

Geografisches Institut Universität Basel

Elektronische Datenverarbeitung in der Physiogeografie

Allgemeine Konzepte und Werkzeuge für Studium und Forschung
sowie ein Entwurf für die computergestützte Erarbeitung der
geoökologischen Karte 1 : 25 000

Diplomarbeit

Martin Huber

Geografie, Botanik, Mathematik, Bodenkunde, Informatik

8. Semester

Leitung: Prof. Dr. H. Leser

Basel 1990

Inhaltsverzeichnis

Inhaltsverzeichnis	2
Abbildungsverzeichnis	4
Tabellenverzeichnis	5
1. Einleitung	7
2. Computereinsatz im Geografiestudium	9
2.1 Weshalb soll der Computer im Geografiestudium eingesetzt werden?	9
2.2 Wie kann der Computer im Geografiestudium eingesetzt werden?	9
2.2.1 Vorlesungsaufbereitung, Anfängerpraktika	9
2.2.2 Seminararbeiten	12
2.2.3 Geländepraktika	13
2.3 Ein Beispiel: Literaturverwaltungsprogramm	15
2.3.1 Installation	15
2.3.2 Allgemeine Angaben zur Bedienung	16
2.3.3 Titel eingeben	18
2.3.4 Stichworte eingeben	21
2.3.5 Datenbank-Abfrage	25
2.3.6 Schnellausdruckverfahren	26
2.3.7 Titel ändern	28
2.3.8 Spezialfunktionen	29
2.3.9 Erweiterungsmöglichkeiten des Literaturverwaltungsprogramms	29
3. Computereinsatz in der physiogeografischen Forschung	34
3.1 Vorschlag für ein Datenverwaltungskonzept in der geoökologischen Forschung	34
3.1.1 Entflechtung von Daten und Datenauswertung	34
3.1.2 Grundbegriffe der Datenmodellierung	35
3.1.3 Datenanalyse	37
3.1.4 Datenkonzept für die geoökologische Forschung	39
3.2 Datenauswertung	47
3.3 Ein Beispiel: Bodendatenverwaltungsprogramm	50
3.3.1 Installation	50
3.3.2 Programmübersicht	50
3.3.3 Neue Datenbank anlegen	52
3.3.4 Bohrprofile eingeben	53
3.3.5 Grubenprofile eingeben	55
3.3.6 Bodendaten für die GÖK 25 eingeben	56

3.3.7	Labordaten eingeben	56
3.3.8	Daten ändern	58
3.3.9	Daten abfragen	59
3.3.10	Profil drucken	61
3.3.11	Labordaten exportieren und darstellen	61
3.3.12	Drucken ohne Bildschirmanzeige	65
3.3.13	Export in Fremdformate	65
3.3.14	Weitere Verarbeitungsmöglichkeiten	66
3.3.15	Datenstruktur	66
3.3.16	Protokolle und Symboltabellen	75
4.	Anmerkungen zu einer digitalen geoökologischen Karte 1 : 25 000	91
4.1	Datenerhebung und Datenspeicherung	91
4.2	Datenauswertung und Datendarstellung	93
5.	Zusammenfassung	94
6.	Quellen	95
6.1	Zitierte Literatur	95
6.2	Bei der Programmierung verwendete Literatur	97
6.3	Für die Arbeit verwendete Software	98
Anhang A: Programmtext Literaturverwaltungsprogramm		
Anhang B: Programmtext Bodendatenverwaltungsprogramm		

1. Einleitung

Die vorliegende Diplomarbeit folgt dem Bedürfnis, die am Geographischen Institut der Universität Basel vorhandenen EDV-Einrichtungen effizienter anzuwenden, als dies bisher getan wurde. Probleme gab es vor allem deshalb, weil kaum jemand Zeit für die Installation und den Unterhalt der Geräte hatte. Die Maschinen liefen mit unterschiedlichen Betriebssystemversionen, die Programmparameter waren nicht überall gleich eingestellt, und die Drucker lieferten unerwartete Ergebnisse, wenn eine Datei auf verschiedenen Computern bearbeitet worden war. Konzeptlos wurden Messresultate mal mit diesem, mal mit jenem Programm ausgewertet. War eine Dissertation oder eine Diplomarbeit abgeschlossen, konnten die dabei erhobenen Daten nur mit grösster Mühe, meist aber gar nicht, weiterverwendet werden. An eine Modellrechnung oder Simulation aufgrund von langjährigen Messreihen konnte nicht gedacht werden.

Eine im Juni 1988 vom Autor geleitete Grundsatzdiskussion zum Thema "Datenverarbeitung am Geographischen Institut" führte zu einem Katalog von dringenden Massnahmen zur Verbesserung der EDV-Infrastruktur und -Organisation. Eine grosse Änderung trat jedoch nicht ein, sei es, weil sich die Beteiligten zu wenig für die Datenverarbeitung interessierten, oder weil die Vorteile eines einheitlichen Systems bei der Datenauswertung nicht erkannt wurden.

Diese Arbeit soll dazu anregen, Lösungen im Informatikbereich gemeinsam zu suchen. Das Ziel ist, dass nach einem allgemeinen Konzept gearbeitet wird, so dass mit der Zeit ein Sortiment von spezifischen Datenauswertungswerkzeugen erstellt werden kann. Die Auswertung wird dadurch vereinfacht, und es bleibt mehr Zeit für die Dateninterpretation übrig.

Der erste Teil der Arbeit behandelt den Computereinsatz im Geografiestudium. Die diversen Seminare und Praktika, die im Verlaufe des Studiums absolviert werden müssen, geben Gelegenheit, den Umgang mit dem Computer einzuüben. Die heute verfügbaren Programme für Textverarbeitung, Statistik, Datenverwaltung oder Grafik sind umfangreiche Hilfsmittel, deren Bedienung eine gewisse Einarbeitungszeit erfordert. Je früher die Studierenden mit diesen Werkzeugen vertraut werden, desto selbstverständlicher werden sie sie einsetzen. Es geht keinesfalls darum, alles mit dem Computer zu erledigen. Aus der Kenntnis der Möglichkeiten heraus kann jedoch einfacher entschieden werden, welche Arbeitsweise am schnellsten zum gewünschten Ziel führt.

Im zweiten Teil wird die Computeranwendung in der physiogeografischen Forschung angesprochen. Es wird ein Datenverwaltungsmodell vorgestellt, mit dem bereits bei der Konzipierung der Feldkampagne methodische Fehler erkannt und vermieden werden können. Während der Datenerfassung können verschiedene Tests Aufschluss geben über Plausibilität der Messresultate und Repräsentativität der Messstandorte. Bei der Auswertung und Interpretation von Messreihen schliesslich geben klar strukturierte Daten den nötigen Überblick auch über grosse Mengen von Daten. Am Beispiel eines Programms zur Verwaltung von Bodenkartierungsdaten soll gezeigt werden, wie das Modell praktisch eingesetzt werden kann.

Der letzte Teil der Arbeit behandelt die Umsetzung der "Kartieranleitung Geoökologische Karte 1 : 25 000 (KA GÖK 25)" (Leser & Klink 1988) zu einem computergestützten Planungsinstrument. Hier ist ein klar strukturiertes Datenmodell von grösster Bedeutung, denn die Daten müssen jederzeit abrufbar und ohne grosses Hintergrundwissen verständlich sein. Die Kartieranleitung muss in einigen Punkten an die elektronische Datenverarbeitung angepasst werden, sie ist aber von ihrer Struktur her gut geeignet für die Übertragung auf den Computer. Das praktische Beispiel einer digitalen GÖK 25 fehlt noch. Einige Vorbereitungen dazu sind bereits abgeschlossen, sie werden aber nicht im Rahmen dieser Arbeit behandelt. Die Darstellung des Konzeptes der digitalen GÖK 25 soll klar werden lassen, dass geowissenschaftliche Daten nicht nur für die einzelne Forscherin oder den Forscher im Rahmen einer Dissertation oder Diplomarbeit Bedeutung haben, sondern dass sie auch in der Planung eingesetzt werden können. Was es dazu braucht, ist ein sauberes Datenkonzept und eine ausreichende Dokumentation der Daten.

Als eigentliches Kernstück dieser Diplomarbeit wurden ein Literatur- und ein Bodendatenverwaltungsprogramm entwickelt. Diese Programme basieren auf "dBase IV" und werden in den Kapiteln 2.3 und 3.3 dokumentiert. Der Programmcode ist im Anhang abgedruckt. Alle Programmdateien sind auf den beigelegten Disketten gespeichert und können frei kopiert werden.

Obwohl diese Arbeit die EDV zum Thema hat, soll sie nicht dazu verführen, den Computer zu einem Hauptinhalt der Geografie zu machen. Beim Computer handelt es sich wie bei einem Pürckhauer oder einem Windschreiber um ein Werkzeug, das – richtig eingesetzt – einiges zu leisten vermag. Beim falschen Gebrauch entsteht viel Ärger und die Arbeit wird eher verzögert als beschleunigt. Es geht also darum, um mit einem Bild zu sprechen, das Messer zum Schneiden am Griff und nicht an der Klinge zu fassen.

3. Computereinsatz in der physiogeografischen Forschung

Die Datenbeschaffung in Form von Feld- und Laborarbeit verschlingt bei der physiogeografischen Forschung den grössten Teil der zur Verfügung stehenden Zeit. Es kann also mit Recht behauptet werden, dass physiogeografische Daten, wie sie z.B. bei der landschaftsökologischen Komplexanalyse anfallen, teuer sind. Damit sich dieser Aufwand lohnt, müssen die Daten sinnvoll und effizient ausgewertet werden können. Mit Vorteil wird deshalb bereits bei der Planung einer Untersuchung ein Konzept für die Datenverwaltung erstellt.

Für die weiteren Ausführungen wird vorausgesetzt, dass die Daten mit EDV-Hilfsmitteln verwaltet werden. Als erstes wird ein Datenverwaltungskonzept für die geoökologische Forschung vorgeschlagen. Danach werden ein paar Gedanken zur Datenauswertung geäußert und zuletzt soll ein praktisches Beispiel das Behandelte vertiefen. Das vorgeschlagene Konzept ist zwar auf die Geoökologie ausgerichtet, es kann aber ohne weiteres auf andere Forschungsbereiche mit ähnlicher Datenlage übertragen werden.

3.1 Vorschlag für ein Datenverwaltungskonzept in der geoökologischen Forschung

Zu Beginn einer geoökologischen Feldaufnahme ist die Art und Menge der anfallenden Daten noch nicht vollumfänglich bekannt, es kann deshalb auch noch nicht bis ins Detail vorausgesagt werden, wie die Daten weiterverarbeitet werden. Die Verwaltung der Daten muss also folgenden drei Punkten Rechnung tragen: Sie muss offen sein für unvorhergesehene Daten (z.B. Einsatz eines neuen Messgerätes), sie muss bereits während der Messphase eine Datenauswertung ermöglichen und sie muss jede Weiterverarbeitung (Statistik, Grafik, Simulation) unterstützen. Dass diese Ansprüche mit den Mitteln der modernen Informationstechnik erfüllt werden können, soll hier gezeigt werden.

3.1.1 Entflechtung von Daten und Datenauswertung

In der Regel sieht der Computeranwender eine enge Verflechtung der Daten mit ihrer Auswertung: Mit einem Textverarbeitungsprogramm werden Textdateien erstellt, die mit dem gleichen Programm weiterverarbeitet und ausgedruckt werden können. In ein Statistikprogramm müssen zuerst Daten eingetippt werden, bevor sie analysiert werden können, und auch das selbstprogrammierte Messdatenauswertungsprogramm muss mit den Messdaten gefüttert werden, damit weitere Werte berechnet und die Resultate dargestellt werden können. Bei kleinen Problemen ist diese Verknüpfung der Datenverwaltung mit der Datenauswertung gerechtfertigt, weil das gestellte Problem meist ohne weiteres gelöst werden kann. Wenn aber für die Verarbeitung der Daten mehrere Programme verwendet werden müssen, führt die anwendungsorientierte Datenhaltung in eine Sackgasse: Die eine Anwendung kann mit den Daten der anderen nichts anfangen, so dass die gleichen Werte mehrmals in verschiedenen Formen eingegeben werden müssen. Es muss deshalb ein allgemeines, logisches Datenmodell gefunden werden. Das ist nur möglich, wenn Daten und Datenzugriff entflechtet werden (Zehnder 1987, 61). So erhalten wir den Leitsatz für die weitere Arbeit:

Die Daten sind unabhängig von ihrer Auswertung zu betrachten.

Mit den theoretischen Modellen, die hinter der Idee der relationalen Datenbank stehen, kann dieser Grundsatz verwirklicht werden (Relationenmodell von Codd 1970, Entity-Relationship-Modell von Chen 1976, siehe Zehnder 1987, 57ff). In der Praxis wird zuerst eine Datenanalyse nach diesen Modellen durchgeführt, worauf die dabei erzielten Resultate auf einem relationalen Datenbanksystem umgesetzt werden. Die Daten, die in diese Datenbank eingegeben werden, können mit einer Datenabfragesprache in jede für die Weiterverarbeitung gewünschte Form gebracht werden.

3.1.2 Grundbegriffe der Datenmodellierung

Damit das Relationenmodell und das Entity-Relationship-Modell angewendet werden können, müssen einige Grundbegriffe erklärt werden. Die hier verwendeten Definitionen stützen sich auf Zehnder (1987) und Vetter (1990).

Ein individuelles und identifizierbares Exemplar von Dingen, Personen oder Begriffen der realen oder der Vorstellungswelt wird als **Entität** (Entity) bezeichnet. Eine Entität könnte zum Beispiel der am 1.8.1990 am Messstandort X gemessene Temperaturwert von 21°C oder das am Standort Y aufgenommene Bodenprofil einer Parabraunerde sein. Entitäten können problemangepasst definiert werden: Während beispielsweise in der Bodenkunde die Farbe nur als Attribut für die Horizontbeschreibung verwendet wird, kann eine Farbe auch als Entität verstanden werden, z.B. in der Farbenchemie.

Eine Entität kann durch bestimmte Eigenschaften näher beschrieben werden. Der Temperaturwert 21°C wird beispielsweise durch die Angabe des verwendeten Messgerätes und dessen Genauigkeit weiter charakterisiert, während bei der Parabraunerde Ausgangsgestein, Humusform und Bodenart zur weiteren Kennzeichnung beitragen. Es muss hier unterschieden werden zwischen **Eigenschaft (Attribut)** und **Eigenschaftswert**. Ein **Attribut** zur Kennzeichnung von Bodenprofilen ist z.B. die Humusform, ein **Eigenschaftswert**, der die Humusform bei einem konkreten Profil beschreibt, wäre z.B. F-Mull. Die Menge aller für die Charakterisierung eines Attributs zugelassenen Eigenschaftswerte nennt man **Domäne** oder **Wertebereich**. Für das Attribut "Humusform" könnte die Domäne <L-Mull, F-Mull, mullartiger Moder, ..., Rohhumus> festgelegt werden. Entitäten, die durch die gleichen Attribute beschrieben werden, bilden zusammen eine **Entitätsmenge**.

Betrachten wir jetzt zwei Entitätsmengen: "Messstandorte" und "Messgeräte". Jeder Messstandort ist mit einer Anzahl Messgeräten ausgestattet, und ein Messgerät ist normalerweise auf einem Messstandort aufgestellt. Die Zuordnung von Messgeräten zu Standorten nennt man **Assoziation**. Auch die Ausstattung des Messstandortes mit Geräten bildet eine **Assoziation**. Eine **Assoziation** legt fest, wieviele Entitäten der einen Entitätsmenge einer Entität der anderen Entitätsmenge zugeordnet sind. Es gibt vier verschiedene **Assoziationstypen**: Die einfache Assoziation ordnet einer Entität der ersten Menge genau eine Entität der zweiten

Menge zu (Bezeichnung: 1), die konditionelle Assoziation (c) keine oder eine, die multiple (komplexe) Assoziation (m) eine oder mehrere und die multipel-konditionelle Assoziation (mc) keine, eine oder mehrere.

Zwei wechselseitige Assoziationen bilden zusammen eine **Beziehung** (relationship). Die Messstandorte und die Messgeräte stehen in einer 1-m-Beziehung, weil jedem Messstandort ein oder mehrere Messgeräte zugeordnet sind und jedes Messgerät genau auf einem Messstandort steht. Werden alle Messgeräte miteinbezogen, d.h. auch die, die im Materiallager sind und nicht benutzt werden, dann ergibt sich eine c-m-Beziehung: Die Messstandorte sind immer noch mit einem oder mehreren Geräten ausgestattet, jedes Gerät steht aber auf einem oder keinem Messstandort.

Durch Kombination aller Assoziationstypen ergeben sich 16 Beziehungsarten, davon sind aber nur zehn wirklich verschieden. Einige Beziehungstypen sind bei der Darstellung in relationalen Datenbanken problematisch: Die konditionellen Beziehungen c-c, m-c und mc-c sowie die netzwerkförmigen Beziehungen m-m, mc-m und mc-mc. Diese können aber durch Einführen einer Beziehungsmenge in die verwendbaren hierarchischen Beziehungen 1-1, 1-c, 1-m und 1-mc umgewandelt werden (Zehnder 1987, 66ff). Eine **Beziehungsmenge** hält Beziehungen zwischen zwei oder mehreren Entitätsmengen fest. Auch eine Beziehungsmenge kann durch Merkmale charakterisiert werden.

Wie sieht nun das ganze in einer relationalen Datenbank aus? Eine Entität wird hier als Tupel dargestellt. Ein **Tupel** ist eine Aneinanderreihung von Werten. Ein Tupel für die Entität Parabraunerde könnte z.B. so aussehen: <Parabraunerde, Mull, IU, Würmlöss, 8°, 160 cm,...>. Eine Menge von gleichartigen Tupeln bildet eine **Relation**. Eine Relation wird meist als Tabelle dargestellt, wobei ein Tupel jeweils eine Tabellenzeile einnimmt, während die Werte der gleichen Domäne in der selben Spalte untereinanderstehen (s. Abb. 5 oder Vetter 1990, 115f). Damit die Tupel in einer Relation voneinander unterschieden werden können, wird ein Relationsschlüssel definiert. Ein **Relationsschlüssel** ist ein Attribut, in dessen Werten sich die einzelnen Tupel (Entitäten) einer Relation eindeutig unterscheiden. Ein Relationsschlüssel kann auch aus mehreren Attributen zusammengesetzt sein, es wird aber darauf geachtet, dass möglichst wenige Attribute am Schlüssel beteiligt sind. Ein Bodenprofil könnte z.B. über die Koordinaten des Profilverpunktes oder über eine Profilvernummer eindeutig identifizierbar gemacht werden.

Zur Wiederholung: Ein relationales Datenbanksystem ermöglicht es, Relationen zu definieren. Die einzelnen Zeilen der Relation entsprechen den Entitäten, die Spalten den Attributen. Ein Attribut oder mehrere werden als Schlüssel verwendet, über den die einzelnen Entitäten eindeutig gefunden werden können. Jedes Attribut (Spalte) hat einen eindeutigen Namen und alle Werte einer bestimmten Spalte entstammen der gleichen Domäne. Die Relationen und deren Verwaltung bilden die Grundlage für die im nächsten Abschnitt beschriebene Datenmodellierung.

3.1.3 Datenanalyse

Vor dem Entwickeln eines Datenkonzepts muss darauf hingewiesen werden, dass die Daten im Computer in den wenigsten Fällen so abgelegt werden, wie sie ein Benutzer am Bildschirm haben möchte. Man unterscheidet daher zwischen einer externen, einer konzeptionellen und einer internen Ebene. Die **externe Ebene** (auch externes Schema) ist die Datensicht des Anwenders (= View). Sie ist unabhängig von Computer und Datenbankprogramm. Auf dieser Ebene fallen die Daten an und auf diese Ebene sollen sie aus der Datenbank auch wieder geholt werden können. Durch die Datenanalyse wird ein **konzeptionelles Schema** entworfen. Hier werden die Daten so dargestellt, wie sie in die Relationen des Datenverwaltungssystems eingetragen werden. Hinter der konzeptionellen Ebene verbirgt sich die interne. Das **interne Schema** entspricht der Art und Weise, wie das Datenbankverwaltungssystem die Daten auf dem Datenträger (Diskette, Harddisk etc.) ablegt (Zehnder 1987, 15f). Mit einer Datenverwaltungs- und -abfragesprache werden die Daten von der externen Ebene auf die konzeptionelle übertragen und umgekehrt. Das Datenbankverwaltungssystem sorgt dafür, dass die Daten bei einer Abfrage auf der konzeptionellen Ebene zur Verfügung stehen.

Sinn der Datenanalyse ist es, eine redundanzfreie, wirklichkeitsgetreue und fehlerfreie Datenabspeicherung zu gewährleisten. Von **Redundanz** spricht man, wenn ein und dasselbe Faktum mehrmals festgehalten wird (Vetter 1990, 38). Wirklichkeitsgetreu heisst, dass die Information, die in eine Datenbank eingegeben worden ist, mit dem gleichen Informationsgehalt wieder aus der Datenbank geholt werden kann. Es dürfen keine Fakten verlorengehen, es dürfen aber auch keine neuen Fakten entstehen (was keineswegs trivial ist). Es gibt wohl kaum eine fehlerfreie Datenbank, denn schon ein Tippfehler bei der Dateneingabe führt zu falschen Informationen. Alle Eingabefehler können nicht abgefangen werden. Es gibt aber Fehler, die besonders schwerwiegend sind, z.B. ein falscher Schlüsselwert. In diesem Fall muss das System eine Eingabe zurückweisen. Wie sollen diese Forderungen erfüllt werden?

Als erstes müssen alle Entitätsmengen beschrieben werden, die in die Datenbank aufgenommen werden sollen. Dies kann auch bereits gemacht werden, wenn noch nicht alle Datenmengen bekannt sind. Zu jeder Entitätsmenge muss ein Entitätsschlüssel definiert werden. Man unterscheidet zwischen natürlichen und künstlichen Schlüsseln. **Natürliche Schlüssel** sind Attribute, die ein Objekt bereits von sich aus hat, z.B. ein Name. Weil meist nicht gewährleistet werden kann, dass sich alle Entitäten im Namen unterscheiden, werden auch **künstliche Schlüssel** definiert. Das sind meist Nummern, z.B. Standortsnummer, Profilnummer usw., die fortlaufend den Entitäten zugeordnet werden. Nachdem die Schlüssel der Entitätsmengen definiert sind, werden die jeweiligen Attribute ermittelt und notiert. Eine Entitätsmenge mit Schlüssel und Attributen kann z.B. in Form eines Tabellenkopfes dargestellt werden (s.z.B. Abb. 5). Wenn ein paar Tupel (Entitäten) in die Tabelle eingetragen werden, wird der nächste Schritt der Datenanalyse leichter vonstatten gehen.

Die Entitätsmengen, die in der vorliegenden Form als Relationen bezeichnet werden können, sollen nun normalisiert werden. Die **Normalisierung** ist eine Analyse der inneren Zusammenhänge einer einzelnen Relation, die dazu führen soll, dass Redundanzen eliminiert werden (Zehnder 1987, 51).

In der **1. Normalform** sind für jedes Attribut der Relation nur einfache Attributswerte zugelassen. Auf dem Schnittpunkt einer Spalte mit einer Zeile darf demnach nur ein Wert zu stehen kommen. Wenn z.B. das Attribut "Koordinaten" gewählt wurde, in das Rechts- und Hochwert eingetragen werden soll, so ist dieses aufzuspalten in die Attribute "Rechts" und "Hoch" (siehe dazu Abb. 5).

Die **2. Normalform** ist dann erreicht, wenn eine Relation in der 1. Normalform ist, und wenn jedes Attribut, das nicht zum Identifikationsschlüssel gehört, vom ganzen Schlüssel abhängt und nicht nur von Teilen desselben. Alle 1NF-Relationen mit einem einfachen Schlüssel sind automatisch auch in der 2. Normalform. Es müssen die Relationen kontrolliert werden, die einen zusammengesetzten Schlüssel haben (Schlüssel aus zwei und mehr Attributen). Es könnte z.B. eine Relation "Horizonte" definiert worden sein mit einem Schlüssel aus "Profilnummer" und "Horizontnummer". Ein Nichtschlüsselattribut "Gründigkeit" wäre von einem Teil des Schlüssels abhängig, nämlich von der Profilnummer, da die Gründigkeit das ganze Profil charakterisiert. Das Attribut "Gründigkeit" sollte deshalb in eine Relation "Profile" eingefügt werden, die die Profilnummer als alleinigen Schlüssel hat. Es kann mitunter vorkommen, dass eine Relation zur Bildung der 2. Normalform in zwei oder mehr Relationen aufgespalten werden muss. Bei genauerer Betrachtung wird man feststellen, dass dadurch eine logische Gruppierung nach Sachgebieten erreicht wird (Zehnder 1987, 53).

Für die **3. Normalform** müssen die Bedingungen der 2. Normalform erfüllt sein und zusätzlich darf kein Nichtschlüsselattribut von einem anderen Nichtschlüsselattribut funktional abhängen. Auch hier kann wieder eine Aufspaltung der Relation nötig werden. Wenn beispielsweise in einer Relation "Geräte" mit Schlüssel "Gerätenummer" die beiden Nichtschlüsselattribute "Standortnummer" und "Standortsname" enthalten wären, dann wäre die 3. Normalform nicht erfüllt, weil der Standortsname von der Standortnummer abhängig ist. Der Standortsname müsste in einer Relation "Standorte" untergebracht werden und dürfte nur in dieser Relation vorhanden sein. Es gibt noch weitere Normalformen (s.z.B. Vetter 1990, 188ff), für die Praxis reicht aber meist die dritte Normalform aus. Die Relationen sind nun für sich allein genommen in einer geeigneten Form für die Abspeicherung (sie entsprechen jetzt dem Relationenmodell von Codd 1970), es kann aber sein, dass Redundanzen oder sonstige Ungeheimheiten zwischen den verschiedenen Relationen bestehen. Im nächsten Schritt müssen auch diese beseitigt werden.

Im folgenden sollen die Beziehungen zwischen den Relationen im Rahmen des erweiterten Relationenmodells oder der globalen Normalisierung untersucht werden. Dazu müssen noch ein paar Begriffe definiert werden: Wir unterscheiden zwischen lokalen und globalen Attributen. Ein **globales Attribut** kommt in mindestens einer Relation im Schlüssel vor, während ein **lokales Attribut** nur in einer Relation und dort nicht im Schlüssel vorkommt (Zehnder 1987, 56). Als **Fremdschlüssel** bezeichnen wir ein Nichtschlüsselattribut einer Relation, das in einer anderen Relation als **Primärschlüssel** auftritt.

Relationen können über globale Attribute miteinander in Beziehung gesetzt werden. Solche Beziehungen ergeben aber nicht von vornherein eine richtige Aussage. Alle möglichen Aussagen im jetzt bestehenden Datenmodell müssen daher auf Primärschlüssel-Fremdschlüssel-Beziehungen hin untersucht werden. Es dürfen nur hierarchische Beziehungen (1-1, 1-c, 1-m

und 1-mc) vorhanden sein, weil nur diese in relationalen Datenbanken direkt darstellbar sind. In der realen Welt können aber sehr wohl auch andere Beziehungen auftreten. Für solche Beziehungen werden Beziehungsmengen definiert. Stehen beispielsweise zwei Relationen in einer m-mc-Beziehung zueinander, so wird von der ersten Relation zur Beziehungsmenge eine 1-m-Beziehung aufgebaut, während von der zweiten Relation zur Beziehungsmenge eine 1-mc-Beziehung ausgeht. Der Schlüssel der Beziehungsmenge ist zusammengesetzt aus den Primärschlüsseln der beiden anderen Relationen. Auch die übrigen konditionellen und netzwerkförmigen Beziehungen können durch solche Beziehungsmengen dargestellt werden.

Wenn im Datenmodell Beziehungsmengen nötig sind, müssen diese definiert und normalisiert werden. Es kann durchaus möglich sein, dass dadurch Attribute überflüssig werden. Es wird z.B. oft der Fehler gemacht, dass bei einer Literaturverwaltung die Stichworte in der gleichen Datei wie die Titel abgespeichert werden, etwa in den Attributen "Stichwort1" bis "Stichwort8". Nach der Normalisierung wird eine Relation für die Stichworte nötig, wodurch die Stichwortattribute in der Buchtitelrelation entfallen. Die Verbindung der Stichworte mit den Buchtiteln erfolgt über eine Beziehungsrelation.

Sind alle Entitäts- und Beziehungsmengen definiert, so müssen auch noch die Primärschlüssel untereinander verglichen werden. Wenn in zwei Relationen der gleiche Schlüssel verwendet wird, und wenn in beiden Relationen Aussagen zu den gleichen Entitäten gemacht werden, können sie zusammengefasst werden. Nach einer Zusammenfassung muss wieder geprüft werden, ob die 3. Normalform erfüllt ist.

Die Überprüfung der Beziehungen der verschiedenen Relationen führt immer wieder zu den Normalisierungsschritten zurück. Es handelt sich hier um einen Kreislauf, der mehrmals durchlaufen werden muss, bis das endgültige Modell bereit ist. Wenn aber das Datenkonzept steht, kann jede neue Entitätsmenge eingefügt werden, ohne dass die bestehende Struktur verändert werden muss. Dazu wird die neue Relation in die 3. Normalform gebracht. Wenn ihre Beziehungen zu den bestehenden Relationen geprüft worden sind, kann sie in die Datenbank eingebracht werden.

Nach diesen nicht unwichtigen theoretischen Ausführungen kann jetzt ein Datenmodell für die geoökologische Forschungspraxis entwickelt werden. Selbstverständlich konnte hier nur ein kleiner Überblick über die Begriffe und Konzepte des relationalen Datenmodells gegeben werden. Für weitere Informationen sei nochmals auf die beiden grundlegenden Werke von Zehnder (1987) und Vetter (1990) verwiesen. Eine auf die Geografie bezogene Darstellung gibt Barthelme (1989, Kap. 4).

3.1.4 Datenkonzept für die geoökologische Forschung

Die geoökologische Untersuchung in der Art der landschaftsökologischen Komplexanalyse strebt eine Landschaftsbeschreibung an. Einerseits soll die flächenhafte Verbreitung von Geoelementen erfasst werden, andererseits sollen Raumeinheiten mit gleichem Stoff- und Energiehaushalt ausgeschieden und mit Stoff- und Energiebilanzen belegt werden (Mosi-

mann 1984, 38). Als Methoden stehen zum einen die Kartierung und die Anlage von Messnetzen, zum andern die Einrichtung von Messstandorten (Tesserae) zur Verfügung (Mosi- mann 1984, 63ff). Aus diesem Ansatz ergeben sich grob gesagt zwei Arten von Daten:

1. einmalig aufgenommene Flächendaten und
2. periodisch erhobene Punktdaten.

Zu den **Flächendaten** gehören z.B. die topografische Grundlagenkarte (ev. auch digitales Höhenmodell und daraus abgeleitete Hangneigungs- und Besonnungskarten), die geomorp- hologische Karte, die Daten der Bodenkartierung (Profil- und Bohrdaten), Mikroklima- und Nutzungskarte etc. Reine geometrische Daten (Unterschied geometrische und Attributdaten s. Barthelme 1989, Kap. 2 und 3) sollten mit einem geografischen Informationssystem (GIS) verwaltet werden. Es stehen heute genügend kommerzielle Programme zur Verfügung, die auch den ausgefallensten Bedürfnissen der geometrischen Speicherung gerecht werden. Bei der Speicherung von Attributdaten können relationale Datenbanken verwendet werden. Die einzelnen Datensätze müssen durch Angabe eines Geometriebezugs eindeutig lokalisierbar sein, wozu z.B. die Koordinaten oder die Zugehörigkeit zu einer Fläche in einer digital abge- speicherten Karte geeignet sind. Durch den geometrischen Bezug können die Datensätze in ein GIS eingespeist und mit Interpolationswerkzeugen als Karte dargestellt werden. Als Bei- spiel für eine Datenstruktur für "flächenhaft" aufgenommene Punktdaten kann das Bodenver- waltungsprogramm in Kapitel 3.3 betrachtet werden. Dieses Beispiel kann auch auf andere Kartierungsarten wie z.B. Vegetationsaufnahmen übertragen werden.

Im Folgenden sollen die mengenmässig überwiegenden **periodisch erhobenen Punktdaten** behandelt werden. Anhand von verschiedenen Messeinrichtungen kann gezeigt werden, wie sich das Datenmodell den Daten anpasst und nicht umgekehrt. Die Beispiele sind als Grund- lage für die eigene Datenanalyse gedacht und können nur bei gleichen Messvoraussetzungen direkt übernommen werden.

1. Beispiel:

Es wird ein Messnetz mit mehreren Messpunkten eingerichtet. An jedem Messpunkt stehen die gleichen Messgeräte. Pro Messgerät und Messperiode wird ein Messwert erhoben. Vor- stellbar wäre z.B. ein Niederschlagsmessnetz, für das an 30 Standorten je ein Hellmann-Re- genmesser aufgestellt wird. Dabei soll der Niederschlag wöchentlich gemessen werden.

Als erstes wird die Relation "Niederschlagsmessung" mit den Attributen Standortnummer, Standortname, Koordinaten, Datum, und Niederschlag entworfen (Abb. 5). Als Schlüssel wird die Kombination von Standortnummer und Datum verwendet (unterstrichene Attribute). Ein Eintrag könnte so aussehen: <1, Höchi, 605300/264400, 4.8.90, 21 mm>. Für die 1. Nor- malform müssen alle Attribute mit mehreren Einträgen (hier: Koordinaten) aufgespalten wer- den. Dies führt zur 1.NF-Relation Standortnummer, Standortname, rechts, hoch, Datum und Niederschlag.

Nun muss geprüft werden, ob einzelne Attribute der Relation von Teilen des Schlüssels abhängig sind. Der Standortname und die Koordinaten sind direkt von der Standortnummer abhängig, während die Niederschlagsmenge von der Standortnummer und vom Datum abhängt. Die Relation wird also aufgespalten in die 2.NF-Relationen "Standort" (Standortnummer, Standortname, rechts, hoch) und "Niederschlag" (Standortnummer, Datum, Niederschlag).

Die Relationen müssen für die 3. Normalform auf transitive Abhängigkeiten hin überprüft werden. Die Relation "Niederschlag" ist bereits in der 3. Normalform, denn für eine transitive Abhängigkeit braucht es mindestens zwei Nichtschlüsselattribute. Für die Relation "Standort" kann man so argumentieren: Der Standortname ist überflüssig, weil er aus den Koordinaten mit Hilfe einer Karte erschlossen werden kann, er ist also wegzulassen. Dieses Argument zählt nur, wenn im Informationssystem eine topografische Karte abgespeichert ist, in der über die Koordinaten alle Flurbezeichnungen ermittelt werden können. Muss aber für den Standortnamen eine Karte aus Papier oder gar ein Stadtplan herangezogen werden, so kann das Weglassen des Standortnamens als Informationsverlust bezeichnet werden. Somit ist auch die Relation "Standort" in der 3. Normalform.

Die beiden Relationen stehen in einer 1-mc-Beziehung (nach der ersten Niederschlagsmessung 1-m), denn zu jedem Standort sind mehrere, einer oder kein Niederschlagswert vorhanden, während ein Niederschlagswert nur auf einen Standort bezogen werden kann. Möchte man alle Niederschlagswerte an einem Standort ermitteln, so sucht man sich in der Standortdatei die Standortnummer heraus und setzt anschliessend einen Filter über die Relation "Niederschlag", so dass nur die Werte mit der betreffenden Standortnummer angezeigt werden. Sollen die Werte aller Standorte nebeneinandergestellt werden, kann die oben beschriebene Abfrage für alle Standorte wiederholt werden, worauf die dabei entstandenen Relationen über das Datum in Beziehung gesetzt werden. Hier tritt aber ein Problem auf: Wenn nicht alle Stationen am selben Wochentag bedient werden, stimmen die Daten nicht überein, was eine Verknüpfung erschwert. Dieses Problem liegt in der Natur des gewählten Schlüsselattributs "Datum". Das Datum ist für die Kennzeichnung von Tagen geeignet, es sollen aber Wochen beschrieben werden. Möchte man trotzdem am Datum festhalten, kann eine Relation "Woche" (Wochenummer, Anfang.Datum) definiert werden. Jedes Datum, das grösser oder gleich dem Anfangsdatum und kleiner dem Anfangsdatum + 7 ist, gehört zu der entsprechenden Woche. Somit können die Werte der gleichen Woche zueinander in Beziehung gesetzt werden, auch wenn sie über das Datum identifiziert werden.

2. Beispiel:

Es werden an mehreren Standorten verschiedene Messgeräte aufgestellt. Alle Standorte sind mit den gleichen Geräten ausgestattet, die Geräte werden aber in unterschiedlichen Perioden abgelesen. Man stelle sich z.B. vier Klimamessstationen vor, an denen Temperatur, Luftfeuchtigkeit, Niederschlag und Windstärke gemessen werden. Die Temperatur- und Luftfeuchtigkeitswerte werden vom Thermohygrografenstreifen für jeden Tag um 7, 14 und 21 Uhr abgelesen, die Windstärke wird jede halbe Stunde automatisch aufgezeichnet, und der Regenschirm wird jede Woche einmal betreut.

Ausgangsrelation:

<u>Niederschlagsmessung</u>	<u>Standortnummer</u>	Standortname	Koordinaten	<u>Datum</u>	Niederschlag
	1	Höchi	605300/264400	4.8.90	21 mm
	2	Chirsbaum	606700/263750	4.8.90	30 mm
	1	Höchi	605300/264400	11.8.90	0 mm

1. Normalform:

<u>Niederschlagsmessung</u>	<u>Standortnr.</u>	Standortname	rechts	hoch	<u>Datum</u>	Niederschlag
	1	Höchi	605300	264400	4.8.90	21
	2	Chirsbaum	606700	263750	4.8.90	30
	1	Höchi	605300	264400	11.8.90	0

2. Normalform und
3. Normalform:

Standort	<u>Standortnummer</u>	Standortname	rechts	hoch
	1	Höchi	605300	264400
	2	Chirsbaum	606700	263750

Niederschlag	<u>Standortnummer</u>	<u>Datum</u>	Niederschlag
	1	4.8.90	21
	2	4.8.90	30
	1	11.8.90	0
	2	11.8.90	0
	1	18.8.90	18

Abb. 5: Normalisierung der Relation "Niederschlagsmessung". (siehe Text Beispiel 1)

Die im ersten Beispiel erarbeitete Relation "Standort" soll hier der Einfachheit halber übernommen werden (Abb. 6). Wir definieren eine zweite Relation "Messwerte" (Standortnummer, Datum, Zeit, Temperatur, Luftfeuchtigkeit, Windstärke, Niederschlag). Diese Relation ist in der 1. Normalform, da zu jedem Attribut nur ein Wert eingegeben werden muss. Werden Werte in die Relation eingetragen, so zeigt sich, dass viele Felder leer bleiben. Im Beispiel auf Abb. 6 wurde die Messstation "Höchi" um 13.00 Uhr betreut. Der Niederschlagswert der ganzen Woche wird diesem Zeitpunkt zugeordnet, während die restlichen 335 Niederschlagsfelder der Woche leer bleiben. Die 2. Normalform ist also noch nicht erfüllt, weil das Attribut "Niederschlag" nicht vom Schlüsselattribut "Zeit" abhängig ist. Wir bilden deshalb zwei Relationen: "Wind_Temp" (Standortnummer, Datum, Zeit, Windstärke, Temperatur, Luftfeuchtigkeit) und "Niederschlag" (Standortnummer, Datum, Niederschlag). Wie be-

reits beim ersten Beispiel ist auch hier das Schlüsselattribut "Datum" für den Niederschlag nicht geeignet, weil nicht Tageswerte, sondern Wochenwerte erhoben werden. Wir definieren deshalb "Niederschlag" (Standortnummer, Wochennummer, Niederschlag).

Wenden wir uns der Relation "Wind_Temp" zu: Alle Nichtschlüsselattribute sind zwar voll vom Schlüssel abhängig, aber irgendetwas scheint noch falsch zu sein. Wenn wir das Schlüsselattribut "Zeit" untersuchen, stellen wir fest, dass für das Attribut "Windstärke" die Zeitwerte aus der Domäne (0.00, 0.30, 1.00, ..., 23.00, 23.30) stammen, während für Temperatur und Luftfeuchtigkeit nur die Werte 7.00, 14.00 und 21.00 realisiert sind. Formal ist der Schlüssel für alle Attribute gleich, die Domäne ihres Schlüssels jedoch nicht. Aus diesem Grund wird auch diese Relation geteilt, wobei die Relationen "Wind" (Standortnummer, Datum, Zeit, Windstärke) und "Thermohygrograf" (Standortnummer, Datum, Zeit, Temperatur, Luftfeuchtigkeit) entstehen. Je nach Fragestellung wäre auch eine Relation "Thermohygrograf" (Standortnummer, Datum, Temperatur7, Temperatur14, Temperatur21, Luftfeuchtigkeit7, Luftfeuchtigkeit14, Luftfeuchtigkeit21) denkbar. In dieser Relation können ebenfalls alle Messwerte dargestellt werden. Zudem ist der Speicherbedarf viel geringer, müssen doch die Zeitwerte nicht abgespeichert werden.

Die 3. Normalform ist bei "Wind" und "Niederschlag" erfüllt, weil nur ein Nichtschlüsselattribut vorhanden ist. In der Relation "Thermohygrograf" sind Temperatur und Luftfeuchtigkeit zwar korreliert, aber nicht funktional abhängig, womit auch hier die 3. Normalform erreicht ist.

Die Werte eines Standortes werden auch bei dieser Datenstruktur durch Herausfiltern der entsprechenden Standortnummern gefunden. Sollte die Korrelation der unterschiedlichen Messwerte bestimmt werden (z.B. Wind und Luftfeuchtigkeit), so müssen die Werte so weit aggregiert werden, bis sie das gleiche Zeitniveau haben.

3. Beispiel:

Es werden verschiedene Messstandorte mit unterschiedlichen Messgeräten ausgestattet. Die Messintervalle sind nicht bei allen Geräten gleich. Man kann sich eine Einzugsgebietsuntersuchung mit drei Tesseræ (mit Thermohygrograf, Niederschlagsmesser, Verdunstungsmesser, 3 Saugkerzen und Lysimeter), einem wandernden Niederschlagsmessnetz (30 Niederschlagssammler) und einer Pegelmessstation vorstellen.

Aus den Erfahrungen der Beispiele 1 und 2 kann folgende Annahme gemacht werden: Alle Messwerte, die in gleichen Zeitintervallen erhoben werden, können in einer Relation zusammengefasst werden. Wir definieren deshalb die Relationen: "Standort" (Standortnummer, Standortname, rechts, hoch), "Tageswerte" (Standortnummer, Datum, Temperatur7, Temperatur14, Temperatur21, Luftfeuchtigkeit7, Luftfeuchtigkeit14, Luftfeuchtigkeit21, Abflussmenge) und "Wochenwerte" (Standortnummer, Wochennummer, Niederschlag, Verdunstung, pH.SK20, Leitfähigkeit.SK20, Phosphat.SK20, Ammonium.SK20, Nitrat.SK20, Kalium.SK20, Calcium.SK20, Magnesium.SK20, pH.SK40, Leitfähigkeit.SK40, Phosphat.SK40, Ammonium.SK40, Nitrat.SK40, Kalium.SK40, Calcium.SK40, Magnesium.SK40, pH.SK60,

Ausgangsrelationen (1. Normalform erreicht):

Standort	Standortnummer	Standortname	rechts	hoch
	1	Höchi	605300	264400
	2	Chirsbaum	606700	263750

(Diese Relation wird nicht mehr verändert. Sie wird deshalb unten nicht mehr aufgeführt)

Messwerte	Standortnummer	Datum	Zeit	Temperatur	Luftfeuchtigkeit	Windstärke	Niederschlag
	1	4.8.90	13.00			5	21
	1	4.8.90	13.30			3	
	1	4.8.90	14.00	23	65	7	

Wind_Temp	Standortnummer	Datum	Zeit	Temperatur	Luftfeuchtigkeit	Windstärke
	1	4.8.90	13.00			5
	1	4.8.90	13.30			3
	1	4.8.90	14.00	23	65	7

2. Normalform und
3. Normalform:

THG	Standortnummer	Datum	Zeit	Temperatur	Luftfeuchtigkeit
	1	4.8.90	07.00	17	95
	2	4.8.90	07.00	16.5	100
	1	4.8.90	14.00	23	65
	2	4.8.90	14.00	23.5	55
	1	4.8.90	21.00	18	85

Niederschlag	Standortnummer	Woche	Niederschlag
	1	1	21
	2	1	30
	1	2	0
	2	2	0
	1	3	18
	2	3	16

Wind	Standortnummer	Datum	Zeit	Windstärke
	1	4.8.90	13.00	5
	2	4.8.90	13.00	2
	1	4.8.90	13.30	3
	2	4.8.90	13.30	2
	1	4.8.90	14.00	7

Abb. 6: Normalisierung der Relation "Messwerte". (siehe Text Beispiel 2)

Leitfähigkeit.SK60, Phosphat.SK60, Ammonium.SK60, Nitrat.SK60, Kalium.SK60, Calcium.SK60, Magnesium.SK60, pH.Ly, Leitfähigkeit.Ly, Phosphat.Ly, Ammonium.Ly, Nitrat.Ly, Kalium.Ly, Calcium.Ly, Magnesium.Ly). Bei dieser Definition wurde bereits darauf geachtet, dass die 1. Normalform eingehalten wird. Die erste Normalform könnte beispielsweise missachtet werden, indem für die chemischen Analysen pro untersuchtem Stoff nur ein Feld zur Verfügung gestellt würde, in das die Werte der Saugkerzen und des Lysimeters in der richtigen Reihenfolge einzutragen sind.

Die Relation "Wochenwerte" ist zwar in der 2. Normalform, doch haben auch hier nicht alle Attribute die gleiche Schlüsselwertdomäne. Die Standortnummer kann für alle chemischen Analysen und für die Verdunstung nur drei Werte annehmen, die Werte der drei Tesseræ. Für die Niederschlagsmessung sind aber 33 Standortnummern vorzusehen, wenn die Messpunkte verschoben werden sogar noch mehr. Deshalb muss eine Relation "Niederschlag" (Standortnummer, Wochennummer, Niederschlag) errichtet werden. Die Relation "Wochenwerte" wird beibehalten, jedoch ohne das Attribut "Niederschlag".

Angenommen, der Boden sei an einem Messstandort nicht so tiefgründig, dass die 60 cm Saugkerze eingesetzt werden könnte. Somit würden also bei jedem dritten Datensatz die Werte für die 60 cm Saugkerze fehlen. Will man dies vermeiden, so muss eine Relation "Saugkerze60" definiert werden. Im Extremfall führt das dazu, dass für jedes Messgerät und jede Aufstellungsart eine Relation eingerichtet wird. Der Vorteil dieser Aufspaltung ist, dass Messwerte aus den verschiedensten Projekten in der gleichen Datenbank verwaltet werden können. Damit die einzelnen Bearbeiter ihre Daten noch überblicken können, sollte eine Personendatei eingerichtet werden, die über eine Beziehungsrelation "Bearbeitet" (Personennummer, Standortnummer, Beginn, Ende) mit der Standortsdatei verknüpft wird. Dadurch kann eine Datenbank nach allen Daten abgefragt werden, die eine Person erhoben hat.

Aus den drei Beispielen können folgende Leitsätze für den Entwurf von raum-zeitlich orientierten Relationen abgeleitet werden:

1. Die Rauminformation (Standortsbeschreibung) wird in einer Relation gehalten.
2. Werte, die in gleichen Zeitintervallen erhoben werden, können zusammengefasst werden.
3. Werte, die nicht an allen Standorten aufgenommen werden, müssen in einer eigenen Relation aufgeführt werden.

Diese Leitsätze sind nicht allgemein verbindlich und erübrigen auch nicht die Normalisierung. Sie führen jedoch zu einer standort- und messgrößenorientierten Datenhaltung, die sehr einfach zu dokumentieren ist. Man braucht nur die Relationen mit all ihren Attributen anzugeben und die Primärschlüssel-Fremdschlüssel-Beziehungen aufzuzeichnen. Die Daten sind somit auch ohne die Gegenwart der Person verwendbar, welche die Daten aufgenommen hat.

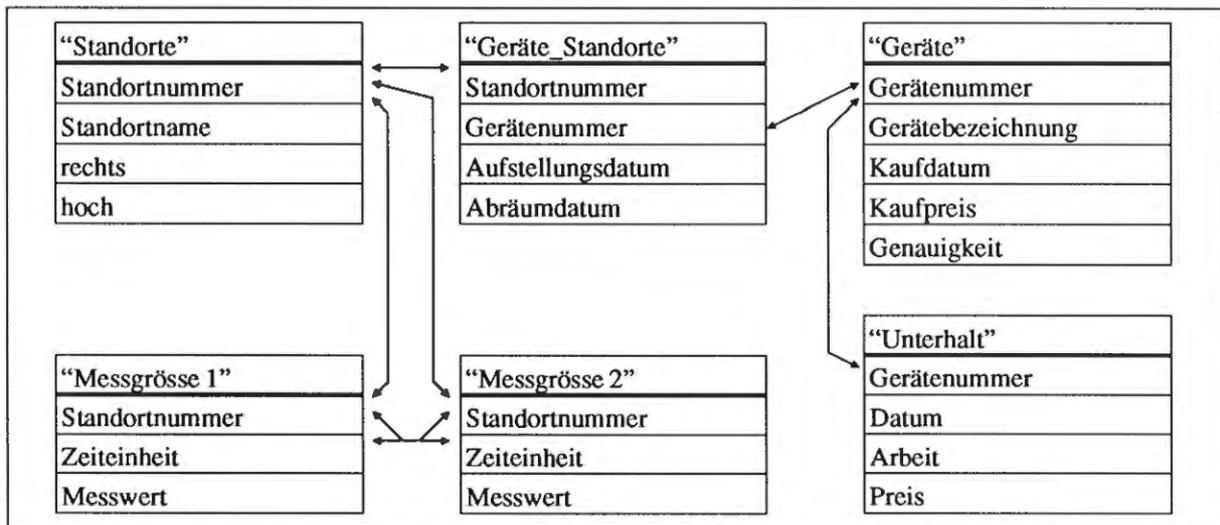


Abb. 7: Erweiterung des Datenkonzepts für eine Geräteverwaltung. Mit den drei neuen Relationen können z.B. folgende Aussagen gemacht werden: Welche Messwerte wurden mit welchem Messgerät gemessen, wie gross war der Fehler, welche Geräte sind wo im Einsatz, wieviel hat ein Messgerät bisher gekostet (Kauf und Unterhalt) usw.

Am Beispiel der Geräteverwaltung soll gezeigt werden, dass das oben entworfene Datenmodell für jede Erweiterung offen ist. Möchte man Informationen über die Messgeräte wie Geräteart, Aufstellungsort, Messgenauigkeit, Aufstellungsdatum, Abräumdatum etc. verwalten, so wird man eine Relation “Geräte” (Gerätenummer, Geräteart, Kaufdatum, Messgenauigkeit, ...) einrichten. Nach der Normalisierung wird sich herausstellen, dass der Geräteeinsatz über eine Beziehungsrelation “Geräte_Standorte” (Gerätenummer, Standortnummer, Aufstellungsdatum, Abräumdatum) registriert werden muss. Mit einer zusätzlichen Relation “Unterhalt” (Gerätenummer, Datum, Arbeit, Preis) kann Buch geführt werden über Revisionen und Reparaturen der Geräte. Die erweiterte Datenarchitektur wird in Abbildung 7 dargestellt. Alle Ergänzungen verändern die bestehenden Daten in keiner Weise. Der Datenbankbetrieb kann auch während der Erweiterung aufrecht erhalten bleiben, und die bisher entworfenen Abfragen sind ohne Einschränkung weiter verwendbar. Diese neuen Relationen können dazu beitragen, einen Teil der Anforderungen zu erfüllen, die Dettling (1989, 39ff) an die Datenprotokollierung stellt.

Um den Datenmodellentwurf in den richtigen Rahmen zu stellen, wird noch ein letztes Beispiel aufgeführt, mit dem gezeigt werden kann, dass durchaus noch andere Datenarchitekturen denkbar sind, die genau den gleichen Sachverhalt darzustellen vermögen wie das oben entworfene Modell. Dazu übernehmen wir die Relation “Standort” von oben und führen eine neue Relation “Messgrössen” (Messgrössennummer, Messgrössenbezeichnung, Dimension) ein (Abb. 8). Alle Messwerte können nun in einer einzigen Relation “Messwerte” (Standortnummer, Messgrössennummer, Datum, Messwert) untergebracht werden. Dieses Datenmodell ermöglicht ohne Einschränkungen die gleichen Auswertungen wie das oben beschriebene. Der “Weg” zu einem Messwert wird jedoch länger, weil dazu drei Angaben erforderlich sind, nämlich Standort, Messgrösse und Datum.

Wie dieses Datenmodell auf einem Computer verwirklicht wird, hängt von der verwendeten Software ab. In dieser Arbeit kann nicht auf die Einzelheiten der verschiedenen Systeme eingegangen werden. Mit Vorteil wird ein System eingesetzt, das den SQL (Structured Query

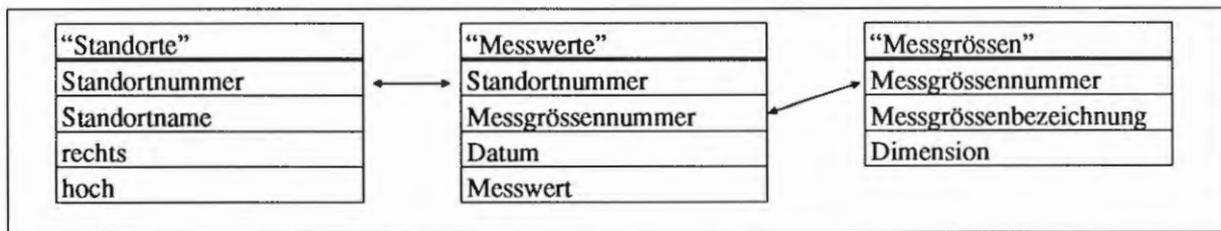


Abb. 8: Alternative Datenarchitektur: Alle Messwerte werden in einer Relation untergebracht. Messwerte, die den gleichen Sachverhalt darstellen, haben die gleiche Messgrößennummer.

Language)-Standard unterstützt. Dadurch wird man unabhängig von den unterschiedlichen Betriebssystemen. (Es gibt inzwischen bereits Betriebssysteme, die SQL integriert haben, so z.B. OS/2 oder AS/400). Hat man sich für ein Datenbankverwaltungsprogramm entschieden, so gibt es neben den entsprechenden Programmdokumentationen genügend Literatur, welche den Weg zur fertigen Datenbank beschreibt. Für “dBase IV” kann u.a. Tiley (1989) empfohlen werden, für “Oracle” das Buch von Finkenzeller, Kracke & Unterstein (1989).

Für die Forschungsarbeit am Geografischen Institut wäre es ein grosser Gewinn, wenn die Daten, die im Laufe der Jahre gesammelt werden, nach einheitlichen Kriterien gespeichert würden. Es könnten so auch langjährige Bilanzen und gut fundierte Modellsimulationen erarbeitet werden. Der Aufwand für die Datensammlung im Gelände ist zu gross, wenn die Daten nur in einer einzigen Diplom- oder Doktorarbeit ausgewertet werden. Es sollte deshalb – nach sorgfältiger Evaluation und einer ausführlichen Diskussion innerhalb der Forschungsgruppe – zumindest ein gemeinsames Datenhaltungskonzept formuliert werden. Im besten Fall entsteht vielleicht sogar eine gemeinsame Datenbank.

3.2 Datenauswertung

In einer Datenbank werden Daten meist so abgespeichert, wie sie anfallen, also als “Rohdaten”. Die Daten können je nach Fragestellung zu einer Datensicht (View) zusammengestellt werden, ohne dass die Datenbasis verändert wird. Es wird kein Auswertungsverfahren dem anderen vorgezogen, weil alle Anwendungen auf die gleichen Daten zugreifen können. Für die weitere Arbeit gibt es eine Fülle von Auswertungsprogrammen, angefangen bei Statistik- und Grafikanwendungen bis hin zu Simulation und Kartendarstellung. Viele Softwarepakete können über Makros und Applikationsgeneratoren an spezielle Bedürfnisse angepasst werden. Wenn trotzdem nicht alle Auswertungen durchgeführt werden können, so gibt es die Möglichkeit, über Programmierschnittstellen selbstprogrammierte Funktionen einzubauen. Es wird in den wenigsten Fällen nötig sein, selber zu programmieren.

Die Schritte zu einer Datenauswertung sind nun folgende: Man stellt sich mit einer Datenbankabfrage oder einer Datensicht die Daten zusammen, die weiterbearbeitet werden sollen. Diese Daten werden exportiert, d.h. sie werden von der Darstellungsform der Datenbank in ein allgemeines Datenformat übersetzt und gespeichert. Beinahe jedes Datenverwaltungssystem unterstützt mindestens ein Exportformat, so dass dieser Vorgang mehr oder weniger automatisch abläuft. Nun kann ins Auswertungsprogramm gewechselt werden, wo die Daten importiert werden. Mit dem Importieren werden die Daten in das Format des Auswertungsprogramms übersetzt. Die Daten können jetzt nach allen Regeln der Kunst analysiert, ver-

rechnet, dargestellt oder ausgedruckt werden; nach allen Regeln der Kunst deshalb, weil die Softwarepakete eine solche Flut von Funktionen anbieten, dass es von Vorteil ist, wenn man weiss, was man mit den Daten machen will.

Bei der herkömmlichen Datenauswertung musste man sich jeweils gut überlegen, ob sich die eine oder die andere statistische Methode besser eignet, wann welches Diagramm am aussagekräftigsten ist usw. Mit den Statistikprogrammen stehen nun alle möglichen Tests und Analysen zur Verfügung, und die Grafikprogramme können jede Zahlenreihe in fast allen Grössen, Farben und Formen darstellen. Allzuoft wird aber vergessen, dass einem die Maschine das Denken nicht abnimmt. Es ist nicht verboten, mit den Möglichkeiten zu spielen. Damit aber die geistige Umwelt nicht noch mehr mit der Publikation von unsinnigen Tabellen und Grafiken verschmutzt wird, sollten Sinn und Aussage einer Datenauswertung geprüft werden. Diese Prüfung kann nur erfolgen, wenn die nötigen Kenntnisse in den involvierten Fachgebieten vorhanden sind. Es muss nochmals betont werden: Der Computer ist nur ein Werkzeug, das zur Lösung von Aufgaben eingesetzt wird. Welche Aufgaben gelöst werden sollen und wie die Lösung auszusehen hat, bestimmen wir.

Mit der Datenauswertung sollte sehr früh begonnen werden, d.h. während der Feldphase. Werden die Messprotokolle laufend in die Datenbank eingetippt, so stehen bereits nach wenigen Wochen genügend Daten zur Verfügung um Datenauswertungsverfahren zu entwerfen. Mit wenigen Daten lassen sich Datensicht und Datenauswertung einfach auf ihre Richtigkeit hin prüfen. Es ist also sinnvoll, in der Anfangsphase die ersten Makros für Datenimport und Datenexport, Grafikschemata usw. zu erstellen. Mit diesen Auswertungswerkzeugen können die Daten laufend dargestellt und überprüft werden, so dass die Felddaten an neue Erkenntnisse aus der Datenauswertung angepasst werden kann (fehlerhafte Geräte austauschen, unnötige Messungen einstellen, neue Messgrössen einbeziehen etc.)

Nach diesen allgemeinen Bemerkungen können wir uns geökologischen Fragestellungen zuwenden. Eines der komplexesten Auswertungsprobleme ist wohl die Erstellung von Gebietsbilanzen. Dabei müssen die geometrischen Daten aus Kartierungen und digitalen Geländeanalysen mit Messreihen verknüpft werden. Eine Wasserbilanz kann dies verdeutlichen:

In einem Einzugsgebiet wurden zehn Regenmesser, vier Tankevaporimeter und eine Abflussmessstation an "repräsentativen" Standorten aufgestellt. Die Geoelemente Relief und Vegetation (Nutzung) wurden im Feld kartiert und anschliessend in digitaler Form in einem Geoinformationssystem abgelegt. Zusätzlich sind eine Hangneigungs-, eine Expositions- und eine Wölbungskarte aus einem digitalen Höhenmodell abgeleitet worden. Nach zwei Jahren Messdauer soll eine Wasserbilanz für das Einzugsgebiet erstellt werden. Unter der Annahme, dass kein unterirdischer Abfluss stattgefunden habe, sollen die Glieder der Gleichung $N = A_0 + ET + (R - B)$ mit Zahlen gefüllt werden (N: Niederschlag; A_0 : Oberflächenabfluss; ET: Evapotranspiration; R: Rücklage; B: Aufbrauch). Die Bilanzglieder sind in mm ($= l/m^2$) anzugeben.

R und B wurden nicht direkt gemessen, sie werden am Ende der Auswertung zur Korrektur der Bilanz verwendet. A_0 wird aus Pegelständen und Abflussgeschwindigkeit in Litern berechnet. Dieser Wert wird auf die ganze Einzugsgebietsfläche bezogen, weil nur eine Ab-

flussmessstation vorhanden ist. Die Flächenberechnung ist mit dem Geoinformationssystem leicht zu bewerkstelligen. Somit fehlen nur noch der Niederschlag und die Evapotranspiration.

Die punkthaften Messungen von Niederschlag und Evapotranspiration müssen mit den Flächen in Verbindung gebracht werden, für die sie repräsentativ sind. Entsprechend der Grösse dieser Flächen können die Einzelwerte bei der Mittelwertbildung gewichtet werden. Durch die Überlagerung der digitalen Karten können diese Flächen ermittelt und berechnet werden. Wird nun angenommen, dass die Niederschlagsverteilung in erster Linie von der Höhe, der Hangneigung, der Exposition und der Wölbung abhängig sei, so können durch Verschneidung der entsprechenden Karten Areale mit gleicher Höhenstufe, Hangneigungsklasse, Exposition und Wölbung dargestellt werden. Diese Areale werden mit der Lage der Messpunkte (vgl. Relation "Standorte") in Verbindung gebracht. Wenn die Messpunkte wirklich repräsentativ sind, dann sollte für jedes Areal oder jede Arealklasse ein Messpunkt vorhanden sein. Natürlich ist die Zahl der Areale von der Wahl der Klassengrenzen und von der Gewichtung der einzelnen Faktoren abhängig. Mit geschicktem Umklassieren und Umbewerten könnten praktisch alle Punkte "repräsentativ" gemacht werden, doch kann hier nur aus der Fachkenntnis heraus entschieden werden, welche Einteilung sinnvoll ist. Wenn zuletzt noch Areale ohne Messpunkt übrigbleiben, dann werden diese den Messpunkten zugeschlagen, denen sie räumlich und inhaltlich am nächsten kommen.

Für die vier Verdunstungsmesspunkte muss nun der gleiche Vorgang wiederholt werden. Bei der Verdunstung spielt die Oberflächenbedeckung eine wichtige Rolle, so dass neben den Reliefaktoren auch die Vegetationskarte in die Kartenverschneidung einbezogen werden kann. Werden die einzelnen Bedeckungstypen mit Verdunstungswerten aus der Literatur belegt – selbstverständlich nach Anpassung an die Gebietswerte –, so resultiert eine differenzierte Verdunstungskarte, aus der ein Mittelwert für das ganze Einzugsgebiet errechnet werden kann. Damit wären alle Glieder der Wasserbilanz beisammen.

Aus dem Beispiel wird sichtbar, dass durch den Einsatz eines geografischen Informationssystems viel präzisere quantitative Aussagen gemacht werden können als dies bisher der Fall war. Dies nicht etwa, weil mehr Information vorhanden ist – Kartierungen und Messungen werden genau gleich durchgeführt wie bisher –, sondern weil Punktdaten (als funktionale Komponente (Leser & Klink 1988, 227)) auf eine neue Weise mit den flächenhaft erfassten Geoelementen (als räumliche Komponente) in Beziehung gebracht werden können. Der Messwert wird mit Hilfe genau derjenigen Grössen extrapoliert, die ihn auch beeinflussen, was für die darzustellende Systemgrösse zu einer exakteren Quantifizierung führt als der Bezug auf Ökotopgrenzen, die nach Homogenitätskriterien festgelegt worden sind (Leser 1978, 211). Die Homogenität wird aufgrund von dominanten Merkmalen bestimmt und muss für die gerade zu berechnende Systemgrösse nicht in jedem Fall erfüllt sein. Dieser Gedanke sollte vermehrt auch in der räumlichen Anordnung der Messgeräte berücksichtigt werden, damit die "Repräsentativität" der einzelnen Messwerte erhöht werden kann. Eine breit abgestützte Kartenanalyse mit einem Geoinformationssystem lässt die nach dem "geoökologischen Gefühl" getroffene Auswahl von Messstandorten objektivieren.

4. Anmerkungen zu einer digitalen geoökologischen Karte 1 : 25 000

4.1 Datenerhebung und Datenspeicherung

Eine geoökologische Karte vereinigt Methoden aus verschiedenen geowissenschaftlichen Fachbereichen. Sofern diese Methoden Aussagen in der Dimension der GÖK 25 (topisch, chorisch) ermöglichen, sollten sie unverändert übernommen werden, damit die Anwender von geowissenschaftlichen Daten nicht verwirrt werden und die Akzeptanz für eine GÖK bei den Nachbardisziplinen gefördert werden kann. In der aktuellen Fassung der Kartieranleitung (Leser & Klink 1988) sind aber Methoden von Nachbardisziplinen ohne ersichtlichen Grund abgewandelt worden, so z.B. die Klassierung des ökologischen Feuchtegrades (Leser & Klink 1988, Tab. 20), bei der eine gegenüber der DIN-Norm (Arbeitsgruppe Bodenkunde 1982, Tab. 58) genau umgekehrte Skala verwendet wurde. Umso mehr erstaunt es, dass in der Bewertungsanleitung des Leistungsvermögens des Landschaftshaushaltes (Marks et al. 1989, Tab. 9), die auf der GÖK aufbaut, wiederum auf die ursprüngliche Klassifikation zurückgegriffen wird. Solche Umwege sind eher verwirrend und bringen keine neuen Erkenntnisse. Für eine weitere Bearbeitung der Kartieranleitung für die geoökologische Karte müssen deshalb unbedingt alle Methoden und Klassifizierungen mit denjenigen der Nachbardisziplinen in Einklang gebracht werden. Dies ist wichtig, wenn die GÖK in ein bestehendes Informationssystem (z.B. Bodenkunde, Lufthygiene etc.) eingebunden werden soll. Dass solche fächerübergreifenden Informationssysteme in Zukunft an Bedeutung gewinnen werden, belegt z.B. der "Vorschlag für die Errichtung eines länderübergreifenden Bodeninformationssystems" (Arbeitsgruppe Bodeninformationssystem 1989).

Eine zweite Anpassung der Kartieranleitung für eine digitale Bearbeitung muss im Bereich der Feldaufnahme und der Berechnung von abgeleiteten Grössen geschehen. Der Grundsatz muss hier lauten, dass alle Daten in der Form abgelegt werden, in der sie aufgenommen wurden. Das Beispiel der Bodenart (Leser & Klink 1988, 69ff) soll dies verdeutlichen.

Die Aufnahme der Bodenart erfolgt nach Tabelle 12 der bodenkundlichen Kartieranleitung (Arbeitsgruppe Bodenkunde 1982). Die anschliessende Klassierung (Leser & Klink 1988, Tab. 1) reduziert den Informationsgehalt massiv. Das wäre tolerierbar, wenn nicht von diesen Klassen weitere Grössen abgeleitet würden (Luftkapazität, nutzbare Feldkapazität, Durchlässigkeit). Dadurch ergibt sich die statistisch nicht erklärbare Situation, dass aus einer Bodenartenklasse zwei Luftkapazitätsklassen entstehen können (Leser & Klink, Tab. 9). Mit Ordinalskalen können keine arithmetische Rechenoperationen durchgeführt werden (Klasse 2 + 1 ergibt nicht Klasse 3!). Diese Regel wird bei der Bestimmung der Klasse der nutzbaren Feldkapazität verletzt (Leser & Klink 1988, 104ff). Es werden dort je nach Humusgehalt und Gründigkeit halbe und ganze Klassenzu- und -abschläge verlangt. Halbe Klassen stellen die gewählte Klassierung selbst in Frage. Ein vertretbares Vorgehen wäre die Berechnung aus den Rohdaten: Für jeden Horizont wird aus der Bodenart die nFK in mm abgeleitet. Dieser Wert wird je nach Humusgehalt und Skelettgehalt variiert. Anschliessend werden die Horizontwerte aufsummiert. Erst dieser letzte Wert, der immer noch in mm oder l/m^2 vorhanden ist, wird einer Klasse zugeordnet. Entsprechende Berechnungsarten finden sich in der bodenkundli-

chen Kartieranleitung (Arbeitsgruppe Bodenkunde 1982, 145ff), in der Geländepraktikumsanleitung des Instituts für Bodenkunde und Waldernährungslehre der Universität Freiburg (Hädrich, Lamparski & Trüby 1989) oder bei Schlichting & Blume (1966, 35f).

Wird das Prinzip der rohdatenorientierten Abspeicherung genau befolgt, so reduziert sich die Boden- und Bodenwasseraufnahme für die GÖK 25 (Leser & Klink 1988, Kap. 2.2 und 2.3) auf folgende Größen: Bodenart (nach Horizonten), Skelettgehalt in %, Gründigkeit in cm, Humusgehalt, Humusform, pH-Wert (nicht Säurestufe), Substrat, Bodentyp, Grundwasserflurabstand in cm, Staunässeinfluss in cm und ökologischer Feuchtegrad. Der ökologische Feuchtegrad stellt einen Sonderfall dar, weil er auf einer Pflanzenaufnahme beruht. Können die Daten der Pflanzenbestimmung in einer passenden Form abgespeichert werden, so muss der ökologische Feuchtegrad weggelassen werden, weil er abgeleitet werden kann.

Der Vorteil einer digitalen GÖK gegenüber fachspezifischen Informationssystemen ist, dass die Redundanz vermindert werden kann. Dazu ein Beispiel aus einem Bodeninformationssystem:

Die Klimaparameter Jahresmitteltemperatur und Jahresniederschlag gelten als wichtige Indikatoren für Bodenbildungsprozesse, weshalb sie bei jeder ausführlichen Bodenprofilaufnahme angegeben werden (s.z.B. Standortaufnahmeformular von T.Mosimann 1986 in Leser & Klink 1988, 311). Weil aber nicht zu jedem Profilstandort langjährige Temperatur- und Niederschlagsmessreihen vorliegen, beziehen sich die Werte meist auf die nächstgelegene Klimamessstation. Werden die Klimawerte mit den Bodenwerten abgespeichert, so wird der gleiche Stationswert in verschiedenen Datensätzen vorhanden sein. Weitere mehr oder weniger überflüssige Größen für die Bodenkennzeichnung sind die Kartenummer (Oelkers 1984, 18), das Forstamt (Oelkers 1984, 19), der Klimaraum (Oelkers 1984, 25), der Naturraum (Oelkers 1984, 25f) etc., weil sie alle aus den Koordinaten abgeleitet werden können. In einer digitalen GÖK können solche Angaben entsprechend der Herkunft der Daten gespeichert werden, also z.B. in einer Relation mit den Klimamessstationen. Diese Relation enthält Angaben über die Lage (Koordinaten, Höhe) der Station, die Messwerte und die Messdauer. Durch Einspeisung der Stationsdaten in ein GIS wird eine grobe Klimakarte erstellt, die durch Überlagerung mit einem Höhenmodell oder mit Satellitendaten weiter differenziert werden kann. Wenn nun Aussagen über die Bodenbildungsfaktoren gemacht werden sollen, dann wird diese Karte bessere Ergebnisse liefern als die Temperatur- und Niederschlagswerte, die zusammen mit den Bodenprofilen abgespeichert wurden.

Zusammenfassend können folgende Grundsätze für die Datenerfassung und Datenspeicherung bei einer digitalen GÖK aufgestellt werden:

1. Die Felddaten werden nach allgemein verbreiteten Methoden erhoben.
2. Daten werden entsprechend ihrer Herkunft abgespeichert: Messwerte in Verbindung mit der Lage des Messpunktes in normalisierten Relationen, Flächendaten nach Informationsschichten getrennt als digitale Karten.
3. Klassierungen werden nach den Regeln der Statistik aus Rohdaten vorgenommen.

4.2 Datenauswertung und Datendarstellung

Werden die oben aufgestellten Forderungen erfüllt, so wird die Datenauswertung sehr vereinfacht. Die Ausscheidung geoökologischer Einheiten könnte z.B. wie folgt ablaufen: Alle punkthaft vorhandenen Daten werden – in entsprechende Klassen eingeteilt oder in ursprünglicher Form – als Karten dargestellt. Diese Karten werden mit den schon bestehenden digitalen Karten verschnitten, wobei sie nach ihrer landschaftshaushaltlichen Bedeutung gewichtet werden können. Damit wäre bereits eine Geoökotopenkarte erstellt. Die Kennzeichnung der einzelnen Areale kann erzeugt werden, indem die Ausgangskarten mit der Geoökotopenkarte überlagert und die Bandbreiten der einzelnen Faktoren innerhalb der Geoökotope ermittelt werden.

Die Ausscheidung geoökologischer Einheiten ist aber nur eine von unzähligen Auswertungs- und Darstellungsmöglichkeiten. Darüber hinaus kann in dieses Datenhaltungsmodell jede weitere punkt- oder flächenhaft erhobene Messgrösse integriert und mit den schon vorhandenen Daten verbunden werden.

Eine digitale GÖK bildet eine gute Grundlage für die Modellierung der Auswirkungen von Planungsentscheidungen. Dazu werden über die eher statischen Ausstattungskarten eine dynamische Entwicklungsebene und eine strategische Entscheidungsebene gelegt. Es lassen sich so verschiedene Szenarien durchspielen. Bereits implementierte Anwendungen beschreiben Grossmann & Schaller 1990.

Für die digitale GÖK wird es notwendig sein, ein Set von Karten zu definieren, das standardmässig vom System bereitgestellt werden soll. Für diese Karten müssen anschliessend sogenannte Makros oder ins GIS eingebundene Prozeduren erstellt werden, damit die Karten quasi auf Knopfdruck erzeugt werden. Die Kartendarstellung lässt sich durch Verwendung verschiedener Symboltabellen noch variieren (s. Barthelme 1989, 208f), so dass auf verschiedene Benutzergruppen zugeschnittene Kartenbilder erstellt werden können.

Für die Interpretation der abgeleiteten Karten sollte auf jeden Fall das verwendete Ableitungsmodell dokumentiert werden. Dies kann entweder in einem Erläuterungsheft oder direkt im Informationssystem geschehen. Denkbar wäre z.B., dass eine Kartenableitung nach Bedarf in Schritte aufgeteilt werden kann. Ein Anwender kann somit nachvollziehen, aus welchen Daten eine abgeleitete Karte entstanden ist. Durch hinzufügen von Algorithmen zur Berechnung der Aussageschärfe bzw. Genauigkeit der Daten könnte das System noch transparenter gemacht werden.

Alle diese Möglichkeiten, die eine digitale GÖK bietet, können aber nur ausgeschöpft werden, wenn einerseits die nötigen Daten vorhanden sind, andererseits aber leistungsfähige Softwareprodukte für Datenverwaltung und -darstellung verwendet werden. Eine denkbare Kombination wäre z.B. die GIS-Software "SPANS" von TYDAC im Zusammenspiel mit dem Datenverwaltungsprogramm "ORACLE" auf dem Betriebssystem OS/2. Es können natürlich andere Softwarepakete eingesetzt werden. Die vorgeschlagene Kombination bietet aber alle in diesem Konzept erwähnten Funktionen und ist zudem auf andere Systeme (z.B. UNIX) übertragbar.

5. Zusammenfassung

Im Rahmen dieser Arbeit wurden zwei Datenverwaltungsprogramme für das Geografische Institut der Universität Basel erstellt. Anhand dieser zwei Beispiele und theoretischer Überlegungen wird der Computereinsatz im Geografiestudium und in der physiogeografischen Forschung diskutiert. Mit dieser Arbeit soll zur vermehrten Zusammenarbeit im Bereich der elektronischen Datenverarbeitung angeregt werden. Es wird eine vereinheitlichte Datenverwaltung und -dokumentation sowie ein EDV-Konzept für langfristige Ansprüche gefordert.

Im zweiten Kapitel werden Vorschläge zum Computereinsatz im Geografiestudium gemacht. In den Anfängerübungen, bei Seminararbeiten oder beim Geländepraktikum kann der Umgang mit dem Computer eingeübt werden, so dass der Computer nicht mehr als undurchschaubare Maschine, sondern als Werkzeug im Arbeitsalltag eines Geografen betrachtet wird. Es wird ein Literaturverwaltungsprogramm vorgestellt, das sowohl im Studium als auch bei der späteren praktischen Arbeit eingesetzt werden kann.

Im dritten Kapitel wird ein Datenverwaltungskonzept für die physiogeografische Forschung vorgeschlagen. Dieses Konzept lässt sich leicht an verschiedene Messanordnungen anpassen, es ermöglicht jede gewünschte Datenauswertung, es erleichtert die Datenübersicht und die Datenbeschreibung, und es verhindert redundante Speicherung. Grundlage dazu bilden das Relationenmodell und das Entity-Relationship-Modell.

Als praktische Anwendung dieses Datenmodells wurde ein Bodendatenverwaltungsprogramm auf "dBase IV" entwickelt. Dabei stellte sich heraus, dass "dBase IV" für grössere Anwendungen schlecht geeignet ist, weil es bei einer grossen Anzahl Relationen viel zu langsam wird. Ein weiteres Problem ergab sich dadurch, dass versucht wurde, drei verschiedene Datenaufnahmeverfahren in einem System zu vereinigen. Eine einheitliche Datenerfassung bei der Bodenkartierung könnte hier Abhilfe schaffen, weshalb Codelisten und ein neues Bohrprofilprotokoll erarbeitet wurden.

Das letzte Kapitel beschäftigt sich mit der digitalen Bearbeitung der geoökologischen Karte 1 : 25 000 (s. Leser & Klink 1988). Das Konzept der GÖK ist sehr gut geeignet für die Einrichtung eines geografischen Informationssystem für die Planung. Es können aber nicht alle Methoden der Kartieranleitung ohne weiteres übernommen werden. Wichtige Schritte zu einer digitalen GÖK sind die Anpassung der Methoden, die Datenbeschaffung, die Definition von Standardkarten, die Implementation von Planungsmodellen und eine umfassende Dokumentation.

6. Quellen

6.1 Zitierte Literatur

- ARBEITSGRUPPE BODENINFORMATIONSSYSTEM: Vorschlag für die Einrichtung eines länderübergreifenden Bodeninformationssystems. - Hannover (1989), 25.
- ARBEITSGRUPPE BODENKUNDE: Bodenkundliche Kartieranleitung. - Hannover (³1982), 331.
- ASHTON-TATE: Befehle und Funktionen. - DBase IV Dokumentation Ashton-Tate Frankfurt a.M. (1989).
- BARCKHAUSEN J., LOOK E., VINKEN R. & VOSS H.: Symbolschlüssel Geologie. Symbole für die Dokumentation und Automatische Datenverarbeitung (ADV) geologischer Feld- und Aufschlussdaten. - Hannover (²1975), 135.
- BARTELME N.: GIS-Technologie - Geoinformationssysteme, Landinformationssysteme und ihre Grundlagen. - Springer Berlin/Heidelberg/New York/Paris (1989), 280.
- CHEN P.P.: The Entity-Relationship-Model - Towards a Unified View of Data. - In: ACM TODS, 1/1 (1976), 9-36.
- CODD E.F.: A Relational Model for Large Shared Data Banks. - In: Comm. ACM 13/6 (1970), 377-387.
- DETLING W.: Die Genauigkeit geoökologischer Feldmethoden und die statistischen Fehler quantitativer Modelle. - In: Pysiogeographica 11 Geogr.Inst.Uni.Basel Basel (1989), 140.
- DREXLER O.: Symbolschlüssel Bodenkunde Bayerische Geologisches Landesamt. - Manuskript Bayerisches Geol. Landesamt München (1990), 18.
- DREXLER O. & SCHRAA: Datenstruktur Bodendatenbank Bayerisches Geologisches Landesamt. - Manuskript Bayerisches Geol. Landesamt München (1990), 9.
- FINKENZELLER H., KRACKE U. & UNTERSTEIN M.: Systematischer Einsatz von SQL-Oracle. - Addison-Wesley Bonn/Reading/Menlo Park/New York (1989), 497.
- GRELL J. & GRELL M.: Unterrichtsrezepte. - Beltz Weinheim/Basel (1987), 330.
- GROSSMANN W.D. & SCHALLER J.: Connecting Dynamic Feedback Models with Geographic Information Systems. - In: Proceedings of the 4th International Symposium on Spatial Data Handling, Vol. I, Zürich (1990), 501-511.
- HÄDRICH F., LAMPARSKI F. & TRÜBY P.: Beurteilung des Bodens im Gelände - Eine Anleitung zum Geländepraktikum. - Manuskript Freiburg i.Br. (1989), 17.

- HARTGE K.H.: Die physikalische Untersuchung von Böden - Eine Labor- und Praktikumsanweisung. - Enke Stuttgart (1971), 168.
- HARTGE K.H. & HORN R.: Die physikalische Untersuchung von Böden. - Enke Stuttgart (²1989), 175.
- HOSANG J.: Die Anwendung der Kartieranleitung Geoökologische Karte 1 : 25 00 (KA GÖK 25) und der Vergleich mit geoökologischen Messungen auf einem Ausschnitt des Blattes 8211 Kandern der TK 25 (Südbaden). - Diplomarbeit Geographisches Institut Universität Basel (1989), 134.
- KILCHMANN A. & DIEBOLD K.: EDV-Datenbanken und Computer-Kommunikationsnetze für Geoökologen. - In: Karlsruher Geoökologische Manuskripte 2 Karlsruhe (1988), 40.
- LESER H. & KLINK H.-J.: Handbuch und Kartieranleitung Geoökologische Karte I : 25 000 (KA GÖK 25). - In: Forsch.z.dt.Landeskunde 228 (1988), 349.
- LESER H., HAAS H.-D., MOSIMANN T. & PAESLER R.: Diercke - Wörterbuch der Allgemeinen Geographie Band 1 und 2. - DTV/Westermann München/Braunschweig (²1985), 422.
- LESER H.: Landschaftsökologie. - UTB 521 Ulmer Stuttgart (²1978), 433.
- MARKS R., MÜLLER M.J., LESER H. & KLINK H.-J.: Anleitung zur Bewertung des Leistungsvermögens des Landschaftshaushaltes (BA LVL). - In: Forsch.z.dt.Landeskunde 229 (1989), 222.
- MOSIMANN T.: Landschaftsökologische Komplexanalyse. - Wiss.Paperbacks Geographie Steiner Wiesbaden Stuttgart (1984), 115.
- MOSIMANN T.: Böden der Region Basel (südlicher Oberrheingraben und Tafeljura). - In: Materialien zur Physiogeographie 8 Basel (1985), 66.
- OELKERS K.-H.: Datenschlüssel Bodenkunde - Symbole für die automatische Datenverarbeitung bodenkundlicher Geländedaten. - Hannover (1984), 100.
- POHL B.: Datenbanken in Schule, Studium und Beruf. - Karlsruhe (1986).
- SCHLICHTING E. & BLUME H.-P.: Bodenkundliches Praktikum. - Hamburg/Berlin (1966), 209.
- TILEY W.E.: dBase IV Programmieretechniken. - Addison-Wesley Bonn/München/Reading (1989), 639.
- VESTER F.: Denken, Lernen, Vergessen - Was geht in unserem Kopf vor, wie lernt das Gehirn, und wann lässt es uns im Stich? - DTV München (⁸1982), 190.

VETTER M.: Aufbau betrieblicher Informationssysteme mittels konzeptioneller Datenmodellierung. - Teubner Stuttgart (6 1990), 510.

WIEDERHOLD G.: Dateiorganisation in Datenbanken. - McGraw Hill Hamburg (1989), 783.

ZEHNDER C.A.: Informationssysteme und Datenbanken. - In: Leitfäden der angewandten Informatik B.G.Teubner Stuttgart (⁴1987), 274.

6.2 Bei der Programmierung verwendete Literatur

ASHTON-TATE: Befehle und Funktionen. - dBase IV Dokumentation Ashton-Tate Frankfurt a.M. (1989).

ESSER W.: Harvard Graphics - Business-Grafiken professionell gestalten. - Markt & Technik Haar bei München (1990), 356.

FINKENZELLER H., KRACKE U. & UNTERSTEIN M.: Systematischer Einsatz von SQL-Oracle. - Addison-Wesley Bonn/Reading/Menlo Park/New York (1989), 497.

JAMSA K.: MS-DOS Batch Files. - Microsoft Press Redmond (1989), 166.

RUBEL M.C.: dBase IV Power Tools - Advanced Functions, Procedures, and Utilities. - Bantam Toronto/New York/London (1989), 963.

STEELE P.: 85 dBase IV User-Defined Functions and Procedures. - Windcrest Blue Ridge Summit (1989), 217.

STEINER F.-J. & CZERWINSKI M.: Turbo-Pascal-Lexikon: von A bis Z: das umfassende Nachschlagewerk für jeden Turbo-Pascal-Programmierer. - Markt & Technik Haar bei München (1989), 525.

TILEY W.E.: dBase IV Programmieretechniken. - Addison-Wesley Bonn/München/Reading (1989), 639.

ZEHNDER C.A.: Informationssysteme und Datenbanken. - In: Leitfäden der angewandten Informatik B.G.Teubner Stuttgart (⁴1987), 274.

6.3 Für die Arbeit verwendete Software

“dBase IV” von Ashton-Tate 1989

“MS-Word 5.0” von Microsoft 1989

“1st Word plus” von Atari 1990

“Turbo Pascal 5.5” von Borland

“Harvard Graphics” von Software Publishing Corp.

“Ventura Publisher 2.0” von Rank Xerox 1988

“Ventura Publisher Professional Extension 2.0” von Rank Xerox 1988