

Syndrome & Co.:  
Qualitative und semi-quantitative Ansätze in der  
Forschung zum Globalen Wandel

von

Hans-Joachim Schellnhuber, Arthur Block, Hermann Held,  
Matthias K.B. Lüdeke, Oliver Moldenhauer und  
Gerhard Petschel-Held

Potsdam-Institut für Klimafolgenforschung (PIK)  
Postfach 60 12 03  
D-14412 Potsdam

# 1 Einleitung: Der Globale Wandel als wissenschaftliche Herausforderung

Es ist heute unumstritten, dass die Erforschung des Globalen Wandels sich insbesondere der Komplexität der Wechselwirkung zwischen Natur und Zivilisation annehmen muss. Komplexität ist charakterisiert durch Multikausalität und Nichtlinearität und lässt keine Vorhersage der künftigen Entwicklung im engeren Sinne zu (Schellnhuber, 1999). Beispielsweise wird man auf Aussagen wie „im Jahr 2010 werden die jährlichen CO<sub>2</sub>-Emissionen 42.3 Gt C betragen“ nicht vertrauen können. Auch viele quantitativ-deterministische Modellierer des Globalen Wandels relativieren zwischenzeitlich ihre Resultate - don't trust the numbers, just trust the trends - wobei allerdings offen bleibt, ob die Abbildung der Komplexität in den Modellen wenigstens hinreichend ist, um diesen Tendenzangaben Glauben schenken zu können.

Eine weitere Konsequenz der oben erwähnten Komplexität ist, dass der Versuch, durch politische Maßnahmen einzelne problematische Entwicklungen („Symptome“) zu beeinflussen, ähnlich wie in der Medizin zu „unerwünschten Nebenwirkungen“ führen kann, die dann in der Folge erneut bekämpft werden müssen. Um im Bild zu bleiben: eine wirksame Therapie setzt die Kenntnis der systemischen Zusammenhänge der unterschiedlichen Symptome - also der Krankheit oder des Syndroms - voraus.

Zu diesen komplexitätsbedingten Problemen (siehe auch Kates und Clark, 1999; Rotmans, 1998) kommen zusätzliche Unsicherheiten, die die einzelnen, das komplexe Beziehungsnetz konstituierenden Wechselwirkungen betreffen. Unsere oft vage, unvollständige oder rein qualitative Kenntnis über diese Wechselwirkungen könnte sich als durchaus resistent gegenüber dem weiteren wissenschaftlichen Fortschritt erweisen.

Man könnte nun argumentieren, dass auf Grund der soweit beschriebenen Eigenschaften des Globalen Wandels eine mathematisch-formale Analyse des Problems nicht durchführbar wäre. Die Mathematik und Physik bisher behandelte komplexer Systeme zeigt jedoch, dass man mit Überraschungen im Systemverhalten konfrontiert wird, die nur auf der Basis formaler Analyse zu verstehen sind (Haken, 1977), während Intuition und rein qualitative

Argumentation hier oft versagen. Daneben ist anzumerken, dass z.B. ohne die mathematisch-physikalischen Computer-Modelle der Klimaentwicklung das Umweltproblem der anthropogenen Atmosphärentransformation sicher nicht den Weg in die öffentliche Diskussion gefunden hätte. Durch solche Beispiele ermutigt wird nun auch im noch weit komplexeren Bereich des Globalen Wandels versucht, „das Unquantifizierbare zu quantifizieren“, also der traditionellen mathematischen Bearbeitung zugänglich zu machen, was unserer Ansicht nach problematisch ist und die gegenwärtigen Fortschritte in der Methodik der *formalen qualitativen Modellierung* ignoriert.

Im zweiten Abschnitt wird zunächst das Syndromkonzept als eine Möglichkeit des Umgangs mit den Komplexitätsaspekten des Globalen Wandels eingeführt, wobei die Begriffe zunächst aus einer abstrakten (und hypothetischen) Beschreibung des Gesamtsystems hergeleitet werden („idealistische Deduktion“) und dann der eher induktive Weg an Beispielen ihrer vorläufigen Konkretisierung nachgezeichnet wird („realistische Induktion“). In den Abschnitten drei und vier werden dann zwei mathematische Konzepte erläutert, die die Berücksichtigung des häufig qualitativen Charakters der Kenntnis relevanter Beziehungen (samt möglicher semi-quantitativer Spezifizierungen) einerseits und der statistisch beschreibbaren Unsicherheit bei quantitativen Beziehungen (samt deren Aktualisierung bei neu auftretenden Beobachtungen) andererseits erlauben.

Im fünften Abschnitt wird das erste dieser Konzepte, die qualitativen Differentialgleichungen, dazu genutzt, die induktive Bestimmung der syndromerzeugenden funktionalen Muster (wie in Abschnitt 2 eingeführt) zu formalisieren und zu systematisieren. Das Verfahren wird am Beispiel der Problematik der kleinbäuerlichen Landwirtschaft in Entwicklungsländern illustriert und rundet die Darstellung des Syndromkonzepts als konsistenter Herangehensweise an den Globalen Wandel ab.

Schließlich wird im sechsten Abschnitt eine Methode eingeführt, wie im Falle der Verfügbarkeit großer relevanter Datenmengen und geringem Zusammenhangswissen (eine Situation, die derzeit nur für kleine Ausschnitte des globalen Gesamtsystems vorliegt) rein datenbasiert Muster identifiziert werden können und so zum Beispiel Klassen gleichermaßen klimatisch verwundbarer Regionen bestimmt werden können.

## 2 Syndrome I: Idealistische Deduktion versus realistische Induktion

Um die systemtheoretischen Grundlagen des Syndromkonzepts zu beleuchten, soll von der (notwendigerweise) hypothetischen Annahme ausgegangen werden, dass man über ein System von gewöhnlichen Differentialgleichungen verfügt, das das globale Gesamtsystem (i.e. alle natur- und sozialräumlichen Aspekte samt ihrer Wechselwirkungen) vollständig beschreibt. Räumliche Verteilungen sind in einer solchen Beschreibung durch Diskretisierung approximiert.

Bei Verwendung eines beliebigen Variablensatzes zur Beschreibung des globalen Zustands ist nun eine sehr große Anzahl von Variablen und eine enge Verkopplung der einzelnen Gleichungen des Systems zu erwarten, was das Problem in dieser Form unbehandelbar macht. Ein möglicher Ausweg ist zu versuchen, das System in untereinander möglichst unabhängige Teilsysteme zu zerlegen, was z.B. durch geeignete Variablentransformationen erreicht werden kann. Ein einfaches Beispiel aus der Physik stellt die Behandlung des Zweikörperproblems dar, die auf dem Übergang von den expliziten Raumkoordinaten der beiden Massen auf deren relative Position und den Ort des Schwerpunktes des Systems beruht und so das komplizierte verkoppelte System der Bewegungsgleichungen in ein Subsystem für die Relativbewegung und eines für die Schwerpunktsbewegung (vollständig) separiert. Die Einführung einer kleinen dritten Masse führt nun zu einer schwachen Kopplung dieser beiden Subsysteme (Berry 1978) - und macht dieses System damit zu einer Illustration für das, was idealerweise bei der Dekomposition des globalen Gesamtsystems zu erwarten wäre. Im Allgemeinen kann man eine solche Entkopplung durch kanonische Variablentransformation erreichen und im Folgenden sollen die resultierenden Submodelle als *funktionale Ursache-Wirkungsmuster* bezeichnet werden.

Da nun - wie der Name „Syndromkonzept“ schon nahe legt - besonderes Augenmerk auf nicht-nachhaltige Entwicklungen gelegt werden soll, werden im Weiteren nur solche funktionalen Ursache-Wirkungsmuster untersucht, die zumindest *eine* nicht-nachhaltige Trajektorie in ihrem Lösungsraum

aufweisen. Diese Ursache-Wirkungsmuster werden syndromgenerierende oder „*riskante funktionale Muster*“ (*Hazarduous Functional Patterns, HFPs*) genannt. Klassen von nicht-nachhaltigen Trajektorien eines HFP werden als „*Syndrome*“ bezeichnet. Wird ein HFP weiter dahingehend spezifiziert, dass es nur noch ein Syndrom zur Lösung hat, dann soll das resultierende Modell „*detailliertes funktionales Muster*“ (*Detailed Functional Pattern, DFP*) genannt werden.

Die transformierten Variablen, die zur Entkopplung und Vereinfachung (durch gleichzeitige Aggregation) des hypothetischen komplexen Differentialgleichungssystems führen, werden im Syndromkonzept als „*Symptome*“ bezeichnet (Schellnhuber et al., 1997). Der Prozess der hypothetischen Deduktion von HFPs und Syndromen ist in Abbildung 1 zusammenfassend dargestellt.

Nun ist der Ausgangspunkt der obigen Argumentation - das vollständige Differentialgleichungssystem - notwendigerweise idealistisch, da das Wissen über die relevanten Zusammenhänge unsicher, teilweise irreduzibel qualitativer Natur sowie teilweise kontrovers ist. Deshalb liegt der Wert der soweit beschriebenen Deduktion in der Klärung der Begrifflichkeit, während die tatsächliche Syndrom- und HFP-Identifikation vom gegenwärtig verfügbaren, limitierten Wissen über die wesentlichen Zusammenhänge sowie von der Kenntnis derzeitiger problematischer Entwicklungen im Umwelt- und sozioökonomischen Bereich auszugehen hat (siehe Abbildung 2, oben).

Die Resultate einer solchen, ersten induktiven Erschließung der globalen Gesamtproblematik liegen in Form eines Kataloges von etwa 80 relevanten Trends oder Symptomen als „Zustandsvariablen“ vor (ursprünglich: WBGU 1997, weiterentwickelt: Petschel-Held et al., 1999a). Projiziert man nun das derzeit verfügbare Zusammenhangswissen über den Globalen Wandel auf Beziehungen zwischen diesen Zustandsvariablen, erhält man ein komplexes Beziehungsgeflecht, das als Ausgangspunkt für die induktive Identifikation typischer riskanter funktionaler Muster (HFPs) und deren nicht-nachhaltiger Trajektorien (Syndrome) dient. Ein wichtiger Aspekt bei der nun durchzuführenden Separation ist, dass nicht alle gesammelten Wechselwirkungen zwischen den Symptomen an allen Orten gültig sind. Die armutsverstärkende Wirkung von Bodendegradation beispielsweise ist in Entwicklungsländern zu beobachten, während diese Beziehung in Industrieländern nicht von Relevanz ist.

**Hypothetisches, umfassendes Weltmodell:**

Mathematische Formulierung	Symbolische Repräsentation
$\dot{x}_1 = f_1(x_1, \dots, x_i, \dots, x_j, \dots, x_{n-2}, x_{n-1}, x_n)$ ... $\dot{x}_i = f_i(x_1, \dots, x_i, \dots, x_j, \dots, x_{n-2}, x_{n-1}, x_n)$ ... $\dot{x}_j = f_j(x_1, \dots, x_i, \dots, x_j, \dots, x_{n-2}, x_{n-1}, x_n)$ ... $\dot{x}_n = f_n(x_1, \dots, x_i, \dots, x_j, \dots, x_{n-2}, x_{n-1}, x_n)$	

**Kanonische Transformation und Aggregation der Variablen:**

$\vec{x} = (x_1, \dots, x_i, \dots, x_j, \dots, x_n)$ $\vec{x}^* = (x_1^*, \dots, x_j^*, \dots, x_k^*)$	
--	--

**Transformation und Aggregation: Schwach gekoppelte Teilmodelle:**

Math. Form.	Symbolische Repräsentation
$\dot{x}_1^* = f_1^*(x_1^*, x_2^*, x_3^*)$ $\dot{x}_2^* = f_2^*(x_2^*, x_3^*, x_4^*)$ $\dot{x}_3^* = f_3^*(x_1^*, x_3^*)$ $\dot{x}_4^* = f_4^*(x_4^*, x_5^*)$ $\dot{x}_5^* = f_5^*(x_4^*, x_5^*)$ ... $\dot{x}_{k-3}^* = f_{k-3}^*(x_{k-3}^*, x_{k-2}^*, x_{k-1}^*)$ $\dot{x}_{k-2}^* = f_{k-2}^*(x_{k-2}^*, x_{k-1}^*)$ $\dot{x}_{k-1}^* = f_{k-1}^*(x_{k-3}^*, x_{k-1}^*, x_k^*)$ $\dot{x}_k^* = f_k^*(x_4^*, x_{k-3}^*, x_{k-2}^*, x_{k-1}^*)$	

**Auswertung der Submodelle: Identifikation HFPs:**

	=> generiert mindestens eine nicht-nachhaltige Trajektorie	=> <b>Riskantes funktionales Muster</b>	=> <b>Weitere Erforschung</b>
	=> generiert keine unakzeptable zeitliche Entwicklung	=> <b>Unproblematisches Subsystem</b>	=> Keine Berücksichtigung im Rahmen der Syndrom - Analyse

Abbildung 1: Hypothetischer Prozess der Deduktion von „riskanten funktionalen Mustern“ (Hazardous Functional Patterns, HFPs), die Syndrome als Klassen nicht-nachhaltiger Zeitverhalten generieren.

Neben diesen räumlichen Separationsansätzen spielen die funktionalen Aspekte eine wichtige Rolle – unterschiedliche Akteursgruppen in derselben Region tragen oft zu unterschiedlichen HFPs bei.

Entlang dieser Kriterien wurde eine erste Liste von 16 Syndromen des Globalen Wandels bestimmt, die in Tabelle 1 zusammen mit einer kurzen Erläuterung aufgeführt sind. Als ein Beispiel ist in Abbildung 3 das SAHEL-SYNDROM-erzeugende HFP dargestellt. Da es sich bei HFPs und deren

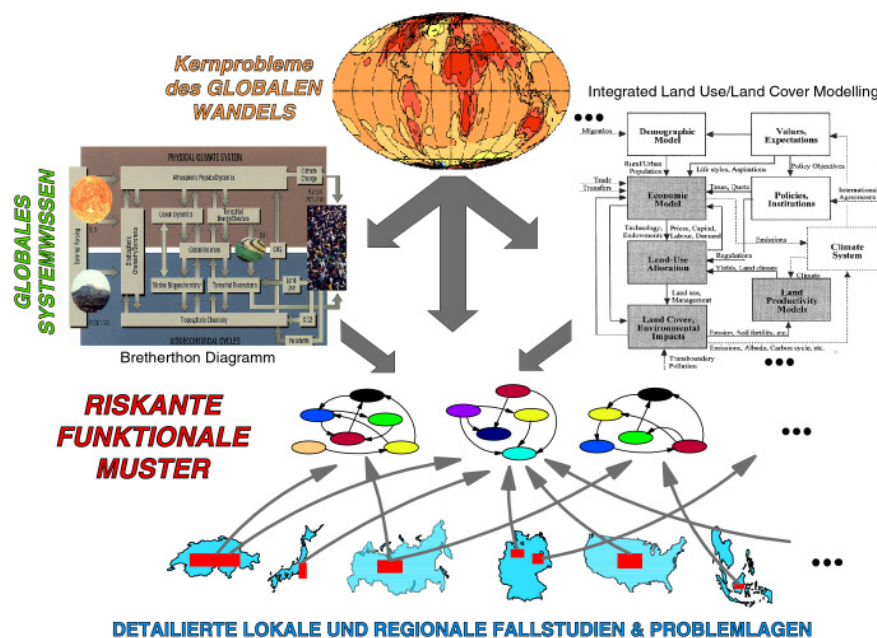


Abbildung 2: Herangehensweise zur Identifikation von „riskanten funktionalen Mustern“ (HFPs) in mittlerer funktionaler Auflösung. **Top-down** (oben): generelle Zusammenhangskenntnis (Bsp.: Bretherton-Diagramm, CIESIN 1992; Landnutzungsänderungsmechanismen, LUCC 1995) und Kenntnis problematischer Entwicklungen (Bsp.: anthropogene Klimabeeinflussung, Cubasch, 1992) werden zu typischen funktionalen Mustern arrangiert (Abschnitt Syndrome I). **Bottom-up** (unten): systematische Verallgemeinerung von detaillierten Fallstudien problematischer Mensch-Natur-Wechselwirkungen führt zu einigen typischen, immer wiederkehrenden funktionalen Mustern.

Syndromen um tiefliegende kausale Konzepte handelt, ist ihre direkte Überprüfung nicht möglich. Stattdessen müssen, wie das bei Überprüfungen solcher Theorien üblich ist, aus ihnen deduzierte Resultate mit beobachteten Phänomenen verglichen werden. In diesen Rahmen kann die Syndrom-Diagnose eingeordnet werden, die im ersten Schritt für jedes HFP Regionen zu identifizieren sucht, die gegenüber nicht-nachhaltigen Trajektorien (Syndromen) des entsprechenden HFP anfällig (*disponiert*) sind. Dies bedeutet, dass (a) die wichtigsten Wechselwirkungen des HFP potentiell gültig sein müssen (etwa die Beziehungen zwischen den gelb gekennzeichneten Symptomen in Abbildung 3) und, dass (b) diese Beziehungen in der betrachteten Region so beschaffen sind, dass nicht-nachhaltige Trajektorien erzeugt würden (im gleichen Beispiel: der Wettlauf zwischen Ressourcendegradation und kurzfristiger Ertragssteigerung durch Intensivierung der Landwirtschaft führt letztlich zu einer schlechteren Ertragssituation).

Zur Bestimmung einer solchen *Disposition* einer Region gegenüber einem bestimmten Syndrom ist nun ein Satz hierarchisch geordneter Indikatoren nötig, die es erlauben, die potentielle Gültigkeit sowie das Verhältnis der relevanten Wechselwirkungen des entsprechenden HFP zu beschreiben.

Als eingängige Darstellung hierfür haben sich Entscheidungsbäume erwiesen, die ausgehend von den zu Grunde liegenden Basisindikatoren die hierarchischen Ebenen samt ihrer logischen Verknüpfung darstellen. Als angemessene Form der mathematischen Formulierung dieser Entscheidungsbäume stellte sich das Fuzzy-Logic Kalkül (siehe z.B. Zimmermann, 1991) heraus, das es ermöglicht, trotz der teilweise qualitativen Natur der Zusammenhänge eine formale Auswertung durchzuführen (siehe z.B. Cassel-Gintz et al., 1997).



a) Syndromgruppe „Nutzung“

**SAHEL-SYNDROM:** Landwirtschaftliche Übernutzung von marginalen Standorten

**RAUBBAU-SYNDROM:** Raubbau an natürlichen Ökosystemen

**LANDFLUCHT-SYNDROM:** Umweltdegradation durch Preisgabe traditioneller Landnutzungsformen

**DUST-BOWL-SYNDROM:** Nicht-nachhaltige industrielle Bewirtschaftung von Böden und Gewässern

**KATANGA-SYNDROM:** Umweltdegradation durch Abbau nicht-erneuerbarer Ressourcen

**MASSENTOURISMUS-SYNDROM:** Erschließung und Schädigung von Naturräumen für Erholungszwecke

**VERBRANNT-ERDE-SYNDROM:** Umweltzerstörung durch militärische Nutzung

b) Syndromgruppe „Entwicklung“

**ARALSEE-SYNDROM:** Umweltschädigung durch zielgerichtete Naturumgestaltung im Rahmen von Großprojekten

**GRÜNE-REVOLUTION-SYNDROM:** Umweltdegradation durch Verbreitung standortfremder landwirtschaftlicher Produktionsverfahren

**KLEINE-TIGER-SYNDROM:** Vernachlässigung ökologischer Standards im Zuge hochdynamischen Wirtschaftswachstums

**FAVELA-SYNDROM:** Umweltdegradation durch unregelmäßige Urbanisierung

**URBAN-SPRAWL-SYNDROM:** Landschaftsschädigung durch gelpante Expansion von Stadt- und Infrastrukturen

**HAVARIE-SYNDROM:** Singuläre anthropogene Umweltkatastrophen mit längerfristigen Auswirkungen

c) Syndromgruppe „Senken“

**HOHER-SCHORNSTEIN-SYNDROM:** Umweltdegradation durch weiträumige diffuse Verteilung von meist langlebigen Wirkstoffen

**MÜLLKIPPEN-SYNDROM:** Umweltverbrauch durch geregelte und unregelmäßige Deponierung zivilisatorischer Abfälle

**ALTLASTEN-SYNDROM:** Lokale Kontamination von Umweltschutzgütern an vorwiegend industriellen Produktionsstandorten

Tabelle 1: Syndrome des Globalen Wandels (Klassen von nicht-nachhaltigen Trajektorien, wie sie durch die „riskanten funktionalen Muster“ erzeugt werden)

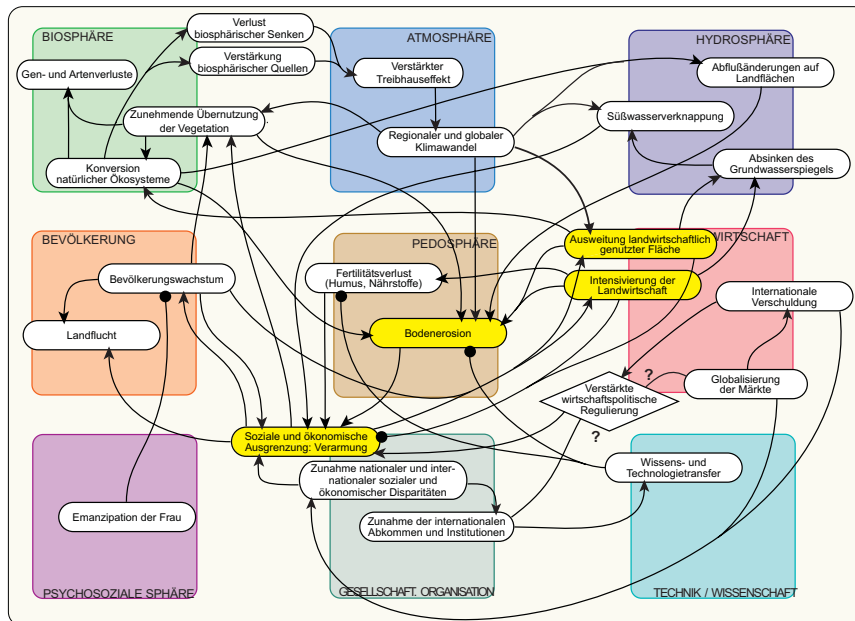


Abbildung 3: Beziehungsgeflecht für das SAHEL-SYNDROM-erzeugende funktionale Muster

Zur Illustration dieses Verfahrens soll hier die Bestimmung der Disposition gegenüber dem Sahel-Syndrom diskutiert werden. Aus dem Studium des syndromerzeugenden HFP (Abbildung 3) identifiziert man die folgenden zentralen Wechselwirkungskomplexe:

1. armutsgetriebene Intensivierung (also mit geringer Kapitalausstattung) und Expansion der Landwirtschaft bewirken Bodendegradation;
2. Ertragsrückgang zwingt die Armen der ländlichen Regionen zu weiterer Intensivierung, da keine ökonomischen Alternativen bestehen.

Bedingung für die erste Wirkungsbeziehung ist nun, dass die betrachtete Region naturräumlich fragil bezüglich ihrer landwirtschaftlichen Nutzung ist, während der zweite Zusammenhang dann gültig ist, wenn ein hoher Anteil von Subsistenzlandwirtschaft in einer auf den ersten Sektor ausgerichteten Wirtschaft vorliegt. In Abbildung 4 ist der Entscheidungsbaum zur Bestimmung und Verknüpfung der oben genannten Bedingungen auf der Basis von verfügbaren globalen Datensätzen und Modellen dargestellt. Zu diesen

gehören z.B. für die naturräumliche Seite (Wirkungsbeziehung 1) die Nettoprimärproduktion der natürlichen Vegetation zur Charakterisierung der generellen Wachstumsbedingungen (relevant für Ackerbau, Viehzucht und z.B. Holzsammelaktivitäten) und die Orographie als Indikator für das Erosionsrisiko.

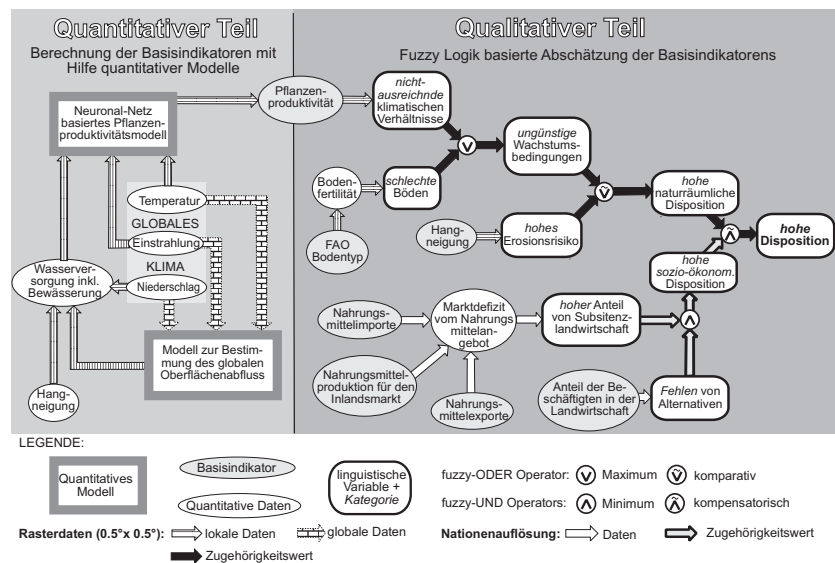


Abbildung 4: Struktur des Algorithmus' zur Bestimmung der Disposition gegenüber dem Sahel-Syndrom unter Benutzung quantitativer und qualitativer Modellelemente.

In die sozio-ökonomische Dimension gehen zum Beispiel Daten zur Wichtigkeit des primären gegenüber den anderen Wirtschaftssektoren ein, um die Alternativlosigkeit der Kleinbauern zu indizieren.

Im Sinne der Fuzzy-Logic werden alle linguistischen Kategorien – in Abbildung 4 als abgerundete Rechtecke dargestellt - durch Wahrheitswerte zwischen 0 („die Aussage trifft in der betrachteten Region überhaupt nicht zu“) und 1 („... trifft sicher zu“) charakterisiert.

Die Kreise bezeichnen entsprechende Fuzzy-Logic-Erweiterungen der Verknüpfungen der Booleschen Algebra (zu Details des gesamten Algorithmus einschließlich seiner Anwendung unter anthropogener Klimaänderung siehe Lüdeke et al., 1999).

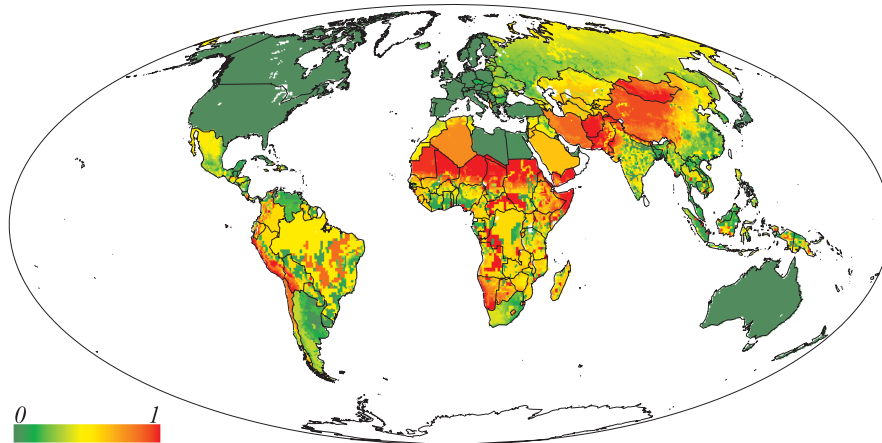


Abbildung 5: Disposition gegenüber dem Sahel-Syndrom. Dargestellt ist der Wahrheitswert für die Aussage „Disposition ist hoch“ (1: trifft sicher zu; 0: trifft sicher nicht zu)

Das globale Resultat des obigen Algorithmus, das in Abbildung 5 dargestellt ist, zeigt auch für naturräumlich sehr fragile Regionen in Industrieländern (z.B. den Westen der USA) keine Disposition gegenüber dem Syndrom an, da hier die sozioökonomischen Bedingungen fehlen. In Regionen wie der Sahelzone, weiteren Teilen Westafrikas, dem Nordosten von Brasilien, der Westküste Südamerikas, der Mongolei und dem Westen von Indien sind sowohl die naturräumlichen als auch die sozioökonomischen Bedingungen erfüllt, die die Regionen als gefährdet gegenüber dem SAHEL-SYNDROM erscheinen lassen.

Ein weiterer Schritt in der Syndromdiagnose ist die Bestimmung der *Syndromintensität*, wobei hier danach gefragt wird, in welchen Regionen ein bestimmtes Syndrom derzeit aktiv ist. Da es sich hierbei um einen Aspekt handelt, der sehr eng mit der Dynamik von Syndromen und der „bottom-up“-Identifikation derselben verknüpft ist, werden die hiermit verbundenen Fragen im Abschnitt „Syndrome II“ behandelt. Zunächst sollen nun jedoch zwei mathematische Konzepte eingeführt werden, die erlauben, qualitative und unsichere Zusammenhänge zu formalisieren, wobei das Kalkül der qualitativen Differentialgleichungen des Weiteren eine wichtige Rolle bei der Systematisierung der Typisierung (ein zentraler Aspekt des Syndromkonzepts) spielt.

### 3 Qualitative Differentialgleichungen

Im folgenden Abschnitt sollen die Grundzüge der qualitativen Differentialgleichungen (QDGI) beschrieben werden. An dieser Stelle geschieht dies anhand eines Beispiels, während die mathematischen Details in der jeweiligen Literatur zu finden sind (Kuipers, 1994; Clancy, 1997; Moldenhauer 2000). Als Beispiel wird hier das aus der theoretischen Ökologie (s. z.B. Wissel, 1989) bekannte logistische Wachstum in der Populationsdynamik erweitert um eine Managementkomponente. Die quantitative Beschreibung des logistischen Wachstums einer Population  $P$  wird typischerweise durch die Differentialgleichung

$$G = dP/dt = \alpha P (P_m - P)$$

angegeben. Hieraus ergibt sich die Klimaxpopulation  $P_m$  und die maximale Wachstumsrate  $G_m = (\alpha P_m^2)/4$ , die bei der Population  $P_0$  auftritt. Der Graph dieser Funktion zeigt eine invertierte U-Form für  $G$  in Abhängigkeit von  $P$ , wie dies in der schwarzen, durchgezogenen Linie in Abbildung 6 dargestellt ist. Bei dieser Parametrisierung werden alle Populationssysteme mit einem Anfangswert von  $P > 0$  auf den stabilen Fixpunkt  $P_m$  zusteuern und dort verharren. Als Verfeinerung dieses Modells führen wir nun eine externe Störung in Form einer konstanten Entnahme ein. D.h. das neue Wachstum  $G'$  ist durch  $G' = G - E$  gegeben. Abhängig von der Relation zwischen  $E$  und  $G_m$  ergeben sich nun 3 qualitativ verschiedene Fälle, wie sie auch in Abbildung 6 dargestellt werden:

- $E < G_m$  (1. und 2. Graph): Der stabile Fixpunkt bei  $P = P_m$  hat sich zu kleineren Werten hin verschoben, d.h.  $P'_m < P_m$ . Außerdem gibt es

nun einen Minimalbestand, dessen Unterschreiten bei externer Entnahme zur Ausrottung führt,

- $E=G_m$  (3. Graph): Der nun instabile Fixpunkt liegt genau bei  $P=P_0$ . Jeder Startwert  $P < P_0$  führt in diesem Fall zum Aussterben der Population.
- $E > G_m$  (4. Graph): Nun ist die Entnahme so groß, dass unabhängig vom Startwert die Population zwangsläufig aussterben muss.

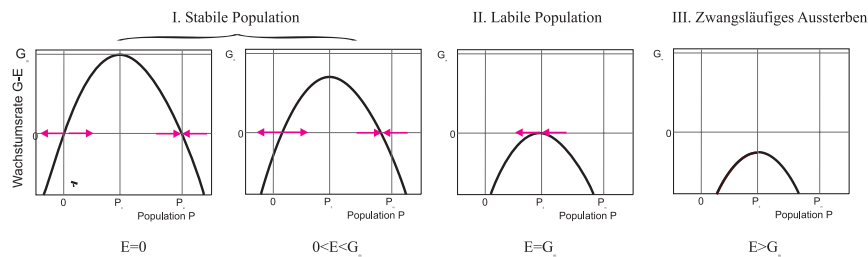


Abbildung 6: Darstellung der verschiedenen Fälle für das Verhältnis des Wachstums und der Entnahme aus dem System in der quantitativen Form unseres didaktischen Modells. Dargestellt sind jeweils das Wachstum als Funktion der Population und die jeweiligen Fixpunkte, wobei die Pfeile die Art des Fixpunktes (stabil/labil) charakterisieren.

Das genaue dynamische Verhalten des Systems hängt natürlich von den Werten für die Parameter  $\alpha$  and  $P_m$  ab, aber die qualitative Struktur der drei oben beschriebenen Klassen scheint eine allgemeine Eigenschaft des logistischen Wachstums und ähnlicher Systeme zu sein, so dass diese auch aus einer rein qualitativen Beschreibung ableitbar sein sollten. Dies hätte den Vorteil, dass so für eine wesentlich allgemeinere Klasse von Systemen die qualitative Natur des Verhaltens bewiesen werden könnte. Das Konzept der qualitativen Differentialgleichungen (QDGI) und ihrer Implementation im QSIM-Paket der University of Texas (Austin) erlaubt die Darstellung des logistischen Wachstums und verwandter Systeme auf recht umfassende Art:

In einem ersten Schritt werden die relevanten Wertebereiche der Variablen durch sogenannte *Landmarks* (eine mögliche deutsche Entsprechung wäre der Begriff „Meilensteine“) eingeteilt. Dies sind Werte, bei denen ein qualitativer Wandel in den Beziehungen zwischen der jeweiligen Variablen und dem Gesamtsystem stattfindet. Bei der Variable Population ( $P$ ) im obigen Beispiel, werden diese *Landmarks* als  $0$ ,  $P_0$  und  $P_m$  mit  $0 < P_0 < P_m$  gesetzt. Wichtig bei der Betrachtung von QDGIs ist, dass die numerischen Werte dieser *Landmarks*

keineswegs bekannt sein müssen, sondern dass nur das Wissen über ihre Existenz und Anordnung notwendig ist. Bei der Wachstumsrate  $G$  sind die Landmarks  $0$  und  $G_m$  mit  $0 < G_m$ .

Der *qualitative Zustand* ( $qval$ ) einer Variablen besteht aus ihrer *qualitativen Größe* ( $qmag$ ) und ihrer *qualitativen Richtung* ( $qdir$ ). Die *qualitative Richtung* ist das Vorzeichen der Änderungsrate der jeweiligen Variablen, kann also steigend ( $\uparrow$ ), konstant ( $^\circ$ ) or fallend ( $\downarrow$ ) sein. Mit *qualitativer Größe* wird das Verhältnis des Wertes einer Variablen relativ zu den Landmarks bezeichnet, also ob der Wert direkt auf einem Landmark liegt oder im offenen Intervall zwischen zwei Landmarks. Ein Beispiel für einen qualitativen Zustand in diesem Sinne für  $P$  ist  $[(P_0, P_m), \downarrow]$ , also eine fallende Population zwischen  $P_0$  und  $P_m$ .

Die Beziehungen zwischen den Variablen werden bei einer QDGI nicht als Gleichungen sondern als sogenannte *Constraints* (Bedingungen, Einschränkungen) angegeben. Im Falle des logistischen Wachstums wird die U-Constraint benutzt um das Verhältnis von  $G$  und  $P$  auszudrücken:

$$((U - P \ G \ (P_0 \ G_m)) \ (0 \ 0) \ (P_m \ 0)).$$

Damit wird festgelegt, dass für  $P < P_0$  die Wachstumsrate  $G$  eine monoton steigende Funktion von  $P$  ist, während sie für  $P > P_0$  monoton fallend ist (also ähnlich der Form eines umgekehrten U). Bei  $P = P_0$  ist  $G = G_m$  festgeschrieben. Außerdem gilt  $G = 0$  für Populationen von  $P = 0$  und  $P = P_m$ . Die Syntax in der obigen Gleichung entspricht derjenigen, die im QSIM-Paket implementiert ist. Analog können auch die anderen Beziehungen, implementiert werden, so z.B. dass  $E$  konstant ist, oder dass  $G'$  die Ableitung von  $P$  ist. Im Einklang mit der rein qualitativen Natur der QDGIs benutzt QSIM keinerlei numerische Werte, sondern nur diskrete Symbolverarbeitung. Abbildung 7 zeigt die qualitativen Zeitverläufe des oben skizzierten qualitativen Modells des logistischen Wachstums: Es gibt drei verschiedene Zeitverläufe, die mit den qualitativen Constraints vereinbar sind, über die in QDGIs die Beziehungen zwischen den verschiedenen Variablen dargestellt werden.

Qualitative Differentialgleichung:  $G=U(P)$        $\dot{P}=d/dt P$   
 $G=\dot{P}+E$        $E=const.$

Qualitative Zeitverläufe:

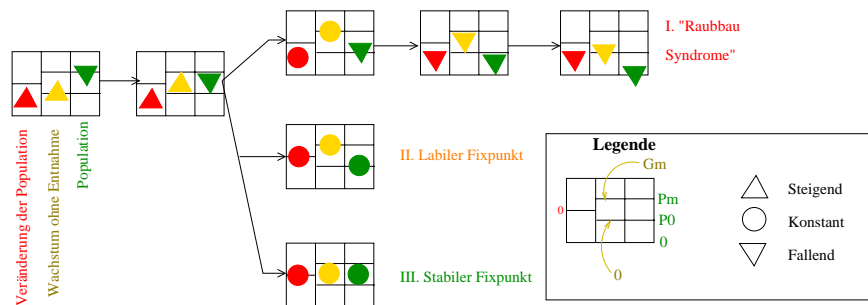


Abbildung 7: Qualitative Zeitverläufe des einfachen didaktischen Populationsdynamikmodell in Verallgemeinerung des logistischen Wachstums.

Diese drei Zeitverläufe entsprechen genau den drei qualitativ unterschiedlichen Fällen, die oben für die quantitative Version des Modells herausgearbeitet wurden - wieder gibt es bei einer maximalen Anfangspopulation zwei Fälle, in denen sich die Population stabilisiert, und einen, in dem sie ausstirbt. Im Vergleich zum quantitativen Modell gilt diese Aussage nun jedoch für eine wesentlich allgemeinere Klasse von Gleichungssystemen, und nicht nur für eine quadratische Abhängigkeit zwischen Population und Wachstum: Sie ist gültig für alle denkbaren Systeme gewöhnlicher Differentialgleichungen, die den Constraints genügen. Gerade übertragen auf die Modellierung ähnlicher, aber nicht gleicher, sozioökonomischer Effekte in verschiedensten Regionen der Erde ist diese Eigenschaft sehr wertvoll, da die größere Allgemeinheit der Aussage die Übertragbarkeit der Ergebnisse wesentlich erleichtert und oft erst ermöglicht.

Neben der Existenz der drei qualitativ unterschiedlichen Lösungen, gibt uns das Ergebnis der qualitativen Simulation noch wichtige Informationen über



Zeitverläufe und Bifurkationspunkte. So ergibt sich in diesem Modell, dass das Ereignis zum Zeitpunkt  $T_1$  bereits vollständig den endgültigen Ausgang bestimmt: Wenn bei  $P=P_0$  die Population weiter abnimmt, so wird sie unweigerlich vollständig verschwinden, wenn man annimmt, dass es keine externen Eingriffe gibt und dass die Modellstruktur sich nicht ändert. Dies wäre ein Beispiel für eine dynamische Eigenschaft, die in Bezug auf ein reales System nicht-nachhaltig genannt werden könnte.

Tabelle 2: Vergleich der wichtigsten Eigenschaften der konventionellen Modellierung mit DGIs (links) mit denen qualitativer Modelle auf der Basis von QDGs (rechts).

<b>Konventionelle Modellierung mit DGIs</b>	<b>Qualitative Modellierung mit QDGs</b>
Reelle Zahlen	<ul style="list-style-type: none"> <li>• <i>Landmark</i>-Werte, die qualitative wichtige Werte in Relation zu den anderen Variablen im System spezifizieren wie z.B. <math>P_0</math></li> <li>• Der <i>qualitative Zustand</i> einer Variablen besteht aus dem Verhältnis ihres Wertes zu den <i>Landmarks</i> und aus dem Vorzeichen ihrer Änderungsrate (<math>\uparrow</math>, <math>\downarrow</math>, oder <math>\circ</math>)</li> </ul>
Reellwertige Funktionen, die die Beziehungen zwischen den verschiedenen Variablen beschreiben.	Nur qualitative Merkmale, z.B. A ist eine monoton steigende Funktion von B, A ist eine „U-förmige“ Funktion mit $B_0$ als Maximumsstelle, etc.
System von Differentialgleichungen	Qualitative <i>Constraints</i> , die die Variablen und deren <i>qdirs</i> miteinander in Beziehung setzen.
Ein eindeutiger Zeitverlauf in expliziter, „metrischer Zeit“	Ein ganzer Baum <i>aller</i> möglichen Zeitverläufe, die mit den <i>Constraints</i> vereinbar sind. Zeit ist eine qualitative Variable, deren <i>Landmarks</i> die Zeitpunkte sind, bei denen sich das qualitative Verhalten des Systems ändert („topologische Zeit“).

In Tabelle 2 werden die Unterschiede zwischen dem qualitativen Ansatz der QDGs und der konventionellen Modellierung mit gewöhnlichen Differentialgleichungen übersichtshalber zusammengefasst. Bisher wurde eine qualitative Modellierung beschrieben, aber oft ist ein Mindestmaß an quantitativer Information vorhanden, und es wäre natürlich sinnvoll, diese auch nutzen zu können. Im Kontext qualitativer Modellierung können verschiedene Formen (semi-)quantitativer Information unterschieden werden:

1. Intervallinformation für alle oder einige *Landmarks* ist vorhanden.
2. Quantitative Informationen über die *Funktionen*, die in einer QDGI auftreten, sind verfügbar. Hier wären z. B. obere und untere Einhüllende denkbar.

Bezogen auf unser Beispiel könnten Intervalle für die Menge der externen Entnahme  $E$  und für das maximale Wachstum  $G_m$  bekannt sein. Wenn dann beispielsweise  $G_m$  in jedem Falle größer wäre als  $E$  (z.B.  $E=[12; 14]$ ,  $G_m=[15; 30]$ ), dann könnte der Zusammenbruch der Population ausgeschlossen werden.

Zusätzlich zu der Möglichkeit einige Zeitverläufe herauszufiltern, können auch umgekehrt die qualitativen Informationen genutzt werden, um die Unsicherheiten in den Intervallen zu verkleinern. Ein Beispiel hierfür wäre, dass, unter der Voraussetzung des qualitativen Verhaltens des Populationszusammenbruchs, die untere Grenze von  $E$  größer oder gleich der unteren Grenze von  $G$  sein muß.

Die numerische Komplexität des Problems semi-qualitativer Modellierung steigt natürlich drastisch, wenn von rein topologischer Zeit wie im rein qualitativen Fall zu metrischer Zeit übergegangen wird, d. h. wenn man etwas über die quantitative Bedeutung der Zeitpunkte in den qualitativen Zeitverläufen wissen will. Neben der direkten intervallarithmetischen Näherung, die allerdings nur sehr grobe Abschätzungen auf der Grundlage des Mittelwertsatzes liefern kann, gibt es verschiedene Methoden, dieses Problem zu lösen, welches sich im wesentlichen auf die Bestimmung der erreichbaren Menge in der Theorie der Differentialinklusionen zurückführen läßt. In diesem Bereich aktueller Forschung arbeitet unsere Gruppe derzeit an einer Methode vom Hamilton-Jacobi-Typ (Moldenhauer et al., 1999) und an einem Ansatz auf der Grundlage von Level-Sets (Sethian, 1999; Moldenhauer, 2000).

## 4 Bayessche Vermutungsnetze: Eine Brücke zwischen qualitativem und quantitativem Schließen

Die Faszination wahrscheinlichkeitstheoretischer Konzepte ist bereits für die alte ägyptische Kultur gut belegt (anhand des Umgangs mit Wettspielen). Gegenwärtig besteht ein zunehmendes Interesse an (teilweise) automatisierter Entscheidungsfindung, weil immer häufiger Beschlüsse großer Tragweite aufgrund unsicherer Informationen über sehr „komplexe“ Systeme zu fassen sind. Eintrittswahrscheinlichkeiten für bestimmte Szenarien sind für Entscheidungsträger in der Regel von zentraler Bedeutung. Wie nun im folgenden ausgeführt wird, erweist es sich als fruchtbar, die im Kontext der Verwaltung und Aufbereitung unsicherer Information auf natürliche Weise auftretenden wahrscheinlichkeitstheoretischen Konzepte mit weiteren, der Graphentheorie entlehnten, zu verschmelzen:


Bevor das Wissen über ein komplexes System durch Wahrscheinlichkeiten ausgedrückt werden kann, ist es erforderlich, ein Netzwerk von Beeinflussungsrelationen der einzelnen Systemkomponenten anzugeben.

### 4.1 Elementare probabilistische Konzepte

Bevor weiter auf die Verschmelzung graphen- und wahrscheinlichkeitstheoretischer Methoden eingegangen wird, sollen in diesem Abschnitt kurz die nötigen (elementaren) wahrscheinlichkeitstheoretischen und statistischen Konzepte rekapituliert werden. In Abbildung 8 wird eine Fundamentalsituation erläutert, die Bayessches Schließen erlaubt: Entscheidend ist, dass zwei hintereinandergeschaltete Zufallsexperimente durchgeführt werden (hier: Auswahl einer Urne, dann Auswahl einer Kugel),

wobei der Ausgang des ersten unbekannt, der des zweiten bekannt ist. Mit Hilfe der Bayesschen Formel schließt man vom Ergebnis des zweiten Experiments auf das (verborgene!) des ersten.

I.  *Wahrscheinlichkeiten für die Urnen ( $P(\text{Gelb}), P(\text{Blau})$ ) und bedingte Wahrscheinlichkeiten für die Kugeln ( $P(\text{Rot}|\text{Gelb}), P(\text{Rot}|\text{Blau}), P(\text{Blau}|\text{Gelb}), P(\text{Blau}|\text{Blau})$ ) sind bekannt, z.B.  $P(\text{Rot}|\text{Blau}) = 2/3$ .*

II.  *Zunächst wird eine Urne zufällig ausgewählt, dann aus dieser Urne eine Kugel.*

III.  *Gegeben die Farbe der Kugel: Aus welcher Urne könnte die Kugel stammen?*

### Bayessche Formel:

$$P(\text{Gelb}|\text{Rot}) = \frac{P(\text{Gelb}) \times P(\text{Rot}|\text{Gelb})}{P(\text{Gelb}) \times P(\text{Rot}|\text{Gelb}) + P(\text{Blau}) \times P(\text{Rot}|\text{Blau})}$$

**= 1/3 (in diesem Beispiel)**

Abbildung 8: Bayessches Schließen - ein Standardbeispiel. Ein „Experimentator“ wählt mit verbundenen Augen eine Urne (in unserem Beispiel sind beide gleich wahrscheinlich) und zieht dann aus dieser Urne eine Kugel, deren Farbe (hier „Rot“) er zur Kenntnis nimmt. Die Bayessche Formel liefert ihm eine Wahrscheinlichkeit dafür, aus welcher Urne die gezogene Kugel stammen könnte. Da diese *nach* dem Experiment zu ermitteln ist, heißt sie *aposteriori* (hier: 1/3 für „Gelb“). Hierzu benötigt er als Vorwissen die Zusammensetzung der Urnen (die bedingten Wahrscheinlichkeiten) und die Wahrscheinlichkeit, anfänglich eine gewisse Urne zu wählen (hier 1/2 für beide Urnen). Letztere „anfängliche“ Wahrscheinlichkeit liegt *vor* dem Experiment fest und heißt daher *a priori*.

Wir beschreiben den ersten Schritt (Urne) durch die *Zufallsgröße A* (die hier die Werte „Gelb“ und „Blau“ annehmen kann). Nachdem eine Urne ausgewählt worden ist, hängt die Wahrscheinlichkeit für eine bestimmte

Kugelfarbe  $B$  von der Urne ab. Dies ist die *bedingte Wahrscheinlichkeit*  $P(B/A)$ . Ferner wird die Wahrscheinlichkeit, dass eine bestimmte Urne *und* zugleich eine bestimmte Kugelfarbe vorkommen, durch die *Produktregel*  $P(A\&B)=P(A) P(B/A)$  angegeben. Durch diese Regel kann jeder Knoten einer *stochastischen Kette* (siehe Abbildung 9a) berechnet werden. Dabei entspricht jedem Knoten in der Kette eine Zufallsgröße. Das Beispiel in Abbildung 8 etwa würde durch nur zwei Knoten repräsentiert.

Liegt umgekehrt eine Situation wie im unteren Teil der Abbildung 8 vor, in der auf die Teilmenge von Ereignissen mit einer bestimmten Realisierung („Gelb“) der nachgeschalteten Zufallsgröße („Farbe der Kugel“) Bezug genommen wird, so wird der Rückschluss auf die voranstehende Zufallsgröße („Urne“) durch die *Bayessche Formel* möglich:

$$P(A_i | B_j) = \frac{P(A_i) \times P(B_j | A_i)}{\sum_k P(A_k) \times P(B_j | A_k)}.$$

Eine Anwendung der Bayesschen Formel (hier in allgemeiner Gestalt) ist in Abbildung 8 angegeben, wobei der Index  $k$  über die Realisierungen „Gelb“ und „Blau“ läuft.

Bis hierher zielten die mathematischen Ausführungen auf eine rein stochastische Situation: Die Wahrscheinlichkeiten bezogen sich auf die Vorhersage von wiederholbaren Experimenten, alle Wahrscheinlichkeiten wurden als bekannt vorausgesetzt. In diesem Sinne liefert die Bayessche Formel in Abbildung 8 für viele Wiederholungen des Vorgangs (II+III) die relative Häufigkeit für „Gelb“ unter der Nebenbedingung „Rot“. Dieselbe, im stochastischen Sinne wohldefinierte Formel lässt sich nun auch erfolgreich im Bereich der *Statistik* anwenden.

Statistik fragt nach Eigenschaften verborgen liegender Verteilungen, die jedoch gewisse Messreihen beeinflusst haben. Aufgrund von Daten wird versucht, Rückschlüsse auf die verborgene Verteilung zu ziehen, etwa ihre Parameter zu schätzen. Die hinzutretenden Gesichtspunkte hinsichtlich der Bayesschen Formel beziehen sich nicht auf ihre algebraischen Ausdrücke, sondern auf deren Interpretation (Berger, 1994). In der statistischen Variante von Abbildung 8 könnten die Urnen zwei Hypothesen über die Wahrscheinlichkeit, "Rot" zu produzieren, bedeuten (1/3 versus 2/3). Die Aprioriverteilung könnte durch

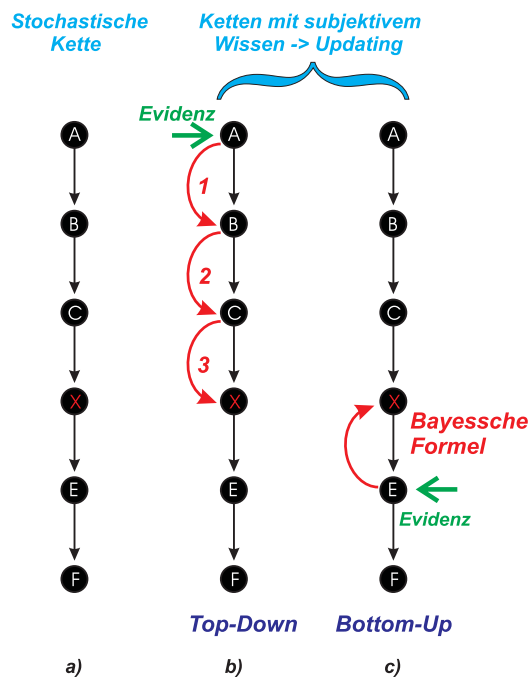


Abbildung 9: Aktualisierung einer Netzwerk-Kette.

eine subjektive Expertenmeinung gegeben sein. Nur eine der Hypothesen (Urnen) wird als wahr (aktualisiert) angenommen.

Unter sehr häufigem Testen derselben Urne und mehrfachem Anwenden der Bayesschen Formel gewinnt man schließlich Sicherheit, um welche Urne es sich handelt (Freedman, 1963). Das anfängliche, subjektiv gefärbte Wissen wäre so sukzessive und wohldefiniert objektiviert worden.

Traditionelle, nicht-Bayessche statistische Verfahren benötigen demgegenüber keine Aprioriverteilung, die in vielen Anwendungen eine Schwierigkeit darstellen könnte. Um dennoch zu nichttrivialen Aussagen zu gelangen, treffen sie im Gegenzug meist nur Aussagen über Ensembles von Messergebnissen, und nicht - wie nach Bayes - unter der Nebenbedingung *einer*, verlässlich vorliegenden Messung. (Ein prominentes Beispiel eines

solchen traditionellen Verfahrens stellt das Schätzen der Parameter einer Normalverteilung aufgrund einer Messreihe dar. Die mitgelieferte Sicherheitswahrscheinlichkeit bezieht jedoch alle zukünftigen Experimente mit ein.) Demgegenüber wertet nur das Bayessche Verfahren *eine*, unbezweifelbar vorliegende Messreihe optimal aus (Berger, 1994), sofern die Aprioriverteilung bekannt ist.

## 4.2 Probabilistische Netzwerke

Nur gerichtete, azyklische Graphen (Beispiele sind in Abbildung 9 und Abbildung 10 gegeben) lassen eine Anwendung der Bayessche Formel zu. (Zyklische Elemente müssen durch Prozessaufspaltung in mehrere azyklische zerlegt werden.) Graphentheorie handelt zum Teil davon, Äquivalenzklassen von Graphen zu identifizieren (Pearl, 1988). Dies hat praktische Bedeutung, weil so numerisch belastender „Wildwuchs“ wirksam zurückgeschnitten werden kann.

Falls in einem gerichteten, azyklischen Graphen jeder Beeinflussungs-Pfeil mit einer bedingten Wahrscheinlichkeit identifiziert und außerdem subjektives Wissen in die Verteilungen eingeflossen ist, spricht man von einem *Bayesschen Vermutungsnetz*, und zwar aus folgendem Grund:

Steht neue Information zur Verfügung, etwa in Abbildung 9b erstmalig über die Größe  $A$ , in Abbildung 9c über die Größe  $E$ , verändert dies die zugehörigen Verteilungen. Dies wirkt sich auf das Wissen über eine gesuchte Größe  $X$  aus. Eine aktualisierte  $X$ -Verteilung kann dann in (b) „top-down“ durch die Produktregel, in (c) „bottom-up“ mit Hilfe der Bayesschen Formel ermittelt werden.

In multivariaten Problemen sind in der Bayesschen Formel oft hochdimensionale, numerisch oder analytisch schwer zugängliche Integrale auszuwerten. Daher finden unter dem Eindruck expandierender Rechnerleistung Bayessche Netzwerke erst in den letzten zehn Jahren breitere Anwendung. Als hilfreich erweist sich vielfach der Einsatz von Markov-Ketten-Monte-Carlo-Methoden. Hierbei wird die Markovkette so eingerichtet, dass ihre Gleichgewichtsverteilung mit der gesuchten Aposterioriverteilung identisch ist (Tierney, 1994).

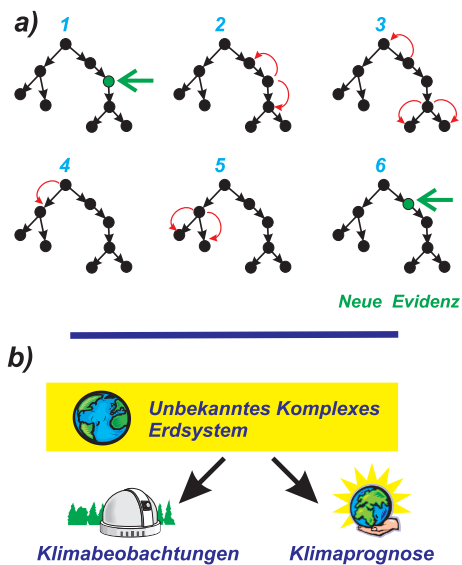


Abbildung 10(a): Baumartiges Netzwerk. Auch komplexere (als lineare) Netze lassen sich nach Bayes aktualisieren. (b): Eine paradigmatische Anwendung von (a) in der Global-Change-Forschung: Von Klimabeobachtungen wird auf das (unbekannte) Erdsystem(modell) geschlossen („bottom-up“). Das verbesserte Wissen (aktualisiertes Netz) lässt eine realistischere Prognose über die zukünftige Klimaentwicklung erwarten („Top-Down“-Schlussweise). In der Praxis ist hier mit einem viel komplexeren Netz zu rechnen.

### 4.3 Robuste Bayessche Statistik

Wie oben ausgeführt wurde, verfügt das Bayessche Verfahren über eine erstaunliche Flexibilität. Paradoxerweise benötigt es jedoch präzise Wahrscheinlichkeitsverteilungen. Hier setzt in der Regel die Kritik von Praktikern an, denn exakt bekannte Verteilungen sind nur in Ausnahmefällen zu erwarten. Dies gilt insbesondere im Falle subjektiver Verteilungen, für die das Bayessche Verfahren ansonsten besonders geeignet ist.

Aus diesem Grund ist in den vergangenen Jahrzehnten die *Robuste Statistik* (Walley, 1991) entwickelt worden. In der einfachsten Variante wird die



unbekannte Verteilung durch eine viel besser zu rechtfertigende *Klasse von Verteilungen* ersetzt. Oftmals erhält man als Endergebnis nicht mehr eine präzise Wahrscheinlichkeit, sondern ein *Wahrscheinlichkeitsintervall* (Berger, 1994). Das Konzept der Wahrscheinlichkeitsintervalle ist bereits in der traditionellen Statistik etabliert: So läßt sich etwa dem Konfidenzintervall des zu schätzenden „ $p$ “ einer Binomialverteilung keine feste Sicherheitswahrscheinlichkeit (wie etwa im Fall einer Normalverteilung) zuordnen, sondern eben lediglich ein Wahrscheinlichkeitsintervall.

Zusammenfassend läßt sich konstatieren, dass Bayessche Vermutungsnetze die Verschmelzung zweier mathematischer Konzepte darstellen: Zum einen stellen sie gerichtete azyklische Graphen dar, die die Möglichkeit bieten, qualitative kausale Information aufzubereiten. Zum anderen erlauben sie es, Unsicherheit, die in Wahrscheinlichkeitsverteilungen gefasst ist, systematisch zu propagieren. Zusätzliche Information kann in wohldefinierter und, in gewissem Sinne, optimaler Weise berücksichtigt werden, indem das Netzwerk lokal mit Hilfe der Bayesschen Formel aktualisiert wird. Bayessche Statistik erlaubt es, subjektives Wissen in den Formalismus einfließen zu lassen. Das Netz im Lichte neuer Information zu aktualisieren bedeutet zugleich, die gespeicherte Information zu objektivieren.

Falls nicht alle erforderlichen Wahrscheinlichkeitsverteilungen genau bekannt sind, kann auf die Robuste Statistik zurückgegriffen werden. Schließlich sind eine Reihe von Methoden entwickelt worden (Graphenverschlangung, Markov-Ketten-Monte-Carlo-Methoden), die es erlauben, auch komplexere Netzwerke numerisch zu realisieren.

## 5 Syndrome II: Formalisierung funktionaler Muster

Im Folgenden soll nun das Syndromkonzept auf der Basis des in Abschnitt 3 vorgestellten Kalküls der qualitativen Differentialgleichungen weiter formalisiert werden, wobei zwei Aspekte im Vordergrund stehen:

1. die systematische Verallgemeinerung von detaillierten Fallstudien im Sinne der Identifikation von riskanten funktionalen Mustern (HFPs), wie im unteren Teil von Abbildung 2 symbolisiert, und
2. die Bestimmung von nicht-nachhaltigen Trajektorien (Syndromen), die von einem solchen HFP produziert werden.

Mit diesen Punkten direkt verbunden sind die Fragen nach der räumlichen Verteilung und dem dynamischen Zustand (dem Stadium) von Syndromen, der eingeschränkten Vorhersage der weiteren Entwicklung einer Region, in der ein bestimmtes HFP Gültigkeit hat sowie der Identifikation von Politikoptionen zur Beeinflussung von syndromaren Entwicklungen. Um die Methodik anschaulich zu erläutern, soll die Diskussion entlang des schon eingeführten SAHEL-SYNDROM-erzeugenden HFP geführt werden (Cassel-Gintz et al., 1997, Schellnhuber et al., 1997; Petschel-Held et al., 1999a, Lüdeke et al., 1999), wobei auf den wichtigsten Regelkreis (im Wesentlichen die gelb markieren Symptome in Abbildung 3) fokussiert werden soll, der schon der Dispositionsbestimmung in Abschnitt 2 zugrunde lag. Dieser *kybernetische Kern* des HFP ist in Abbildung 11 noch einmal dargestellt, wobei zur Verdeutlichung die Hilfsvariable „Ertrag“ hinzugefügt wurde.

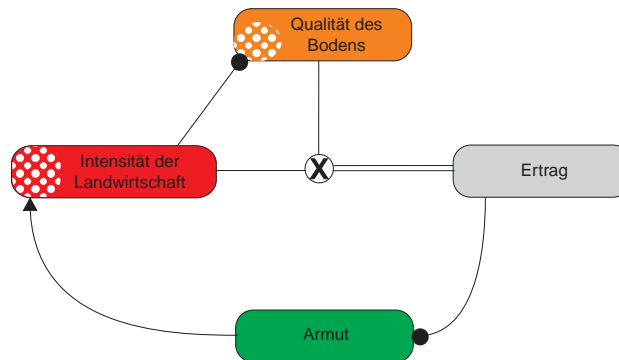


Abbildung 11: Kernmechanismus oder kybernetischer Kern des SAHEL-SYNDROM-erzeugenden HFP, der den wichtigsten Regelkreis im Gesamtbeziehungsgeflecht (Abbildung 3) darstellt. Die Symbole (zur Definition siehe Anhang) charakterisieren im Rahmen des Kalküls qualitativer Differentialgleichungen die Zusammenhänge zwischen den Zustandsvariablen (abgerundete Rechtecke)

## 5.1 Generalisierung von Erklärungsmustern in Fallstudien

Die Voraussetzung für die Verallgemeinerung von einzelnen Ursache-Wirkungsbeziehungen, wie sie in unterschiedlichen Fallstudien in unterschiedlichen Regionen identifiziert werden, ist die Identifikation von Überbegriffen (in unserer Terminologie *Symptomen*) für die jeweils verursachenden und beeinflussten Faktoren.

Im Falle unseres Beispiels könnte etwa in einer Fallstudie die Größe des Ziegenbestands als Ursache für Bodenverdichtung nachgewiesen werden, während in einer anderen Region der Wirkungszusammenhang zwischen Brachezeit und Fertilitätsverlust identifiziert wird. Es liegt nun nahe, die Ursachen (Größe des Ziegenbestandes und „Kürze“ der Brachezeit) unter dem Begriff „Intensität der Landwirtschaft“ zusammenzufassen und die beeinflussten Größen (Verdichtung und Fertilitätsverlust) unter „Bodendegradation“.

Ob dies ein zielführender Schritt war, hängt nun davon ab, wie die Wirkungszusammenhänge in den beiden Fallstudien genauer spezifiziert sind - zeigen sie die gleichen qualitativen Eigenschaften (im einfachsten Fall: die Wirkung ist eine monoton steigende Funktion der Ursache), gilt dieser Zusammenhang auch auf der Ebene der aggregierten Variablen oder Symptome. In Abbildung 12 ist dieser Zusammenhang (hypothetisch) für zwei Fallstudienregionen dargestellt (blaue Koordinatensysteme), wobei hier von in den Fallstudien explorierten *quantitativen* Zusammenhängen ausgegangen wird, die dann auf der aggregierten Ebene der Symptome und deren Beziehungen als eindeutige *qualitative* Beziehung (M+ für monoton steigend) erscheinen.

Das hier vorgestellte Verfahren ist nun nicht von solch hochgesteckten Erwartungen an Fallstudien abhängig, da natürlich auch qualitative Aussagen schon in den Fallstudien in die Aggregation integriert werden können. Als Beispiel für eine kompliziertere qualitative Beziehung ist ein umgekehrt U-förmiger Zusammenhang (U-) zwischen den Symptomen Armut (als Verallgemeinerung von Einkommen und Armut) und Intensität dargestellt – eine Beziehung die darin begründet liegt, dass unterhalb einer bestimmten Minimalversorgung auch eine armutsgetriebene landwirtschaftliche Intensivierung nicht mehr möglich ist (gelbe Koordinatensysteme in Abbildung 12).

Insgesamt illustrieren diese Beispiele, dass die begriffliche Aggregation (Symptomdefinition) unauflöslich mit den Wirkungsbeziehungen auf der Fallstudienebene verknüpft ist. Weiterhin stimmen die hier gefundenen qualitativen Beziehungen nun genau mit den Definitionen im Kalkül der qualitativen Differentialgleichungen überein, das damit optimal für die Abbildung verallgemeinerter Ursache-Wirkungsbeziehungen geeignet ist. Auf die soweit beschriebene Art können aus Sätzen von thematisch verwandten Fallstudien Ursache-Wirkungsnetze widerspruchsfrei abstrahiert werden, wobei einzelne Fallstudien durchaus nur Teile des gesamten Beziehungsgeflechts widerspiegeln können. Zum Themenkreis des SAHEL-SYNDROM-erzeugenden HFP findet man eine beachtliche Fülle von Fallstudien (beispielhaft sei hier nur die Übersicht von Kates und Haarman, 1992, sowie das Buch von Blaikie und Brookfield, 1987, genannt). Hier ist die forschungslogische Notwendigkeit von Konzepten der systematischen Generalisierung (wie dem Syndromkonzept) besonders augenfällig.

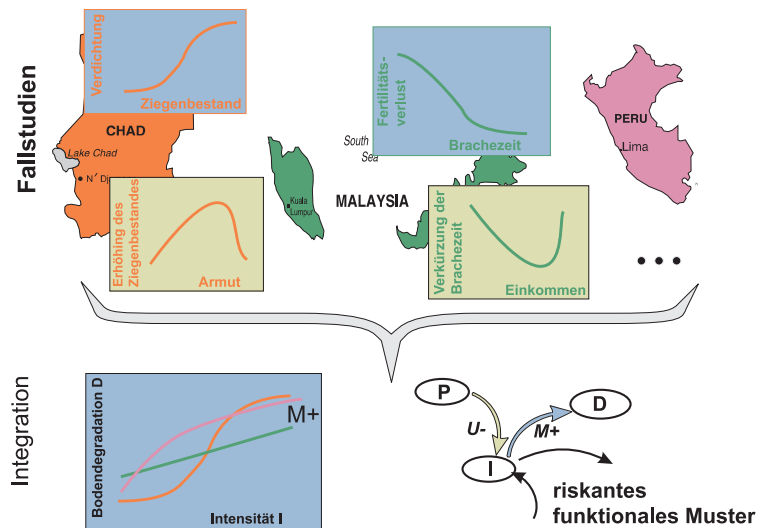


Abbildung 12: Generalisierungsschema zur Formulierung einer Klasse von Zivilisation-Natur-Wechselwirkungen. Fallstudien (hier: hypothetisch Tschad, Malaysia und andere) können genutzt werden, die regional gültigen Beziehungen zwischen relevanten Variablen (hier: Verdichtung/Ziegenbestand/Armut bzw. Fertilitätsverlust/Brachezeit/Einkommen) zu spezifizieren. Subsumiert man diese unter allgemeinere Begriffe (hier: Bodendegradation, Intensität der landwirtschaftlichen Nutzung, Armut) können auf dieser Ebene *qualitative Relationen* definiert werden, die dann für alle *Fallstudienregionen* gleichermaßen gültig sind. In diesem Fall ergibt sich: Bodendegradation ist eine monoton wachsende Funktion der Intensität - die Intensität hängt von der Armut gemäß einer umgekehrt U-förmigen Funktion ab.

Das in Abbildung 11 gezeigte Beziehungsgeflecht stellt nun genau so eine Verallgemeinerung von fallstudienbasierten Ursache-Wirkungshypothesen dar. Die zugrunde liegenden Fallstudien enthalten allerdings über die angestellten Ursache-Wirkungshypothesen hinaus oft auch Beobachtungen über den zeitlichen Verlauf der diskutierten Variablen, die häufig quantitativer (z.B. Zeitverläufe des Bodenabtrags), aber oft auch qualitativer Natur (z.B. „die Erdrutschhäufigkeit nahm bis 1950 zu, ist aber derzeit eher konstant“) sind. Die Methode der qualitativen Differentialgleichungen erlaubt nun die Validierung der abstrahierten HFPs an diesen Beobachtungen.

## 5.2 Auswertung von HFPs: Validierung, Nachhaltigkeit und Vorhersage

Auch dieser Schritt soll wieder am SAHEL-SYNDROM-erzeugenden HFP illustriert werden. Das qualitative Modell, wie es in Abbildung 11 dargestellt ist, definiert nahezu vollständig ein qualitatives Differentialgleichungssystem, das mit Hilfe des in Abschnitt 3 dargestellten Kalküls zu bearbeiten ist (zur Bedeutung der graphischen Charakterisierung der Beziehungen siehe Appendix). Hinzugefügt werden lediglich ein „landmark“ – d.h. die Existenz eines ausgezeichneten Wertes - der Bedeutung „maximal nachhaltig“, *ms*, für die Intensität der Landwirtschaft, oberhalb dessen Bodendegradation einsetzt (unterhalb dieses Wertes wird Bodenerholung angesetzt) sowie ein „landmark“ der Bedeutung „existentielle Armut“, *ex*, von dem an armutsgetriebene Intensivierung einsetzt (beides durch viele Fallstudien gestützte Eigenschaften, siehe z.B. Schal, 1998; Kates und Haarman, 1992). Abbildung 13 stellt nun alle - mit Hilfe des Kalküls der qualitativen Differentialgleichungen bestimmten - möglichen Zeitverläufe dar, die mit dem qualitativen Modell aus Abbildung 11 (unter der angegebenen Anfangsbedingung) verträglich sind (zur Nomenklatur vergleiche Abschnitt 3).

Es sei an dieser Stelle noch auf einige Eigenschaften dieser Klasse von Lösungen hingewiesen: oft existieren mehrere Nachfolgezustände für einen gegebenen qualitativen Zustand (dargestellt als Verzweigungen im Verhaltensbaum). D.h. unter der Annahme der Gültigkeit des Modells für eine gegebene Region sind u.U. mehrere zukünftige Verläufe möglich – die hier eingeführte Methodik stellt also eine Systematisierung der sogenannten „schwachen Prognose“ (oft auch als „Szenarienbildung“ bezeichnet) dar. Des weiteren erkennt man bei genauerer Inspektion der möglichen Zeitverläufe, dass es „Äste“ des Verhaltensbaums gibt, die zu akzeptablen Zuständen führen, d.h. sowohl in der sozioökonomischen als auch der naturräumlichen Dimension anstrebenswert sind, während andere Äste eindeutig inakzeptable Entwicklungen indizieren (hohe und ansteigende Armut, beständige Verschlechterung des Zustands der natürlichen Ressourcen). Nach dieser Einordnung der unterschiedlichen Trajektorien bezüglich ihrer (Nicht)-Nachhaltigkeit lassen sich beispielsweise kritische Verzweigungspunkte in der Systemdynamik identifizieren, an denen über das Einschwenken in einen

möglicherweise irreversibel nicht-nachhaltigen Pfad entschieden wird, was natürlich von großer Bedeutung für die Entwicklung wirksamer Gegenmaßnahmen ist. In der Nomenklatur des Syndromkonzepts stellen letztere Trajektorien das SAHEL-SYNDROM dar.

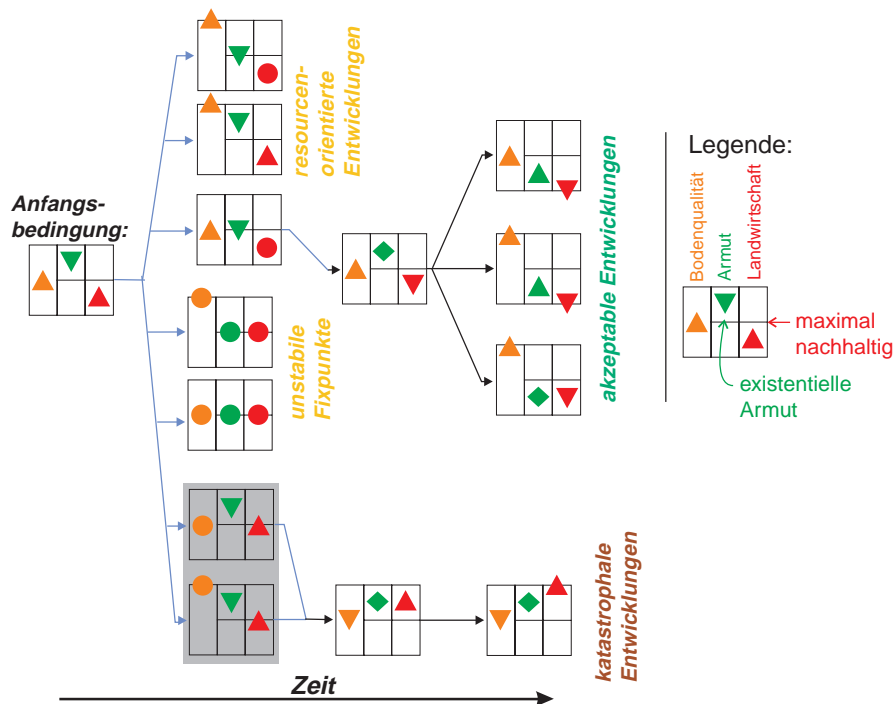


Abbildung 13: Vollständiger Verhaltensbaum des kybernetischen Kerns des SAHEL-SYNDROM-erzeugenden funktionalen Musters aus Abbildung 11. Die Notation entspricht der in Abbildung 7. Man beachte auch hier die Verzweigung des „Baumes“, die zeigt, daß unter den allgemeinen Annahmen des qualitativen Modells (des HFP) unterschiedliche qualitative Trajektorien möglich sind - insbesondere solche, die als akzeptabel (geringe Armut, guter Naturzustand) oder als inakzeptabel (große Armut, schlechter werdender Naturzustand) zu kennzeichnen sind. Diese Trajektorien stellen das SAHEL-SYNDROM dar.

Bezüglich der Validierungsfrage ist die Bedeutung des aus dem qualitativen Modell deduzierten Verhaltensbaums offenkundig: Wird in einer Region, für die die Gültigkeit des Modells behauptet wird, ein zeitliches Verhalten

beobachtet, das sich nicht unter die berechneten qualitativen Trajektorien subsumieren läßt, bedeutet dies die Falsifizierung des Modells für die entsprechende Region – andernfalls liegt eine notwendige Bestätigung vor. Diesem Verfahren liegt die klare Trennung zwischen Zusammenhangshypothesen und Beobachtungen des Zeitverlaufs relevanter Variablen zu Grunde – beides Informationen, die aus detaillierten Fallstudien zu gewinnen sind.

Im nächsten Abschnitt wird dieses Verfahren am Vergleich eines etwas erweiterten Modells mit einem Satz von Fallstudien illustriert.

### 5.3 Illustration: Validierung eines HFP am Fallstudiensatz eines DFG-Programms

Im Folgenden soll die Validierung eines um die explizite Modellierung der Lohnarbeitsalternative erweiterten kybernetischen Kerns des Sahel-Syndrom-erzeugenden HFP an einem Satz von Fallstudien durchgeführt werden, die im Rahmen des DFG-Schwerpunktprogramms „Human Dimensions of Global Environmental Change“ erstellt wurden (Lohnert und Geist, 1999). Die untersuchten Fallstudienregionen sind in Abbildung 14, jeweils zusammen mit ihrem jeweiligen Zustand der Bodendegradation nach Oldeman (1991) dargestellt. Diese Fallstudien wurden nun in zweierlei Hinsicht ausgewertet. Einerseits wurden die in ihnen enthaltenen Zusammenhangshypothesen analysiert, was zu einer Erweiterung des (ansonsten zutreffenden) ursprünglichen HFP (Abbildung 11) um die Lohnarbeitsperspektive führte (siehe Abbildung 15, blaue Ellipse). Als Hypothese zur nun zu beschreibenden Allokationsentscheidung für die kleinbäuerliche Arbeitskraft wurde die Orientierung am Vergleich der Arbeitsproduktivität der derzeitigen landwirtschaftlichen Arbeit mit der der Lohnarbeit herangezogen (graue Ellipsen, zur Illustration der Gültigkeit siehe auch Weber, 1993). Weiterhin wurde zur Charakterisierung der ökonomischen Situation der relative Preis für landwirtschaftliche Güter und Arbeitszeit berücksichtigt.



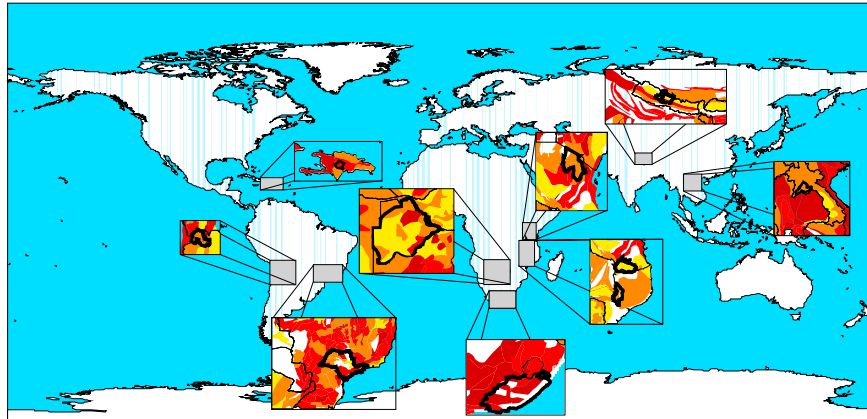


Abbildung 14: Fallstudienregionen (untersucht im DFG-Programm „Human Dimension of Global Change“), die als Informationsgrundlage für das funktionale Muster in Abbildung 15 dienten. Die Ausschnitte zeigen den Zustand der anthropogenen Bodendegradation nach GLASOD (Oldeman, 1991) in diesen Regionen, wobei die Farbskala von weiß (keine Degradation) über gelb und orange zu rot (sehr starke Degradation) reicht.

Auf der anderen Seite erlaubten die Fallstudien die Rekonstruktion der qualitativen Zeitverläufe der im Modell enthaltenen Variablen während der jeweiligen Untersuchungsperioden – d.h. es war möglich, zu Validierungszwecken die Konsistenz von Beobachtungen und Modellresultaten zu vergleichen.

Letztere sind als Ergebnis des Algorithmus zur Lösung qualitativer Differentialgleichungen für die gegebene Anfangsbedingung (und unter konstanter gesamtwirtschaftlicher Situation) in Abbildung 16 dargestellt, wobei nun einige interessante neue Trajektorien entstehen, die darauf beruhen, dass sich für den Kleinbauern die Möglichkeit eröffnet, bei beständigem Rückgang der Bodenqualität (und damit seiner landwirtschaftlichen Arbeitsproduktivität) auf Lohnarbeit auszuweichen, was im Idealfall zu einer Erholung des Bodens und einer Wiederaufnahme der Landwirtschaft führen kann. D.h., das im ursprünglichen HFP als mögliche Lösung nahegelegte Ausweichen auf alternative Einkommensquellen zur Entlastung der Ressourcenbasis - und damit das Ausbrechen aus dem Teufelskreis - ist nun mit in das Modell einbezogen. Der unakzeptablen Entwicklung in Abbildung 13 entspricht die in Abbildung 16 grau-gelb gekennzeichnete Entwicklung (steigende

Bodendegradation bei gleichzeitiger Intensivierung der Landwirtschaft), die nun jedoch in den Satz der akzeptablen Entwicklungen (den naturräumlichen Bedingungen angepaßter Mix aus landwirtschaftlicher Arbeit und Lohnarbeit) übergehen kann.

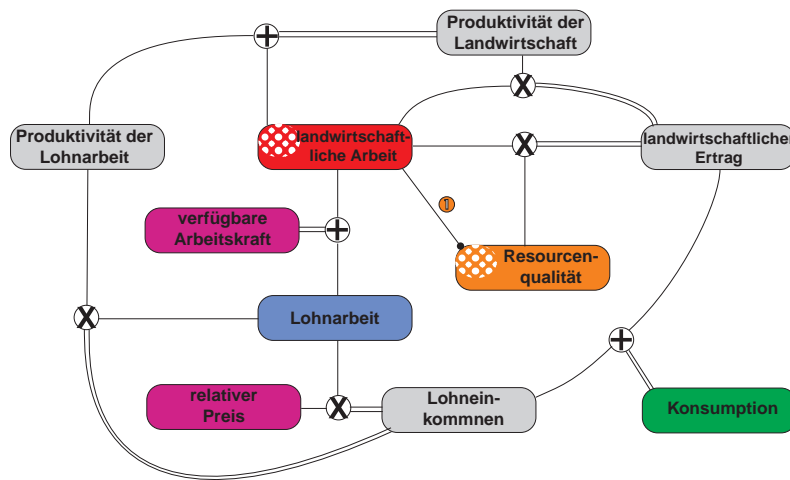


Abbildung 15: Erweitertes Ursache-Wirkungsschema zum SAHEL-SYNDROM.

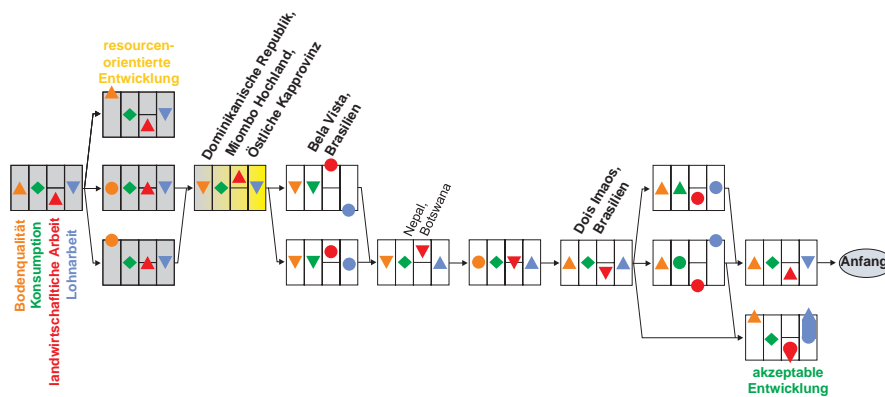


Abbildung 16: Verhaltensbaum zum erweiterten Ursache-Wirkungsschema aus Abbildung 15.

Über diese allgemeinen Resultaten zur Systemdynamik hinaus wurde nun geprüft, ob die in den Fallstudienregionen beobachteten qualitativen Dynamiken mit gewissen modellgenerierten Entwicklungen übereinstimmen. Dies war für sechs Fallstudienregionen, die in Abbildung 16 über den entsprechenden Stadien der Entwicklung angeführt sind, möglich. Prognosen (im Fall von Bifurkationen: schwache Prognosen) und daraus abgeleitete Politikempfehlungen werden nun möglich (siehe Petschel-Held et al., 1999b).

#### 5.4 Iteratives Schema zur systematischen Identifikation von HFPs

Die bisher erläuterten Schritte der Generierung von Hypothesen über funktionale Muster aus Fallstudien sowie der Validierung dieser Hypothesen an beobachteten qualitativen Trajektorien kann nun in einem Schema zur iterativen Fallstudienverallgemeinerung zusammengefaßt werden (Petschel-Held und Lüdeke, 2000). Der erste Schritt besteht in der Bestimmung von generalisierten Variablen, die die Subsumption der fallstudien-spezifischen Variablen erlauben. Auf der Basis dieser generalisierten Variablen können dann aus den (quantitativen oder qualitativen) Beobachtungen in jeder Fallstudienregion qualitative Trajektorien bestimmt werden (Schritt 2). Andererseits enthalten die Fallstudien Zusammenhangshypothesen, die bei geschickter Wahl der generalisierten Variablen das Aufstellen eines qualitativen Beziehungsgeflechts zwischen diesen erlauben, das mit dem gesamten Satz der Fallstudien in Einklang steht (Schritt 3). Auf dieses Beziehungsgeflecht kann dann der Algorithmus der qualitativen Differentialgleichungen angewandt werden – mit der Erzeugung aller mit dem qualitativen Modell verträglichen Zeitentwicklungen (Schritt 4). Damit ist eine „Falsifikationsmöglichkeit“ (Schritt 5) gegeben: kommen nicht alle beobachteten Verläufe des Fallstudien-satzes in diesem Verhaltensbaum vor, stellt das hypothetisierte Beziehungsgeflecht keine gültige Verallgemeinerung der Dynamik der Fallstudienbasis dar. In diesem Fall ist eine weitere Iteration nötig, die zunächst auf der Basis der mit den Fallstudien verträglichen Modifikation des Beziehungsgeflechtes durchgeführt wird. Erst im Falle, dass

diese Modelländerungen nicht ausreichend sein sollten, wird man die Aggregation der fallstudienspezifischen Variablen verändern.

Konvergiert dieses iterative Verfahren (d.h., erfordert das weitere Hinzufügen von Fallstudien nicht immer neue Modellvariationen) ist die dem Syndromkonzept zugrunde liegende Generalhypothese „typischer riskanter funktionaler Muster“ nachgewiesen.

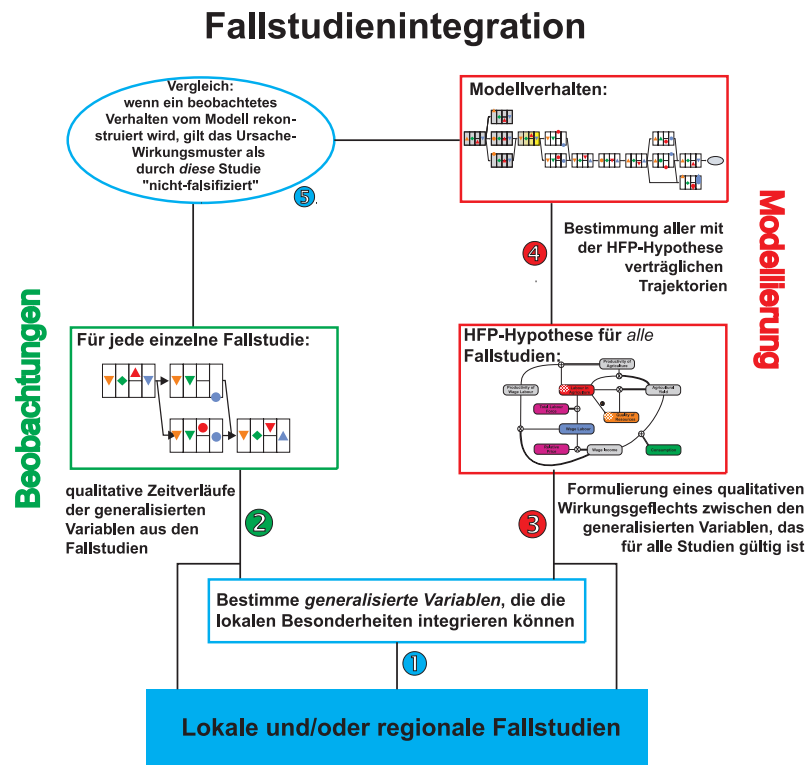


Abbildung 17: Allgemeines Schema zur iterativen Generalisierung von Fallstudien mit dem Ziel der Identifikation von allgemein gültigen funktionalen Mustern.

## 6 Self-Organized Maps (SOM) zur Musteridentifikation in hochdimensionalen Datensätzen

In vielen Bereichen der Untersuchung von komplexen realen Systemen, in denen umfangreiche Datenreihen zu einzelnen Aspekten des Systemverhaltens vorliegen, haben moderne Techniken des „Data Mining“ zur Synthese oder Integration dieser Daten erstaunliche Fortschritte erzielen lassen. Diese Verfahren, die aus der rasanten Entwicklung der Informationsverarbeitung und *-verdichtung* im Rahmen der Untersuchungen zur Modellierung künstlicher Intelligenz entstanden, sind der Wirkungsweise realer neuronaler Strukturen im Gehirn nachempfunden. Die biologischen Vorbilder - Neuronale Netzwerke als reale Strukturen im zentralen Nervensystem oder auch der Retina - werden in der Computertechnologie als sogenannte *konnektionistische Systeme* aus einer mittleren bis sehr großen Anzahl von einfachen Verarbeitungseinheiten nachgebildet, die in einem Netzwerk miteinander verschaltet sind. Diese „neuronalen Netzwerke“ dienen in der Erforschung der künstlichen Intelligenz zur Simulation der Leistungen natürlicher Nervenverbände. Ihre Arbeitsweise und die Möglichkeiten des Netzwerkes werden durch seinen Aufbau - die Topologie des Netzes - bestimmt. Sie gibt an, wie die Einheiten eines Netzes verbunden sind, wie diese Struktur verändert werden kann und wie sich die Aktivierung von Nervenimpulsen im Netzwerk ausbreitet (Kohonen, 1993).

Je nach der gewählten Topologie des Netzes lassen sich unterschiedliche Anwendungen realisieren. Die Verarbeitungseinheiten als Knoten des Netzes können beispielsweise die Zentren oder Zielpunkte einer Clusteranalyse in Form einer Mustererkennung bilden, wobei die Stärken der Verbindungen sich im Zeitverlauf der Simulation nach dem strukturellen Zusammenhang der zu analysierenden Eingangsdaten organisieren. Durch die algorithmische Ausformulierung dieser Strategie gewinnt man ein adaptives numerisches

Verfahren, das als sogenannte selbstorganisierende neuronale Karte ein Problem löst, welches sowohl in lebenden Organismen als auch in technischen Anwendungen häufig vorkommt: *die Reduktion einer überreichlichen Datenflut auf das Wesentliche*. Die Reduktion auf das "Wesentliche" ist hier so zu verstehen, daß die Dimensionalität (Zahl der Freiheitsgrade) reduziert wird und die hochdimensionale Ausgangsdatenstruktur auf die wichtigsten Informationsachsen abgebildet wird. Im Gegensatz zu herkömmlichen Clusteranalyseverfahren findet also eine „Informationsverdichtung“ statt und – besonders wichtig - es entsteht *keine* Datengruppierung, die nur auf einfachen linearen Strukturzusammenhängen beruht.

Wenn also die Eingangsdaten hinsichtlich einiger Freiheitsgrade eine gewisse Regelmäßigkeit in der Art zeigen, daß einige Kombinationen häufiger vorkommen als andere, so organisiert sich die Karte in der Weise, daß diese Datenzusammenhänge betont und als eine neue Ergebnisdatenstruktur herausgearbeitet werden. Dies geschieht ohne Rückgriff auf irgendeine Theorie, die vielleicht für den gesetzmäßigen Zusammenhang unter den Daten existiert, sondern nur aufgrund der nachbarschaftlichen Zusammenhänge in den Ausgangsdaten.

## 6.1 Anwendungsbeispiel für SOM: Identifikation gleichermaßen klimatisch verwundbarer Gemeinden in NRW

Um mit Hilfe dieser Technik die Gemeinden in NRW zu bestimmen, die sich durch eine ähnliche Gesamtverwundbarkeit hinsichtlich zunehmender Extremwetterereignisse auszeichnen (Wesselhoeft et al.), wurden alle vorhandenen Indikatoren zum klimasensitiven Inventar als Datensatz für alle Gemeinden zusammengestellt. Dieser Eingangsdatensatz<sup>1</sup> umfaßt damit 24 Freiheitsgrade

---

<sup>1</sup> Dies waren als Datensatz für jede Gemeinde: Einwohnerdichte, Landwirtschaftliche Fläche, Landwirtschaftliche Dauerkulturen, Landw. Fläche heterogen, Nadelwald, Obstanbauflächen, Gewässeranteil, Boden-Erodierbarkeitsbewertung, KFZ-Bestand, Unfallhäufigkeit, Bevölkerung (insgesamt), Bevölkerungsanteil (>50 Jahre), Anzahl der touristischen

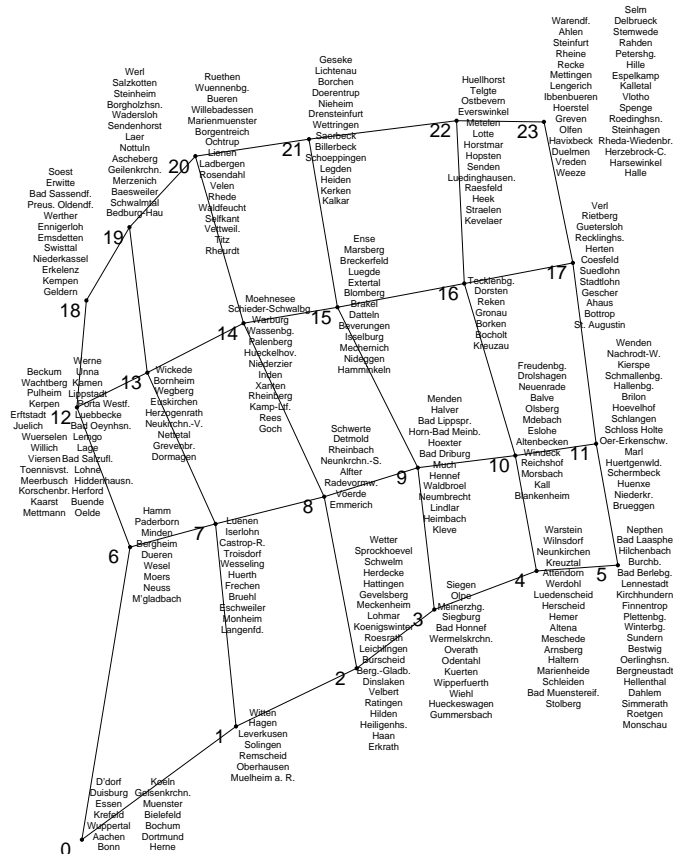


Abbildung 18 : Darstellung der Anpassung des Netzwerkes an die Mannigfaltigkeit der Daten und die den Knoten zugeordneten Gemeinden. Die Länge der Netzwerkkanten ist ein Maß für die Ähnlichkeit der gefundenen Verwundbarkeitskategorien. Es ist deutlich erkennbar, daß sich die industriellen Zentren durch besondere Merkmale auszeichnen (Knoten 0).

Übernachtungsmöglichkeiten, Fremdenverkehr/übernachtungen, Fremdenverkehr/  
 Aufenthaltsdauer, Beschäftigte am Arbeitsort, Einpendler, Arbeitslosenzahlen, Saisonaler  
 Arbeitslosenindex, Land- u. Forstwirtsch. Beschäftigte, Baugewerbe - Beschäftigte, Handel,  
 Dienstleister - Beschäftigte, Kaufkraftkategorien/Wohlstand, Länge der oberird. Stromleitungen

und bildet 396 Punkte (Gemeinden) in einem 24-dimensionalen Zustandsraum ab.

Wenn zwischen den Eingangsdaten ein topologischer Zusammenhang besteht, läßt sich unter Umständen die Dimensionalität entscheidend verringern. Die Rahmenbedingung für diese Art der Informationsverdichtung ist die Erhaltung der Nachbarschaft der Datenpunkte im Ergebnisraum. Mittels einer selbstorganisierenden neuronalen Karte wird dies im hohen Maße erreicht: Eine Vielzahl von Testläufen ergab eine optimale Informationsverdichtung in einem Dateneinbettungsraum mit einer Dimension, die bei 2 liegt. Dies galt für verschieden Knotenzahlen. Das beste Ergebnis fand sich für eine 2-dimensionale Geometrie des Netzes mit 6x4 Knoten. Es existiert allerdings eine Verzerrung in einer Dimension, die durch die Datenelemente des ersten Netzknotens (Knoten 0, siehe Abbildung 18) verursacht wird. Dennoch zeigt sich, daß eine Datenreduktion in Form einer Informationsverdichtung auf eine ebene Topologie möglich ist: Die 24-dimensionalen Eingangsdaten für alle Gemeinden konnten adaptiv auf eine 2-dimensionale Basisstruktur reduziert werden. Hinsichtlich ihrer wesentlichen Verwundbarkeitscharakteristika lassen sie sich in 24 Typen mit jeweils topologisch ähnlichen Eigenschaften einordnen .

Abbildung 18 zeigt die so bestimmte Gruppierung verschiedener Gemeinden. Das Ergebnis dieser Merkmalskartierung erscheint auf den ersten Blick schwer deutbar, da intuitiv manche Einordnungen befremdlich erscheinen. In den folgenden Diskussion wird jedoch klarwerden, daß das benutzte objektive Verfahren einer Clusterung mit Informationsverdichtung sehr wohl wichtige Einsichten erlaubt. Zumindest wird schon auf den „zweiten“ Blick deutlich, daß einzelne Verwundbarkeitsklassen als sinnvoll erscheinen: Der Knoten mit der Nummer 0 repräsentiert beispielsweise die industriellen Zentren in Nordrhein-Westfalen, die sich sicherlich durch vergleichbare Eigenschaften auszeichnen dürften.

Des weiteren bildet diese Informationsstruktur auch ein *Datenmodell* mit der Möglichkeit, weitergehende Sensitivitätsanalysen hinsichtlich der Übergänge einzelner Clustermitglieder in andere Cluster durchzuführen: Leichte Variationen der Eingangsdaten einer Gemeinde können beispielsweise dazu führen, daß sich die „jetzige Clusterzugehörigkeit“ verändert und ab einer zu bestimmenden Größenordnung der Veränderung die Gemeinde als Element eines anderen Clusters auftritt. Das theoretische Konzept der



Symbolischen Dynamik leistet hier Erstaunliches: Es lassen sich Eingriffe, Planungsmaßnahmen und sonstige Veränderungen an den einzelnen Verwundbarkeitskategorien hinsichtlich ihrer Wirkungen selektiv *und* in Kombination bewerten sowie als symbolische Datenreihen erfassen. Die Cluster, die als besonders witterungsrobust beurteilt werden, können Zielkategorien definieren, die im Rahmen der Durchführung von Vorsorgemaßnahmen erreicht werden sollten. Hierbei können dann komplexe Kontrollstrategien theoretisch untersucht und evaluiert werden, indem deren Wirkung auf die Überführung („Symbolische Trajektorien“) der vulnerablen Cluster in die Zielkategorie geprüft werden.

Natürlich sollte hier einschränkend darauf hingewiesen werden, daß dieses Konzept insbesondere hinsichtlich einer breiteren Datenbasis erweitert werden müßte. Dies gilt insbesondere deshalb, da eine Reihe von Faktoren, die nur indirekt die Verwundbarkeit einer Gemeinde beeinflussen, im Rahmen dieser Untersuchung nicht betrachtet werden konnten: Der immens wichtige Bereich der Vorsorgeinstrumente - auch in öffentlicher Trägerschaft - wie Feuerwehr, Katastrophendienste oder Straßenbauämter, die durch den Indikator der monetären Ausgaben in diesem Bereich gemessen werden, könnte die Verwundbarkeit hinsichtlich zunehmender Extremwetterereignisse signifikant senken. Aber grundsätzlich kann schon jetzt die oben skizzierte Strategie einer Symbolischen Dynamik die theoretische Grundlage für weitergehende Aussagen bilden und insbesondere die Wirkungsweise von Handlungsstrategien fundiert evaluieren helfen.

Die Einordnung der Gemeinden in sogenannte *charakteristische Verwundbarkeitsklassen*, die mit Hilfe des neuronalen Netzes in Form einer Clusteranalyse erreicht wurde, stellt jedoch nur einen ersten Schritt in Richtung einer *integrierten Verwundbarkeitsbewertung* dar. Es wurden zwar 24 Verwundbarkeitsklassen für die Gemeinden in Nordrhein-Westfalen identifiziert, doch noch fehlt ein Kriterium, diese Klassen nach der verschiedenen „Gesamtverwundbarkeiten“ einordnen zu können. Grundsätzlich würden sich hier verschiedene Vorgehensweisen anbieten, die von Experteneinschätzung bis hin zum „Ranking“ der verschiedenen Verwundbarkeitsklassen auf der Basis eines systemnahen „objektiven“ Kriteriums erfolgen könnten. Für die Formulierung einer solchen Maßfunktion bieten sich beispielsweise die Fläche der Spektren (siehe Abbildung 19) an.

Sie repräsentiert die integrale Verwundbarkeit, d.h. die Summe aller Verwundbarkeitsaspekte einer Verwundbarkeitsklasse.

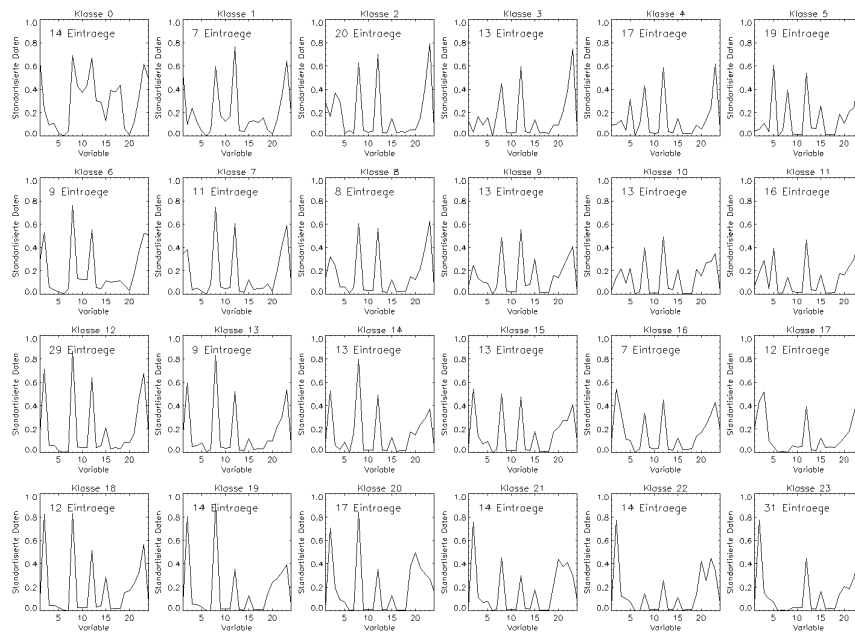


Abbildung 19: Die spektrale Verteilung der Ausgangsdaten in den einzelnen Verwundbarkeitsklassen

Die Spektren in Abbildung 19 stellen die „normierten“ Anteile der Eingangsdaten dar, wie sie auf die verschiedenen Knoten abgebildet werden. Jeder Knoten des neuronalen Netzes (oder damit gleichbedeutend jede Verwundbarkeitsklasse der Clusteranalyse) ist durch eine bestimmte Zusammensetzung der Eingangsdaten repräsentiert. Jedes Spektrum ist damit charakteristisch für die jeweilige Verwundbarkeitsklasse und die gesamte Fläche stellt somit die „Zusammenfassung“ aller Verwundbarkeitsaspekte für diese Klasse dar. Als Arbeitshypothese gilt also: Je größer die Fläche, desto größer ist die integrierte Verwundbarkeit dieser Klasse von Gemeinden. Mit Hilfe dieser *heuristischen* Vorgehensweise läßt sich eine reproduzierbare integrative Bewertung der vielfältigen und facettenreichen Verwundbarkeiten

einer Region gegenüber einer Zunahme von Extremwetterereignissen erreichen.

Es zeigt sich insbesondere, daß sich die Klasse 0 hinsichtlich dieses *integralen* Verwundbarkeitsmaßes signifikant von den anderen unterscheidet. Die meisten Elemente dieser Klasse repräsentieren industrielle Zentren in Nordrhein-Westfalen (siehe Abbildung 18), so daß diese Gruppe sehr wohl sehr viel stärker von einer Zunahme von Extremwetterereignissen betroffen sein könnte, wenn man - als Voraussetzung - die in dieser Studie gewählten Verwundbarkeitsaspekte zugrunde legt. Die Schadenspotentiale, die mit dem individuellen Wohlbefinden oder beispielsweise dem Straßenverkehr verknüpft sind, scheinen die integrierte Verwundbarkeit urbaner Zentren stark zu erhöhen und deshalb ihre Anfälligkeit zu betonen.

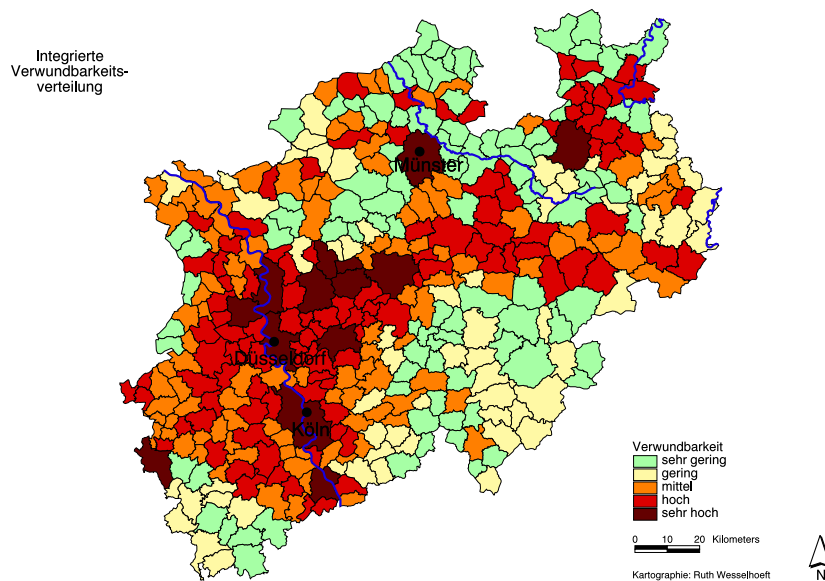


Abbildung 20: Darstellung der integralen Verwundbarkeiten auf Gemeindeebene, wobei als Maß für die Reihung die Flächen aus Abbildung 18 bestimmt wurden

Es sollte hier nochmals hervorgehoben werden, daß die benutzte *heuristische* Bewertungsvorschrift keinesfalls exakte, quantitative

Abschätzungen liefert. Das berechnete Flächenmaß der einzelnen Knoten, das sich aus dem Spektrum der Eingangsdaten bestimmen läßt, kann nur einen *qualitativen* Hinweis darauf geben, welche Klasse verwundbarer ist und welche robuster erscheint. Die exakten Zahlenwerte sollten deshalb auch nicht weitergehend interpretiert werden. Aus diesem Grund wurde der gesamte Zahlenbereich aller Flächenwerte in fünf geeignete Intervalle unterteilt und mit qualitativen Verwundbarkeitskategorien bezeichnet, die von "sehr gering" bis "sehr hoch" reichen. Abbildung 20 zeigt die Verteilung der dadurch gewonnenen Einschätzung einer integrierten Verwundbarkeit. Das Ergebnis stellt in dieser Form eine *erste Abschätzung* dar, die sicherlich noch weiter verfeinert werden sollte. Insbesondere konnten - wie schon mehrfach erwähnt - nicht alle Verwundbarkeitsaspekte datenmäßig in hinreichender Auslösung erfaßt werden, so daß gegenwärtig die Stabilität der integrierten Bewertung bei einer weitergehenden Einbeziehung anderer Aspekte noch unklar ist. Dennoch ist auf der Basis der bisher benutzten Verwundbarkeitsaspekte eine approximative Identifikation der verwundbarsten Gemeinden in Nordrhein-Westfalen gelungen. Dieses Ergebnis stellt somit eine tragfähige Diskussionsgrundlage und einen Ausgangspunkt für weitergehende Analysen dar.

## 7 Schlußbemerkungen

Die Schwierigkeiten bei der Behandlung des Globalen Wandels ergeben sich, wie in Abschnitt 1 erläutert, einerseits aus der Komplexität des Systems (Vielzahl und Nichtlinearität der Beziehungen), andererseits aus der „modularen“ Unsicherheit (oft nur vage, unsichere oder qualitative Kenntnis der einzelnen Beziehungen). Als Möglichkeit der Behandlung des ersten Aspekts wurde in Abschnitt 2 das Syndromkonzept eingeführt, dessen Grundhypothese die Möglichkeit der Dekomposition des komplexen globalen Gesamtsystems in einzelne, nur schwach gekoppelte Subsysteme ist, wobei solche, die unakzeptable Trajektorien („Syndrome“) produzieren, Gegenstand der Forschung sein sollen (sogenannte HFPs oder „riskante funktionale Muster“).

In den Abschnitten 3 und 5 wurde dann ein Konzept vorgestellt, mit dem es möglich ist, solche HFPs auf systematische Weise aus der Vielzahl der vorliegenden detaillierten Fallstudien über problematische Mensch-Naturwechselwirkungen iterativ herzuleiten. Hier wird der Ansatz der qualitativen Differentialgleichungen verwendet, der den Aspekten der Typisierung von quantitativem Fallstudienwissen einerseits und der Aufnahme von qualitativem Wissen andererseits gerecht wird. Hiermit konnte ein wesentlicher Schritt zur methodischen Systematisierung des Syndromkonzeptes geleistet werden.

In Abschnitt 4 wurde ein Konzept diskutiert, das einen kritischen Aspekt der modularen Unsicherheit in der Erdsystemanalyse berührt: In vielen Fällen ist die Vagheit des gegenwärtigen Wissens über einzelne Beziehungen statistisch beschreibbar, und mit zunehmender Beobachtungserfahrung oder Modellqualität kann das Wissen „schärfer“ werden - die Bayesschen Vermutungsnetze formalisieren diesen „update“-Prozess. Ob und wie diese sicherlich sehr relevante Facette der Unsicherheit unseres Systemwissens mit

der Methode der qualitativen Differentialgleichungen verknüpft werden kann, bleibt eine wichtige Forschungsfrage.

Bisher kreisten die methodischen Überlegungen um Vagheit und qualitativen Charakter unseres Systemwissens sowie die Typisierung unter diesen Bedingungen. Nun gibt es in einigen Bereichen des Globalen Wandels auch andersgeartete Situationen, die durch große Datenmengen und geringes oder völlig unklares Zusammenhangswissen charakterisiert sind. Die hierfür verfügbaren mathematisch-methodischen Möglichkeiten wurden in Abschnitt 6 am Beispiel der Anwendung des Kohonen-Algorithmus dargestellt. Diese datengeleitete, und damit zunächst nur auf den direkten Untersuchungsgegenstand bezogene Analyseverfahren bietet viele Ansatzpunkte zur *Generierung* allgemeinerer Zusammenhänge (etwa durch die Bestimmung der Anzahl der Freiheitsgrade des betrachteten Systems), die dann in die oben diskutierten Methoden einfließen können.

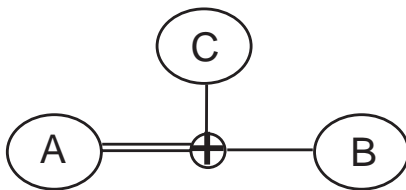
Insgesamt sollte klar geworden sein, dass qualitative und semi-quantitative Verfahren zur Beschreibung komplexer Dynamiken (wie der des Globalen Wandels) ein noch schlummerndes Potenzial darstellen, das möglichst bald aktiviert werden muss.

## 8 Appendix: Symbole für die graphische Repräsentation von qualitativen Modellen

Notation zur Charakterisierung der Symbole:

qdir: Änderungsrichtung der Variablen (positiv/0/negativ)

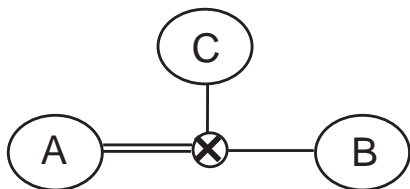
qmag: qualitativer Wert der Variablen, d.h. ein Landmark oder ein Intervall zwischen zwei Landmarks



Dieses Zeichen codiert die qualitative Addition von C und B mit dem Resultat A. Qualitative Addition ist u.a. durch die folgenden Eigenschaften spezifiziert:

- Die Änderungsrichtungen werden addiert, d.h., wenn  $qdir(C) > 0$  und  $qdir(B) > 0$ , dann ist  $qdir(A) > 0$ ; falls jedoch  $qdir(C) > 0$  und  $qdir(B) < 0$ , dann kann  $qdir(A)$  entweder positiv, negativ, oder 0 sein.
- Für  $qmag(B) = 0$  und  $qmag(C) = 0$  gilt  $qmag(A) = 0$ .
- Für  $qmag(B) = 0$  und  $qmag(C) \neq 0$  gilt  $qmag(A) = qmag(C)$

Qualitative Subtraktion, i.e.  $A = B - C$ , wird als qualitative Addition, i.e.  $A + C = B$ , ausgedrückt.

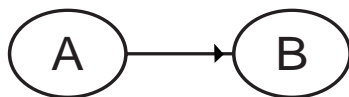


Dieses Zeichen codiert die qualitative Multiplikation von B und C mit dem Resultat A, i.e.

- Die Änderungsrichtungen kombinieren sich nach der Rechenregel der

Differentialrechnung, i.e.  $qdir(A) = qmag(B) \cdot qdir(C) + qmag(C) \cdot qdir(B)$

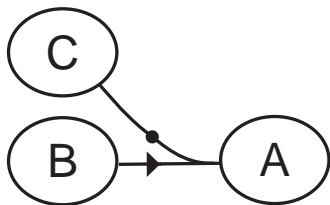
- Falls  $qmag(B)=0$  oder  $qmag(C)=0$ , folgt  $qmag(A)=0$ .



B ist eine monotone Function von A, was z.B. bedeutet, dass, wenn A wächst, auch B wächst. Dies entspricht der Bedingung

$$\frac{\partial B}{\partial A} > 0.$$

Im Falle eines Punktes anstatt eines Pfeils hat die partielle Ableitung ein negatives Vorzeichen.



Dies ist eine multivariate Verknüpfung,  $A(B,C)$ , die folgende Eigenschaften besitzt

$$\frac{\partial A}{\partial B} > 0, \frac{\partial A}{\partial C} < 0.$$

Hier bedeutet der Punkt eine negative und der Pfeil eine positive partielle Ableitung.



Hier wird nicht die Variable A selbst, sondern ihre erste Ableitung, i.e.  $dA/dt$ , in den Zusammenhang einbezogen.



## 9 Literatur:

- Berger, J.O., 1994. An Overview on Robust Bayesian Analysis, Test 3, 5.
- Berry, M.V., 1978. Regular and irregular motion. In S. Jona, ed., Topics in non-linear dynamics. AIP conference proceedings 46, La Jolla.
- Blaikie, P. and Brookfield, H., 1987. Land Degradation and Society. Methuen, London, New York.
- Cassel-Gintz, M. A., Lüdeke, M. K. B., Petschel-Held, G., Reusswig, F., Plöchl, M., Lammel, G., 1997. Fuzzy-Logic Based Global Assessment on the Marginality of Agricultural Land Use. Climate Research 8, 135-150
- CIESIN, 1992. Pathways of Understanding: The Interactions of Humanity and Global Environmental Change. Consortium for International Earth Science Information Network (CIESIN).
- Clancy, D.J., 1997. Solving complexity and Ambiguity problems with qualitative simulation. PhD thesis, Department of Computer Science, University of Texas, Austin.
- Cubasch, U., 1992. Das Klima der nächsten 100 Jahre. Physikalische Blätter, 48, 85-89.
- Freedman, D. A., 1963. On the Asymptotic Behavior of Bayes' Estimates in the Discrete Case, Ann. Math. Statist. 34, 1386.
- Haken, H., 1977. Synergetics. Springer-Verlag, Berlin Heidelberg New York.
- Kasperson, J. X., Kasperson, R.E., Turner II, B. L. (eds), 1995. Regions at Risk. United nations Unoversity Press, Tokyo.
- Kates, R.W. and Haarman, V., 1992. Where the Poor Live: Are the Assumptions Correct? Environment, vol. 34, pp.4-11, 25-8.
- Kates, R.W. and Clark, W. C., 1999 (eds). Our Common Journey. Board on Sustainable Development - Policy Devision - National Research Council. National Academy Press, Washington, D. C.
- Kohonen, T., 1993. Physiological interpretation of the self-organizing map algorithm. Neural networks 6, 895-905
- Kuipers, B., 1994. Qualitative Reasoning: Modeling and Simulation with incomplete Knowlegde. MIT Press, Cambridge.
- Lohnert, B. und Geist, H. (Hsg.), 1999. Coping with changing environments. Ashgate, London.
- LUCC, 1995. Land-Use and Land-Cover Change. Science Plan. IGBP-Report No. 35, Stockholm.
- Lüdeke, M. K. B., Moldenhauer, O., G., Petschel-Held, 1999. Rural poverty driven soil degradation under climate change: the sensitivity of disposition towards the SAHEL SYNDROME with respect to climate. Environmental Modeling and Assessment 4, 295-314

- Moldenhauer, O., Bruckner, Th., Petschel-Held, G., 1999, in: Conference Proceedings of Computational Intelligence for Modelling, Control and Automation (CIMCA) 99, The use of semi-qualitative reasoning and probability distributions in assessing possible behaviors of a socio-economic system, ed. M. Mohammadian, IOS Press, London, 1999.
- Moldenhauer, O., 2000 Robust Modeling of Complex Environmental Systems under Uncertainty. PhD-Thesis Universitaet Potsdam. in Preparation.
- Oldeman, L.R., Hakkelling, R.T.A., Sombroek, W.G., 1991. World map of the status of human-induced soil degradation: an explanatory note. International Soil Center, Wageningen.
- Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems, Morgan Kaufmann Pub., San Mateo, California.
- Petschel-Held, G., Block, A., Cassel-Gintz, M., Kropp, J., Lüdeke, M.K.B., Moldenhauer, O., Reusswig, F., Schellnhuber, H.J., 1999a. Syndromes of Global Change. A qualitative modelling approach to assist global environmental management. *Environmental Modeling and Assessment* 4, Nr. 4, 315-326
- Petschel-Held, G., M. K. B. Lüdeke and F. Reusswig, 1999b. Actors, Structures and Environment. A Comparative and Transdisciplinary View on Regional Case Studies of Global Environmental Change. In: B. Lohnert and H. Geist Coping with changing environments pp. 255-291. Ashgate, London.
- Petschel-Held, G. and Lüdeke, M. K. B., 2000. Integration of Case Studies on Global Change by Means of Artificial Intelligence. submitted to Integrated Assessment.
- Rotmans, J., 1998. Methods for IA: The challenges and opportunities ahead. . *Environmental Modeling and Assessment* 3, 155-179.
- Schal, P., 1998. Humusvorräte und nutzbare Wasserspeicherkapazitäten der Böden als differenzierende Faktoren in der traditionellen hochandinen Landwirtschaft in der Charazani-Region (Bolivien). *Forstw. Cbl.* 177, 176-188.
- Schellnhuber, H. J., 1999. 'Earth system' analysis and the second Copernican revolution. *Nature* 402, C19-C23.
- Schellnhuber, H.J., Block, A., Cassel-Gintz, M., Kropp, J., Lammel, G., Lass, W., Lienenkamp, , Loose, C., Lüdeke, M.K.B., Moldenhauer, O., Petschel-Held, G., Plöchl, M., Reusswig, F., 1997. Syndromes of Global Change, *GAIA* 6, Nr. 1, 19-34
- Sethian, J. A. , 1999. Level set methods and fast marching methods. Cambridge University Press.
- Tierney, L., 1994. Markov Chains for Exploring Posterior Distributions, *The Annals of Statistics* 22, 1701.
- Walley, P., 1991. Statistical Reasoning with Imprecise Probabilities, Chapman and Hall.
- WBGU - German Advisory Council on Global Change, 1997. World in Transition: The Research Challenge, Springer, Berlin.
- Weber, G., 1993. Entwicklung einer Angebotskomponente und einer Budgetkomponente für SPEL. Wehle, Witterschlick, Bonn.
- Wesselhoft, R., Block, A., Reusswig, F., Kropp, J., Lienenkamp, R., Regionale Auswirkungen des anthropogen bedingten Treibhauseffekts: Die Anfälligkeit Nordrhein-Westfalens für einen möglichen Klimawandel. Unveröffentlichte Studie im Auftrag des Ministeriums für Umwelt, Raumordnung u. Landwirtschaft d. Bundeslandes NRW.
- Wissel, C., 1989. Theoretische Ökologie, Springer Verlag, Berlin Heidelberg New York
- Zimmermann, H.J., 1991. Fuzzy set theory and its applications. 2nd revised edn. Kluwer Academic Publishers, Boston.