

Elektronische Datenverarbeitung in der Geographie

W.-D. Rase, Vancouver-Burnaby, Kanada

Zusammenfassung: Die Möglichkeiten der Anwendung von Computern in der Geographie werden untersucht und beschrieben. Es werden unterschieden kartographische und statistisch-taxonomische Fragestellungen sowie die Anwendung von Simulationsmodellen. Es ergibt sich, daß es für den Laien überraschend viel Anwendungsmöglichkeiten von Computern in der Geographie gibt, z. T. sogar mit recht einfachen Mitteln.

Summary: The possibilities of applying computers in geography are discussed and described. There are three main areas of computer application: cartographic problems, statistical/taxonomical problems, and simulation models. The author demonstrates the fact that there are many possible and real applications of computers in geography.

1. Geographie und Informationsverarbeitung

Die Aufgabe der geographischen Wissenschaft ist, ganz grob definiert, die Analyse räumlicher Beziehungen, in den meisten Fällen unter speziellen Gesichtspunkten. Ein sehr wesentlicher Teil der analytischen Arbeit besteht aus Informationsverarbeitung und Informations-transformation: Die mit unseren Sinnesorganen erfassbaren Elemente der geographischen Wirklichkeit werden gesammelt, kategorisiert und miteinander in Beziehung gesetzt, das generalisierte Bild der Wirklichkeit dann zu einem Modell synthetisiert, etwa in Form einer Tabelle, einer Karte, einer verbalen Beschreibung oder von Kombinationen aller Darstellungsmethoden.

In Bezug auf Datenspeicherung und Verarbeitungsgeschwindigkeit sind dem menschlichen Gehirn Grenzen gesetzt. Der Schluß liegt daher nahe, elektronische Datenverarbeitungsanlagen als "Denkverstärker" einzusetzen, die den Menschen bei der Informationsverarbeitung unterstützen. Allerdings sind die formalen Voraussetzungen für eine automatisierte Verarbeitung, ein widerspruchsfreier Satz von Regeln und Vorschriften, nur für Zahlen erfüllt, nämlich im Lehrgebäude der Mathematik. Deshalb ist vorerst der Anwendungsbereich geographischer Datenverarbeitung auf in Zahlen verschlüsselbare Geoelemente beschränkt. Darunter fallen geometrische Beziehungen, Wetter- und Klimadaten, quantifizierbare Bereiche der Geomorphologie, demographische und ökonomische Größen.

Es ergeben sich zwei Aspekte für den Einsatz von Computern:

1. Kostensenkung durch Verlagerung bisher manuell durchgeführter Arbeiten auf eine Datenverarbeitungsanlage (DVA). Darunter fallen hauptsächlich elementare statistische Berechnungen, Diagramme und kartographische Arbeiten.

2. Erweiterung der analytischen Möglichkeiten durch Verwendung statistischer Methoden, die aus Gründen des Speicher- und Rechenaufwandes nur auf einer DVA möglich sind. Faktorenanalyse, multivariate Klassifizierungen und Simulation sind in diese Kategorie einzureihen.

2. Computer-Kartographie

BOARD [2] beschreibt einen Kreislauf der geographischen Analyse als iterativen Prozeß zur Verbesserung eines Modells (Bild 1).

Bei manueller Kartenherstellung kann dieser Kreis aus finanziellen Gründen nicht sehr oft durchlaufen werden. Mit den Möglichkeiten der automatisierten Kartierung ist aber der Kostenzuwachs bei jedem Zyklus vernachlässigbar klein, und unbrauchbare Karten kann man daher ohne schlechtes Gewissen dem Papierkorb überantworten. Es gibt nun schon eine Reihe von Computerprogrammen und Programmsystemen für die automatische Kartenherstellung. Ein Programmsystem jedoch, das SYMAP-System des Laboratory for Computer Graphics and Spatial Analysis der Harvard University in Cambridge (Mass.), ist so ausgearbeitet und gleichzeitig so anpassungsfähig, daß es als Beispiel für die Möglichkeiten und Probleme automatisierter Kartierung dienen soll.

2.1. Das SYMAP-Programm

Das Kernstück des Harvard-Systems ist SYMAP (für SYnagraphic MAPping), das inzwischen in der Version V existiert [7]. Als Ausgabemedium dient ein Zeilendrucker, wie er an jeder DVA vorhanden ist. Das macht

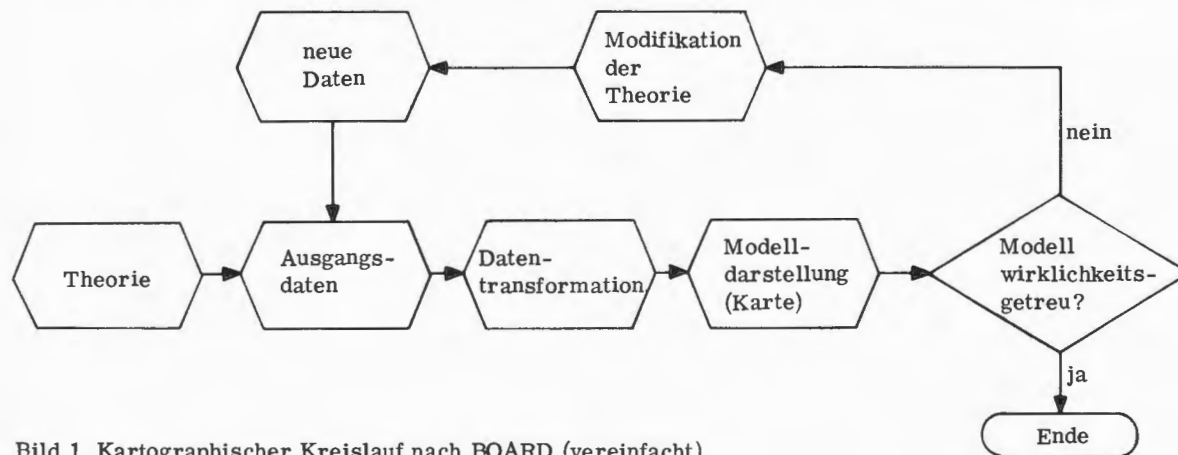


Bild 1 Kartographischer Kreislauf nach BOARD (vereinfacht)

das Programm sehr flexibel, denn computergesteuerte Zeichengeräte oder Kathodenstrahlröhren sind nicht allgemein verfügbar. Der Nachteil der Zeilendrucker-Ausgabe liegt im verhältnismäßig groben Raster; es hat sich aber herausgestellt, daß es für den praktischen Gebrauch vollständig ausreicht. Da das Programm beliebig große Karten herstellen kann, indem mehrere Papierbahnen nebeneinander geklebt werden, läßt sich durch photographische Verkleinerung das Raster verfeinern. SYMAP liefert drei Darstellungsmethoden: Isolinienkarten, Choroplethenkarten, Proximitätskarten.

Isolinienkarten verwendet man zur Darstellung kontinuierlicher Phänomene, wie z. B. Luftdruck oder Temperatur oder einer topographischen Oberfläche. Verschiedene Interpolationsalgorithmen, darunter eine dreidimensionale Ausgleichsrechnung zur räumlichen Trenddarstellung, bauen aus einem Satz von unregelmäßig verteilten Punkten ein dreidimensionales Modell auf, das durch Linien gleicher Intensität sichtbar gemacht wird (Bild 2).

Choroplethenkarten werden verwendet, um einen Sachverhalt pro abgegrenzte Flächeneinheit zu zeigen, z. B. die Bevölkerungsdichte einer Ortschaft. Die Berechnung von Klassengrenzen geschieht automatisch, es können aber auch Schwellenwerte vorgegeben werden.

Proximitätskarten dienen zur Darstellung von Nachbarschaftsbeziehungen aller Art, wie Entfernung, Dichte und Zentralität (Bild 3).

Durch mehrmaliges Drucken auf die gleiche Zeile erreicht man unterschiedliche Schwärzungsgrade, was die Lesbarkeit einer Karte verbessert. Auch Versuche für farbige Karten mittels SYMAP sind schon erfolgreich durchgeführt worden, der allgemeinen Anwendung stehen

allerdings die erhöhten Kosten entgegen. Das Programm enthält noch sehr viele andere Möglichkeiten, wie Einfügen von Legenden, Häufigkeitsdiagramme, Ausplanimetrieren usw., deren Beschreibung zu weit führen würde. Schwierig ist bisher noch, insbesondere für deutsche Verhältnisse, die Digitalisierung der Grundkarte. Ein automatischer Digitizer, der das Grundmuster der Karte, wie Gemeindegrenzen oder Meßpunkte, in eine numerische, Computer-lesbare Form bringt, ist selten zugänglich. Wenn aber von einer Grundkarte viele thematische Karten produziert werden müssen, ist einmaliges manuelles Kodieren immer noch wirtschaftlicher als vielfaches manuelles Kartenzeichnen.

2.2. Perspektivische Darstellungen

Eine Karte dient in den meisten Fällen dazu, dreidimensionale Sachverhalte in eine zweidimensionale Form zu bringen: Bei Isolinien werden Punkte gleicher Intensität verbunden, bei Choroplethen Häufigkeitsklassen durch Signaturen (Grauwerte, Farben, Symbole) dargestellt. Das "Lesen" einer Karte ist daher ein Transformationsprozeß, nämlich die Übersetzung von zwei Dimensionen in drei Dimensionen, der ein gewisses Maß an Erfahrung voraussetzt. So gelingt es nicht jedem, sich bei Betrachtung einer Isolinienkarte die wahre Oberfläche vorzustellen.

Die Alternativen, Blockbilder und andere perspektivische Darstellungen, sind bei manueller Herstellung mit sehr großem Aufwand verbunden. Hier beschleunigt eine DVA die Arbeit sehr, weil die Koordinatentransformation numerisch und mit sehr vielen Punkten durchgeführt werden kann. Wenn die Größe der Projektionswinkel variabel ist wie beim Programm SYMVU (aus

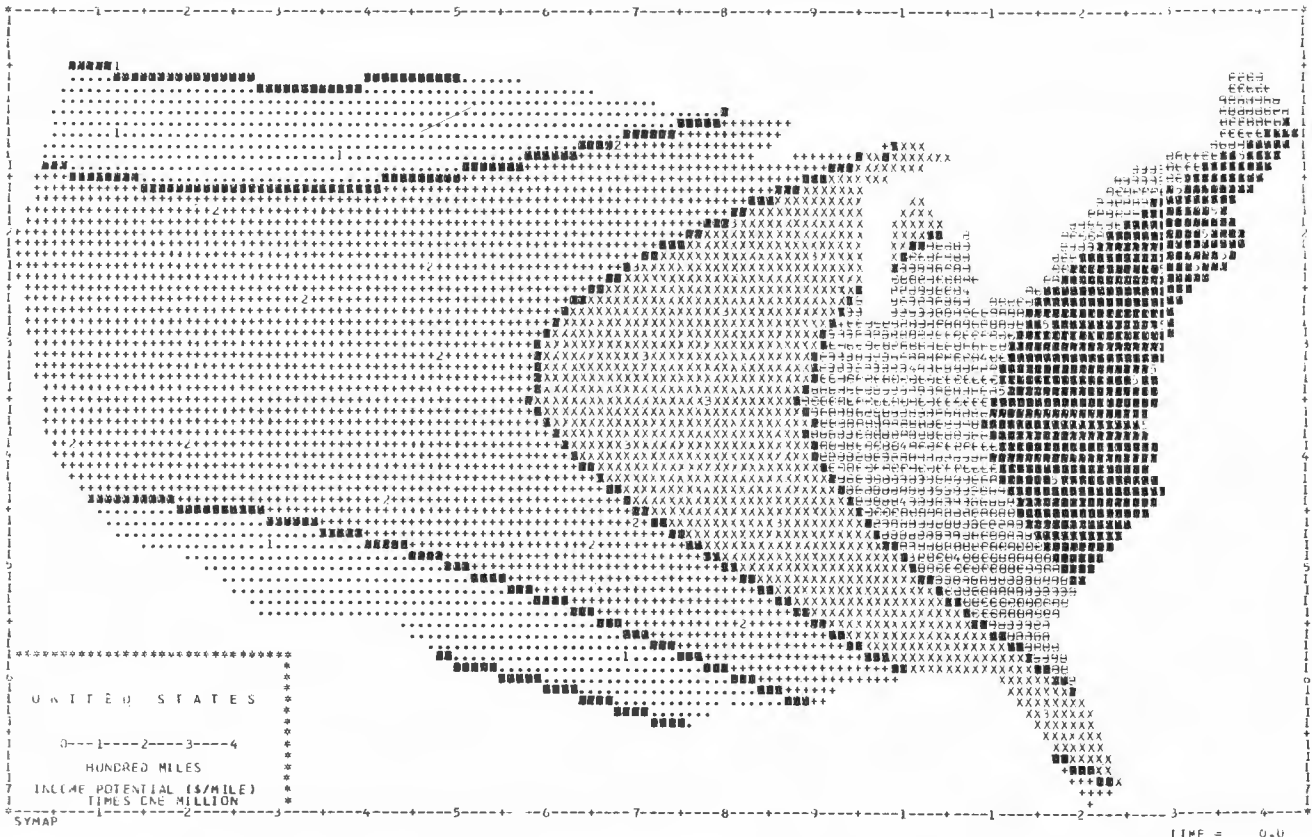


Bild 2 SYMAP-Isolinien-Karte (Einkommenspotential der USA)

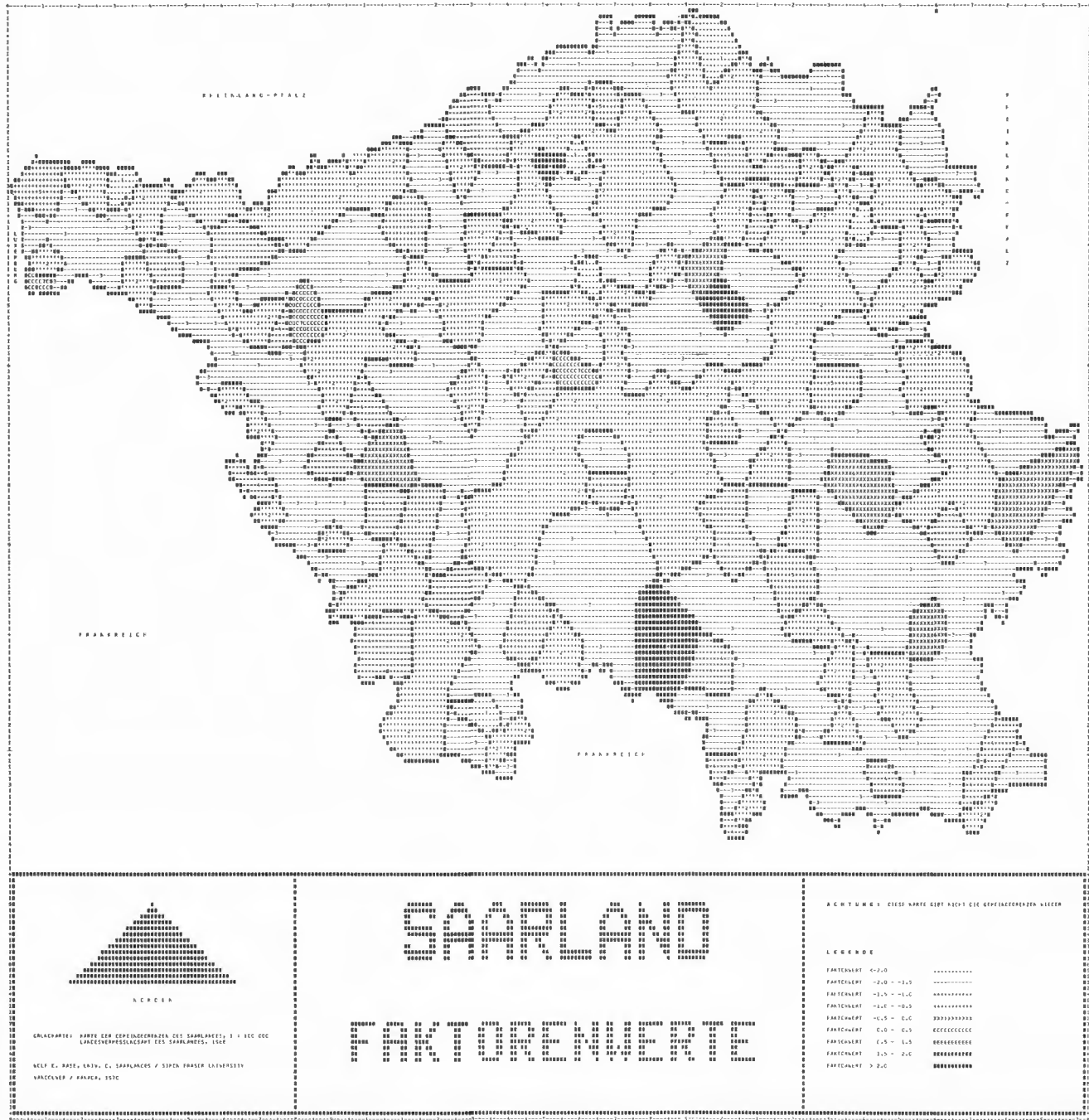


Bild 3 Proximitätskarte von Gemeindetypen im Saarland mit zehn Graustufen und Legende

der SYMAP-Familie), kann man praktisch einen Rundflug über der Fläche simulieren. In SYMVU ist auch das Problem der verdeckten Linien gut gelöst, was für den räumlichen Eindruck des Blockbildes sehr wichtig ist. Das Programm enthält noch andere Möglichkeiten, die den praktischen Gebrauch erleichtern, wie Glättung der Oberfläche, verschiedene Projektionsverfahren, Einfügen von Punkten auf der Oberfläche usw. (Bild 4). Das Analogon zur Choroplethenkarte ist ein räumliches Histogramm: Über Grundflächen, z. B. Gemeindeflächen, repräsentieren Säulen verschiedener Höhe einen Wert, etwa Bevölkerungsdichte oder Baulandpreise. Auch diese Darstellungsart ist in SYMVU vorhanden.

OCTOBER 1969
 AZIMUTH = 31
 ALTITUDE = 30
 WIDTH = 2,00
 SMOOTHING = -1

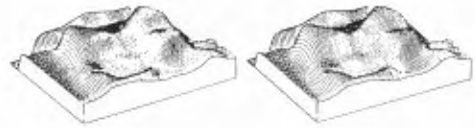


Bild 4 Stereo-Blockbildpaar einer Oberfläche; Programm SYMVU, computergesteuertes x-y-Zeichengerät. Durch ein geeignetes Stereoskop betrachtet, wirkt das Bild plastisch.

THEORETICAL SURFACE, PITS, PEAKS AND PALES

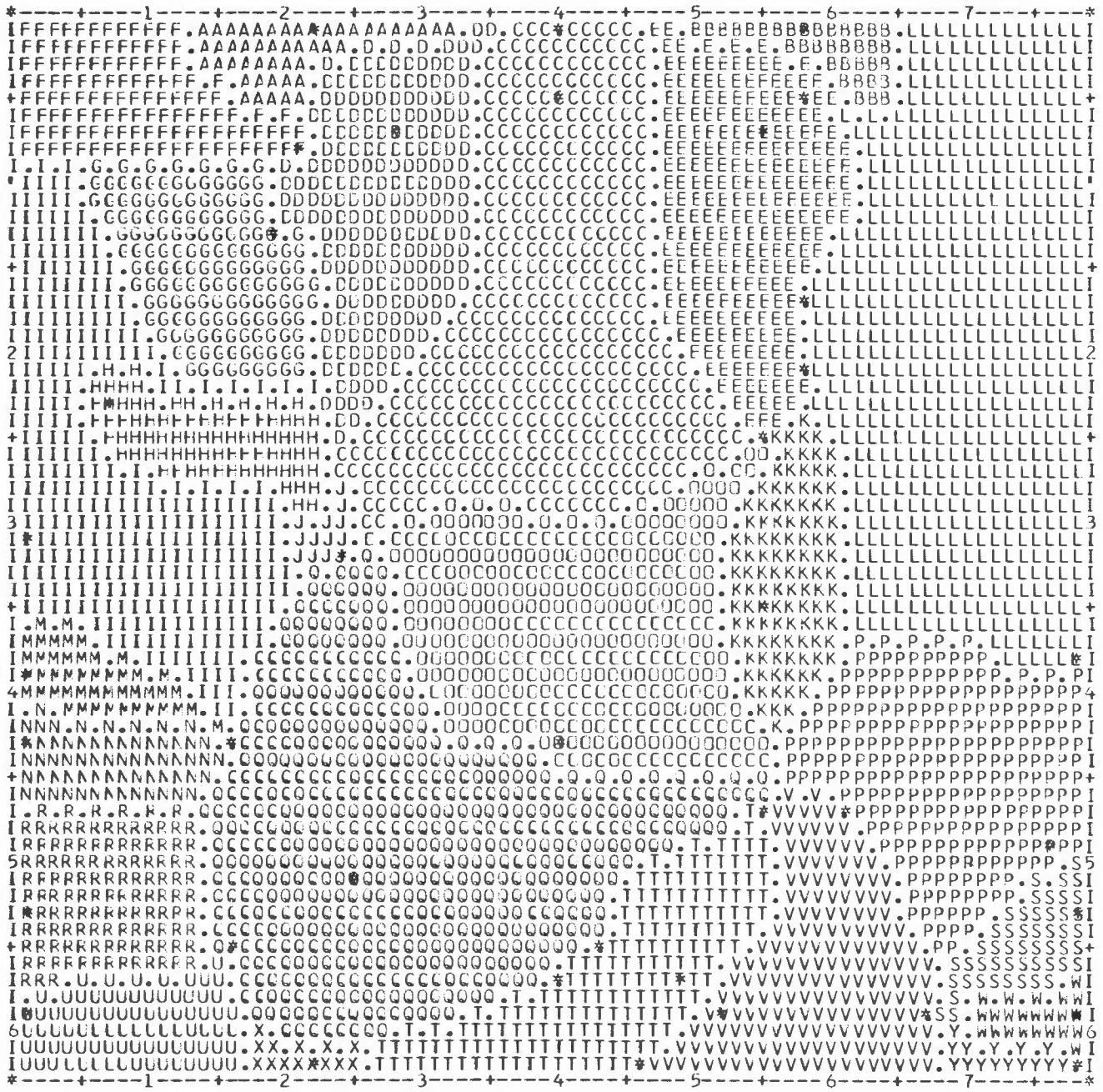


Bild 5 Analyse einer Oberfläche (die gleiche wie in Bild 4) mit WATERSHED in Bezug auf Täler. Die mit einem Stern überdruckten Buchstaben repräsentieren den tiefsten Punkt eines Beckens (gleiche Buchstaben), das von Rückenlinien (Punkte) eingefaßt wird. Die mit einem Stern überdruckten Punkte sind die Gipfel

--DALE--
PITS OVERPRINTED WITH *
PKS SHOWN BY *
RDGE LINES SHOWN BY .

2.3. Automatische Analyse von Oberflächen

Ein Problem der Computer-Kartographie ist die optimale Speicherung der zu kartierenden Oberflächen. Eine Möglichkeit ist (sie wird bisher am meisten verwendet), die Karte zu rastern und die Höhe jedes Rasterpunktes zu speichern. Das ergibt ein sehr genaues Bild, erfordert allerdings einen sehr hohen Speicheraufwand und lange Verarbeitungszeiten. Eine 50 x 50 cm-Karte, in 1 mm-Abständen gerastert, bestände aus 250 000 Punkten. Bei der zweiten Möglichkeit, einzelne Isolinien nach x-y-Koordinaten abzuspeichern, könnte die Anzahl

der Punkte verringert werden, aber man müßte größere Ungenauigkeiten in Kauf nehmen. Die dritte Möglichkeit besteht darin, die Oberfläche nach ihren Eigenschaften zu analysieren und diese Eigenschaften zu speichern, aus denen das Original mit großer Zuverlässigkeit wieder rekonstruiert werden kann [6]. So sucht das Programm WATERSHED (auch aus der SYMAP-Familie) die Gipfel, Täler, Rücken und Pässe aus einer gerastert-digitalisierten Oberfläche (in den meisten Fällen ein Ausgabe-Datensatz von SYMAP), bestimmt die Hang-

THEORETICAL SURFACE, PITS, PEAKS AND PALES

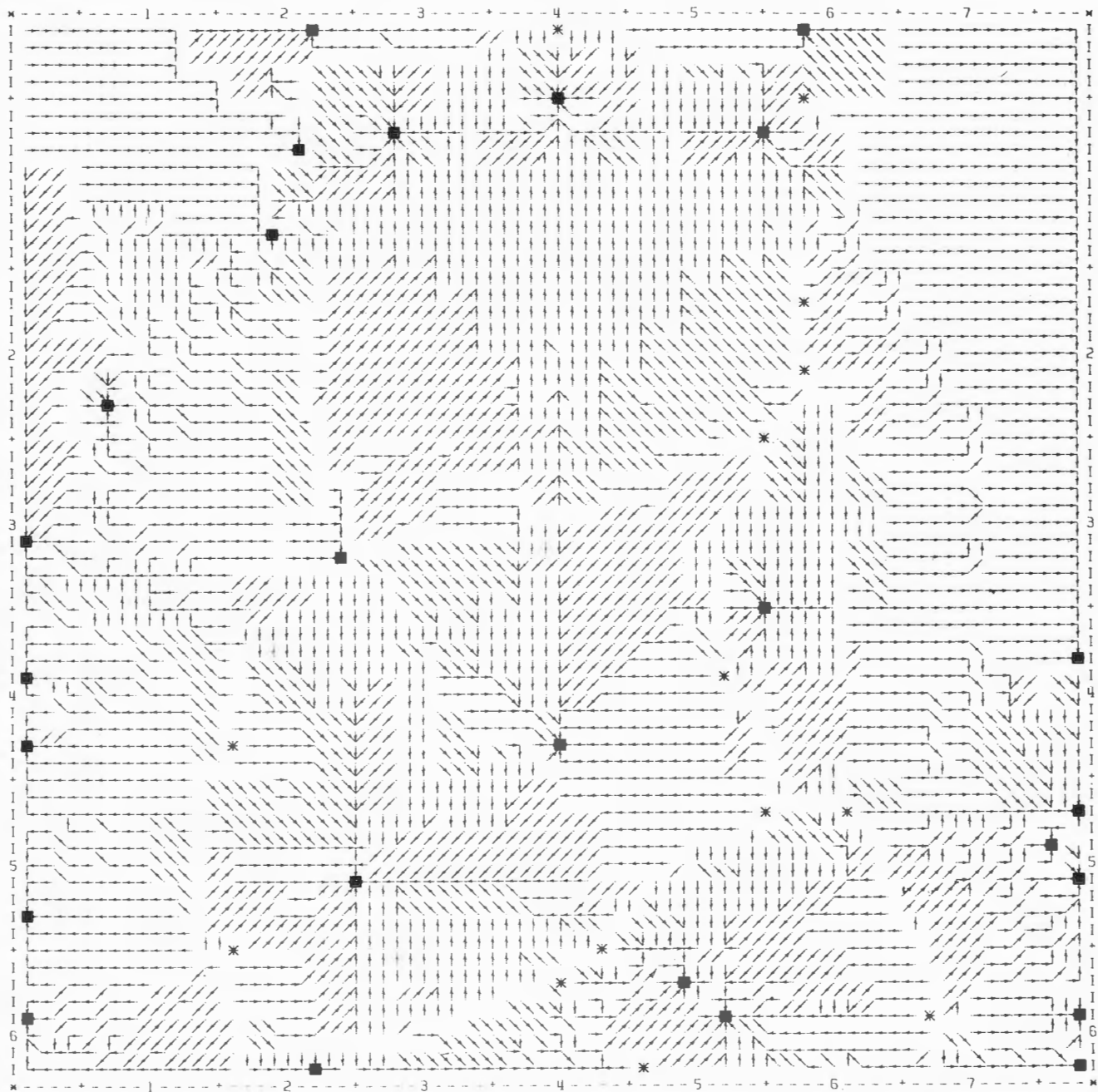


Bild 6 Darstellung der Hangexposition und -neigungsrichtung mittels x-y-Zeichengerät; gleiche Oberfläche und gleiches Programm wie in Bild 5. Die schwarzen Quadrate entsprechen den tiefsten Punkten, die Sterne den Gipfeln

neigungswinkel und die Exposition. Diese Eigenschaften können dann auf dem Ausgabedruker oder einem x-y-Plotter graphisch fargestellt oder auf einem Magnetband zur späteren Verwendung gespeichert werden (Bild 5 und 6).

Diese Technik ist nicht nur für Kartographen von Interesse, weil geomorphologische Großformen, wie etwa Wasserscheiden, automatisch erkannt werden können; auch die Berechnung von Neigungswinkeln und Expositionen kann für den physischen Geographen sehr hilfreich sein.

3. Statistisch-taxonomische Fragestellungen

Im statistischen Sinne ist der Forschungsgegenstand der Geographie ein n-dimensionaler Raum, der von n nichtorthogonalen Vektoren, repräsentiert durch die einzelnen Geoelemente, aufgespannt wird. Schon in den zwei Dimensionen der Erdoberfläche lassen sich mit den Methoden multivariater Statistik geographische Eigenschaften feststellen, etwa räumliche Mittel, Mediane und Schwerpunkte, Nachbarschafts- und Wegebeziehungen, etwa der kürzeste Weg von einem beliebigen Punkt zu einem anderen, kürzeste Rundreise usw.

Eine verhältnismäßig breite Anwendung hat ein Verfahren bivariater Statistik gefunden, die "Trend Surface Analysis": Durch eine Anzahl von Punkten, die durch ihre Lage auf der Erdoberfläche und einen weiteren Wert, etwa topographische Höhe oder auch Bevölkerungsdichte, bestimmt sind, wird eine Fläche (meist ein Polynom n-ter Ordnung) nach dem Kleinste-Quadrate-Kriterium berechnet. Aus der kartierten Fläche läßt sich dann ein räumlicher Trend ablesen, etwa die Abnahme der Bevölkerungsdichte mit der Entfernung von der Großstadt.

Rein formal kann man den geographischen Raum in Form einer Matrix darstellen, bei der die Örtlichkeiten die Spalten und die Geoelemente die Zeilen bilden:

Element	1	2	3	4	5	6
Ort 1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}
2	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}
3	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}
4	x_{41}	x_{42}	x_{43}	x_{44}	x_{45}	x_{46}
5	x_{51}	x_{52}	x_{53}	x_{54}	x_{55}	x_{56}

Die Aufgabe für den Wissenschaftler besteht nun darin, bestimmte Eigenschaften des Systems zu finden, und das geschieht meistens durch Klassifikation. Zwei verschiedene Arten lassen sich unterscheiden, einmal die Gruppenbildung in Richtung der Spalten (R-mode cluster analysis), bei der Variablen zu Faktoren zusammengefaßt werden, zum anderen in Richtung der Zeilen, indem man Lokalitäten ähnlicher Eigenschaften in Klassen, Typen oder Regionen gliedert (Q-mode cluster analysis).

3.1. Faktorenanalyse

Da sehr viele Variablen die Eigenschaft einer Örtlichkeit darstellen, ist es angebracht, den n-dimensionalen Variablenraum auf m Faktoren zu reduzieren, indem man die Variablen mit den höchsten Korrelationskoeffizienten zu einem neuen Faktor zusammenzieht. Die Beziehungen zwischen Variablen und Faktoren werden durch die Faktorladungen ausgedrückt, die von -1. bis +1. reichen: Die Ladung -1.0 oder +1.0 sagt aus, daß der Informationsgehalt der Variablen völlig im neuen Faktor enthalten ist, 0. bedeutet keinerlei Anteil der Variablen am betreffenden Faktor. Geometrisch ist die Faktorladung der Cosinus des Winkels zwischen den Vektoren der Variablen und des Faktors (Bild 7). Faktorenanalytische Verfahren machen den Wirkungszusammenhang zwischen verschiedenen geographischen Variablen quantifizierbar. Z. B. besteht in einem bestimmten Gebiet eine hohe Korrelation zwischen der Zahl der Beschäftigten in der Landwirtschaft, Selbständigen und mithelfenden Familienangehörigen oder zwischen der Anzahl der Beschäftigten im produzierenden Gewerbe und dem Prozentsatz der Arbeiter.

Durch Transformation der ursprünglichen Variablen auf die neuen Faktoren wird das Beziehungssystem orthogonal, d. h. die Faktorenachsen stehen senkrecht aufeinander und sind nicht miteinander korreliert. Das erleichtert, neben der Reduktion der Variablen auf eine angemessene Anzahl von Faktoren, die Klassifikation der Objekte. Der numerische Ausdruck des Faktorenräume in Bezug auf die Objekte, also die Umformung der ursprünglichen Variablenwerte auf den Faktorenräum, sind die Faktorenwerte [8, 9].

Da faktorenanalytische Methoden schon lange für medizinische und psychologische Untersuchungen benutzt werden, sind eine Reihe von Computerprogrammen für diesen Zweck verfügbar, die von wissenschaftlichen Institutionen und Computerherstellern entwickelt wurden.

Es muß aber auch vor unüberlegter Anwendung der Faktorenanalyse in der geographischen Wissenschaft gewarnt werden. Im amerikanischen Bereich sind Arbeiten erschienen, die auf Faktorenanalyse begründet sind, wo aber die Autoren rein mechanisch eine Technik adaptiert und ziemlich grobe Fehlschlüsse aus den berechneten Zahlen gezogen haben. Damit erweist man den quantitativen Methoden, die ohnehin um ihre Anerkennung in der Geographie ringen müssen, einen Bärendienst.

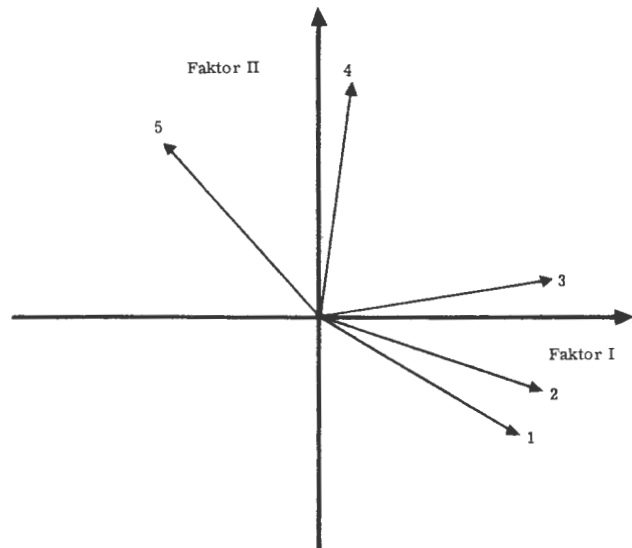


Bild 7 Geometrische Darstellung einer Faktorenstruktur mit zwei Faktoren. Die Variablen 1 bis 3 haben hohe Ladungen auf Faktor I, d. h. die Winkel zwischen Variablen- und Faktorenvektoren sind relativ klein. Die Variablen 4 und 5 sind sehr hoch mit Faktor II korreliert.

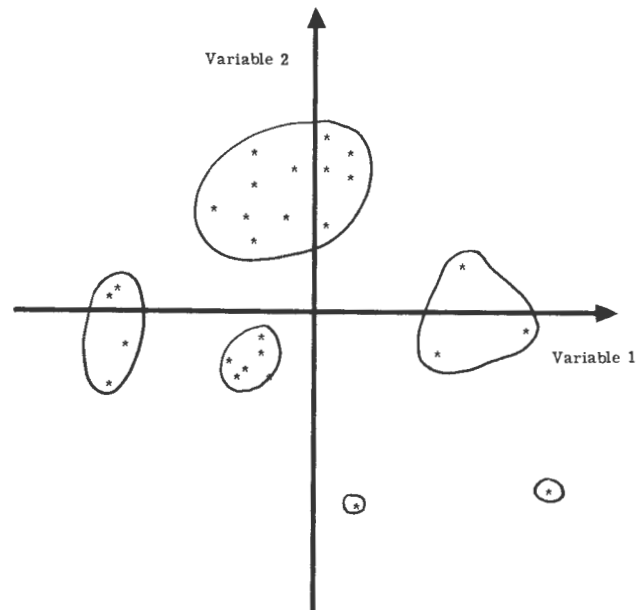


Bild 8 Klassifikation einer Probe mit zwei Variablen. Benachbarte Punkte werden zu Gruppen (umrandet) zusammengefaßt.

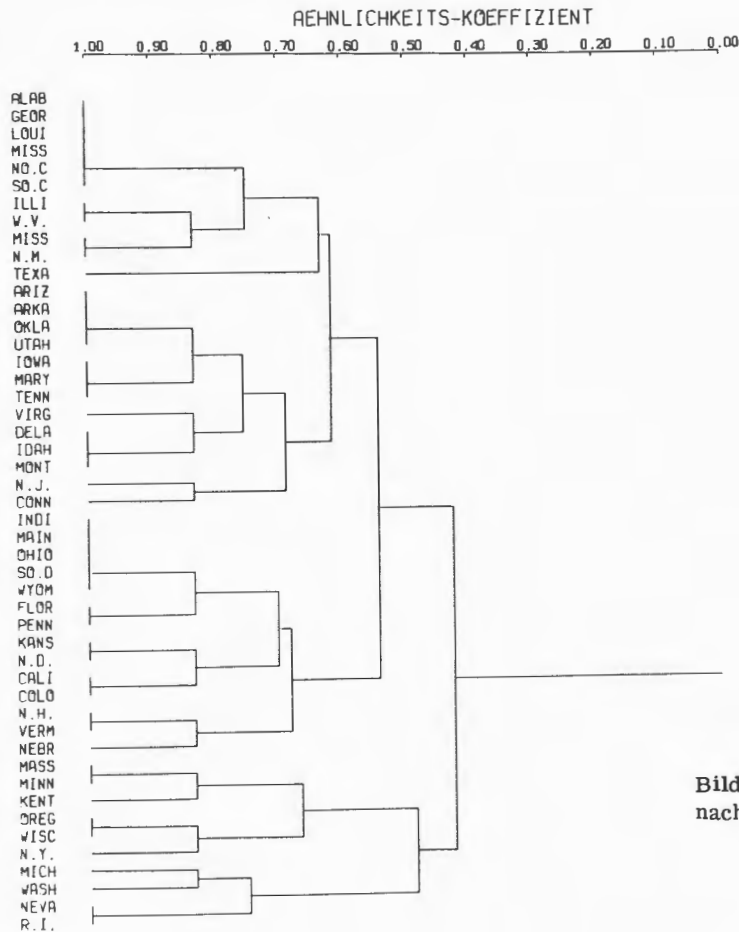


Bild 9 Dendrogramm einer Klassifikation der US-Staaten nach qualitativen Merkmalen.

3.2. Geographische Klassifikation

Eine Klassifikation ist die Zusammenfassung funktional ähnlicher Objekte in Gruppen. Im m -dimensionalen orthogonalen Raum, der von den Variablen oder Faktoren der geographischen Örtlichkeiten aufgespannt wird, läßt sich die Ähnlichkeit als Entfernung der jeweiligen Raumpunkte angeben. Die Punkte werden nun aufgrund eines Ähnlichkeitskoeffizienten, der aus den m -dimensionalen Distanzen errechnet wird, gegliedert (Bild 8) [5].

Diese Klassifikationsmethode hat den Vorteil, daß nicht die vorgegebenen Grenz- und Schwellenwerte für bestimmte Variablen ausschlaggebend sind, sondern daß die Lokalitäten allein aufgrund ihrer Nachbarschaftsbeziehungen im gesamten Variablen- oder Faktorenräum klassifiziert werden. Eine solche Klassifikation entspricht damit viel besser der komplexen Struktur der geographischen Wirklichkeit.

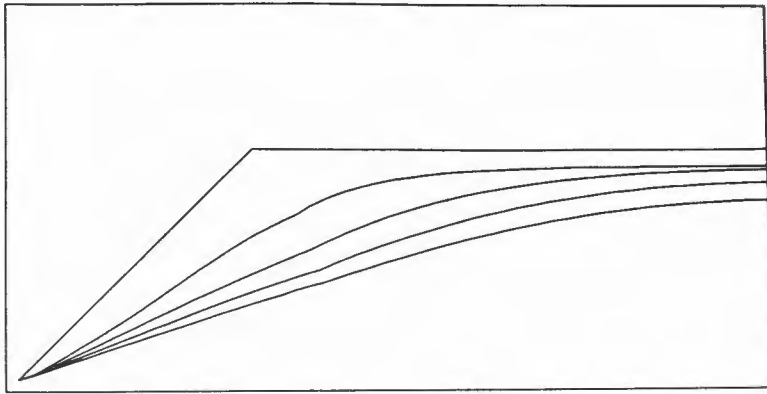
Qualitative Daten, z. B. Anwesenheit einer bestimmten Pflanze in biogeographischen Einheiten, können ebenfalls mit Ähnlichkeitskoeffizienten quantifiziert und dann einer Clusteranalyse unterworfen werden [4]. Das eröffnet neue Perspektiven für die naturräumliche Gliederung, wo die Erfassung der multivariaten Korrelationen zwischen den Geofaktoren eine dominierende Rolle spielt.

Durch Anwendung verschiedener statistischer Prüfmethoden (z. B. Diskriminanzanalyse, Bayessche Formel) läßt sich eine Klassifikation auf iterativem Wege noch weiter verbessern [4].

4. Simulation

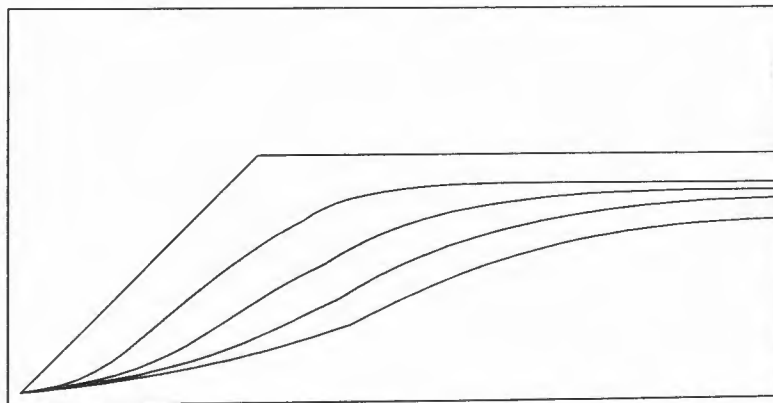
Anders als in den Naturwissenschaften lassen sich in der Geographie die natürlichen Vorgänge nur selten als Experimente im Labor untersuchen, weil die geographische Wirklichkeit zu komplex und die Zeitabläufe sehr langsam sind. Eine Datenverarbeitungsanlage gestattet nun, ein Modell hohen Komplexitätsgrades aufzubauen und die Vorgänge in diesem Modell zeitlich beschleunigt und mehrfach unter veränderten Ausgangsbedingungen ablaufen zu lassen.

Ein anschauliches Beispiel ist die Simulation der Hangdenudation von AHNERT [1]. Durch ein Computerprogramm werden die Vorgänge bei der Abtragung eines Hanges aus anstehendem Fels, hauptsächlich Verwitterung und Abrutschung, simuliert. Je nach Gesteinswiderständigkeit, Vorherrschen von chemischer oder physikalischer Verwitterung und Schichtlagerung ergeben sich unterschiedliche Hangformen (Bild 10).



DENUATIONSMODELL NR. 1

LINEARE VERWITTERUNG



DENUATIONSMODELL NR. 3

EXPONENTIELLE VERWITTERUNG

Bild 10
 Plotter-Ausgabe zweier Denudations-
 Simulationen mit verschiedenen
 Ausgangsvariablen

Diese Technik ließe vielleicht bei Übertragung in drei Dimensionen Schlüsse auf umstrittene geomorphologische Phänomene wie Schichtstufen und Rumpfflächen zu.

Ein schon beinahe klassisches Modell ist die Simulation von Innovationerscheinungen von HÄGERSTRAND; weiterhin wären die Standortmodelle der Wirtschaftsgeographie zu erwähnen.

Literatur

- [1] F. Ahnert: Zur Rolle der Elektronischen Rechenmaschine und des mathematischen Modells in der Geomorphologie. Geogr. Zeitschr. 54 (1966), S. 118-122
- [2] C. Board: Maps as Models, in: Chorley, Hagget, Physical and Information Models in Geography, London 1968, S. 671-725
- [3] G. F. Bonham-Carter: FORTRAN IV Program for Q-Mode Cluster Analysis of Nonquantitative Data Using IBM 7090/7094 Computers. Computer Contribution 17, State Geological Survey, The University of Kansas, Lawrence 1967
- [4] F. Demirmen: Multivariate Procedures and FORTRAN IV Program for Evaluation and Improvement of Classifications. Computer Contribution 31, State Geological Survey, The University of Kansas, Lawrence 1969
- [5] J. W. Harbaugh, D. F. Merriam: Computer Applications in Stratigraphic Analysis. New York 1968
- [6] Th. Peucker: Some thoughts on optimal mapping and coding of surfaces. Paper delivered at the CAG Meeting, St. John's, Newfoundland, August 1968
- [7] Reference Manual for Synagraphic Computer Mapping "SYMAP", Version V, Draft 2, 1968
- [8] D. Steiner: Die Faktorenanalyse, ein modernes statistisches Hilfsmittel des Geographen für die objektive Raumgliederung und Typenbildung. Geographica Helvetica 20 (1965), S. 20-34
- [9] K. Überla: Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler. Berlin 1968