

Hybride Modellierung und Analyse von unsicheren Daten

Michael Glemser, Ulrike Klein
Institut für Photogrammetrie
Universität Stuttgart

1 Einführung

Die Verfügbarkeit digitaler räumlicher Daten hat sich mit steigender Anzahl an Anwendungsbereichen ständig verbessert. So kann heute auf Datenbanken zugegriffen werden, die Basisdaten für vielfältige Zwecke bereitstellen. Beispiele sind amtliche Daten zu den Themen Topographie (z. B. ATKIS), Liegenschaften (z. B. ALK mit ALB), Geologie (z. B. Digitale Geologische Karte, Digitale Bodenkarte), Umwelt (z. B. Umweltinformationssysteme der Länder und Kommunen) und kommerzielle Daten wie Straßendaten (z. B. NavTech, Scout), Satellitenbilder (Landsat, Spot, MOMS) und Marketinginformationen (Kaufkraftdaten, Daten zum Konsumverhalten)). Ein potentieller Anwender kann prinzipiell alle diese Daten nutzen, um seine spezifische Aufgabe zu lösen. Im Idealfall könnte er vollständig auf eine eigene Erfassung verzichten und seine Analysen ausschließlich mit Fremddaten betreiben. Für Anwendungen mit komplexen Fragestellungen ist dazu die Kombination vieler unterschiedlicher Datenquellen notwendig. Betrachtet man das Wachstum der Computernetzwerke, so kann man sich vorstellen, daß Datenaustausch und Zugriff auf externe Datenbanken in naher Zukunft direkt und unmittelbar erfolgen werden. Entsprechende technische Spezifikationen werden momentan dazu erarbeitet (McKee, Kuhn, 1997; OpenGIS, 1999). Der problemlose Zugriff auf an beliebigen Orten und mit unterschiedlichsten Inhalten bereitgestellten Daten birgt aber auch ein beträchtliches Risiko in sich. Es können Daten fehlerhaft genutzt und zu falschen oder widersprüchlichen Ergebnissen kombiniert werden. Aus diesem Grund sind Überprüfungen der Qualität der Daten und darauf basierender Ergebnisse von unerläßlicher Bedeutung.

Von Anwenderseite lassen sich dabei zwei Problemstellungen unterscheiden. Der *Angebotsaspekt* ist für alle Anwender von Bedeutung, die fremde Daten nutzen wollen. Sie haben zu entscheiden, ob die verfügbaren Daten für ihre eigenen Zwecke geeignet sind. Treten konkurrierende Anbieter auf (z. B. im kommerziellen Bereich), so sind die besten Daten zu ermitteln. Dies gelingt durch Ergänzung der Daten um zusätzliche Qualitätsbeschreibungen (Metadaten), anhand denen der Anwender eine Prüfung durchführen kann. Es liegt in der Verantwortung des Anbieters, die entsprechenden Informationen bereitzustellen. Der *Nutzungsaspekt* wird dann relevant, wenn Daten analysiert werden. Die neu erzielten Informationen sind qualitativ zu beurteilen. Dazu ist die Qualität der Eingangsdaten über die Analyse hinweg auf das Ergebnis zu übertragen und als weitere Information mit darzustellen. So können Informationen nach Wert und Güte unterschieden und eventuelle Fehlentscheidungen vermieden werden. Die folgenden Ausführungen konzentrieren sich auf den Nutzungsaspekt, als das schwierigere der beiden Probleme.

2 Geometrische und thematische Unsicherheit

Vektor- und Rasterdaten als wichtigste Datenarten in GIS setzen sich jeweils aus geometrischen und thematischen Anteilen zusammen (Bill, Fritsch, 1991). Bei den Rasterdaten wird die Geometrie durch das Enumerationsverfahren bestimmt, das eine gleichmäßige Unterteilung des Raumes in Rasterzellen vorsieht. Als Verwaltungsstruktur bietet sich eine Matrixform an. Der

Raumbezug ist diskret durch die Position der Zelle im Raster gegeben. Bei den Vektordaten setzt sich die Geometrie auf der Basis geometrischer Primitive (Punkt, Linie, Fläche) zusammen. Dabei kommt das Prinzip der Randbeschreibung zum Einsatz, bei dem ein höherdimensionales Element immer durch seinen Rand repräsentiert wird (z. B. eine Fläche durch ihre Umringslinien). Die Thematik wird bei beiden Datenarten durch Attribute beschrieben. Sie besitzen entweder diskrete oder kontinuierliche Werte. Entsprechend können sie der nominalen oder ordinalen Skala (bei diskreten Werten) bzw. der Intervall- oder Ratio-Skala (bei kontinuierlichen Werten) zugeordnet werden.

Sowohl die Geometrie- als auch die Thematikkomponente sind Einflüssen unterworfen, die Unsicherheiten in den Daten hervorrufen. In den nachfolgenden Abschnitten wird getrennt für beide Komponenten auf Ursachen, Größen und mögliche Parameter eingegangen. Ziel ist dabei, Modelle für jede Komponente zu formulieren, mit denen eine Beschreibung der Unsicherheit in GIS zu realisieren ist.

2.1 Geometrische Unsicherheit

Die Unsicherheit in der Geometrie eines Objektes kann visuell veranschaulicht werden, indem man die Erfassung mehrmals unabhängig voneinander durchführt (Abbildung 1). Überlagert man anschließend alle Erfassungen, so zeigen sich deutliche Unterschiede zueinander. Keine Erfassung gleicht der anderen. Es entsteht vielmehr ein Bereich, in dem die Geometrie streut. Dieser Bereich repräsentiert die Variation und damit die Unsicherheit in der Geometrie des Objektes.

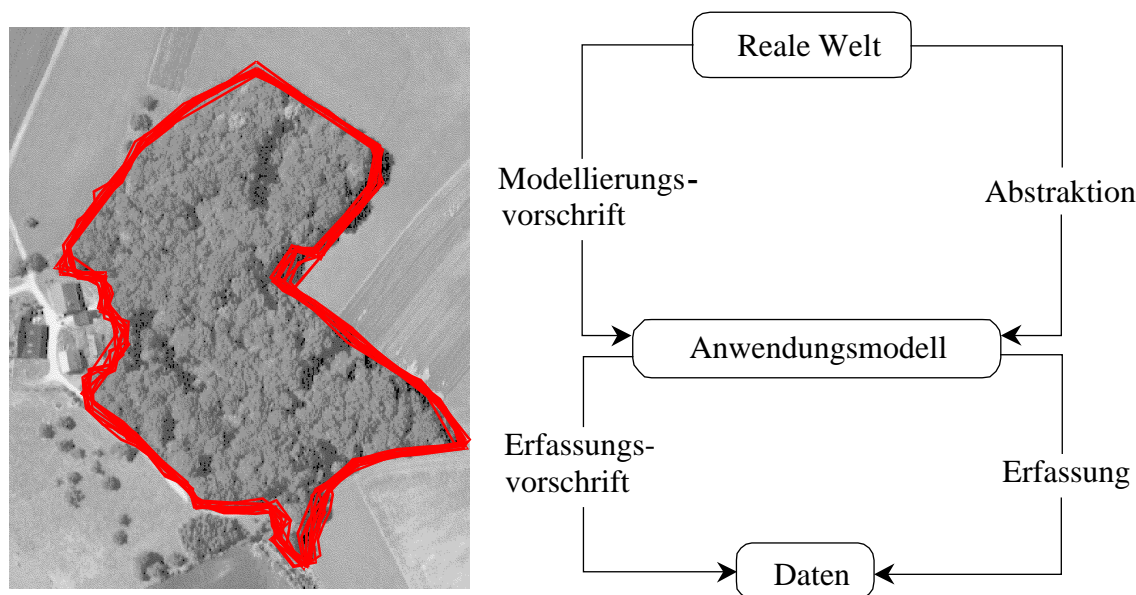


Abbildung 1: Erfassungsvorgang als Quelle für Datenunsicherheit

Es existieren verschiedene Ursachen, die auf die Größe der Variation einwirken. Im besonderen trägt der Erfassungsvorgang hierzu bei. Er definiert eine Abbildung der realen Welt in ein GIS-Modell (Abbildung 1). Wie bei allen Modellbildungen findet auch hier eine Abstraktion statt, die dazu führt, daß Modell und Wirklichkeit Differenzen aufweisen. Der Gesamtvorgang läßt sich in verschiedene Stufen untergliedern, die jeweils charakteristische Einflüsse beinhalten (Caspary, 1992). Ausgehend von der realen Welt ist als erstes das zu erfassende Objekt räumlich von Nachbarobjekten abzugrenzen. Im Anschluß daran ist der gedachte Verlauf der Grenzlinie durch die Bestimmung markanter Punkte in diskreter Form festzulegen. In der letzten Stufe werden dann die gefundenen Punkte digitalisiert. Digitalisierung ist ein Meßprozeß, dessen Genauigkeit von den eingesetzten Geräten und Methoden abhängig ist. Diese Einflüsse

sind im allgemeinen zufälliger Natur. Das Ergebnis der Diskretisierung hängt von der Form der zu erfassenden Grenzlinie (kontinuierlich oder diskret) ab. Ihr Einfluß wirkt sich systematisch auf die Unsicherheit aus und kann bei angepaßter Anzahl und Auswahl der Punkte vernachlässigbar klein gehalten werden. Bei der Abgrenzung spielt die Art des Objektes eine wichtige Rolle. Künstliche (von Menschen geschaffene) Objekte lassen sich besser von anderen Objekten unterscheiden, als dies für natürliche Objekte der Fall ist. Zwischen natürlichen Objekten existiert im allgemeinen ein fließender Übergang, der eine Zone definiert, innerhalb der die Geometrie variieren kann. Weitere Einflüsse ergeben sich durch mögliche Verdeckungen von Informationen oder durch die erfasserspezifische Interpretation.

Wichtig festzuhalten ist, daß alle Geometriedaten zu einem gewissen Grad mit Unsicherheit behaftet sind. Die Berücksichtigung der Unsicherheit erfordert die Bereitstellung eines geeigneten Modells. Eine Reihe möglicher Modelle sind bereits definiert (Glemser, 1994). Als Grundlage für diese Arbeit dient das *stochastische Modell*. Darin wird eine Linie als stochastische Variable (Zufallsvariable) angesehen, deren Variation sich durch eine Verteilung charakterisieren läßt (Caspary, Scheuring, 1992; Glemser, 1996). Als mögliche Verteilung kann die Normalverteilung angenommen werden. Zu deren Beschreibung ist als Parameter die Varianz zu bestimmen. Mit Hilfe der Verteilungsfunktion können dann Wahrscheinlichkeiten für die Zugehörigkeit beliebiger Punkte zum Objekt ermittelt werden.

Die Varianz als wichtigster Parameter der Verteilung kann prinzipiell als ein zusätzliches Attribut angesehen und so verwaltet werden (Drummond, 1995). Eine anschauliche Darstellung ist zu erzielen, wenn auf die erwähnten Wahrscheinlichkeiten übergegangen wird (Kraus, Haussteiner, 1993). Sie definieren eine räumlich-kontinuierliche Wahrscheinlichkeitsfunktion, die zu diskretisieren ist. Die Größe des dazu einzusetzenden Diskretisierungsintervalls Δ ist abhängig von der Varianz σ^2 der Liniengeometrie. Die Festlegung zu $\Delta = \sigma/4$ gewährleistet eine ausreichende Genauigkeit für eine eventuelle Rekonstruktion der Funktion. Mit Hilfe des Intervalls Δ wird über das Objekt ein gleichmäßiges Raster gelegt, dessen einzelne Pixel den Punkten entsprechen, für die die Wahrscheinlichkeiten berechnet werden müssen.

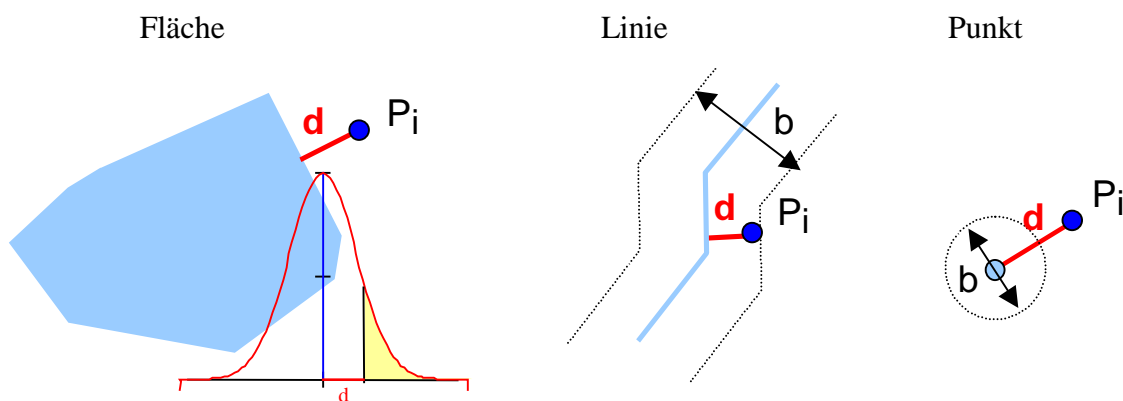


Abbildung 2: Berechnung der Wahrscheinlichkeiten für verschiedene Objekttypen

Die Berechnung der Wahrscheinlichkeiten erfolgt jeweils in Abhängigkeit vom Objekttyp (Abbildung 2):

- Für Flächenobjekte ergibt sich die Wahrscheinlichkeit $p(x, y)$ mit der ein Punkt $P(x, y)$ zum Objekt gehört aus

$$p(x, y) = p(d) = \int_{-\infty}^d f(t) dt = F(d)$$

mit

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}.$$

Bei $f(t)$ handelt es sich um die Dichtefunktion der Normalverteilung und bei $F(d)$ um deren Verteilungsfunktion. Für den Abstand d zur mittleren Grenzlinie gilt, daß er innerhalb des Objektes positiv und außerhalb negativ sein soll.

- Linien- und Punktobjekte benötigen zusätzlich eine Annahme über die Breite b des Objektes in der Wirklichkeit. Entsprechend ergibt sich die Wahrscheinlichkeit $p(x, y)$ bezüglich eines Linienobjektes zu

$$p(x, y) = p(d, b) \cdot p(-d, b) = F\left(\frac{b}{2} + d\right) \cdot F\left(\frac{b}{2} - d\right)$$

mit

$$p(d, b) = \int_{-\infty}^{\frac{b}{2}+d} f(t) dt = F\left(\frac{b}{2} + d\right)$$

mit Breite b des Objektes und Abstand d zur mittleren Linie.

- Bei Punktobjekten berechnet sich die gesuchte Wahrscheinlichkeit $p(x, y)$ aus

$$p(x, y) = p(d, b)^2 \cdot p(-d, b)^2 = F\left(\frac{b}{2} + d\right)^2 \cdot F\left(\frac{b}{2} - d\right)^2,$$

wobei die Wahrscheinlichkeiten $p(d, b)$ und $p(-d, b)$ wie beim Linienobjekt berechnet werden.

Die so entstehenden Wahrscheinlichkeitsmatrizen können auch zum Zwecke der Verwaltung genutzt werden. Sie ersetzen die Varianz als Parameter der Unsicherheit. Ein solcher Ansatz hat bei der Anwendung innerhalb eines hybriden GIS einige Vorteile, auf die in Abschnitt 3 näher eingegangen wird. Abbildung 3 enthält beispielhafte Darstellungen von Wahrscheinlichkeitsmatrizen für die verschiedenen Objekttypen.

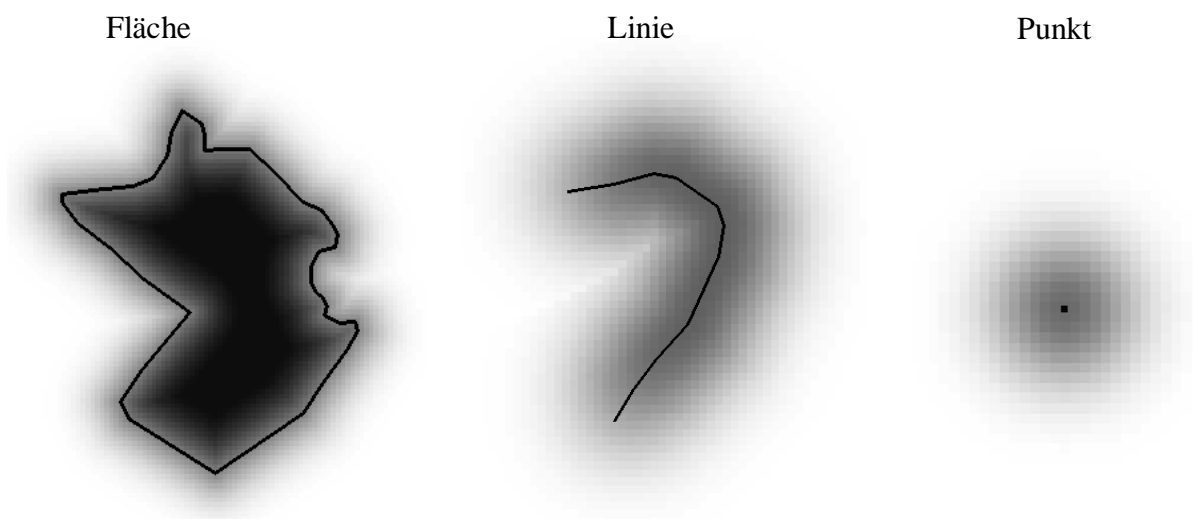


Abbildung 3: Beispiele zur Darstellung der Wahrscheinlichkeiten

2.2 Thematische Unsicherheit

Die Beschreibung der Thematik erfolgt durch Attribute, die entweder diskrete oder kontinuierliche Werte annehmen können. Eine wichtige Datenquelle für die Erstellung thematischer Karten sind multispektrale Fernerkundungsaufnahmen, die als Ausgangsdaten für die nachfolgenden Untersuchungen dienen. Die Auswertung von Fernerkundungsdaten beinhaltet die Definition von Klassen, z. B. Landnutzungsklassen, und die Bestimmung der Zugehörigkeit der flächenhaft erfaßten Ausschnitte der Erdoberfläche zu diesen Klassen. Fehlereinflüsse in beiden Schritten führen zur thematischen Unsicherheit. Bei der Definition der Klassen besteht das Problem, aus der Vielzahl an möglichen Objektarten eine endliche Anzahl von Klassen festzulegen. So stellt sich beispielsweise im Grenzbereich zwischen Wald und Wiese die Frage, ab welcher Häufigkeit von Bäumen die Klasse Wald vorliegt und ob es nicht sinnvoll sein könnte, Mischformen zu definieren.

Durch Klassifizierungsverfahren, z. B. die Maximum-Likelihood-Methode (Richards, 1993), werden dann Maßzahlen für die Zugehörigkeiten der betrachteten Bildelemente zu den einzelnen Klassen berechnet. Bei den gegenwärtigen Klassifizierungsmethoden wird mit Hilfe eines Entscheidungskriteriums - im Falle des genannten Verfahrens ist es die größte Wahrscheinlichkeit - jedem Bildelement eine einzige Klasse zugewiesen. Ausschließlich diese ermittelte Landnutzungsklasse wird als Ergebnis der Klassifizierung für jedes Bildelement festgehalten, jedoch nicht der Grad der Zugehörigkeit. Diese Vorgehensweise ist mit einem Informationsverlust verbunden. So ist nicht möglich, zu beurteilen, ob eine starke Zugehörigkeit (z. B. 0.90 Wald und 0.10 Wiese) oder eine schwache Zugehörigkeit (z. B. 0.51 Wald und 0.49 Wiese) vorliegt. Darüber hinaus kann keine Aussage über eine potentielle alternative Klasse gemacht werden. Abbildung 4 zeigt ein Testgebiet (a) mit dem Ergebnis der Klassifizierung (b) der Landnutzungsklassen Acker, Siedlung und Wald.

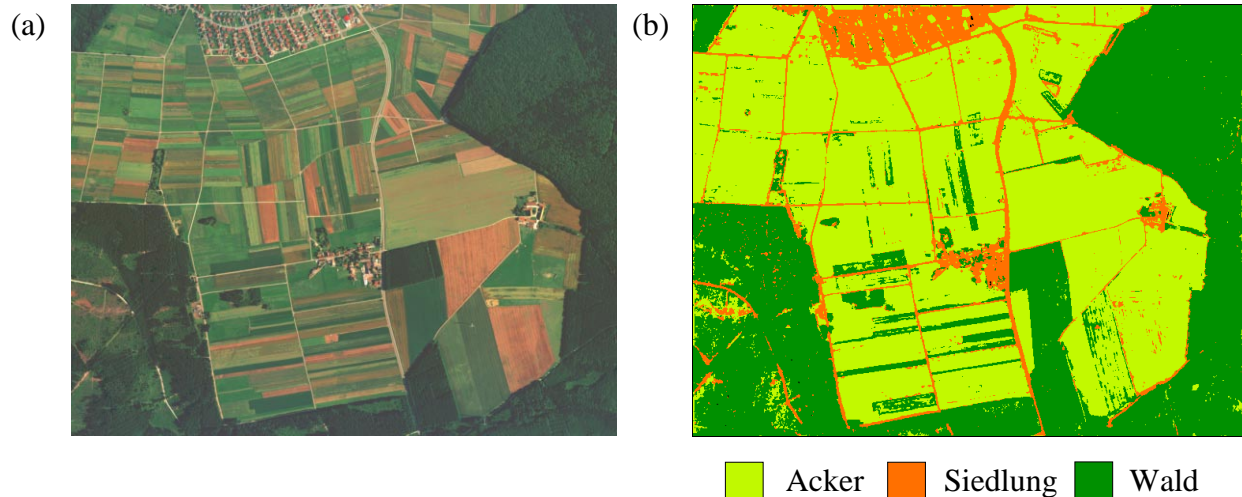


Abbildung 4: Testgebiet (a) und Ergebnis der Klassifizierung (b)

Zur Beurteilung der Qualität der Klassifizierung ist es üblich, Gebiete bekannter Thematik (Kontrollgebiete) mit dem Ergebnis der Klassifizierung zu vergleichen und das Ergebnis dieses Vergleichs in der Konfusionsmatrix bereitzustellen. Zusätzlich werden aus der Konfusionsmatrix verschiedene Kenngrößen abgeleitet, die die Genauigkeit der Klassifizierung zusammenfassend beurteilen (Congalton, 1991, Stehman, 1997). Zur globalen Beurteilung eignen sich der Anteil der korrekt klassifizierten Pixel und der Kappa-Koeffizient. Zur klassenweisen Beurteilung der Güte werden die Produzenten- und die Nutzergenauigkeit verwendet. Diese individuellen Maße können zu sehr unterschiedlichen Beurteilungen der Genauigkeiten der Klassen führen. Diese Genauigkeitsmaße fassen die in der Konfusionsmatrix enthaltene Information unterschiedlich zusammen, aber keines dieser Maße ist für die Beurteilung der erreichten Ge-

naugigkeit universell am besten geeignet. Die Konfusionsmatrix und die abgeleiteten Kenngrößen können als externe Genauigkeitsmaße bezeichnet werden, da zu ihrer Berechnung externe Informationen verwendet werden.

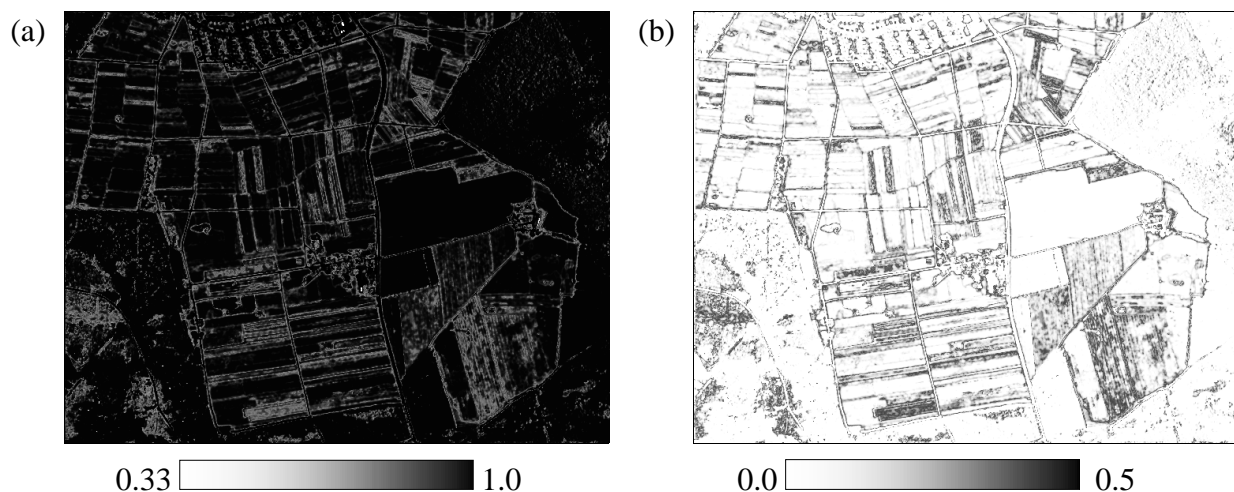


Abbildung 5: Darstellung der größten (a) und zweitgrößten (b) Wahrscheinlichkeiten

Alternativ können die im Rahmen der Klassifizierung berechneten Maßzahlen für die Zugehörigkeit der Bildelemente zu den einzelnen Klassen als Indikatoren für die Genauigkeit der Zuweisung der Thematik verwendet werden. Bei der Maximum-Likelihood-Klassifizierung lassen sich die Wahrscheinlichkeiten jedes einzelnen Bildelementes in einem sogenannten Wahrscheinlichkeitsvektor zusammenfassen. Er enthält die Wahrscheinlichkeiten, mit denen das zugehörige Bildelement den definierten Landnutzungsklassen zugewiesen werden kann. Die Dimension des Wahrscheinlichkeitsvektors entspricht der Anzahl n der Klassen. Der Wertebereich der größten Wahrscheinlichkeit liegt im Intervall $[1/n, 1.0]$, der der zweitgrößten Wahrscheinlichkeit im Intervall $[0.0, 0.5]$ und der der drittgrößten Wahrscheinlichkeit im Intervall $[0.0, 0.33]$. Die Summe sämtlicher Wahrscheinlichkeiten pro Bildelement ergibt 1.0. Im Gegensatz zu den externen Genauigkeitsmaßen werden die Wahrscheinlichkeiten für die Klassenzugehörigkeiten als interne Genauigkeitsmaße bezeichnet. In Abbildung 5 sind die jeweils pro Bildelement größten (a) und zweitgrößten (b) Wahrscheinlichkeiten durch Grauwerte dargestellt. Die dunklen Bereiche zeigen eine hohe Wahrscheinlichkeit an, während die hellen Regionen einen hohen Grad an Unsicherheit aufweisen. Linienhafte Unsicherheit tritt entlang des Grenzverlaufes zwischen benachbarten Objekten auf und kann durch die geometrische Auflösung der Fernerkundungsdaten begründet werden. Flächenhafte Unsicherheit kommt im Beispiel hauptsächlich in der Landnutzungsklasse Acker vor und läßt eventuell auf eine unzureichende Anzahl an Objektklassen in der durchgeführten Klassifizierung schließen.

In den unsicheren Bereichen stellt sich die Frage, welche Klasse alternativ mit welcher Wahrscheinlichkeit zugewiesen werden kann. Eine ausreichende Beschreibung der inneren Genauigkeit ist nur dann gewährleistet, wenn die Alternativklassen mitsamt ihren Wahrscheinlichkeiten als zusätzliche Information bereitgestellt und in nachfolgenden Analysen auch mitverwendet werden. Eine weitere Möglichkeit besteht darin, sämtliche Wahrscheinlichkeiten pro Bildelement, die größer als ein vorgegebener Grenzwert sind, als Informationen zur Genauigkeitsbeschreibung der jeweiligen Landnutzungsklasse zu verwenden. Ergeben sich z.B. für ein Bildelement die Wahrscheinlichkeiten für Acker und Siedlung zu 0.60 und 0.40, so kann dieses Bildelement mit unterschiedlichen Wahrscheinlichkeiten sowohl der Klasse Acker als auch der Klasse Siedlung zugehören. Dies führt dazu, daß sich die Landnutzungsklassen räumlich überlappen können. Dies ist als ein Vorteil zu sehen, da dadurch Mischpixel, die längs der Grenzlinie zwischen verschiedenen Landnutzungsklassen auftreten, modelliert werden können, ohne zusätzliche Klassen definieren zu müssen.

3 Hybrides Datenmodell mit Modellierung der Datenunsicherheit

3.1 Hybrides Datenmodell

Eine einfache Charakterisierung eines GIS-Systems wird dann erhalten, wenn die Datenart des Systems bekannt ist. Sind dies Rasterdaten, so spricht man von einem Raster-GIS, während es sich bei Vektordaten um ein Vektor-GIS handelt. Beide Typen basieren auf einem unterschiedlichen Datenmodell, das entsprechend Vektor-Modell oder Raster-Modell genannt werden kann (Abbildung 6). Eine Vielzahl von Anwendungen läßt sich mit dem einen oder anderen System erledigen. Doch gibt es eine Reihe von z. T. neuen Bereichen, die eine gemeinsame Verarbeitung von Raster- und Vektordaten in einem System erfordern. Dies führte zum Begriff des hybriden GIS (Fritsch, 1988).

Ein hybrides GIS erlaubt die Integration sowohl von Vektor- als auch von Rasterdaten. In der Form der Integration bestehen jedoch Unterschiede. Sie kann in verschiedenen Stufen ausgeführt werden (Ehlers et al., 1989):

- Die erste Stufe verfügt über die Möglichkeit, die jeweils andere Datenart einzulesen und als Zusatzinformation zu visualisieren, z. B. das Darstellen eines Rasterbildes im Hintergrund in einem Vektor-GIS. Es handelt sich dabei um eine einfache Überlagerung der Daten, ohne daß weitergehende Verarbeitungsoperationen möglich sind.
- In der zweiten Stufe wird zusätzlich eine Konversion bereitgestellt, mit deren Hilfe die systemfremde Datenart in die eigene Datenart umgewandelt werden kann. Somit können beispielsweise Vektordaten zum Zwecke der Analyse in einem Raster-GIS über den Weg einer Vektor-Raster-Konversion in Rasterdaten konvertiert werden. Sie lassen sich dadurch wie Rasterdaten im Datenmodell integrieren.
- Die dritte und höchste Stufe der Integration wird erreicht, wenn das im System angewandte Datenmodell selbst hybrid ist. Dazu sind die beiden bisher getrennten Datenmodelle für Vektor- und Rasterdaten zu vereinigen. Es entsteht ein gänzlich neues, jetzt hybrides Modell. Nur wenn die dritte Integrationsstufe in einem System realisiert wurde, kann die Integration als vollständig bezeichnet werden. Die darin enthaltenen Funktionen arbeiten unabhängig von der Datenart und erlauben somit die Ausnutzung der Vorteile beider Datenmodelle. Die Umsetzung dieser Integrationsstufe bedeutet für bestehende Systeme häufig eine komplette Neuentwicklung der Funktionalitäten. Aus diesem Grund ist kein heute gängiges System dieser Stufe zuzuordnen.

Im folgenden wird ein hybrides Datenmodell beschrieben, das als Basis für eine solche Neuentwicklung dienen kann. Ein hybrides Datenmodell ergibt sich aus der Vereinigung der beiden Modelle für Raster- und Vektordaten. Schwierigkeiten treten hierbei aufgrund der folgenden Unterschiede auf. Das Rasterdatenmodell verbindet Geometrie und Thematik durch die Zuweisung genau eines Attributwertes zu jeder Rasterzelle. Zelle und Wert definieren dabei eine untrennbare Einheit. Im Vektormodell sind Thematik und Geometrie über den Begriff des Objektes gekoppelt. Eine direkte Verbindung wie im Rastermodell besteht nicht mehr. Es ist durchaus denkbar, daß Geometrie und Thematik unabhängig voneinander gespeichert und verarbeitet werden. Ein Beispiel hierfür ist die häufig anzutreffende Aufteilung in der Datenhaltung. Sachdaten (Attribute) werden dabei in einer Datenbank, Geometriedaten in einem Dateisystem verwaltet. Ein weiterer Unterschied besteht darin, daß im Rastermodell im allgemeinen nur ein Attributwert (bzw. eine Thematik) jeder Rasterzelle beigefügt werden kann. Dagegen erlaubt das Vektormodell die Beschreibung beliebig vieler Thematiken zu jedem Objekt.

Als Grundlage für das nachfolgend definierte hybride Modell wird auf ein bereits bestehendes Konzept aufgebaut (Molenaar, Fritsch, 1991). Ansatzpunkt für die Vereinigung stellt dabei die Ausdehnung des Objektbegriffes auch auf Rasterdaten dar. Dazu ist es notwendig, daß innerhalb des Rasterdatenbestandes eine Objektbildung durchgeführt wird. Damit wird eine gemeinsame Basis durch den Oberbegriff des Objektes geschaffen, mit deren Hilfe sich dann die entsprechenden Daten strukturieren lassen.

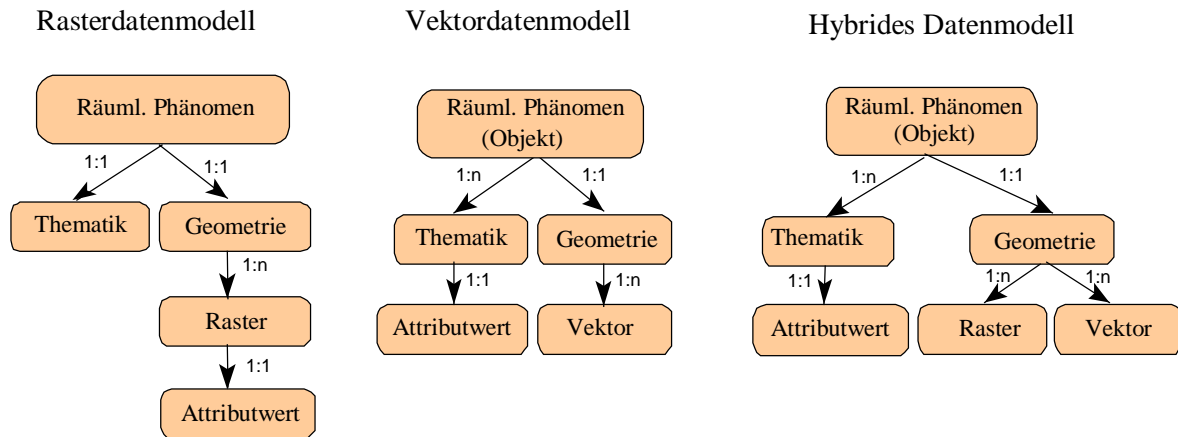


Abbildung 6: Gegenüberstellung des Raster-, Vektor- und hybriden Datenmodells

Abbildung 6 veranschaulicht schematisch die entwickelte Modellierung. Sie besitzt deutliche Ähnlichkeiten zum Vektormodell. Erweiterungen sind in der Geometriekomponente vorhanden. Die geometrische Beschreibung eines Objektes kann nun zusätzlich in Form einer Rasterdarstellung erfolgen. Raster- und Vektorgeometrie stellen dabei völlig gleichberechtigte Möglichkeiten dar. Die Themattikkomponente wurde unverändert aus dem Vektormodell übernommen.

Beispiele für die Modellierung von Raster- bzw. Vektordaten im hybriden Modell sind der Abbildung 7 zu entnehmen.

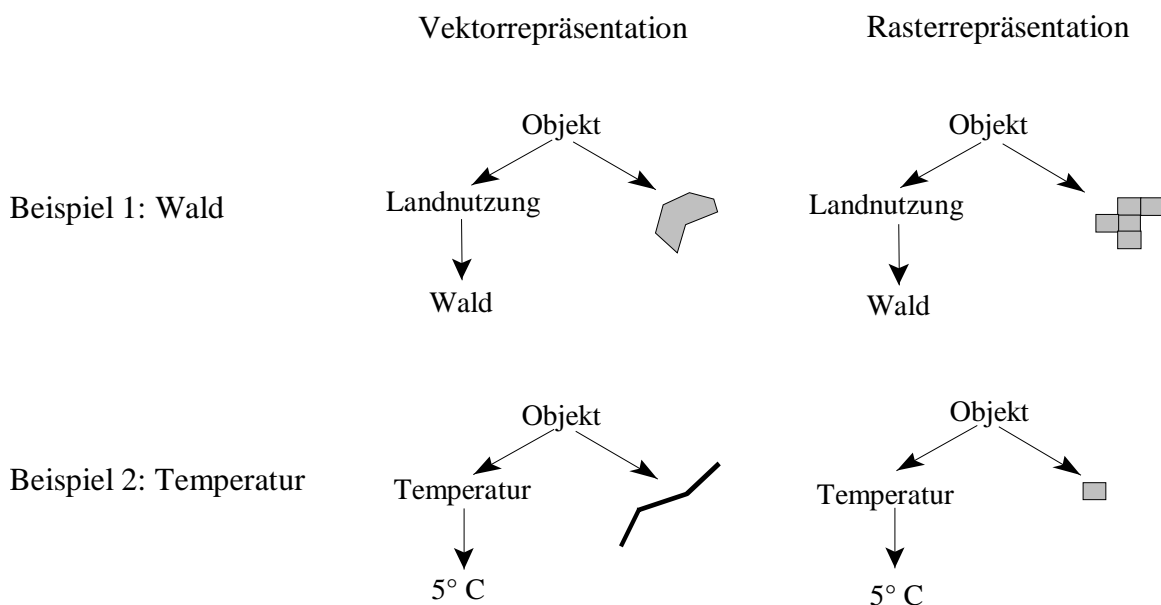


Abbildung 7: Beispiele zur hybriden Modellierung

Sehr anschaulich und für viele Zwecke geeignet kann das Modell dem Stand der heutigen Technik entsprechend in objektorientierter Form beschrieben werden. Sie erfolgt hier (Abbildung 8) unter Nutzung der bekannten OMT-Notation (Object Modelling Technique

(Rumbaugh et al., 1991)). Objektorientierte Modelle müssen aber nicht notwendigerweise mit Hilfe der objektorientierten Programmierung oder auf einer objektorientierten Datenbank implementiert werden (Fritsch, Anders, 1996). Sie definieren vielmehr allgemeine Beschreibungsmodelle, die sich in vielfältiger Weise umsetzen lassen.

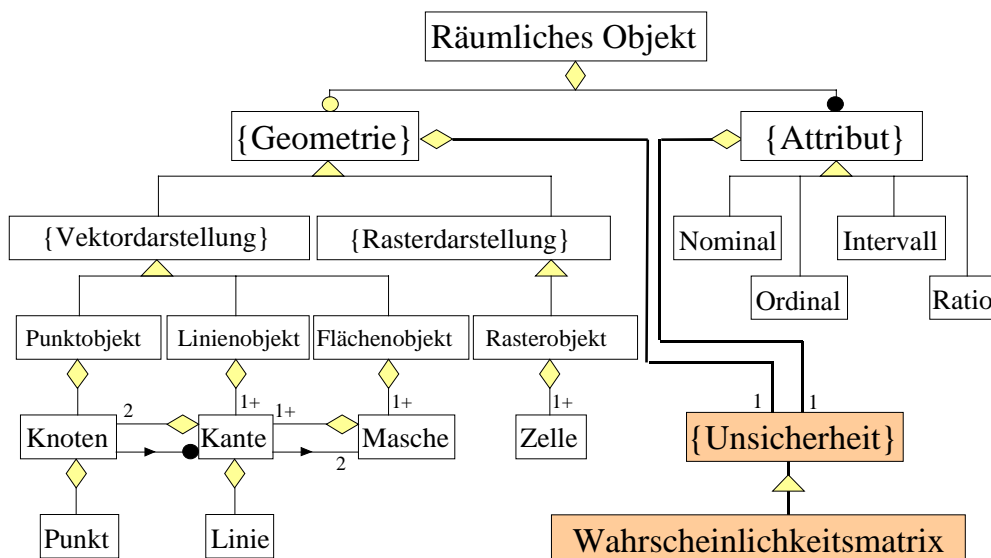


Abbildung 8: Hybrides Datenmodell mit Unsicherheitsmodellierung

3.2 Modellierung der Datenunsicherheit in einer hybriden Systemumgebung

Raster- und Vektordaten enthalten beide jeweils Anteile zur Geometrie und zur Thematik. Damit wirken prinzipiell auf beide Datenarten Einflüsse zur geometrischen und thematischen Unsicherheit ein. Dies führt dazu, daß im hybriden Datenmodell Beschreibungen zur Unsicherheit der beiden Komponenten, Geometrie und Thematik, unabhängig von der Datenart eingefügt werden müssen. Untersucht man die Annahmen, die zur Erfassung der Daten getroffen werden, so ergibt sich hier eine vereinfachte Situation. Generell gilt bei der Erfassung das Prinzip, daß eine der beiden Komponenten als fest, die andere als variabel angenommen wird (Chrisman, 1991). Die jeweils variable Komponente ist dann zu messen. Vektordaten entstehen, indem zu vorgegebenen Attributwerten die Ausdehnung dieser Attribute erfaßt wird. Rasterdaten geben dagegen eine fest definierte Form der Geometrie vor (nämlich die Vektorform), so daß die entsprechenden Attributwerte zu jeder Zelle zu bestimmen sind. Da feste Werte keine Unsicherheiten beinhalten, müssen nur jeweils die Unsicherheiten in der gemessenen Komponente berücksichtigt werden. Für Rasterdaten ist somit die Unsicherheit der Thematik, für Vektordaten die Unsicherheit der Geometrie zu modellieren. Entsprechende Erweiterungen sind dem hybriden Datenmodell beizufügen.

Die Unsicherheit der Geometrie wird innerhalb dieser Arbeit durch die Varianz der Liniengeometrie ausgedrückt, die einfach in Wahrscheinlichkeitswerte umgewandelt werden kann. Wahrscheinlichkeiten bewerten die Zugehörigkeit von beliebigen Positionen zum Objekt und können in Form von Wahrscheinlichkeitsmatrizen strukturiert werden. Mit der Umrechnung der Wahrscheinlichkeiten in diskrete Grauwerte läßt sich die enthaltene Variation besonders anschaulich visualisieren.

Der Parameter für die Unsicherheit der behandelten diskreten Attributwerte ist die Wahrscheinlichkeit des Zutreffens des Wertes. Jedem Wert ist eine Wahrscheinlichkeit beizufügen. Da nur Rasterdaten von der thematischen Unsicherheit beeinflußt werden, können die Wahrscheinlichkeitswerte wiederum in Matrizenform (passend zum Raster) strukturiert werden.

Durch die Nutzung von Wahrscheinlichkeiten ist ein gemeinsamer Parameter definiert, der sowohl die Unsicherheit von Vektor- als auch von Rasterdaten beschreibt. In beiden Fällen können Wahrscheinlichkeitsmatrizen zur Strukturierung eingesetzt werden. Es bietet sich somit an, diesen Parameter zur Beschreibung der Unsicherheit einzusetzen. Eine diesbezügliche Erweiterung des vorgestellten hybriden Datenmodells (Fritsch et al., 1998) ist ebenfalls in Abbildung 8 enthalten.

Die Nutzung desselben Parameters weist einige wichtige Vorteile auf. Die Erweiterung des hybriden Datenmodells läßt sich deutlich vereinfachen, da identische Konzepte innerhalb der Geometrie- und Thematikkomponente einzusetzen sind. Zusätzlich können gleiche Methoden zur Verwaltung, Bearbeitung und Visualisierung genutzt werden. Dies wirkt sich besonders positiv auf den Aufwand bei der Implementierung des Modells aus. Ein weiterer Vorteil ergibt sich erst im Zusammenhang mit der in Abschnitt 4.1 diskutierten Vektor-Raster- und Raster-Vektor-Konversion. Die ebenfalls notwendige Konversion der Unsicherheiten wird durch die Nutzung des gleichen Parameters mit Hilfe einer einfachen Übertragung der Wahrscheinlichkeitsmatrizen realisiert. Dabei werden keine weiteren Berechnungen notwendig.

4 Anwendungen

Nachdem zunächst auf Modellbildung und Verwaltungsaspekte eingegangen wurde, soll in diesem Abschnitt der Einfluß der Unsicherheit auf verschiedene Funktionen eines GIS diskutiert werden. Dazu werden beispielhaft einige wichtige Grundfunktionen eines hybriden Systems ausgewählt.

4.1 Konversionen zwischen Vektor- und Rasterdaten

Zwei wichtige Komponenten innerhalb der hybriden Datenverarbeitung stellen die Vektor-Raster-Konversion und die Raster-Vektor-Konversion dar. Sie ermöglichen es, Daten vom einen in das andere Format zu transformieren. Dies hat den Vorteil, daß Daten, die in verschiedenen Formaten vorliegen, in eine einheitliche Beschreibung umgeformt werden können. Nur so lassen sich die in einem GIS bereitgehaltenen Analysemethoden auch weiterhin nutzen, da sie im allgemeinen das Vorliegen der Daten in einem bestimmten Format voraussetzen und speziell dafür konzipiert wurden (z. B. ist die Flächenberechnung im Vektor-GIS nur mit Vektordaten ausführbar). Ein weiterer Vorteil ist, daß der Nutzer die Operationen, die er auf seine Daten anwenden möchte, unabhängig von dem benötigten Format wählen kann. Somit ist er in der Lage, die am besten geeignete Methode für sein Problem auszusuchen und zu nutzen. Ein Beispiel, das diesen Vorteil verdeutlicht, ist die geometrische Verschneidung. Bei Anwendung der Verschneidung im Rasterbereich ergibt sich ein einfach durchzuführender Prozeß, bei dem nur die entsprechenden Rasterzellen überlagert werden müssen. Im Gegensatz dazu benötigt die Lösung im Vektorbereich aufwendige geometrische Berechnungen, um ein identisches Ergebnis zu erzielen. Netzwerkanalysen sind dagegen ein Bereich, in dem die Vektordatenverarbeitung Vorteile besitzt.

Die Konversion ist, je nachdem in welcher Richtung sie erfolgen soll, unterschiedlich schwierig durchzuführen. Die Erzeugung von Raster- aus Vektordaten (Vektor-Raster-Konversion) stellt sich als der einfachere Prozeß dar, da sie in einem Schritt ablaufen kann. In der Praxis wird diese Form der Konversion recht häufig eingesetzt. Zum Beispiel verfügt jedes Raster-GIS standardmäßig über eine solche Möglichkeit, die als wichtige Eingabeschnittstelle genutzt wird. Desweiteren hält im Prinzip jede Graphikkarte eines Computers eine solche Funktionalität bereit. Fast alle Bildschirme basieren heutzutage auf Rastertechniken, während in den Anwendungsprogrammen die graphischen Darstellungen meistens in vektorieller Art gegeben

werden (siehe auch graphische Standards wie z. B. Phigs). Aus diesem Grund ist hier eine schnelle Konversion einzusetzen.

Demgegenüber bereitet die umgekehrte Richtung von Raster- zu Vektordaten (Raster-Vektor-Konversion) mehr Probleme. Sie kann nur als mehrstufiger Prozeß ausgeführt werden, der oftmals mehrdeutige Lösungen erzeugt. Die Eindeutigkeit muß dann anhand von Parametereinstellungen problemspezifisch hergestellt werden. Dazu sind Expertenwissen und interaktive Eingriffe notwendig, die einer Automatisierung des Ablaufes im Wege stehen. Deshalb ist es nicht verwunderlich, daß bisher noch kein allgemeingültiger Algorithmus entwickelt wurde, der wie für die Vektor-Raster-Konversion stets zuverlässige Ergebnisse liefert. Vielmehr gibt es eine Reihe problemorientierter Ansätze, die entsprechend nur im jeweiligen Anwendungsumfeld optimierte Lösungen erzielen.

4.1.1 Vektor-Raster-Konversion

Da es sich bei der Konversion von Vektor nach Raster um einen häufig benutzten Weg handelt, existieren auch zahlreiche Algorithmen zur Lösung des Problems. Sie weisen zumeist deutliche Ähnlichkeiten auf und liefern im allgemeinen identische Ergebnisse. Einer der bekanntesten ist der Bresenham-Algorithmus (Jäger, 1990). Er transformiert nach Festlegung der Rastergröße zuerst Anfangs- und Endpunkt einer Linie in das Raster und schreitet danach zeilenweise voran und bestimmt zu jeder Zeile, welche Spalte von der Linie durchquert wird. Das entsprechende Pixel zählt zu den gesuchten Linienpixeln. Alle Linienpixel zusammen definieren die gesuchte Rastergeometrie der Linie. Die beschriebene Standardaufgabe für eine Linie vereinfacht sich, wenn ein punktförmiges Objekt konvertiert werden soll (entsprechend nur Übertragung eines Punktes ins Raster). Bei flächenförmigen Objekten müssen zuerst alle Begrenzungslinien einzeln konvertiert werden. In einem weiteren Schritt ist dann das entstandene Randpolygon auszufüllen. Diese Aufgabe kann z. B. dadurch realisiert werden, daß für jedes Pixel einzeln geprüft wird, ob es sich innerhalb des Vektorpolygons befindet (bekanntes Punkt-im-Polygon-Problem im Vektorbereich).

4.1.2 Raster-Vektor-Konversion

Die einfachste Lösung zur Raster-Vektor-Konversion ergibt sich, wenn man jedes Pixel gesondert betrachtet und einzeln in die Vektorgeometrie umwandelt. Dabei wird die Ausdehnung der Rasterzelle durch Umringslinien beschrieben (z. B. ein quadratisches Pixel durch vier Kanten des Quadrats). Dadurch entsteht aus jeder Rasterzelle ein flächenförmiges Objekt mit Vektorgeometrie. Zusätzlich lassen sich in einem weiteren Schritt Nachbarobjekte mit gleichen Attributen zu größeren Aggregaten zusammenfassen. Dieser einfache Ansatz ist allgemein einzusetzen und wurde vielfach als Funktion in bestehende GIS-Produkte implementiert. Er bietet aber nur eine eingeschränkte Einsatzmöglichkeit. So ist es damit nicht möglich, zwischen punkt-, linien- und flächenförmigen Objekten zu unterscheiden und entsprechend eine angepaßte Lösung zu erzeugen.

Welcher Typ aus der Konversion erhalten werden soll, ist meist durch die jeweilige Anwendung festgelegt. Wichtigste Einflußfaktoren für die Wahl sind zum einen das Verhältnis des Maßstabbereiches zur Größe und zum anderen die Bedeutung der Objekte. Die getroffene Festlegung gilt zumeist für alle Objekte einer Objektart und wird innerhalb des thematischen Datenmodells gespeichert. Da alle drei Typen in gleichem Maße in Anwendungen vertreten sein können, muß auch jeweils eine typspezifische Transformation bereitgehalten werden. Die Raster-Punkt-Konversion entspricht einer einfachen Schwerpunktbildung der Rasterobjektfläche. Dagegen gestaltet sich die Raster-Linien-Konversion bedeutend aufwendiger. Sie kann in die Teilschritte Distanztransformation, Skelettierung, Linienextraktion und Linienglättung unterteilt werden (Cramer, 1993). Darauf aufbauend lassen sich auch Flächen erzeugen, indem

das Linienverfahren für den Rand der Rasterobjektfläche durchgeführt wird. Beispiele für Konversionen von verschiedenen Objekttypen zeigt die Abbildung 9.

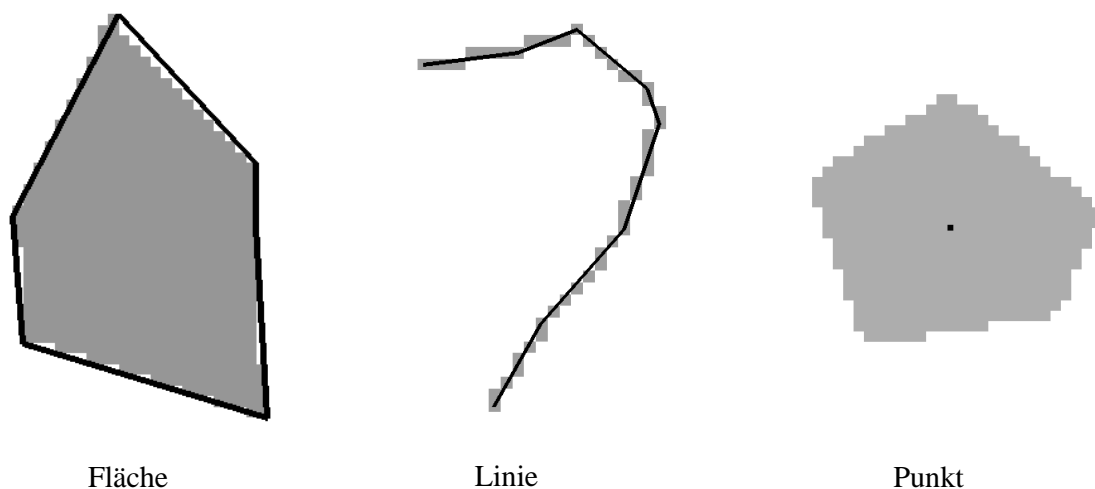


Abbildung 9: Beispiele zur Raster-Vektor-Konversion

4.1.3 Qualität der Konversionen

Die Geometrie und damit die Ausdehnung eines Objektes wird durch die Konversion unweigerlich verändert. Deshalb stellt sich die Frage nach der Qualität (Genauigkeit) der Konversion. Ziel ist ein Verfahren, das verlustfreie Konversionen ermöglicht, so daß diese als Hintergrundoperationen ohne direkte Einflußnahme des Anwenders ablaufen können. Innerhalb der Vektor-Raster-Konversion wird die Qualität durch die Festlegung der Rastergröße (Auflösung) beeinflusst. Je kleiner die Rastergröße ist, desto genauer kann die Übertragung ins Raster erfolgen. Einzelne formelmäßige Abschätzungen werden hierzu in Burrough (1986) gegeben. Sie gelten aber nur für die Fälle, in denen die Geometrie als fest (sicher) angenommen wird. Betrachtet man zusätzlich die mit der Geometrie verbundene Unsicherheit, dann sollte die Rastergröße kleiner als die Genauigkeit der Liniengeometrie sein. Damit wird erreicht, daß der zusätzliche Einfluß gegenüber der bestehenden Unsicherheit der Geometrie eine untergeordnete Rolle spielt und damit vernachlässigt werden kann. Dies läßt sich erreichen, indem die Rastergröße identisch zur Rastergröße der Wahrscheinlichkeitsmatrix gewählt wird. Wie schon in Abschnitt 2.1 beschrieben, ist dort die Auflösung ein Vielfaches geringer als die Standardabweichung σ der Liniengeometrie. Die identische Wahl der Rastergröße in der Wahrscheinlichkeitsmatrix und der Rastergeometrie hat nebenbei den Vorteil, daß sofort jedem Objektpixel genau ein Wahrscheinlichkeitswert aus der Wahrscheinlichkeitsmatrix zugeordnet werden kann. Das hybride Datenmodell fordert eine solche Zuordnung, die nun direkt ohne zusätzlichen Aufwand erzielt wird (vgl. Abschnitt 4.1.4). Da die Festlegung der Rastergröße für die Wahrscheinlichkeitswerte schon bei der Erfassung der Objekte (in Abhängigkeit der Liniengenauigkeit) erfolgt, liegt somit auch die Rastergröße bereits vor der Konversion fest. Dadurch kann die Vektor-Raster-Konversion vollständig automatisiert ablaufen.

Einflußgröße bei der Raster-Vektor-Konversion auf die Qualität der erzielten Vektorgeometrie ist der Faktor der Glättung. Hierbei gilt, je kleiner der Schrankenwert, desto genauer wird die Rastergeometrie in die Vektorgeometrie übertragen. Betrachtet man in diesem Zusammenhang wiederum die Unsicherheit, die diesmal in den Attributen vorliegt, so wird die Ausdehnung des Vektorobjekts stark von der Genauigkeit des für die Konversion verantwortlichen Attributes abhängen. Ist nun der Glättungsfaktor klein im Vergleich zur Unsicherheit in den Attributen, so

kann sein Einfluß auf die Genauigkeit der Geometrie vernachlässigt werden. Der hierfür gewählte Wert orientiert sich an der Rastergröße r . Um das aufgrund der Diskretisierung entstandene treppenförmige Erscheinungsbild von Linien zu beseitigen, wird der Glättungsfaktor zu $f = \sqrt{2}/2 \cdot r$ festgelegt. Die maximal mögliche Variationsbreite der Geometrie beträgt dann $\sqrt{2} r$. Im Normalfall ist die Variation aufgrund der Attributunsicherheit um ein Vielfaches höher (meist mehrere Rasterzellen), so daß die Glättung keinen wesentlichen Beitrag zur Unsicherheit darstellt. Die Speicherung der Unsicherheit bleibt durch Übertragung der Wahrscheinlichkeitsmatrix voll erhalten.

4.1.4 Konversion der Unsicherheit

Die gegenseitige Konversion in die jeweilige andere Datenart bewirkt innerhalb einer hybriden Systemumgebung den Wechsel in der geometrischen Darstellungsart. Nachdem die Geometrie eines Rasterobjekts als fest und damit frei von Unsicherheiten definiert wurde, die Geometrie eines Vektorobjekts aber Unsicherheiten beinhaltet (vgl. Abschnitt 2, Abschnitt 3.2), ist mit der Konversion der Darstellungsart auch immer eine Konversion der Unsicherheit verbunden. Wie diese Konversion zu erfolgen hat, ist in den Eigenheiten der beiden Datenarten begründet.

Betrachtet man ein Vektorobjekt, dann stellt seine Ausdehnung die variable Komponente dar, während die Attribute als fest und über den gesamten Bereich als sicher gelten. Man kann also behaupten, daß überall dort wo man sich innerhalb des Objektes befindet, die Attribute sicher anzutreffen sind. Durch die Variation in der Geometrie ist aber gleichzeitig auch die Aussage, daß man sich innerhalb des Objektes befindet, unsicher. Diese Unsicherheit läßt sich nun auf die sichere Gültigkeit der Attribute übertragen. Wenn man sich in manchen Bereichen (in den Randbereichen) nur unsicher innerhalb des Objektes befindet, dann können in diesen Bereichen die mit dem Objekt verbundenen Attribute auch nur unsicher zutreffen. Damit gilt die Behauptung, daß an einer festen Stelle die Attribute unsicher festgelegt sind. Dieser Wechsel in der Aussage von der Unsicherheit in der Geometrie zur Aussage von der Unsicherheit in den Attributen findet genau innerhalb der Vektor-Raster-Konversion statt. Dort werden eine Reihe von festen Stellen (Rasterzellen) definiert, für die jeweils nun die Aussage, inwieweit die Attribute zutreffen, quantitativ bewertet werden muß. Die Variabilität in der Geometrie wird damit in eine Variabilität der Thematik umgewandelt. Entsprechend werden die Unsicherheiten in der geometrischen Beschreibung in Unsicherheiten der Thematik transformiert. Da sich nur die zugehörige Interpretation ändert, bleiben die Größenwerte (Wahrscheinlichkeiten) unverändert erhalten (Abbildung 10). Der gleiche Umstand trifft auch für die Konversion vom Raster- in den Vektorbereich zu. In einem Rasterobjekt ist für eine sichere Objektgeometrie das Zutreffen der jeweiligen Attributwerte unsicher. Soll nun ein Attribut als sicher vorausgesetzt und dafür die Ausdehnung des Objektes bestimmt werden (dies entspricht dem Ansatz der Raster-Vektor-Konversion), dann ist für jedes Pixel unsicher, ob es weiterhin zum Objekt hinzugezählt werden soll. Die Ausdehnung und damit die Geometrie werden unsicher.

Basierend auf den erläuterten Erkenntnissen definiert sich die Konversion der Unsicherheit innerhalb der Vektor-Raster-Konversion durch eine einfache Übertragung der Wahrscheinlichkeitsmatrizen von der Qualitätsmodellierung der Geometrie auf die Qualitätsmodellierung der einzelnen Attribute und umgekehrt im Falle der Raster-Vektor-Konversion. Damit verbunden ist ein Wechsel in der Interpretation der Wahrscheinlichkeiten.

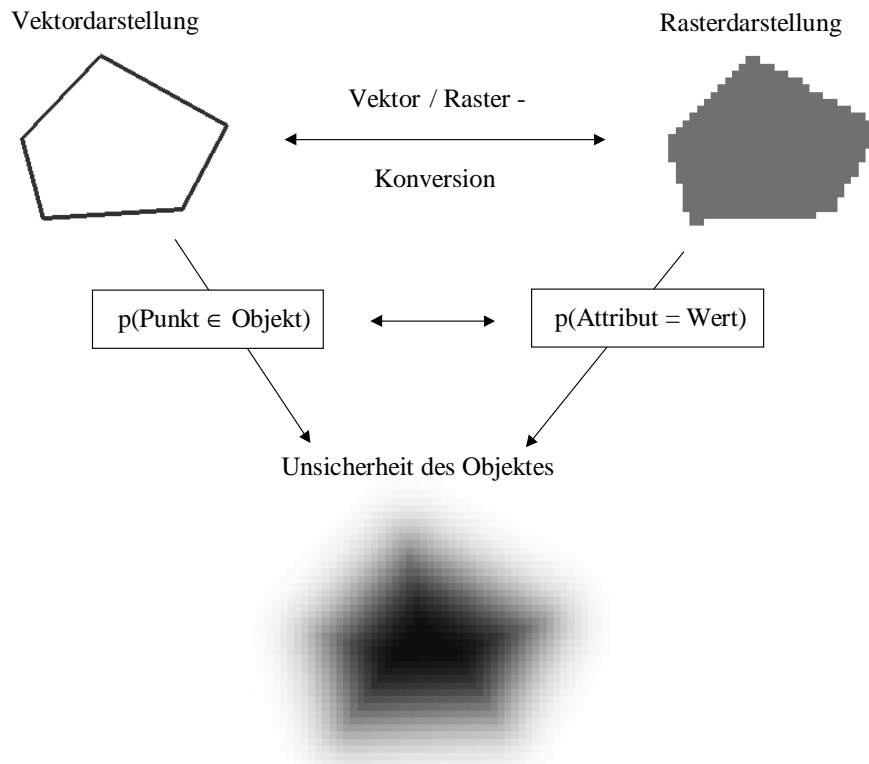


Abbildung 10: Vorgänge bei der Vektor-Raster- und der Raster-Vektor-Konversion

4.2 Verschneidungsanalyse

Die Verschneidung zweier Objektmengen gehört zu den häufig angewendeten Analysemethoden eines GIS. Die Aufgabe ist im Vektorbereich aufgrund geometrischer Schnittbildungen nur aufwendig zu realisieren. Die Rasterstruktur, die bereits geometrische Basiszellen vorgibt, ist hier im Vorteil. Die Verschneidung besteht aus einer logischen Verknüpfung von überlagerten Zellen. Der Vorteil läßt sich in einem hybriden System so ausnutzen, daß die Verschneidungsmethode nur für Rasterdaten implementiert ist. Deshalb sind vorab alle beteiligten Objekte (sofern notwendig) in eine Rasterrepräsentation umzuformen.

Als Ergebnis der Verschneidung entstehen neue Objekte, die die Eigenschaften der Ursprungsobjekte aufweisen. Die Verknüpfung der Eigenschaften erfordert, daß die Unsicherheiten der einzelnen Eigenschaften zusammenzuführen sind, so daß sich die Unsicherheit des gemeinsamen Auftretens ausdrücken läßt. Dieser Vorgang entspricht einer Fortpflanzung der Unsicherheiten auf das Ergebnisobjekt. Durch den Zusammenhang der Verschneidung mit der logischen UND-Operation kann die Fortpflanzung durch die entsprechende Formel aus der Wahrscheinlichkeitstheorie:

$$p(A_1 \cap A_2) = p(A_1) \cdot p(A_2)$$

definiert werden (Shi, 1994). Mit $p(A_1)$ bzw. $p(A_2)$ werden die Wahrscheinlichkeiten, daß eine Rasterzelle die jeweiligen Attribute aufweist, bezeichnet. Einschränkend wird hierbei vorausgesetzt, daß die beiden Thematiken unabhängig voneinander sind. Durch Erweiterung der Verschneidungsmethode um die Fortpflanzung der Unsicherheit können automatisch zu allen neuen Objekte wieder Wahrscheinlichkeitsmatrizen berechnet werden, so daß diese für weiterführende Zwecke in gleicher Weise im System zu benutzen sind.

Abbildung 11 verdeutlicht die Verschneidung an einem Beispiel. Ein flächenhaftes Objekt (Stadtgebiet) und ein linienhaftes Objekt (Stromleitung) werden miteinander verschnitten

(Abbildung 11a). Sie besitzen unterschiedliche geometrische Genauigkeiten, wie aus den visualisierten Wahrscheinlichkeitsmatrizen ersichtlich ist (Abbildung 11b). Das Ergebnis der Verschneidung (Abbildung 11c) zeigt, daß nur ein kleiner Teil der Leitung über das Stadtgebiet führt. Dies entspricht dem traditionellen Ergebnis der Verschneidung. Durch die Fortpflanzung der Unsicherheit wird das Ergebnis um die Wahrscheinlichkeitsmatrix ergänzt, die als Hintergrund in Abbildung 11c visualisiert ist. Daraus läßt sich ableiten, daß weitere Regionen existieren, in denen ein Schnitt wahrscheinlich ist. Durch die Angabe einer Mindestwahrscheinlichkeit (hier: $p=0.25$) läßt sich die Objektgeometrie an ein gewähltes Vertrauensniveau anpassen. Im Gegensatz zu der ursprünglich kleinen Schnittlinie entstehen dadurch eine Reihe von Bereichen, innerhalb denen die Stromleitung das Stadtgebiet passieren kann (Abbildung 11d).

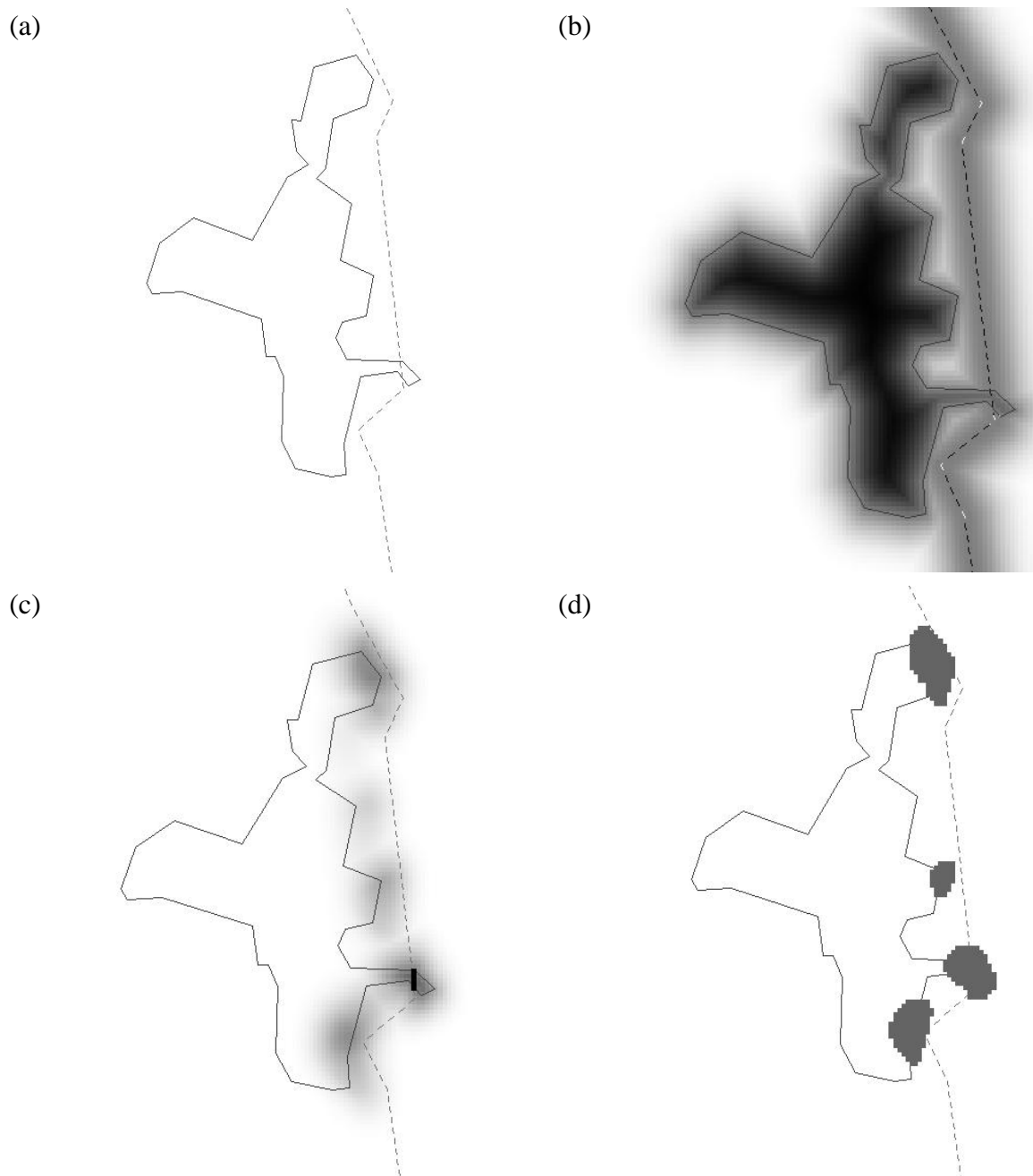


Abbildung 11: Verschneidung zweier Objekte (a), deren Unsicherheit (b), Ergebnisobjekt mit Unsicherheit (c) und Objektgeometrie für $p=0.25$ (d)

5 Ausblick

Die Berücksichtigung der Genauigkeit in einer hybriden Systemumgebung hat wesentlichen Einfluß auf die Funktionsweise und das Erscheinungsbild bisheriger Analysemethoden. Eingangswahrscheinlichkeiten sind mit Hilfe zu definierender formelmäßiger Zusammenhänge auf das Ergebnis fortzupflanzen. Dazu ist es erforderlich, alle Daten mit Genauigkeitsangaben zu versehen. Beim Datenaustausch ist darauf zu achten, daß mit den Daten auch die Genauigkeiten übertragen werden. Nur so läßt sich die steigende Verfügbarkeit an Daten auch zum Vorteil des Anwenders ausnutzen. Die Auswirkung der hybriden Datenmodellierung mit Berücksichtigung der Genauigkeit wurde im Beitrag anhand von zwei Anwendungen diskutiert. Es genügt jedoch nicht, sich auf ausgewählte Methoden zu beschränken, sondern alle vorhandenen Analysen sind so zu modifizieren, daß sie Genauigkeit mitprozessieren können. Das Ziel ist, daß als Ergebnis eines jeden Verarbeitungsschrittes nicht nur die Daten selbst stehen, sondern immer eine Aussage über deren Genauigkeit mitgeliefert wird.

6 Literatur

- Bill, R., Fritsch, D. (1991): Grundlagen der Geo-Informationssysteme. Band 1: Hardware, Software und Daten. Wichmann, Karlsruhe.
- Burrough, P. A. (1986): Principles of Geographical Information Systems for Land Resources Assessment. Oxford Science Publications, Clarendon Press, Oxford.
- Caspary, W. (1992): Genauigkeit als Qualitätsmerkmal digitaler Datenbestände. In: Grünreich, D., Buziek, G. (Hrsg.): Gewinnung von Basisdaten für Geo-Informationssysteme. DVW-Schriftenreihe, Heft 4, 157-166.
- Caspary, W., Scheuring, R. (1992): Error-bands as Measures of Geometrical Accuracy. EGIS '92, Vol. 1, 226-233.
- Chrisman, N. R. (1991): The Error Component in Spatial Data. In: Maguire, D., Goodchild, M., Rhind, D. (Eds.): Geographical Information Systems – Principles and Applications. Longman Scientific & Technical, 165-174.
- Congalton, R. G. (1991): A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of the Environment, Vol. 37, 35-46.
- Cramer, M. (1993): Implementation von Raster-Vektor-Konvertierungsbausteinen als Basis für eine GIS-Teachware. Diplomarbeit am Institut für Photogrammetrie, Universität Stuttgart (unveröffentlicht).
- Drummond, J. (1995): Positional Accuracy. In: Guptill, S., Morrison, J. (Eds.): Elements of Spatial Data Quality. Pergamon Press.
- Ehlers, M., Edwards, G., Bedard, Y. (1989): Integration of Remote Sensing with Geographic Information Systems: A Necessary Evolution. PE&RS, Vol. 11, No. 11, 1619-1627.
- Fritsch, D. (1988): Hybride Graphische Systeme – eine neue Generation von raumbezogenen Informationssystemen. GIS, Vol. 1, Heft 1, 12-19.
- Fritsch, D., Anders, K.-H. (1996): Objektorientierte Konzepte in Geo-Informationssystemen. GIS, Vol. 9, Heft 2, 2-14.
- Fritsch, D., Glemser, M., Klein, U., Sester, M., Strunz, G. (1998): Zur Integration von Unsicherheit bei Vektor- und Rasterdaten. GIS, Vol. 11, Heft 4, 26-35.
- Glemser, M. (1994): Behandlung der Genauigkeit räumlicher Daten in Geo-Informationssystemen. In: Die benutzte Erde, Alfred-Wegener-Stiftung (Hrsg.). Ernst&Sohn, Berlin.

- Glemser, M. (1996): Integration geometrischer Datenqualität in GIS-Funktionen. In: Proceedings Workshop Datenqualität und Metainformation in Geo-Informationssystemen, Universität Rostock.
- Jäger, E. (1990): Untersuchungen zur kartographischen Symbolisierung und Verdrängung im Rasterdatenformat. Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen, Universität Hannover, Nr. 167.
- Kraus, K., Haussteiner, K. (1993): Visualisierung der Genauigkeit geometrischer Daten. GIS, Vol. 6, Heft 3, 7-12.
- McKee, L., Kuhn, W. (1997): The OpenGIS Consortium's Purpose and Technical Approach. In: Fritsch D., Hobbie, D. (Eds.): Photogrammetric Week '97, Wichmann, Heidelberg.
- Molenaar, M., Fritsch, D. (1991): Combined Data Structures for Vector and Raster Representations in Geographic Information Systems. GIS, Vol. 4, Heft 3, 26-32.
- OpenGIS (1999): Web Site of the OpenGIS Consortium, <http://www.opengis.org>.
- Richards, J. A. (1993): Remote Sensing Digital Image Analysis. Springer-Verlag.
- Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., Lorenzen, W. (1991): Object-Oriented Modelling and Design. Prentice Hall.
- Shi, W. (1994): Modelling Positional and Thematic Uncertainties in Geographic Information Systems. ITC Publication, No. 22, Enschede.
- Stehman, S. V. (1997): Selecting and Interpreting Measures of Thematic Classification Accuracy. Remote Sensing of the Environment, Vol. 62, 77-89.