren Anwendung findet. Die Planung interessiert die logische Struktur der Modelle nur mittelbar. Ihr unmittelbares Interesse gilt dem Output einer Modellanwendung. Die Resultate O einer Modellapplikation sind jedoch als logische Folgerungen aus I konditionale Aussagen, d.h. wenn- dann- Aussagen: Wenn I richtig ist, dann gilt O. Wir nehman dabei an, dass das oben erwähnte Deduktionsverfahren logisch korrekt arbeitet. Diese Annahme ist eine Minimalforderung an diskutable Modelle - wenngleich sie auch nicht bei allen gebräuchlichen Modellen erfüllt ist. Modelle sind daher als Planungsinstrumente nur brauchbar

- wenn O effektiv benötigte Informationen enthält
- wenn das verwendete Deduktionsverfahren korrekt ist
- wenn die in I verlangten bzw. enthaltenen Aussagen für den Planer greifbar und akzeptabel sind.

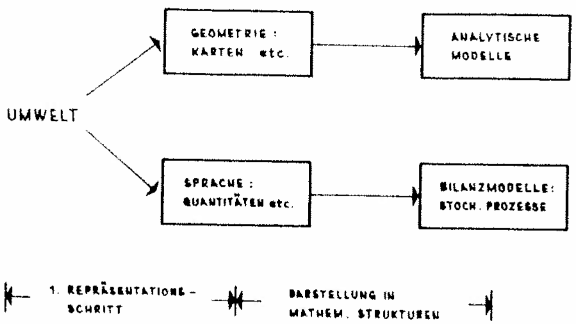
Uns beschäftigen hier die Forderungen an die Aussagenmenge I auf der die Resultate O beruhen. Zunächst ergibt sich eine Anforderung an die Form der Aussagen: Da bei planerischen Aufgaben meist Aussagen zu künftigen Zeitpunkten benötigt werden, d.h. also O sich auf die Zukunft beziehen soll, müssen in I ebenfalls Aussagen enthalten sein, die Künftiges betreffen. Wenn es sich dabei nicht um normative Aussagen handelt, d.h. um willentlich Gesetztes, sondern um Aussagen, die unser Wissen oder unseren Glauben beschreiben, so werden sie immer dann probabilistischen Charakter haben, wenn nicht ein einziger Sachverhalt denkbar ist, sondern mehrere, die für mehr oder weniger wahrscheinlich gehalten werden. So sind z.B. für eine Grösse wie die Geburtenrate im Jahre 2000 verschiedene Werte denkbar. Die einzige heute brauchbare Art, solches Wissen oder solchen Glauben zu codieren, ist die Verwendung der Konzepte der Wahrscheinlichkeitstheorie. Dort fasst man beispielsweise die Geburtenrate als Zufallsgrösse auf. Dies erlaubt Aussagen wie 'Mit einer Wahrscheinlichkeit von 90% ist die Geburtenrate im Jahre 2000 mindestens a und maximal b'. Hier wird ein subjektiver Wahrscheinlichkeitsbegriff verwendet. Das bedeutet: Um die Inputaussagen I für den Planer akzeptierbar zu machen, ist es - von Spezialfällen abgesehen - nötig, dass das Modell eine probabilistische Form dieser Aussagen zulässt. Andernfalls sind nämlich auch kaum Angaben über mögliche Fehler der Resultate möglich. Resultate jedoch, von denen man weiss, dass sie nicht sicher richtig sind, ohne gleichzeitig über Angaben zu verfügen, wie wahrscheinlich eine gewisse Abweichung ist, sind (für die Planung)

völlig nutzlos. Das Rechnen von Varianten oder sonstige Manipulationen des Deduktionsverfahrens führen zwar oft zu probabilistisch aussehenden Resultaten - aber auch häufig zu einer Verletzung der zweiten der oben aufgeführten Forderungen: das so entstandene Deduktionsverfahren ist logisch nicht mehr zu rechtfertigen.

Konfrontiert man die existierenden Modelle mit diesen Forderungen, so erkennt man leicht, dass viele ihnen nicht genügen: Obwohl unexakte Angaben über Modellgrössen für die Praxis typisch sind, verlangen sie exakte Werte. Weitere Schwierigkeiten ergeben sich, wenn zur Herleitung der Ergebnisse neben den Inputinformationen weitere, oft fragwürdige Hypothesen Verwendung finden.

2. Bilanz - oder Datenmodelle

Wir beschreiben nun ein Modellschema, d.h. einen Rahmen, der es erlaubt, für viele Probleme ein passendes Modell zu erstellen, das seinerseits eine gut handhabbare und mathematische Repräsentation besitzt. Wie bereits angedeutet gehen wir ähnlich vor wie die "geometrischen" Modelle, die von der geometrischen Repräsentation der Umwelt ausgehen und zu mathematischen Modellen fortschreiten. Wir benutzen jedoch nicht die geometrische Repräsentation der Umwelt, sondern die Darstellung der Umwelt in der Form von (Daten-) Aussagen zu Indikatorvariablen, d.h. die sprachlichen Elemente zur Erfassung des Quantitativen als Ausgangspunkt. Aus diesem Teil unserer Sprache wählen wir einige Konzepte aus und verbinden sie so, dass die derart sichtbar werdenden Strukturen in diskreten stochastischen Prozessen eine mathematische Darstellung besitzen.



Verwendet wird zunächst das Konzept der Beschreibung der Umwelt in einem bestimmten Zustande. Dieser Zustandsbeschreibung beigeordnet ist - sofern man sich für die Dynamik des Systems interessiert - eine analoge Beschreibung der Veränderungen oder Mutationen. Zustands- und Mutationsbeschreibung können derart aufeinander abgestimmt werden, dass eine bilanzmässige Erfassung der Systemdynamik vorliegt. Das nächste Konzept nutzt die Tatsache, dass sprachliche Begriffe mannigfach miteinander verknüpft sind (wie Pferd, weiss, Schimmel): Die Modellgrössen werden so stark miteinander verknüpft, wie es ihre

sprachliche Bedeutung erlaubt bzw. verlangt. Nimmt man schliesslich noch das Konzept der Zufallsvariable zur Codierung der i.a. unscharfen Daten hinzu, so ist die genannte Basis gelegt, d.h. eine erste Abbildung der Umwelt in den Bereich einer rekonstruierten quantitativen Sprache - wir wollen sie im folgenden Systemsprache nennen - geleistet. Von hier aus ist es einfach, zu stochastischen Prozessen vorzustossen, welche ihrerseits Elemente einer leistungsfähigen Theorie sind.

Zustands- und Mutationsbeschreibung

Das Konzept der Zustandsbeschreibung wird von fast allen Theorien und Modellen benutzt. Es besteht darin, einen Objektbereich mit Indikatorbegriffen oder Variablen zu beschreiben. Die Variablen werden meist als zeitabhängige Grössen angesehen. Wenn nun T einen bestimmten Zeitpunkt bezeichnet, so nennt man die Gesamtheit der Werte, die die einzelnen Variablen zu diesem Zeitpunkt annehmen, Zustand des Systems zur Zeit T. Wir bezeichnen ihn mit X(T).

Die Aenderungen des Zustandes zwischen zwei Zeitpunkten T und T' können als eine andere Art von Objekten aufgefasst werden. Beschreibt z.B. X(T) die Bevölkerung einer Stadt, so werden die Aenderungen zwischen T und T' 'Geburten', 'Immigranten'etc. genannt. Diese Objekte können, wie oben mit Variablen beschrieben werden. Die Gesamtheit der Variablenwerte einer solchen Mutationsbeschreibung notieren wir im folgenden mit M(T,T').

Dieser Schritt - die Beschreibung des Systems mit Variablen - ist der wichtigste: Er ermöglicht die übrigen Schritte und enthält sie bereits in nuce.

Zustands- und Mutationsbeschreibungen sollen so aufeinander abgestimmt erfolgen, dass eine Buchungs- oder Bilanzgleichung gilt:

neuer Zustand = alter Zustand "plus" Mutationen.

Diese Begriffe liefern die Sprachbasis für die Bilanzmodelle. Sie geben an, was mögliche Zustände bzw. Aenderungen sind. (Jedes Modell kann nur das ausdrücken oder reflektieren, was in ihm an Strukturelementen vorkommt. In unserem Falle sind die Indikatorvariablen, die der Zustands- und Mutationsbeschreibung dienen, vom Modell her gesehen elementar: auf ihrer Basis wird alles weitere aufgebaut.) Wir haben damit folgendes erreicht:

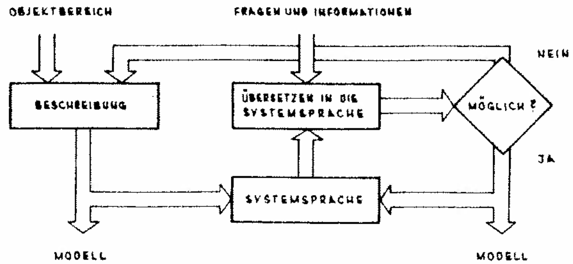
- Eine Sprachbasis, Elementarbegriffe, zusammengefasst im Vektor X für die Zustandsbeschreibung und im Vektor M für die Mutationsbeschreibung.
- Eine Bilanzgleichung $X(T')=X(T)$"plus"$M(T,T')$, oder formal:
  (1) $X(T')=F(X(T),M(T,T'))$.

  (Wir nennen diese Vorschrift (1) gemäss der die Veränderungen $M(T,T')$ zum Zustand $X(T)$ zu verbuchen sind, Buchungs- oder Bilanzgleichung.)

Jedermann kennt solche Buchungsregeln, speziell in Zusammenhang mit Finanzen, aber auch von sogenannten Statistiken her (z.B. Bevölkerungsstatistik). Viele quantitativen Modelle enthalten solche Bilanzgleichungen.

Definition von Input- und Outputgrössen

Wie oben schon gesagt soll das Modell aus Inputdaten gesuchte Daten erzeugen. Statt Daten verwendeten wir auch den Begriff Teilinformationen. Jede Information hat zwei Komponenten: eine sprachliche in der Form einer wohldefinierten Indikatorvariable und einen Datenteil im engeren Sinne in der Form einer Aussage zu den Werten dieser Indikatorvariable. Das bedeutet, dass wir für jede Input und Outputgrösse zunächst eine entsprechende Indikatorvariable ins Modell einführen müssen. Dies soll hier nun so geschehen, dass - wenn immer möglich - diese mit Hilfe der bereits vorhandenen Systemsprache gebildet wird. Zu Beginn besteht die Systemsprache nur aus den Variablen der Zustands- und Mutationsbeschreibung. Durch jede gelungene Definition oder Rekonstruktion einer Input- oder Outputvariable wird sie reicher. Natürlich kann es vorkommen, dass eine Variable nicht auf der Basis der bereits vorhandenen Systemsprache rekonstruierbar ist. In diesem Falle muss die Zustands- und/oder Mutationsbeschreibung passend erweitert werden - oder auf die Variable im Modell verzichtet werden. Nachfolgende Figur fasst diesen Sprachbildungsprozess zusammen.



Ein Beispiel möge diesen Prozess veranschaulichen: Beschreibt man etwa die Bevölkerungsentwicklung eines Gebietes, so ist eine wichtige Inputgrösse die sog. Geburtenrate. Dieser zunächst etwas vage Begriff kann in der Systemsprache z.B. so definiert (oder wenn man lieber will, rekonstruiert) werden: "Anzahl Geburten dividiert durch Anzahl Mütter in einer bestimmten Altersklasse". Man erhält so eine vom Alter der Mutter abhängige Geburtenrate. Die entsprechende Quantität in der Systemsprache erhält man wie folgt:

- Die Anzahl der Frauen in einer bestimmten Altersklasse a ist als Element oder als Summe von Elementen des Zustandsvektors X gegeben. Wir bezeichnen sie mit $W_a(X)$.
- Die Anzahl der Geburten mit Müttern der Altersklasse a im Zeitraum T bis T' ist entsprechend als Element oder Summe von Elementen des Vektors M gegeben und soll mit $B_a(M)$ bezeichnet werden.
- Dann ist die Geburtenrate der Frauen der Altersklasse a durch folgende Quantität definiert:

$$R_a := \frac{B_a(M)}{W_a(X)}$$

d.h. R_a ist eine Funktion von X und M.

Wird dieses Verfahren nun für alle interessierenden Outputgrössen durchgeführt, so erhält man eine Liste von Definitionen:

(2) $Y_i := G_i(X,M)$, $i=1,2,\dots,n$,

oder in Vektorschreibweise:

(2') $Y := G(X,M)$.

Dasselbe soll nun für die Inputsgrössen geschehen, d.h. für jene Grössen, über die dem Modell etwas mitgeteilt wird, um auf dieser Basis Schlüsse über die Grössenordnungen der Outputvariablen zu ermöglichen. Dies führt zu einer weiteren Liste von Definitionen:

(3) $Z_j := H_j(X,M)$, $j=1,2,\dots,m$,

oder in Vektorform:

(3) $Z := H(X,M)$.

Bis dahin haben wir nur Sprache rekonstruiert. Es gibt hier nichts zu beweisen. Nur Fragen der Zweckmässigkeit oder Adäquatheit der vorgeschlagenen Rekonstruktionen sind zu klären. (Viele quantitative Modelle enthalten solche Definitionsgleichungen. Der Rekonstruktionsprozess wird allerdings meist nicht in der hier geforderten systematischen Weise zu einem Kernstück des Modellbaus.)

Daten

Die Definitionen (3) werden nun zu Gleichungen, wenn wir Aussagen über die Grössenordnungen der entsprechenden Quantitäten machen. Dies soll im folgenden deutlich werden.

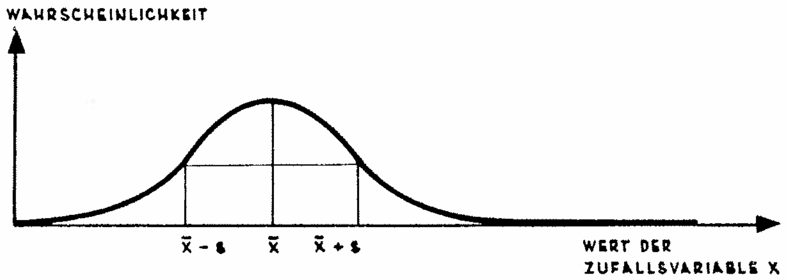
Um rechnen zu können, werden folgende Daten benötigt:
- Evtl. Daten zum Anfangszustand $X(T_0)$ und
- Daten zu den Inputquantitäten (3) für verschiedene Zeitpunkte $T_k: Z(T_k)$, $k=0,1,\dots,r$.

(Alle diese Aussagen sind jeweils Aussagen über eine Quantität und einen Zeitpunkt. Daher nennen wir sie

Daten - im Gegensatz zu "Gesetzen", die Aussagen für alle Zeitpunkte etc. machen).

In der Praxis sind diese Aussagen allerdings meist unpräzise: Was erhältlich ist sind etwa "optimistische" und "pessimistische" Werte für eine Quantität zu einem bestimmten Zeitpunkt (z.B. für Geburtenraten). Der einzige bekannte und brauchbare Weg, solche unpräzisen Aussagen über Quantitäten zu erfassen ist das Konzept der Zufallsvariablen. Deshalb betrachten wir im folgenden die Vektoren X, M, Y und Z als Zufallsvektoren.

Eine Zufallsvariable ist durch eine Wahrscheinlichkeitsverteilung charakterisiert. Diese zu finden, ist allerdings häufig nicht einfach. Man ist daher meist mit der Kenntnis der ersten Momente, etwa von Mittelwert und Varianz, zufrieden. Man nimmt dann oft weiter an, dass die Normalverteilung eine gute Approximation darstellt, um unser Wissen oder Glauben über den Wert der Variablen zu beschreiben. (Detailliertere Ausführungen zu diesem Punkt sind zu finden in: Schmid, 1978.) Die Normalverteilungsannahme ist jedoch nicht unbedingt notwendig.



Betrachten wir nun eine Inputvariable Z_i genauer. Wir nehmen an, ihr Mittelwert sei $\overline{Z}_i$ und ihre Varianz $s^2(Z_i)$. Wir können dann schreiben:

$Z_i=\overline{Z}_i-V_i$, wobei V_i eine Zufallsgrösse mit Mittelwert 0 und Varianz $s^2(Z_i)$ ist.

Zusammen mit der Definition (3)

$Z_i:=H_i(X,M)$

erhält man eine Gleichung:

(4) $\overline{Z}_i=H_i(X,M)+V_i$.

Wenden wir dieses Verfahren auf alle Inputvariable in (3) an, so erhalten wir folgende Vektor-Gleichung:

(4') $\overline{Z}=H(X,M)+V$, wobei V den Mittelwert 0 und die Kovarianzmatrix S_{VV} hat.

Entsprechende Ueberlegungen führen zu einem Vektor $\overline{X}(T_0)$ und einer Kovarianzmatrix $S_{XX}(T_0)$, die unser Wissen über den Anfangszustand beschreiben (Schmid, 1978):

(5) $\overline{X}(T_0), S_{XX}(T_0)$.

Das Modell

Nachdem nun alle Elemente als Zufallselemente aufgefasst werden, kann X(T) als stochastischer Prozess betrachtet werden - und zwar wegen den Beziehungen (1) und (4') als diskreter Markov-Prozess:

$X(T')=F(X(T),M(T,T')$

$\overline{Z}(T')=H(X(T'),M(T,T'))+V$

Unser Modell besteht also aus den stochastischen Gleichungen (1) und (4') und den "Anfangsbedingungen" (5).

Man beachte jedoch: Weder (1) noch (4') - und offensichtlich auch nicht (5) - enthalten hinsichtlich ihrer Form irgendwelche Hypothesen, die empirischer Rechtfertigung bedürften. Die Gleichung (1) ist rein analytisch, d.h. ihre Wahrheit ergibt sich allein aus der Semantik der verwendeten Variablen in X und M (oder (1) ist Bestandteil dieser Semantik). (4') ist bezüglich seiner Form von (3') abgeleitet. Aber (3') enthält ausschliesslich Definitionen, welche nicht falsch, höchstens unzweckmässig sein können. Daraus folgt: Der einzige Teil des Modells mit empirischem Gehalt sind die Daten $\bar{Z}$ und S_{VV} in (4') und die Angaben zu $X(T_0)$ in (5). Wir können diese Modelle daher Datenmodelle nennen, im Gegensatz zu den üblichen nomologischen Modellen. Sie enthalten nämlich offensichtlich keine nomologischen Aussagen, die für alle Zeitpunkte, für alle Situationen einer bestimmten Art Geltung beanspruchen. Daten, wie sie oben eingeführt wurden, sagen etwas über eine Grösse für einen Zeitpunkt aus. Sie sind - im Gegensatz zu nomologischen Aussagen - verifizierbar. (Natürlich kann der Benutzer seine Daten erarbeiten, indem er Theorien, Hypothesen etc. verwendet. Aber er tut und kontrolliert dies, das Modell fügt nichts mit empirischem Gehalt hinzu.)

Algorithmen

Was nun noch fehlt, ist ein Verfahren, welches das Modell auswertet, um Fragen bezüglich der Outputgrössen (2) zu beantworten. Man sieht leicht ein, dass dies möglich wird, wenn es gelingt, die Verteilungen oder wenigstens Momente davon für die Vektoren X und

M für (künftige) Zeitpunkte T_k zu bestimmen. Das Problem besteht also darin, "möglichst gute" Schätzungen für $X(T_k)$ und $M(T_k,T_{k+1})$ zu finden, indem (1) (4') und (5) ausgewertet werden. Das ist ein Problem der Theorie der stochastischen Prozesse bzw. der statistischen Schätztheorie und kann auf verschiedene Weisen gelöst werden.

Gesucht sind also Schätzungen für $X(T_k)$ und $M(T_k,T_{k+1})$. Sie sollen mit $\hat{X}(T_k)$ bzw. $\hat{M}(T_k,T_{k+1})$ bezeichnet werden. Im folgenden beschränken wir uns auf $\hat{X}(T_k)$, da für $\hat{M}(T_k,T_{k+1})$ dasselbe Verfahren gilt.

Die gesuchten Schätzungen basieren auf (1), (4') und (5). Wir haben also eine Funktion $S_{F,H}$ (F und H erinnern an die entsprechenden Funktionen in (1) und (4')) zu bestimmen, so dass

(6) $\hat{X}(T_k)=S_{F,H}(\hat{X}(T_0),Z(T_0),Z(T_1),\ldots,Z(_k))$.

Diese Funktion soll möglichst gute Schätzwerte liefern, d.h. optimale Schätzwerte. Normalerweise definiert man Optimalität im Sinne des "least square": man bestimmt $\hat{X}(T_k)$ so, dass

(7) $E[(X(T_k)-\hat{X}(T_k))^2]$ Minimum, d.h. die Varianz des Schätzfehlers soll möglichst klein werden.

(Eine andere Möglichkeit besteht im Einführen einer (anderen) Verlustfunktion, die die "Strafe" für den Schätzfehler misst. Man definiert dann Optimalität dadurch, dass der erwartete Verlust minimiert werden soll. Unter recht schwachen Annahmen führt das jedoch in unserem Falle zu den gleichen Ergebnissen; vgl. Jazwinski, 1970, S.148ff. oder Meditch, 1969, S.162.)

Die Gleichung (1) wird in der Regel linear sein. Falls (5') auch linear oder linearisierbar ist, existieren sehr effiziente Filter-Algorithmen zur Lösung des Problems. Aber auch für den nicht-linearen Fall gibt es oft recht gute Algorithmen. Für diese algorithmischen Probleme sei jedoch auf die Literatur verwiesen (z.B. Jazwinski 1970).

Resultate

Mit Hilfe der Definitionen (2) und den Schätzungen $\hat{X}(T_k)$ und $\hat{M}(T_k,T_{k+1})$ lassen sich Fragen zu den Outputvariablen beantworten. Im Falle, dass geeignete Annahmen gelten,

- erhalten wir optimale Schätzungen für $\hat{X}(T_k)$ und $\hat{M}(T_k,T_{k+1})$;
- Wenn $X(T_0)$ und die $V(T_k)$ je voneinander unabhängig und normalverteilt sind und (1) und (4') linear sind, so erhalten wir sogar die Verteilungen von $X(T_k)$ und $M(T_k,T_{k+1})$.

Dann sind wir in der Lage, Aussagen wie diese zu machen: "Wenn der Input mit Wahrscheinlichkeit p gilt, dann liegt die Outputvariable $Y_i(T_k)$ mit der Wahrscheinlichkeit q im Intervall I_{ik}", d.h. wir können für die Outputvariable Konfidenzintervalle angeben. Oder - was für die Praxis einfacher zu verstehen ist-: der Output hat denselben Konfidenzgrad wie der Input.

Offensichtlich sind also Bilanzmodelle sehr flexibel, d.h. auf irgendeinen Gegenstandsbereich anwendbar, der mit Variablen (oder Indikatoren) beschreibbar und also mit irgendeiner Art von Buchhaltung oder "Statistik" erfassbar ist. Sie können unscharfe Informa-

tionen verarbeiten und haben keinen fix vorgegebenen Input oder Output; diese können vielmehr den Bedürfnissen des Benützers angepasst werden. Sie enthalten keine nomologischen Aussagen mit empirischem Gehalt und ermöglichen daher dem Benützer eine vollständige Beurteilung der Resultate (gleicher Konfidenzgrad wie Input). Diese Modelle sind also wie gefordert lediglich Informationstransformatoren. Ein Einwand könnte sein, dass die Intervalle I_{ik} in typischen Planungssituationen unbrauchbar gross würden. Bisher durchgeführte Anwendungen weisen in die gegenteilige Richtung: die Intervalle wurden meist kleiner als erwartet.

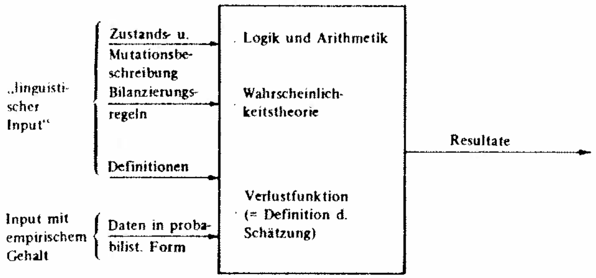
Zusammenfassung

Bilanzmodelle bestehen also nach dem oben Gesagten aus

- einer (problemorientierten) Zustandsbeschreibung der interessierenden Objekte und einer Beschreibung der möglichen Veränderungen dieser Zustände,
- Bilanzierungsregeln (1),
- Definitionen von Input- und Outputvariablen (2) und (3),
- Daten für den Anfangszustand $X(T_0)$ und die Inputvariablen zu verschiedenen Zeitpunkten, d.h. zu $Z(T_k)$, $k=0,1,\ldots,r$,
- einem zugeordneten Schätzproblem mit einem dazugehörenden Algorithmus, um Aussagen über die Outputgrössen zu erhalten.

Der Benützer von Bilanzmodellen hat neben seinen Daten nichts Zusätzliches zu akzeptieren als das Konzept der statistischen Schätztheorie. Während diese Aussagen dem Planenden keine Probleme bereiten dürften, kann er sein Vertrauen in die Daten nötigenfalls erhöhen, indem er sie abschwächt, d.h. die Streubereiche vergrössert.

Nachfolgende Figur veranschaulicht die logische Struktur eines Bilanzmodelles:



<u>Literatur</u>

Jacobi, H.; Lubicz, Ch.; Maurer, J.; Schmid, B.: Donaubereich Wien, Entwicklung der an die Donau angrenzenden Quartiere. Schätzung von Daten für das Wohnungswesen und die Demographie mittels Bilanzmodellen. ORL-Institut der ETH Zürich, Zürich.

Jacobi, H. und Tomasoni, W., 1978: Utilisation du sol. La prise en compte de l'incertitude à l'exemple de l'estimation de l'utilisation du sol dans la ville de Zurich, Studienunterlagen zur Orts-, Regional- und Landesplanung Nr.35, ETH Zürich, Zürich.

Jazwinski, A.H., 1970: Stochastic Processes and Filtering Theory.

Kyburz, R. und Schmid, B., 1979: Accounting Models - A new Tool in Planning. Studienunterlagen zur Orts-, Regional- und Landesplanung, Nr.40, ETH Zürich, Zürich.

Meditch, J.S., 1969: Stochastic Optimal Linear Estimation and Control, Mc Graw Hill, New York, N.Y.

Schaerer, M., 1978: Bilanzverfahren als Hilfsmittel zur Anwendung der Entscheidungslogik in der Raumplanung, DISP Nr.53, ORL-Institut ETH Zürich, Zürich.

Schaerer, M., 1978: Eine Methode zur Quantifizierung der Ungewissheit der demographischen Vorausschau in der Raumplanung, getestet am Beispiel Urserental. Berichte zur Orts-, Regional- und Landesplanung Nr.38, ETH Zürich, Zürich.

Schaerer, M., 1979: Demographische Vorausschau für das Urserental mittels eines Bilanzmodells. DISP Nr.58, ORL-Institut ETH Zürich, Zürich.

Schmid, B., 1978: Bilanzmodelle - Simulationsverfahren zur Bearbeitung unscharfer Teilinformationen. Berichte zur Orts-, Regional- und Landesplanung Nr.40, ETH Zürich, Zürich.

Schmid, B., 1980: Eine Alternative zu nomologischen Modellen als Planungsinstrumente. Jahrb. für Regionalwiss., 1.Jahrgang.

Tomasoni, W., 1979: Schätzen der künftigen Nutzung des Bodens mittels eines Bilanzmodells, DISP Nr.55, ORL-Institut, ETH Zürich, Zürich.

Zürich, 16.10.1981/BS/el