

# Vergleichende Bewertung und Darstellung der Grundwasserbeschaffenheit mit der Methode der Hassediagrammtechnik (HDT)

Rainer Brüggemann<sup>1</sup>, Bernhard Luther<sup>1</sup> und Hans-Peter Lühr<sup>2</sup>

(veröffentlicht in Wasser&Boden 1+2/2001)

1. Institut für Gewässerökologie und Binnenfischerei  
Abteilung Ökohydrologie  
Müggelseedamm 310  
12587 Berlin-Friedrichshagen

2. Technische Universität Berlin  
Fachgebiet Wasserbau und Wasserwirtschaft  
Edelhofdamm 33  
13465 Berlin

## Zusammenfassung

Die Daten- und Faktendarstellung für die Bewertung von Grundwässern erscheint bisher unzureichend, da sie nicht einheitlich ist, Lücken aufweist und in der Auswahl der Parameter oft nicht nachvollziehbar ist. Insbesondere wird ein erheblicher Handlungsbedarf in der Bewertung und in ihren vorbereitenden Schritten gesehen.

Das typische Arbeitsmittel für georeferenzierte Informationssysteme ist die Kartendarstellung. Karten lassen sich thematisch einfärben, wobei aber eine Beziehung zwischen Parametern des Grundwassers und des Farbcodes erarbeitet werden muss. Üblicherweise werden Parameter aggregiert, um eine einfache Skala zu gewinnen, der dann die Farben zugeordnet werden.

In dieser Arbeit wird ein Instrument vorgestellt, das zumindest ergänzend zu der oft kontrovers diskutierten numerischen Aggregation genutzt werden kann, um zu einer Kartendarstellung und zu einer Visualisierung des Zustands von Probenahmeorten bezüglich der Grundwasserbeschaffenheit zu gelangen. Dabei wird die numerische Aggregation von Daten verschiedener Qualität vermieden. Im Mittelpunkt dieses Verfahrens steht die Hassediagrammtechnik (HDT). Diese wird ausführlich erläutert und eine beispielhafte Bewertung für 74 Grundwasserprobenahme-Orte anhand von zehn Messgrößen durchgeführt.

## Summary

The documentation of facts about groundwaters in order to perform a sitespecific evaluation is still an open topic of research. Typically the visualization of the state of groundwaters is a coloured map. However, to get a relation between parameters, describing the status of groundwater and a scale, from which the colour code can be derived, a numerical aggregation of parameters is needed. This is a point which is often controversially discussed and by which often arbitrariness is put into the documentation process.

This paper shows a path, how to avoid numerical aggregation of parameters. The central point is the concept of partially ordered sets and their visualization by Hasse diagrams.

The so-called Hassediagrammtechnik (HDT) will be explained and it is shown, how a colouring of maps can be established. As an example 74 groundwater sample sites were selected and the evaluation is based on 10 typical parameters .

## **1 Problemlage**

Trotz 25 Jahre intensiver Umweltpolitik in der Bundesrepublik Deutschland und in Mitteleuropa ist festzustellen, dass eine vergleichende und bewertende Daten- und Faktendokumentation, sowie die entsprechende Informationsdarstellung spärlich, lückenhaft und vor allem sehr uneinheitlich sind. Dieses gilt insbesondere für den Bereich „Grundwasser“. Es ist sehr schwer, wenn überhaupt möglich, nachzuvollziehen, warum welcher Parameter für welchen Sachverhalt/Wirkungszusammenhang in welcher Grundwasserregion relevant ist.

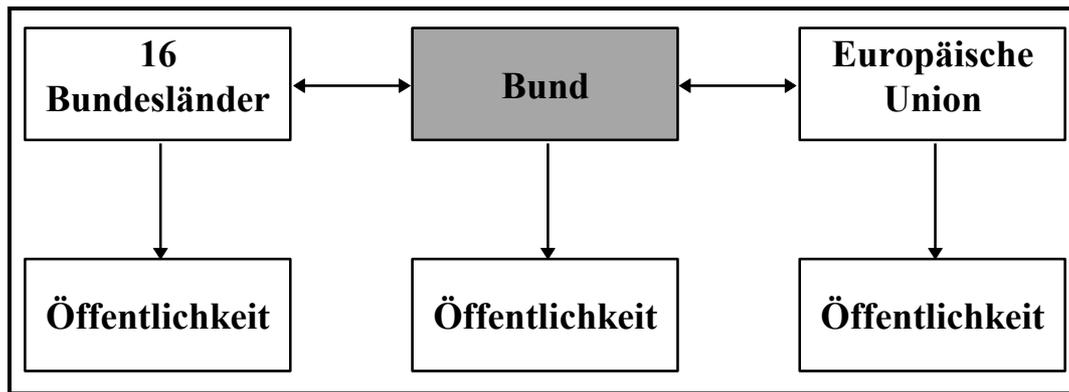
Beispielsweise blieb der Versuch des 1982 für den Umweltschutz zuständigen Bundesministers des Innern über den Wasserversorgungsbericht (N.N. 1982) einen flächendeckenden Situationsbericht zu erlangen, inhaltlich in den Anfängen stecken. Während für die mengenmäßige Darstellung noch einigermaßen vergleichbare Daten erzielt werden konnten, konnte für die gütemäßige Situation des Grundwassers in der Bundesrepublik Deutschland kein einheitliches Bild erstellt werden. Auch die LAWA-Empfehlungen (N.N. 1993) konnten keine zusammenhängende Situationsdokumentation für das Grundwasser realisieren. Grundwasserdarstellungen sind länderspezifisch und weichen in Inhalt und Form teilweise weit voneinander ab, so dass verallgemeinernde und generelle Aussagen einschließlich ihrer Dokumentation über die Grundwassersituation für die Bundesrepublik Deutschland nicht erfolgen konnten und damit insgesamt fehlen.

Zur Zeit gibt es fast ausschließlich thematische Darstellungen der Grundwasserqualität, die sich an Einzelparametern wie Nitrat oder Pestizide festmachen. Dadurch entstehen Kartendarstellungen, die teilweise viele weiße Flächen haben, was aber über die zusammenfassende Grundwasserqualität in der Region/dem Land zunächst einmal nichts aussagt.

## **2 Notwendigkeiten der Darstellung**

Aus Effektivitäts- und Effizienzgründen ist es geboten, einen Datenpool mit harmonisierten Formaten für alle Informations- und Dokumentationsbedürfnisse (Berichterstattung) zu schaffen, aus denen eine objektive Situationsbeschreibung, eine problemangepasste Maßnahmendefinition und eine sachlich begründete Zeitachse für die Realisierung von Maßnahmen abgeleitet werden kann. Dieses hat in einfachen, verständlichen Darstellungen (grafisch und textlich) zu erfolgen, um auch dem Laien (Bürger) ein realistisches, verständliches Bild seiner Umweltsituation und der Prognose zu ermöglichen.

Auf Grund der politischen Struktur hat die Bundesrepublik Deutschland unterschiedliche Aufgaben bei der Daten- und Informationsgewinnung und -weitergabe zu erfüllen (Abbildung 1). Der Bund, vertreten durch den Bundesumweltminister, braucht eine Datenzulieferung aus den jeweiligen Bundesländern, um seinen Verpflichtungen gegenüber der Europäischen Union nachzukommen.



**Abbildung 1: Struktur der Informationspolitik**

Für die Realisierung dieser Struktur ist eine gemeinsame Basis erforderlich. Diese Basis muss einerseits rechtlich fixiert und andererseits fachlich, d.h. problemorientiert sein. Während die rechtliche Fixierung in der Regel unstrittig ist, muss die fachliche Seite erst noch definiert werden. Hierbei ist zwischen wünschbaren und machbaren Aspekten zu unterscheiden. Hinzu kommt, dass das, was für das eine Bundesland von Bedeutung ist, für das andere aber keine oder nur eine untergeordnete Bedeutung hat. Und dieses pflanzt sich auf der EU-Ebene fort.

Unter diesen Aspekten sind im Rahmen der Informations- und Berichtspflichten auf der fachlichen Seite zwei Ebenen zu unterscheiden:

1. Ebene (Mindestinformation): Festlegung der zu charakterisierenden Grundwasserkörper, Mindestinformation zur Charakterisierung der Grundwassersituation für einen bestimmten Grundwasserkörper,
2. Ebene (weitergehende Information): Zusatzinformation für begründete Phänomene in einem bestimmten Grundwasserkörper.

Auf der **rechtlichen Seite** sollte zukunftsorientiert die EG-Wasser-Rahmen-Richtlinie (WRR) die Basis bilden (N.N., 2000). Bei der Definition der fachlichen 1. Ebene kommt es darauf an, die Anforderungen der WRR, sowie die darauf aufbauenden Spezifikationen der Europäischen Umweltagentur (EEA) zu beachten. Die WRR hat erstmalig eine flächige Erfassung und Beschreibung des Grundwassers eingeführt, womit der in der Europäischen Union beschlossene, flächendeckende Grundwasserschutz realisiert wird.

Die verschiedenen Arten der Auswertung und Berichterstattung basieren alle auf der Überwachung des Grundwasserzustandes hinsichtlich des chemischen und mengenmäßigen Zustandes gemäß Art. 8 und 15 der WRR. Die konkreten Anforderungen, die als Mindestanforderungen zu betrachten sind, sind in den Anhängen II und V definiert. In der 2. Ebene können darüber hinaus selbstverständlich aus nationalen Bedürfnissen heraus weitere Zusatzinformationen begründet werden.

Allein das Sammeln von Daten und Informationen, auch wenn man darüber verfügt, sind keine Begründung, diese in den komplizierten Informationslauf aufzunehmen. Hier ist also von allen Beteiligten strengste Disziplin zu üben! Das Informationssystem muss letztlich den flächendeckenden Grundwasserschutz abbilden. Dabei geht es darum, die rechtlich fixierten Pflichten effizient und kostenoptimal zu realisieren.

### **3 Bisherige Formen der Darstellung**

Betrachtet man die Publikationen zur Grundwassersituation, so werden z.B. in den jeweiligen Jahresberichten der Länder, den Berichten der Europäischen Umweltagentur (N.N. 1999a, N.N. 1999b), der Daten zur Umwelt des Umweltbundesamtes (HÖLTER 1998), aber auch anderer Veröffentlichungen grundsätzlich die gewonnenen Primärdaten in Form von Histogrammen, Zeitreihen mit Mittelwertbildung, Trendanalysen, Schwankungsbreiten etc. für die Einzelparameter dargestellt. Hieraus lassen sich z.Zt. keine Aussagen und Darstellungen entsprechend der WRR unmittelbar ablesen und entwickeln. Mit dieser konservativen Auswertung und Darstellung ist eine Vergleichbarkeit der Situation und das Aufzeigen von Entwicklungen in Deutschland und Europa fast unmöglich, so dass Vollzugsdefizite vorprogrammiert sind. Eine Aggregierung zu Grundwassergüteklassen (z.B. guter chemischer Zustand des Grundwassers), wie es die WRR vorsieht, existiert nicht.

Mit den Arbeiten von HANNAPPEL (1996) wird ein erster Schritt in die Richtung unternommen. Danach werden solche Grundwässer anhand von Konzentrationen von Messgrößen sowie Ionenquotienten klassifiziert, die als typisch für eine bestimmte Beeinflussung interpretiert werden (Beeinflussungstypen). Dazu wurden sieben Beeinflussungsklassen definiert, die nach einem hierarchischen Abfragealgorithmus durchlaufen werden, um anhand von Abweichungen von definierten Schwellenwerten den für die Messstelle/den Grundwasserkörper maßgeblichen Beeinflussungstyp herauszufinden. Die Beeinflussungsklassen dabei sind: Geogener Chlorid-Typ, Nährstoff-Typ, Organogener Typ, Versauerungstyp, Sulfat-Typ, diffus beeinflusster Typ, NaHCO<sub>3</sub>-Typ und unbeeinflusster Grundwasser-Typ. Diese Vorgehensweise hat aber noch keinen Einzug in die Praxis genommen.

Mit der Hassediagrammtechnik (HDT) steht eine Methode zur Verfügung, die die Anforderungen an eine vergleichende und bewertende Darstellung von komplexen Sachverhalten erfüllt und gleichzeitig die Primärinformationen in Form der Konzentrationswerte von Einzelparametern bereithält und die Einzelinformationen nicht durch konsensuale numerische Aggregierung miteinander vermischt. Damit ist die HDT u.a. ein geeignetes Instrument, um die Anforderungen der WRR zu erfüllen.

### **4 Methode der HDT**

#### ***4.1 Einordnung des Verfahrens***

Bewertung ist die Verknüpfung von Sachverhalten<sup>1</sup> mit Wertvorstellungen (N.N. 1995). Wertvorstellungen werden in einem Kontext entwickelt, der nicht notwendig in einem naturwissenschaftlichen Sinn stringent ist. Das heißt aber auch, dass sich Wertvorstellungen im Zuge weiterer Erkenntnis und in Abhängigkeit gesellschaftlicher Prioritätensetzungen ändern können. In diesem Sinne ist Bewertung überwiegend konsensual und nicht immer zwingend aus naturwissenschaftlicher Sicht ableitbar. Sind also Objekte, wie im vorliegenden Fall Grundwasser-Messstellen anhand gewisser Messgrößen<sup>2</sup> zu bewerten, so kann man den Bewertungsvorgang in zwei Teile aufspalten:

1. Unabhängig von gesellschaftlichen Wertvorstellungen werden Objekte nach „gut“ und „schlecht“ allein auf der Basis charakteristischer Messwerte bezogen auf Einzelkriterien<sup>3</sup> eingeordnet.
2. Gemäß gegebener Wertvorstellungen werden geeignete Gewichtungen vorgenommen. Das Ziel, das mit der Gewichtung von Einzelkriterien erreicht werden soll, ist, dass anstelle vieler Einzelkriterien eine integrale Maßzahl erarbeitet wird, anhand derer Objekte eindeutig beurteilt werden können.

Der zweite Teil in einem solchen zweistufigen Verfahren ist wichtig, um Entscheidungsprozesse effizient und eindeutig zu gestalten, hat aber die Nachteile

- dass die Bedeutung der einzelnen charakteristischen Messgrößen nicht mehr direkt erkennbar ist und
- dass in die Gewichtung der momentane gesellschaftliche und begrenzte naturwissenschaftliche Wissenstand einfließt.

Dagegen ist die Hassediagrammtechnik (Abkürzung HDT) eine Methode, die basierend auf Gesetzen der Diskreten Mathematik jenen Kern herauszuarbeiten erlaubt, der unabhängig von Wertvorstellungen im Rahmen von Schritt 2 der Bewertung ist und alleine auf der Basis von naturwissenschaftlichen Daten abgeleitet werden kann (BRÜGGEMANN et al. 1996 – 2000). Allerdings erfolgt ein Verlust an metrischer Information, der aber teilweise durch statistische Methoden kompensiert werden kann (Brüggemann et al., 1999).

Die aus der numerischen Aggregation resultierende Größe, die wir  $\Gamma$  nennen wollen, stellt die integrale Maßzahl für ein Objekt  $x$  dar, anhand derer das alle Kriterien umfassende „Gut“- oder „Schlecht“-Urteil im Schritt 2 der Bewertung abgeleitet wird.

---

<sup>1</sup> Unter einem Sachverhalt verstehen wir eine objektive Tatsache wie die allgemeine Gültigkeit des Gravitationsgesetzes oder den Wert einer Messgröße (s.u.), die über ein eindeutig definiertes Messverfahren ermittelt wurde (der Wert der Raumtemperatur wird z.B. ebenfalls als objektive Tatsache angesehen). So ist ein Sachverhalt – im Gegensatz zu einer Wertvorstellung (siehe Text) – *unabhängig* von individuellen oder gesellschaftlichen Präferenzen.

<sup>2</sup> Um von vornherein jedes Mißverständnis zu vermeiden:

Eine *Messgröße* (synonym zu verwenden sind die Begriffe: Attribut, Merkmal) ist z.B. der Alkoholgehalt einer Lösung, ein möglicher *Messwert* dieser Messgröße ist beispielsweise 40%.

<sup>3</sup> Einzelkriterien sind Messgrößen (s.o.), mit denen sehr einfache Wertvorstellungen assoziiert werden können, so dass sie einzeln (jede dieser Messgrößen für sich) als Kriterium für eine einfache Bewertung fungieren können. Die Arsenkonzentration ist etwa zunächst einfach eine Messgröße, bezogen auf Grundwasser-Messstellen würde sie in unserem Sinn zum Einzelkriterium.

## 4.2 Einführende Bemerkungen zur HDT

Zunächst ist sicherzustellen, dass mit einem Satz von Messgrößen  $q_i$  die für die Bewertung charakteristischen Informationen zusammengetragen werden. Dieser Satz von Messgrößen wird daher auch Informationsbasis der Bewertung (Abkürzung: IB) genannt. Anschließend ist zu fragen, wie man Objekte vergleichen kann, die mit einer ganzen Liste von Messgrößen (bzw. Einzelkriterien) beschrieben sind (z.B. für das Grundwasser: Konzentrationen verschiedener Inhaltsstoffe, Leitfähigkeit, Redoxpotenzial u.s.w.). Eine prinzipielle Beschränkung auf Eigenschaften des Wassers besteht nicht. Würde man für die Bewertung einer Messstelle bezüglich des Zustands des Grundwassers auch nicht – wasserbezogene Informationen als wichtig erachten, so müßten diese einfach der Informationsbasis hinzugefügt werden. Wichtig ist nur, dass der Bezug zur Messstelle gewährleistet ist.

Würde gemäß Schritt 2 der Bewertung numerisch aggregiert werden, so stünde zwar eine skalare Beurteilungsgröße in Gestalt einer integralen Maßzahl  $\Gamma$  zur Verfügung, aber diese wäre nicht neutral gegenüber Variationen in Wertvorstellungen oder gegenüber Fortschritten in der wissenschaftlichen Erkenntnis.

Die zur HDT führende Verallgemeinerung auf der Basis eines multidimensionalen Vergleiches ist einfach zu bewerkstelligen:

Es wird gefordert:

1. Alle Messgrößen  $q_i$  haben die gleiche „Gut-Schlecht“-Orientierung.
2. Unterstellt, dass große Zahlenwerte „schlecht“ bedeuten, so soll weiter gelten: Sind in einem Objekt  $x$  *alle* Messwerte  $q_i(x)$  größer als die entsprechenden zu Objekt  $y$  gehörenden Messwerte  $q_i(y)$ , so ist Objekt  $x$  „schlechter“ als Objekt  $y$ . Es werden auch Gleichheiten zugelassen; damit ein Objekt  $x$  schlechter als Objekt  $y$  ist, muß dann mindestens ein Messwert  $q_i(x)$  echt größer als  $q_i(y)$  sein.  
(Kurz:  $x > y$  falls  $q_i(x) \geq q_i(y)$  für *alle*  $i = 1, \dots, n$  und  $q_j(x) > q_j(y)$  für *mindestens ein*  $j$ )

Die beiden Forderungen sind gleichbedeutend damit, dass  $x$  mit  $y$  verglichen werden kann, mit dem Ergebnis:  $x$  ist schlechter als  $y$ . Entscheidend ist aber in 2. Die Forderung dass  $q_i(x) \geq q_i(y)$  für *alle*  $i = 1, \dots, n$  gelten soll. Dadurch sind nicht mehr alle Objekte miteinander vergleichbar, d.h. sie lassen sich auf einer „Gut-Schlecht“-Skala nur noch partiell anordnen, wie folgendes Beispiel zeigt:

Gegeben seien drei Objekte und zwei Messgrößen  $q_1, q_2$ , die gleichsinnig orientiert sind und bei denen große Zahlenwerte einen Hinweis auf „schlecht“ darstellen (Tabelle 1).

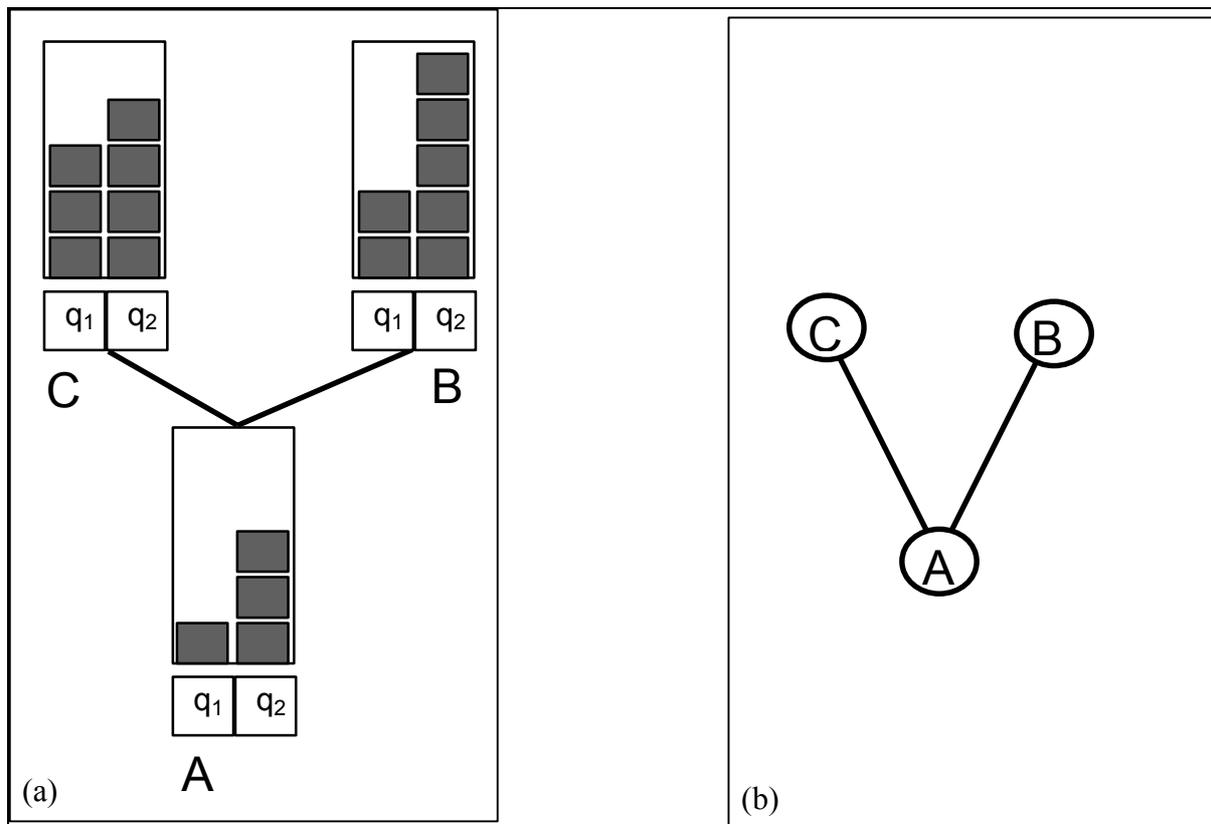
**Tabelle 1: Drei Objekte A, B und C (z.B. Messstellen) und 2 Messgrößen ( $q_1$  und  $q_2$ )**

Objekt \ Messgröße	$q_1$	$q_2$
A	1	3
B	2	5
C	3	4

Das Objekt B ist in allen beiden Messgrößen schlechter als A und analog ist C schlechter als A. Dagegen sind B und C aber nicht vergleichbar: Denn B ist schlechter als C in Bezug auf  $q_2$ , während das umgekehrte gilt für  $q_1$ .

Diese Unvergleichbarkeit weist auf eine charakteristische Datensituation hin. Unabhängig von jeder Wertvorstellung ist aber A besser als B und besser als C.

Das zur Tabelle 1 gehörende sog. Hassediagramm sei ohne weitere Erläuterungen zunächst als ein System von geschickt angeordneten Balkendiagrammen gezeigt, und dann in der üblichen Hassediagrammdarstellung, dem Hassediagramm im engeren Sinn (Abbildung 2). Die graphischen Möglichkeiten, derartige partielle Ordnungen darzustellen, sind ausführlich an anderen Stellen erörtert worden. Für neuere Publikationen hierzu, siehe (BRÜGGEMANN et al. 1999) oder (BRÜGGEMANN et al. 1997).



**Abbildung 2: Hassediagramm mit „organisierten“ Balkendiagrammen (a) und in der üblichen Darstellung (b).**

Die Objekte B und C unterscheiden sich nur durch eine Ausprägungsstufe (ein „Kästchen“), wobei sich dieser Unterschied jedoch entscheidend in der Umwelt auswirken kann. Solange kein kausales mathematisches Modell zur Verfügung steht, das menschliche Gesundheit und Integrität der Umwelt als Funktion von  $q_1$  und  $q_2$  explizit berechnet, so lange ist es besser, diesen Unterschied auszuweisen, als ihn durch konsensuale numerische Aggregation zu verschleiern.

Wie bereits angedeutet, sind Unvergleichbarkeiten ein wesentliches Merkmal der HDT, weil hier die unterschiedliche Qualität oder unterschiedliche Ausprägungen in verschiedenen Dimensionen (hier also in Bezug auf  $q_1$  und  $q_2$ ) evident werden. Unvergleichbarkeiten sind demnach oft wichtiger als die Vergleichbarkeiten, die sich längs der Linien in einem Hassediagramm ableiten lassen (z.B.  $A < C$ , weil in Abbildung 2 von A nach oben eine Linie nach C führt). Eine partiell geordnete Teilmenge von Objekten, für die gilt, dass jedes Objekt mit jedem vergleichbar ist, wird Kette genannt, siehe z.B. (TROTTER 1992) sowie (BRÜGGEMANN & STEINBERG 2000).

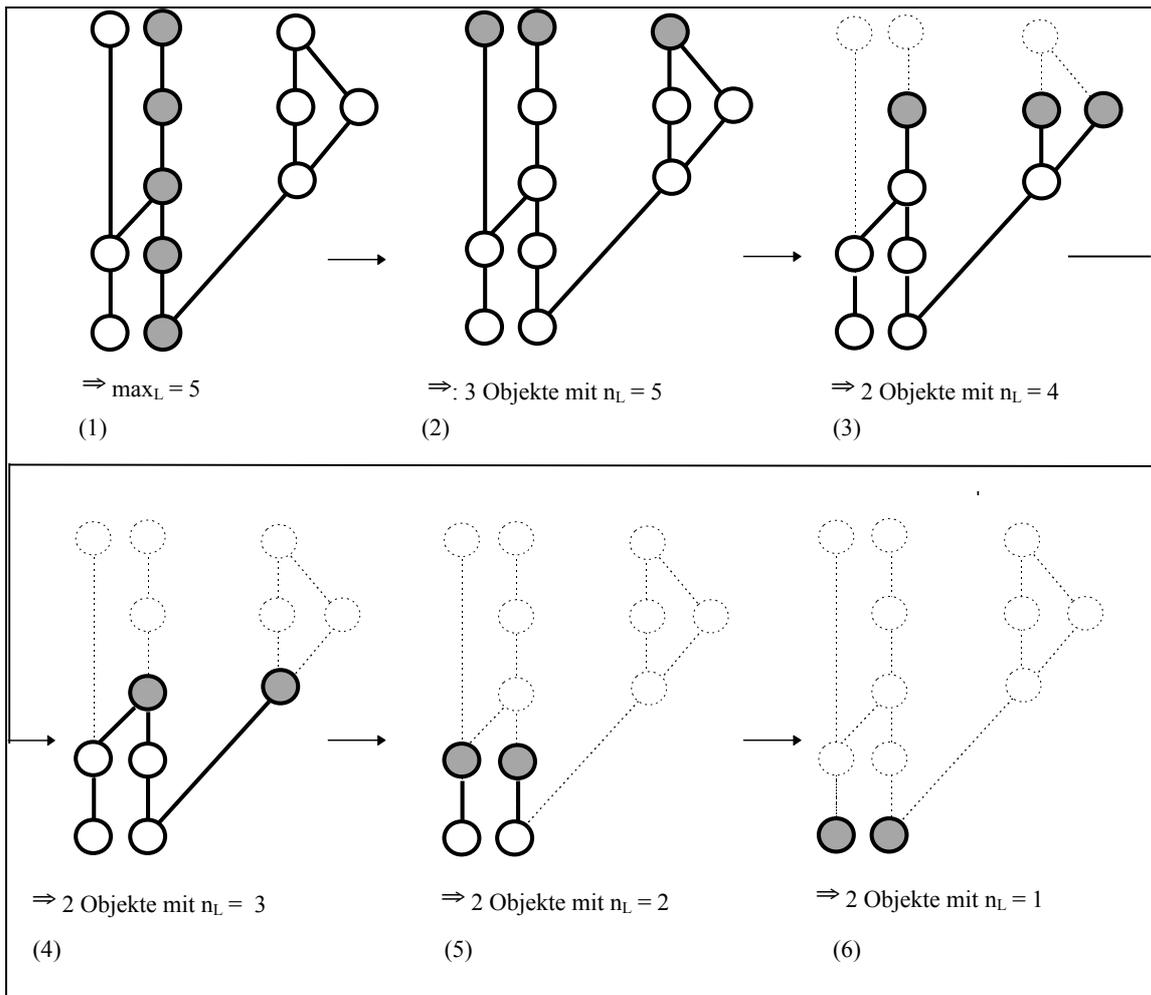
In diesem Sinn stellt ein Hassediagramm den Kern (das Kerndiagramm) dar, der die Sachverhalte für eine Bewertung offenlegt. Zweifelsohne kann man mit zusätzlichem Wissen jederzeit aus einem Hassediagramm eine totale Ordnung via einer integralen Maßzahl  $\Gamma$  erhalten, oder auch nur einige Kriterien gewichtet zusammenfassen. Die Willkür, die in der numerischen Aggregation liegt, kann aber vermieden werden, indem man eine totale Ordnung mithilfe des Konzepts „ordnungserhaltender Abbildungen“ ableitet. Dieses ist Gegenstand des nachfolgenden Kapitels und Grundlage der hier dargestellten Anwendung.

### **4.3 Bewertungsinstrument: Ordnungserhaltende Abbildung**

In der Folge soll ein Verfahren skizziert werden, das ohne numerische Aggregation eine totale Ordnung erzeugt. Ausgangspunkt für diese Bewertungsmethode ist das Hassediagramm, das die partiell geordnete Menge visualisiert. Aus diesem Diagramm wird eine totale Ordnung abgeleitet, die die  $\leq$ -Relationen des Kerndiagramms reproduziert (ordnungserhaltende Abbildung). Dieses stellt eine Kette dar, in der definitionsgemäß alle Objekte mit allen verglichen werden können, was für die Entscheidungsfindung von wesentlicher Bedeutung ist. Diese Methode sieht folgenden Ablauf vor:

- In Kerndiagramm wird die (bzw. *eine*) längste Kette herausgesucht (1 in Abbildung 3).
- Die Objekte dieser Kette werden von unten nach oben durchnummeriert. Das oberste Objekt erhält die Nummer  $\max_L$  (1 in Abbildung 3).
- Alle Objekte, die keine oberen Nachbarn haben, erhalten die Nummer  $n_L = \max_L$  (2 in Abbildung 3).
- Anschließend werden diese Objekte gestrichen (3 in Abbildung 3).
- Im resultierenden Hassediagramm wird es wieder Objekte geben, die keine oberen Nachbarn haben, diese erhalten die Nummer  $n_L = \max_L - 1$  (3 in Abbildung 3).
- Dies wird fortgesetzt, bis mit  $n_L = 1$  alle Objekte erfaßt sind (6 in Abbildung 3).
- Objekte, die die gleiche Nummer  $n_L$  aufweisen, werden in Leveln zusammengefaßt.

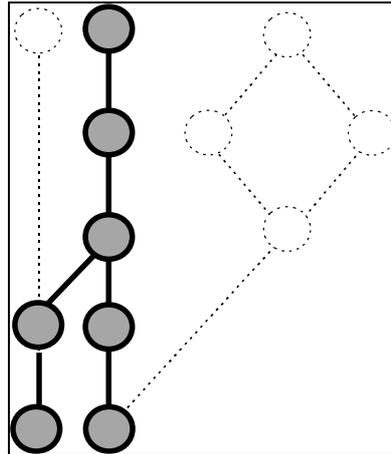
Man kann dieses Verfahren auch mit den untersten Objekten beginnen, dann werden einige Objekte eine andere Nummer erhalten. In diesem Sinn ist das Verfahren, von oben nach unten fortzuschreiten, Konvention.



**Abbildung 3: Zur Konstruktion einer ordnungserhaltenden Abbildung. Die Objekte sind eingeteilt gemäß der  $n_L$  – Zahl in 5 Leveln mit aufsteigendem  $n_L$ -Wert.**

Im hier dargestellten Verfahren werden auf jeden Fall die jeweils ungünstigsten, d.h. also die höchstmöglichen  $n_L$ –Nummern vergeben. Diese für die praktische Arbeit unwesentliche, aber für die Gesamtkonzeption doch willkürliche Vorgehensweise (nämlich entweder von oben nach unten oder umgekehrt zu numerieren) kann vermieden werden. Dazu sucht man die Objekte heraus, deren  $n_L$ -Nummern eindeutig sind. Wir erhalten dann das sog. maximale „graduierte“ Sub-Hasse-Diagramm – zum Begriff der Gradiertheit siehe (BIRKHOFF 1984). In dem Beispiel der Abbildung 3 würden 7 Objekte von 12 übrigbleiben (Abbildung 4).

Die Bewertung im Rahmen von Teil 2 (Kap. 4.1) würde also vom Grad 7/12 willkürfrei sein, 5/12 wären dagegen mit einer gewissen Variabilität zu versehen. Die verbleibende Willkür besteht aber nicht darin, unterschiedliche Qualitäten durch konsensuale Gewichtung miteinander zu verknüpfen, sondern darin, dass man eine pessimistische (von oben nach unten) und eine optimistische (von unten nach oben) ordnungserhaltende Abbildung konstruieren kann.



**Abbildung 4: Maximales graduerte Sub-Hassediagramm**

#### **4.4 Datenaufbereitung in der HDT: Klassifikation der Objekte**

Prinzipiell kann man vermöge des angegebenen Verfahrens aus dem letzten Abschnitt schon die „Rohdaten“ relativ „willkürfrei“ einem ersten Ranking unterwerfen (d.h. *ohne* selbige irgendwie zu „bereinigen“), dies führt aber zu gewissen Problemen. Hinderlich ist dabei nicht nur, dass u.U. ein großer Teil der Objekte verlorengelht (dies muß ja nicht so sein), ein Grundproblem der Hassediagrammtechnik ist, dass geringe numerische Unterschiede in den Daten die gleiche Wirkung haben, wie große Unterschiede. Es sind Verfahren vonnöten, die ähnliche Objekte zu Klassen zusammenfassen – man spricht von Klassifikationsverfahren – um

1. zu robusteren Resultaten zu gelangen und
2. die Hassediagramme zu vereinfachen.

Eine ausführliche Diskussion in Frage kommender Verfahren würde den Rahmen dieser Arbeit sprengen, der geneigte Leser sei auf (BOCK 1974) verwiesen. Hier seien lediglich zwei vom Ansatz her grundverschiedene Wege angesprochen:

1. **Attributweise Klassifizierung:** Man kann für jedes Attribut *einzel*n eine Klassifizierung durchführen. Hierzu zerlege man für Attribut  $q$  das sich aus dem Minimalwert  $\min_q$  und dem Maximalwert  $\max_q$  ergebende jeweilige Intervall  $[\min_q, \max_q]$ , „geeignet“ in Teilintervalle, nummeriere dieselben aufsteigend durch, und ersetze die gemessenen Werte durch die entsprechenden Intervallnummern.
2. **Attributübergreifende Klassifizierung:** Man kann für alle Messgrößen *simultan* eine Klassifizierung durchführen, z.B. durch die Ableitung von Distanzen aus den Rohdaten. Dabei können (und müssen) Unterschiede in den Größenordnungen der einzelnen Messgrößen ebenso ins Kalkül gezogen werden, wie ihre Korrelationen untereinander.

Es bedarf wohl kaum der Erwähnung, dass mit dem ersten Punkt die einfachere Alternative vorliegt, und wir werden ihn nicht zuletzt deswegen beschreiten. In erster Linie soll ja die Hassediagrammtechnik diskutiert werden. (Insbesondere im Zusammenhang mit dem zweiten Punkt sei auf PUDENZ, BRÜGGEMANN, LUTHER, KAUNE und KREIMES (2000) verwiesen.) Dennoch sollte man einen gegebenen Datensatz grundsätzlich auch statistisch charakterisieren, um eine bestimmte Klassifizierungstechnik zu rechtfertigen bzw. zu verdeutlichen, was man von einer solchen zu erwarten hat. In dieser Beziehung werden wir in dieser Arbeit relativ einfach vorgehen: Wir ermitteln für jedes Attribut gewisse wohlbekannte statistische Kennzahlen (Mittelwert, empirische Varianz, e.t.c.) und beschreiben anhand dieser die Einteilungen und die Hassediagramme. Es ist klar, daß neue Objekte (z.B. weitere Messstellen) und zusätzliche Messgrößen eine Neuerstellung des Hassediagramms erforderlich machen. Dies kann aber „mit Knopfdruck“ anhand des EDV-Programmes WHASSE erfolgen und stellt in der Bewertungsanalyse sicher nicht den Engpass dar. Das Programm WHASSE kann beim Erstautor zu Testzwecken und nichtkommerzieller Nutzung angefordert werden. Es kann unter WINDOWS 2000 oder WINDOWS NT installiert werden; ein vereinfachtes Programm, geschrieben in der Programmiersprache JAVA wird derzeit entwickelt.

## **5 Beispiel für Anwendung**

### **5.1 Rohdaten**

Als Beispiel für die Anwendung der HDT zur Grundwasserbewertung lagen anonymisierte Daten des UBA zugrunde, und zwar wurden an 74 Messstellen zehn Messgrößen des Grundwassers vor Ort erfasst: Die Leitfähigkeit (LEITF), sowie die Konzentrationen von  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{NH}_4^+$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{HCO}_3^-$  und  $\text{NO}_3^-$ .

Bei den Messstellen handelt es sich um ein Teilgebiet in Deutschland. Eine absolute Zuordnung dieser Messstellen liegt den Autoren zur Stunde nicht vor. Anhand der Gauß-Krüger-Koordinaten kann jedoch die relative Lage der Messstellen zueinander in einer „Pseudokarte“ dargestellt werden.

### **5.2 Statistische Charakterisierung**

Betrachtet man die Menge  $P = \{1, \dots, 74\}$  der gegebenen Messstellen als Stichprobe einer imaginären Grundgesamtheit (und dies ist natürlich ein heuristischer Ansatz, da kein wahrscheinlichkeitstheoretisches Modell vorliegt), so können für jede Messgröße  $q$  standardmäßig die folgenden statistischen Kennzahlen erhoben werden:

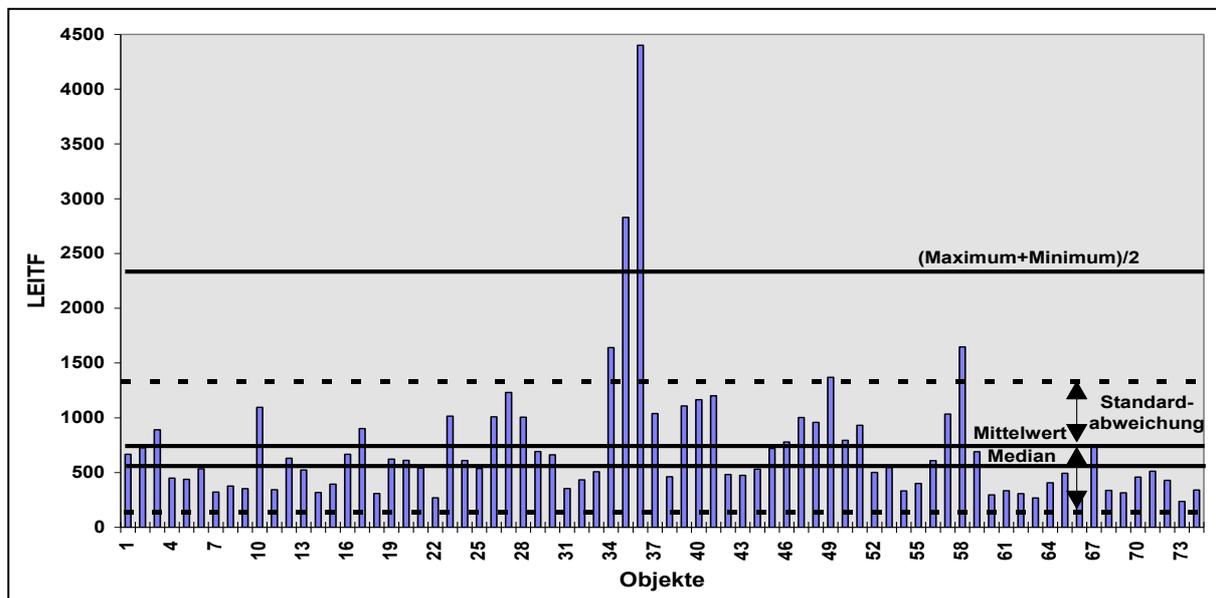
1. das Maximum  $\max_q$ ,
2. das Minimum  $\min_q$ ,
3. der Mittelwert  $m_q$ ,
4. den Median  $\text{med}_q$ ,
5. der Wert  $m_{i_q} = (\max_q + \min_q) / 2$  und
6. die empirische Standardabweichung  $s_q$ .

Die statistischen Kennzahlen Mittelwert  $m_q$ ,  $med_q$ , und  $mi_q$  zerteilen das Intervall  $[\min_q, \max_q]$ , in dem sich alle gemessenen Werte für Attribut  $q$  befinden, an jeweils unterschiedlichen Stellen, so dass man mit dem Ziel einer Robustifizierung eine Vergrößerung des Datenmaterials dadurch erreichen kann, dass man jeden konkreten Messwert  $q(x)$  durch 0 oder 1 ersetzt und zwar je nachdem, ob er sich über- oder unterhalb von  $m_q$  ( $med_q$ ,  $mi_q$ ) befindet (vertiefte Diskussionen allgemeinerer Klassifizierungsverfahren sind z.B. in (BOCK 1974) zu finden).

Die resultierenden drei möglichen Einteilungen der Objekte lassen sich wie folgt charakterisieren:

1. **Einteilung mit  $mi_q$ :**  $[\min_q, \max_q]$  wird in 2 Teilintervalle gleicher Länge geteilt, da  $mi_q$  sich genau in der Mitte zwischen  $\min_q$  und  $\max_q$  befindet. Der Wert  $mi_q$  spielt natürlich im Zusammenhang mit **Ausreißern**<sup>4</sup> eine große Rolle, da er nur von den beiden extremsten Werten bezüglich eines Attributes abhängt.
2. **Einteilung mit dem Mittelwert  $m_q$ :** Im oberen Intervall liegen die überdurchschnittlichen Werte, im unteren Teil die unterdurchschnittlichen.
3. **Einteilung mit dem Median  $med_q$ :** Im oberen Intervall liegt die obere Hälfte der Objekte (einige mit unterdurchschnittlichen Werten) im unteren Teil der Rest.

Im vorliegenden Datensatz sind recht viele obere Ausreißer „mit bloßem Auge“ zu erkennen, für alle Attribute  $q$  gilt zudem die Relation  $med_q < m_q < mi_q$ . In Abbildung 5 sind exemplarisch alle Messwerte für die Leitfähigkeit aufgeführt, um die hier typische Konstellation darzustellen.



**Abbildung 5: Messwerte für die Leitfähigkeit (Balkendiagramm)**

<sup>4</sup> Ein Ausreißer  $x$  bezüglich eines Merkmales  $q$  (kurz:  $q$ -Ausreißer) ist ein Objekt mit extrem hohen (oberer  $q$ -Ausreißer) oder extrem niedrigen (unterer  $q$ -Ausreißer)  $q$ -Wert. Dieser Begriff ist natürlich zunächst rein intuitiv, ebenso wie die Formulierung „extrem hoch/niedrig“.

Die Ermittlung der bivariaten Pearson-Korrelationen unserer „Stichprobe“ stellt eine wohlbekannte Möglichkeit dar, Zusammenhänge zwischen verschiedenen Größen zu entdecken – in Tabelle 2 sind dieselben aufgeführt. Signifikanztests auf den Niveaus 0,01 und 0,05 lassen erkennen, dass „signifikante“ Korrelationen hier durchweg positiv sind (die größte Korrelation ergibt sich für die Chloridionen-Konzentrationen und die Leitfähigkeit). Dies ist ein wichtiges A-priori-Resultat, mit dem sichergestellt werden soll, dass keine stark antagonistischen Messgrößen die ordinale Analyse beeinträchtigen (eine signifikante negative Korrelation einer Messgröße mit allen anderen spricht dafür, diese Messgröße in der ordinalen Analyse um der gleichen Orientierung Willen mit  $-1$  zu multiplizieren, bzw. die Berücksichtigung in der Informationsbasis noch einmal zu überdenken).

Pearson-Korrelation der Attribute										
	LEITF	CA	MG	NA	K	CL	SO4	HCO3	NO3	NH4
LEITF	1	0,549	0,801	0,878	0,294	0,921	0,366	0,475	-0,026	0,184
CA	0,549	1	0,777	0,104	0,297	0,235	0,86	0,312	-0,036	-0,006
MG	0,801	0,777	1	0,504	0,368	0,582	0,564	0,546	-0,023	0,182
NA	0,878	0,104	0,504	1	0,136	0,955	-0,052	0,428	-0,053	0,21
K	0,294	0,297	0,368	0,136	1	0,126	0,3	0,29	0,325	0,069
CL	0,921	0,235	0,582	0,955	0,126	1	0,046	0,315	-0,067	0,177
SO4	0,366	0,86	0,564	-0,052	0,3	0,046	1	-0,011	0,027	-0,069
HCO3	0,475	0,312	0,546	0,428	0,29	0,315	-0,011	1	-0,121	0,266
NO3	-0,026	-0,036	-0,023	-0,053	0,325	-0,067	0,027	-0,121	1	-0,098
NH4	0,184	-0,006	0,182	0,21	0,069	0,177	-0,069	0,266	-0,098	1

**Tabelle 2: Pearson-Korrelationen der Attribute**

**schwarz:** Die Korrelation ist auf dem Niveau von 0,01 signifikant.

**Grau:** Die Korrelation ist auf dem Niveau von 0,05 signifikant.

### 5.3 Hassediagramme der Messstellen

Davon ausgehend, dass die gegebenen Messgrößen sämtlich die gleiche „Gut-Schlecht-Orientierung“ besitzen, (Abschnitt 4.2), wobei große Zahlenwerte „schlecht“ bedeuten, kann die Objektmenge  $P = \{1, \dots, 74\}$  von Messstellen anhand dieser Messgrößen (gemäß Abschnitt 4.2) partiell geordnet werden – für die Rohdaten ergibt sich allerdings ein ziemlich unübersichtliches Diagramm, das hier nicht aufgeführt werden soll. Es wird daher eine äquidistante Zweiteilung der „Attributintervalle“  $[\min_q, \max_q]$  (siehe Abschnitt 5.2) durchgeführt, um eine erste Beschreibung für obere Ausreißer zu gewinnen.

Diese Einteilung in zwei äquidistante Intervalle liefert das Hassediagramm auf Abbildung 6, mit den nichttrivialen (mehrelementigen) Äquivalenzklassen aus Tabelle 3, die im Diagramm eingeschwärzt sind. Es besteht aus fünf Leveln, so dass die Grafik aus „geschickt angeordneten Balkendiagrammen“ übersichtlich dargestellt werden kann (sogar in der „pessimistischen“ Darstellung, wo jedes Objekt auf das „höchstmögliche“ Level gesetzt wird). Die Objekte bzw. Äquivalenzklassen, die *nicht* zum maximalen graduierten Subhassediagramm gehören, sind gestrichelt gezeichnet, von den 74 gegebenen Objekten können also 68 willkürfrei auf ein Level gesetzt werden (etwa 92%) nur 6 Objekte sind zu streichen.

Die weiteren beiden in Abschnitt 5.2 angesprochenen Zweiteilungen der Intervalle führen auf wesentlich unübersichtlichere Hassediagramme mit 10 (Mittelwert) bzw. 8 Leveln (Median), und diese lassen sich zudem schlecht mit dem Hassediagramm auf Abbildung 6 vergleichen. Die *äquidistante* Zweiteilung hat den Vorteil, dass sie sich auf einfache Weise verfeinern läßt – z.B. indem man zur äquidistanten *Dreiteilung* der Intervalle  $[\min_q, \max_q]$  übergeht (haben wir bei einer Zweiteilung nur zwei mögliche Werte für ein Attribut (0/1 bzw. „klein“/„groß“) so kommt jetzt ein Zwischenwert („mittel“) hinzu). Das entsprechende Hassediagramm besitzt in unserem Fall 6 Level und 59 Objekte im maximalen graduierten Subhassediagramm (80%) stehen 15 Objekte gegenüber, deren Level *nicht* eindeutig ist – das Verhältnis verschlechtert sich also etwas gegenüber der Intervallzweiteilung. Zum Vergleich: Das maximale graduierte Subhassediagramm des ursprünglichen Hassediagramms (das sich aus den Rohdaten ergibt) besitzt vier Level und enthält nur noch 17 Objekte (23%).

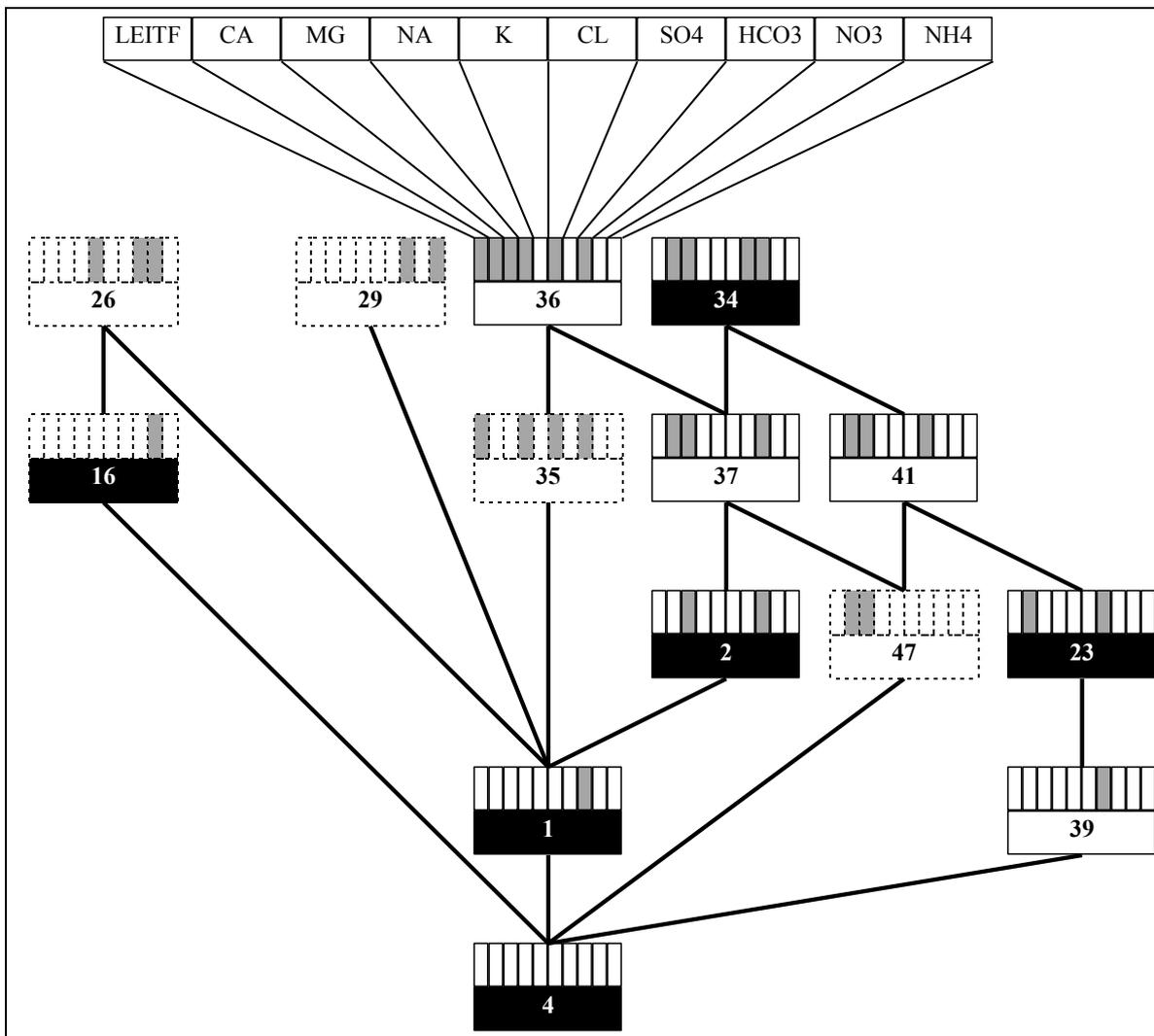
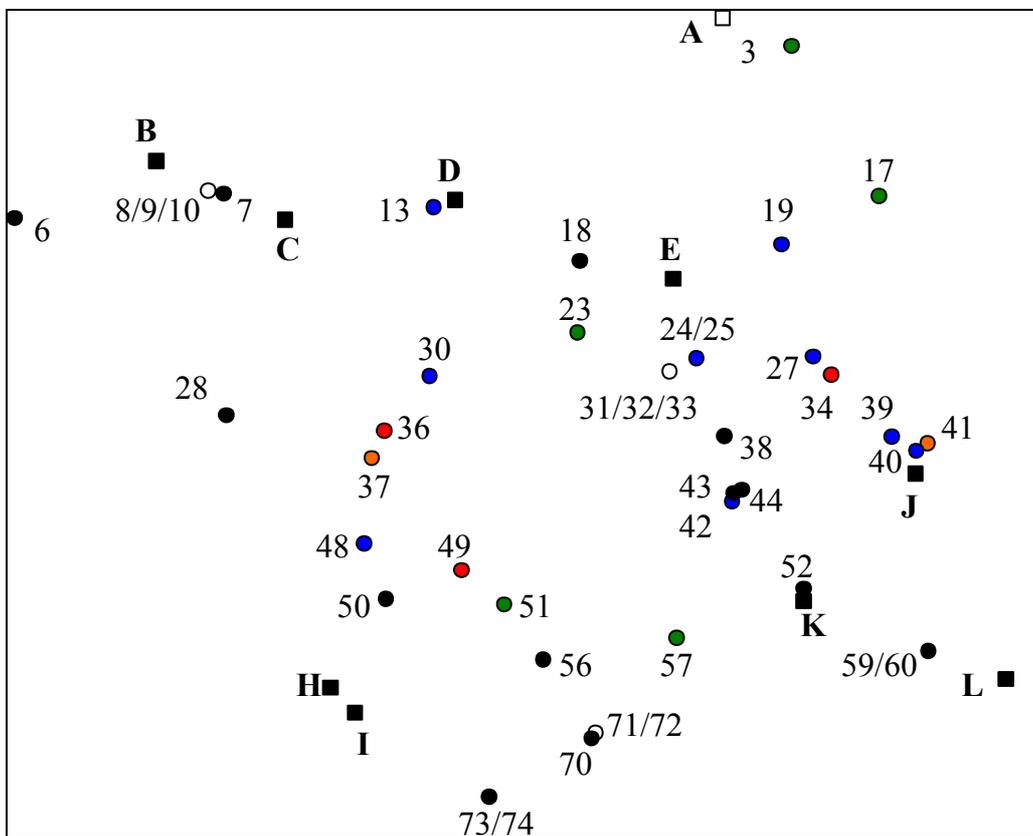


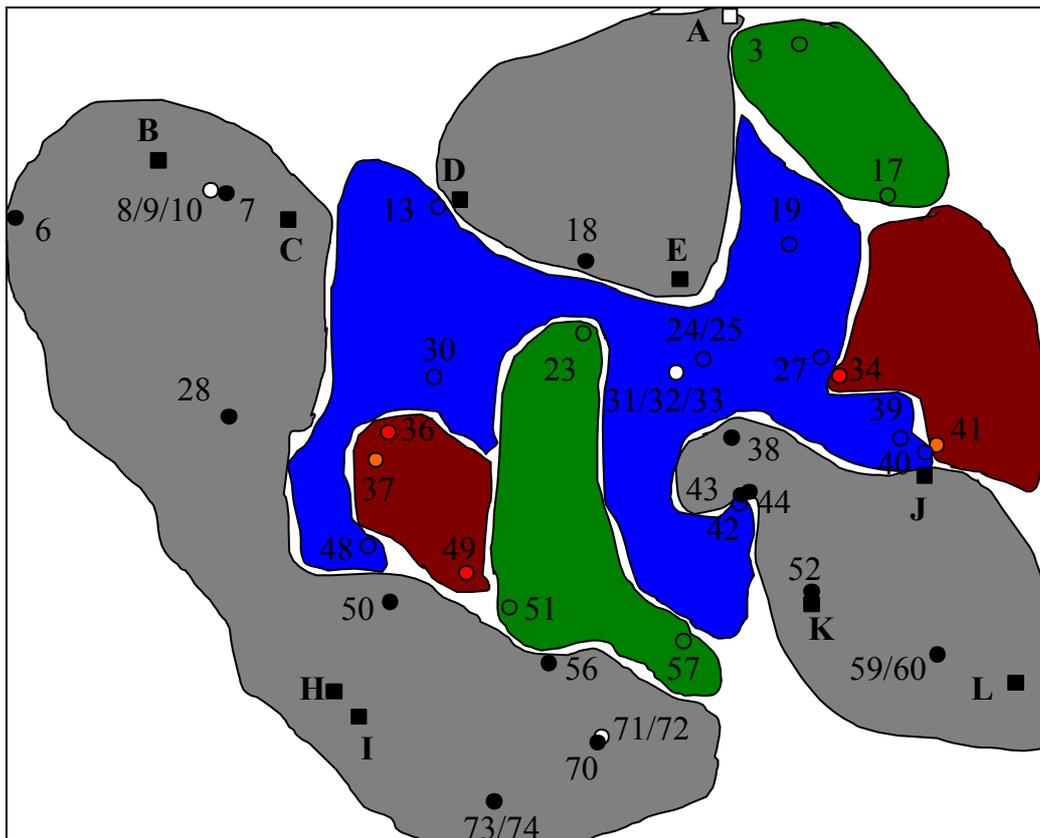
Abbildung 6: Hassediagramm für die Einteilung der Attribute in 2 äquidistante Intervalle (Supergrafik)

Attributliste (Bitmuster)	Anzahl der Elemente	nichttriviale Äquivalenzklasse
HCO3	13	{1, 10, 13, 19, 24, 25, 27, 30, 33, 40, 42, 48, 71}
MG, HCO3	3	{2, 3, 17}
-	42	{4, 5, 6, 7, 8, 9, 11, 12, 14, 15, 18, 20, 21, 22, 28, 31, 32, 38, 43, 44, 45, 46, 50, 52, 54, 55, 56, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 72, 73, 74}
NO3	2	{16, 53}
CA, SO4	3	{23, 51, 57}
CA, MG, SO4, HCO3	3	{34, 49, 58}

**Tabelle 3: Äquivalenzklassen bei Einteilung der Attribute in 2 äquidistante Intervalle (Supergrafik)**



**Abbildung 7: Anhand der Angaben aus dem Hassediagramm (Abbildung 6) eingefärbte Karte. Die Buchstaben bezeichnen nahe beieinander liegende Messstellen, die in dieser Darstellung nicht übersichtlich angeschrieben werden könnten.**



**Abbildung 8: Eine von vielen möglichen „approximativen Gebietseinteilungen“ nach der Vorgabe von Abbildung 7**

Eine „gewohnte“ Kartendarstellung läßt sich aus den graduierten Subhassediagrammen problemlos gewinnen, indem man zunächst die Messstellen gemäß ihres Levels einfärbt (Abbildung 7 setzt die fünf Level des Hassediagramms auf Abbildung 6 wie folgt um: schwarz für das erste, blau für das zweite, grün für das dritte, orange für das vierte und rot für das fünfte Level – die Farbe weiß markiert Messstellen, die wegen konkurrierender Farben innerhalb der gewählten Kartenauflösung nicht gefärbt werden können<sup>5</sup>) und dann per Extrapolation die resultierende Karte in zusammenhängende Gebiete einteilt, in denen jeweils nur Messstellen desselben Levels vorkommen, die die Färbung besagter Gebiete definieren – exemplarisch greift Abbildung 8 die Karte auf Abbildung 7 auf, indem solch eine approximative Einteilung der Pseudokarte angedeutet wird, die die Einzelpunkte als „Stützstellen“ auffaßt. Ein weniger heuristisches Verfahren könnte die Semivariogrammanalyse liefern, die in konventionellen Geostatistikprogrammen zugänglich ist. Das Problem, aus einer Punktmessung auf die Eigenschaften einer umgebenden Fläche zu schließen ist nicht spezifisch für die HDT, sondern gilt auch für alle konventionellen Darstellungen.

<sup>5</sup> Z.B. befindet sich Messstelle 8 auf Level 2, Messstelle 9 hingegen auf Level 3, obwohl beide Messstellen die gleichen Koordinaten haben.

Selbstverständlich würde man mehr Datenmaterial (u.a. topologische Informationen) bezüglich der Messstellen benötigen, um eine solche Einteilung in Gebiete auf einer empirischen Grundlage durchzuführen, doch die unten erstellte Karte mag speziell die Möglichkeit einer geographisch bezogenen Bewertung illustrieren. Die darin unterstellte Homogenität bzw. Heterogenität sollte sich geochemisch/geologisch interpretieren lassen. Zusätzlich könnten auch die Objekte einbezogen werden, die nicht dem maximalen graduierten Subposet angehören. Diese müssen dann mit der zusätzlichen Angabe der möglichen Levelalternativen charakterisiert werden.

## 6 Ausblick

Dokumentation und Berichterstattung stellen eine wesentliche Grundlage für fachliche und politische Entscheidungen sowie für die Information der Öffentlichkeit dar. Diese müssen einfach, durchschaubar und informativ sein, um zu überzeugen. Insbesondere in Bereichen vielfältiger Datenkategorien und komplexer Zusammenhänge, wie es für das Grundwasser gilt, kommt den methodischen Ansätzen für eine objektive und vergleichende Darstellung besondere Beachtung zu. Angesichts des großen Stoffspektrums, der unterschiedlichen geologischen Strukturen, sowie der regionalspezifischen Hydrologie können verschiedene Grundwasserkörper oder Teile davon eine gleich gute oder schlechte Qualität aufweisen, die aber jeweils durch unterschiedliche Stoffe, geologische Strukturen oder hydrologische Gegebenheiten bedingt sind. Im Rahmen einer flächendeckenden, vergleichend bewerteten Darstellung der Grundwasserbeschaffenheit reichen thematische Einzelkartendarstellungen dafür nicht mehr aus.

Auf der fachlichen Ebene sind dazu noch eine Reihe von Arbeiten zu leisten, die aber mit dem methodischen Ansatz nichts zu tun haben. Hierzu zählen u.a. insbesondere die Definitionen für die Klassenbildungen bei den einzelnen in die Bewertung eingehenden Parametern als auch über die zusammenfassende Bewertung des Zustandes des Grundwassers insgesamt, wie es die EU-Richtlinie vorsieht.

Die HDT kann ein wertvolles Instrumentarium in der vergleichenden Bewertung und explorativen Datenanalyse sein. Es können Verwandtschaften bzw. Unterschiede herausgearbeitet werden, die nicht durch konsensuale numerische Aggregationen verschleiert werden. Man muss sich aber klar sein, dass eine ordinale Interpretation der Messgrößen vorgenommen wird. Für die vergleichsweise junge Methode der HDT sind in Konsequenz eine Reihe von Fragen noch zu klären. So ist beispielsweise das geeignete Verfahren der Clusteranalyse herauszufinden, um die geeignete Intervalleinteilung für die Gruppierung der vorliegenden Objektmenge vornehmen zu können. Sie hängt jeweils von der Problemstellung ab. Ferner wurde nicht behandelt, dass mit der HDT auch *nicht*numerische Daten ordinal interpretiert werden können, was wiederum etwa in Gestalt der formalen Begriffsanalyse zu neuen Ansätzen über die logische Struktur der untersuchten Objekte führen kann (GANTER & WILLE, 1996). Eine Zweiteilung der Intervalle, in denen sich die Messwerte numerischer Messgrößen bewegen (so wie sie in Abschnitt 5.2 beschrieben und in Abschnitt 5.3 zur Auswertung herangezogen wurde) stellt beispielsweise (neben einer einfachen Clusteranalyse) auch eine Überführung numerischer Merkmale in binäre Merkmale dar, die auch Fragen nach Implikationen erlauben.

Abschließend kann festgestellt werden, dass mit der Hassediagrammtechnik (HDT), wie sie zuvor an einer realen Situation erprobt wurde, eine Methode zur Verfügung steht, die die Ansprüche einer objektiven und vergleichenden Bewertung erfüllt. Sie erscheint insbesondere geeignet, die Berichtspflichten im Rahmen einer zukünftigen EU-Wasserrahmen-Richtlinie realisieren zu können.

## 7 Literatur

1. Bock, H.H. (1974): *Automatische Klassifikation*. Göttingen:Vandenhoeck&Ruprecht, pp. 6-480.
2. Brüggemann, R., & Bartel, H.-G. (1999): A Theoretical Concept to Rank Environmentally Significant Chemicals. *J Chem Inf Comp Sc*, 39, 211-217.
3. Brüggemann, R., & Halfon, E. (1997): Comparative Analysis of Nearshore Contaminated Sites in Lake Ontario: Ranking for Environmental Hazard. *J Environ Sci Health*, A32(1), 277-292.
4. Brüggemann, R., & Steinberg, C. (2000): Einsatz der Hassediagrammtechnik zur vergleichenden Bewertung von Analysendaten – am Beispiel der Umweltuntersuchungen in den Regionen Baden-Württembergs. In H. Günzler (Ed.), *Analytiker Taschenbuch 21*. (pp. 3-33). Berlin: Springer-Verlag.
5. Brüggemann, R., Bücherl, C., Pudenz, S., & Steinberg, C. (1999): Application of the concept of Partial Order on Comparative Evaluation of Environmental Chemicals. *Acta hydrochim hydrobiol*, 27, 170-178.
6. Brüggemann, R., Voigt, K., and Halfon, E. (1996): Partially Ordered Sets – A computerized Tool to Compare Environmental Databases. In: Environmental Software Systems, Proceedings of the International Symposium on Environmental Software Systems, edited by Denzer, R., Schimack, G., and Russell, D. London:Chapman&Hall, p. 185-195.
7. Hölter, C., Haan (1998): Daten zur Umwelt. Umweltbundesamt, Berlin oder ÖKOBASE 98
8. Ganter, B. and Wille, R. (1996): *Formale Begriffsanalyse Mathematische Grundlagen*. Berlin: Springer-Verlag, pp. 1-286.
9. Hannappel, S. (1996): Die Beschaffenheit des Grundwassers in den hydrologischen Strukturen der neuen Bundesländer. Berliner Geowissenschaftliche Abhandlungen, Reihe A, Band 182, Selbstverlag Fachbereich Geowissenschaften der FU Berlin
10. N.N. (1982): WBB-82: Bericht über die Wasserversorgungsbericht in der Bundesrepublik Deutschland, Herausgeber Bundesminister des Innern, Mai 1982, E. Schmidt Verlag, Berlin
11. N.N. (1993) : LAW-93: Grundwasser-Richtlinien für Beobachtung und Auswertung – Teil 3 – Grundwasserbeschaffenheit, Herausgeber Länderarbeitsgemeinschaft Wasser, 1993
12. N.N. (1995): Methodik der produktbezogenen Ökobilanzen, Wirkungsbilanzen und Bewertungen, In: Texte des Bundesumweltamtes Band 23/95, Berlin 1995
13. N.N. (1999a): EAR-99: Environment in the European Union at the turn of the century, Environmental Assessment Report No. 2, EEA Copenhagen, 1999
14. N.N. (1999b): GQQ-99: Groundwater quality and quantity in Europe Environmental Assessment Report No. 3, EEA Copenhagen, 1999
15. N.N. (2000): Richtlinie 2000/60/EG des Europäischen Parlaments und des Rates vom 23.10.2000 zur Schaffung eines Ordnungsrahmens für Maßnahmen der Gemeinschaft im Bereich der Wasserpolitik, Abl. der EG L327/1, 22.12.2000
16. Pudenz, S., Brüggemann, R., Luther, B., Kaune, A., and Kreimes, K. (2000): An algebraic/graphical tool to compare ecosystems with respect to their pollution V: cluster analysis and Hasse diagrams. *Chemosphere* 40:1373-1382.
17. Trotter, W.J. (1992): Combinatorics and Partially Ordered Sets Dimension Theory. Baltimore, Maryland: The Johns Hopkins University Press, pp. 1-307.