

Untersuchungen zu den Fortpflanzungseigenschaften der geometrischen Genauigkeit bei der Flächenverschneidung

Michael Glemser, Christian Henneberg
Institut für Photogrammetrie
Universität Stuttgart

1 Einführung

Flächenverschneidung (Polygon Overlay) ist eine der bekanntesten Analysemethoden in Geo-Informationssystemen (GIS). Sie zählt zu den grundlegenden Basisoperatoren und wird nahezu von allen verfügbaren Systemen bereitgestellt. Im allgemeinen Fall werden dabei zwei sich überlagernde Objekte geometrisch miteinander verschnitten und dadurch ein neues Objekt (Schnittobjekt) gebildet. An der Schnittbildung sind die geometrischen Repräsentationen der Objekte beteiligt. Im Falle der hier behandelten Vektordaten hat sich die Randdarstellung als geometrisches Modell etabliert. Der Rand als räumliche Abgrenzung von der Umgebung wird dabei durch geometrische Primitive (Punkte, Linien, Flächen) definiert. Die eingesetzten Primitive besitzen stets eine um eine Stufe geringere Dimensionalität. Z.B. wird ein flächenhaftes Objekt (2-dimensional) durch seine umgebenden Linien (1-dimensional) abgegrenzt. Bei der Verschneidung werden dann alle möglichen geometrischen Schnitte der Primitive gebildet und anschließend eine neue Randbeschreibung der überlagerten Teile zusammengefügt. Z.B. entsteht aus zwei flächenhaften Objekte im Normalfall wiederum ein flächenhaftes Objekt, das sich aus Teilflächen und Teilen der Randlinien der Eingangsobjekte zusammensetzt.

Ein wichtiger Aspekt bei der Interpretation des Ergebnisses der Flächenverschneidung ist, daß die Daten aller Objekte nur eine begrenzte Genauigkeit aufweisen (z.B. Caspary, 1992a). Dies gilt im besonderen auch für die Geometrie, die bei der Flächenverschneidung die zentrale Rolle spielt. Werden die geometrischen Informationen z.B. aus bestehenden Karten mittels Digitalisierung erfaßt, dann bewirken verschiedene Einflüsse (z.B. Geräte, Operateur, Vorlage), daß die gemessenen Koordinaten eine Variation beinhalten. Bei Durchführung von Wiederholungsmessungen würden stets andere Ergebnisse erzielt werden. Die Variation der Koordinatenwerte als eine Art Streuung verdeutlicht, wie sich die *geometrische Genauigkeit* auf Objekte auswirkt. Die Beschreibung der Genauigkeit gelingt mit Hilfe eines stochastischen Ansatzes. Koordinaten z.B. werden in diesem Ansatz als Zufallsvariablen aufgefaßt und einzeln durch Mittelwert und Varianz beschrieben. Funktionale Abhängigkeiten zwischen den Variablen lassen sich durch Kovarianzen ausdrücken. Der Ansatz wird in gleicher Weise wie für Koordinaten auf alle geometrischen Primitive angewendet. Z.B. für Flächen sind dazu alle zugehörigen Randlinien als einzelne Zufallsvariablen einzuführen und mit Varianzen zu versehen.

Nachdem man die Basisdaten mit ihrer Genauigkeit vorliegen hat, stellt sich die Frage nach der Verarbeitung und Berücksichtigung innerhalb von Analysen. Im speziellen auf die hier untersuchte Aufgabe der Flächenverschneidung angewendet, ist folgendes für den Nutzer von Interesse: *Wie genau ist das Ergebnis bei einer Flächenverschneidung?* Die Beantwortung gelingt nur, wenn die vorhandene Genauigkeit der Eingangsdaten auf das Ergebnis übertragen wird. Dieser Vorgang wird im allgemeinen als Fortpflanzung der Genauigkeit oder Fehlerfortpflanzung bezeichnet. Aus der Fortpflanzung erhält man wiederum eine Angabe zur Vari-

anz der Ränder, so daß eine äquivalente Darstellung der Genauigkeit zu den Eingangsdaten auch für das Ergebnis vorliegt. Mit Hilfe dieser Angabe wird der Nutzer in die Lage versetzt, eine Bewertung der Güte des Ergebnisses durchzuführen.

Das Problem der Fortpflanzung wird im Bereich der Geodäsie durch die Ausnutzung des funktionalen Zusammenhangs f zwischen Eingangswerten x und Ergebnis y gelöst. Die Fortpflanzung ist dann durch das *allgemeine Fehlerfortpflanzungsgesetz* definiert (z.B. Höpke, 1980):

$$D(y) = A D(x) A^T$$

mit

$$y = f(x) = Ax$$

$D(x)$ bzw. $D(y)$ sind die Dispersionsmatrizen der Werte x bzw. y und halten auf der Hauptdiagonalen deren Varianzen und auf den Nebendiagonalen die jeweiligen Kovarianzen. Dabei wird ein linearer bzw. linearisierter Zusammenhang ($y = Ax$) vorausgesetzt.

Problem bei der Flächenverschneidung ist, daß der geforderte funktionale Zusammenhang nur in komplexer Form und stets individuell für den jeweils konkret vorliegenden Fall formuliert werden kann. Daher kommt die Anwendung des allgemeinen Fehlerfortpflanzungsgesetzes nicht in Frage. Aus diesem Grund wird hier auf *Simulationen* übergegangen, um Aussagen zur Genauigkeit abzuleiten. Diese Simulationen stützen sich auf die zufällige Erzeugung von großen Mengen an Eingangsdatensätzen, für die dann jeweils die anzuwendende Operation (hier: Flächenverschneidung) durchgeführt wird. Es entsteht wiederum eine große Lösungsmenge, die mit Hilfe der mathematischen Statistik ausgewertet werden kann.

Ziel der vorliegenden Untersuchungen ist es, Aussagen zum Fortpflanzungsverhalten von unterschiedlichen Eingangsdaten bei der Flächenverschneidung mit Hilfe von Monte-Carlo-Simulationen zu treffen. Im einzelnen werden dazu die folgenden Bestimmungen durchgeführt:

- Wahrscheinlichkeit des Schnitts,
- Wahrscheinlichkeit der Objektzugehörigkeit eines beliebigen Punktes zum Schnittobjekt,
- Wahrscheinlichste Geometrie des Schnittobjektes und Verschiebungsbetrag zur traditionellen Lösung

Zum Einsatz kommen jeweils nur synthetisch erzeugte Datensätze, die speziell die Aufdeckung von wesentlichen Eigenschaften erleichtern sollen.

Die notwendigen programmtechnischen Umsetzungen wurden auf Basis des GIS-Paketes MapInfo mit Hilfe der Programmierschnittstelle MapBasic ausgeführt. Zur graphischen Aufbereitung der statistischen Ergebnisse kann das Modul Xprism des Graphikpaketes Khorus und das Microsoft-Programm Excel zum Einsatz.

2 Verschneidung

Die Verschneidung dient dazu, neue Objekte geometrisch aus sich überlappenden Objekten zu bilden (Abb. 1). Nach der Verschneidung erhalten die neuen Objekte die thematischen Eigen-

schaften der Eingangsobjekte übertragen. Es soll beispielsweise geklärt werden, welcher Bodentyp auf welchen Teilen eines Flurstückes anzutreffen ist. Alle Bodenobjekte und das betreffende Flurstück bilden die Eingangsobjekte. Ergebnisobjekte sind die einzelnen Teile des Flurstücks, die jeweils einen anderen Bodentyp aufweisen.

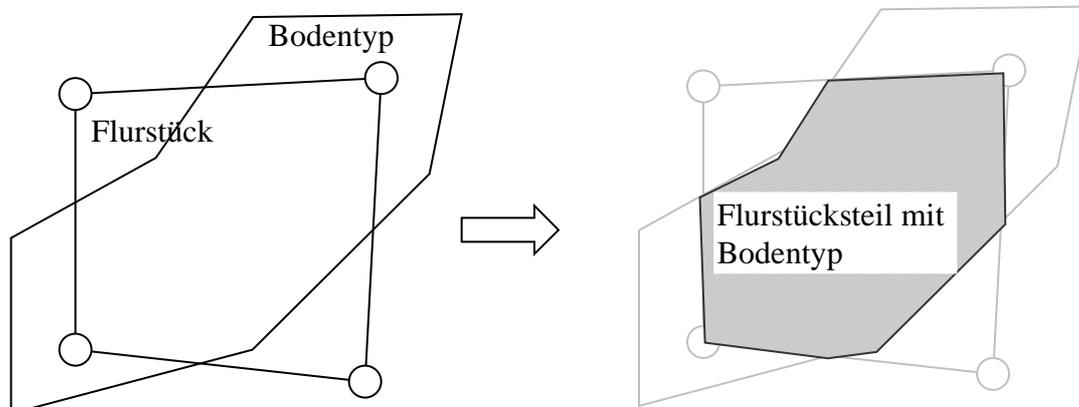


Abb. 1: Beispiel zur Flächenverschneidung

Betrachtet man Objekte als Punktmenge, dann läßt sich die Verschneidung in mengentheoretischer Form als logische Operation formulieren. Dabei können durchaus verschiedene logische Operatoren (Durchschnitt, Vereinigung, Differenz) angewendet werden. Innerhalb der Arbeit erfolgt die Einschränkung auf die Untersuchung des Durchschnitts. Die abgeleiteten Erkenntnisse lassen sich aber in einfacher Form auch auf andere Fälle übertragen.

Die Verschneidung hat für Vektordaten in typspezifischer Weise zu erfolgen. Folgende Fälle sind dabei zu unterscheiden:

1. Punkt – Punkt (→ Punkt)
2. Punkt – Linie (→ Punkt)
3. Punkt – Fläche (→ Punkt)
4. Linie – Linie (→ **Punkt**, Linie)
5. Linie – Fläche (→ Punkt, **Linie**)
6. Fläche – Fläche (→ Punkt, Linie, **Fläche**)

In Klammern ist vermerkt, welcher Typ als Ergebnis in Frage kommt. Zusätzlich hervorgehoben wird jeweils der am häufigsten und damit normalerweise erzeugte Typ. Natürlich kann die Verschneidung auch leere Ergebnismenge liefern. In vielen Fällen kann die leere Menge ebenfalls ein wichtiges Ergebnis darstellen (z.B. Ausschluß des gemeinsamen Auftretens zweier kritischer Eigenschaften). In der vorliegenden Arbeit erfolgt eine Konzentration auf den Verschneidungstyp 6 (Fläche – Fläche), der als *Flächenverschneidung* bezeichnet wird. Zu beachten ist, daß der Begriff der Flächenverschneidung in der Literatur häufig mißbräuchlich als Überbegriff eingesetzt wird, nicht jedoch im Rahmen dieser Arbeit.

Die Flächenverschneidung kann algorithmisch in drei Schritt unterteilt werden (Bill, 1996):

1. *Bildung der Kantenschnitte*: Es werden Schnittpunkte der sich schneidenden Kanten ermittelt. Dabei handelt es sich um einen arbeits- und rechenintensiven Schritt, da alle vorhandenen Kanten potentielle Schnittkandidaten darstellen. Grobtests verringern den Aufwand. Als Ergebnis erhält man die kleinste gemeinsame Geometrie der Eingangsdaten, die sich in Form von Knoten- und Kantenlisten topologisch strukturieren läßt.

2. *Bilden der neuen Polygone:* Es werden die Einzelkanten zu Polygone zusammengeführt. In Abhängigkeit von der gewünschten mengentheoretischen Operation entstehen neue Objektgeometrien.
3. *Übertragen der Attribute:* Die gültigen Attribute der Eingangsobjekte werden auf die neuen Ergebnisobjekte übertragen, so daß diese die gleichen Eigenschaften aufweisen, wie die Objekte, aus denen sie gebildet wurden.

In dieser Weise wurde die Flächenverschneidung für die nachfolgenden Untersuchungen realisiert.

3 Geometrische Genauigkeit

Unter der geometrischen Genauigkeit wird im allgemeinen die Positionsgenauigkeit der benutzten geometrischen Primitive (Punkte, Linien, Flächen) verstanden. Es existieren eine Reihe von Einflüssen, die aus allen Komponenten eines GIS herrühren, die zur Entstehung der geometrischen Genauigkeit beitragen (Caspary, 1992b). Von Bedeutung für einen Anwender ist, daß alle Daten nur mit beschränkter Genauigkeit vorliegen, d.h. es wird niemals ein Datensatz anzutreffen sein, der exakt die Verhältnisse der realen Welt wiedergibt.

3.1 Genauigkeitsmodell: Stochastischer Ansatz

Es existieren eine Reihe von Modellen, die in der Lage sind, die mit der Geometrie verbundene Genauigkeit zu modellieren (Glemser, 1994). Grundlage für die vorliegenden Untersuchungen bildet ein stochastischer Genauigkeitsansatz (Glemser, 1996). Dieser eignet sich insbesondere, um die angestrebte Fortpflanzung der Genauigkeit mit Hilfe von Simulationen durchzuführen, da sich direkt zufallsbasierte Eingangsdatensätze erzeugen lassen. Zur Beschreibung der Genauigkeit stützt sich der Ansatz auf die Angabe der Parameter *Mittelwert* und *Varianz* der statistischen Verteilung der zu beschreibenden Größe. Die Größe wird dabei als *Zufallsvariable* angesehen. Im Falle einer Fläche sind alle Randlinien als solche Größen aufzufassen und müssen einzeln beschrieben werden. Zu jeder Linie L_i sind also ein Mittelwert μ_{L_i} und eine Varianz $\sigma_{L_i}^2$ anzugeben. Der Mittelwert ist dabei gleich der mittleren Position der Linie selbst (kurz: mittlere Geometrie). Die Varianz verdeutlicht die Variation in der Position der Linie. Für die Variation soll gelten, daß sie in jedem Linienpunkt nur senkrecht zur Richtung der Linie wirken soll. Da innerhalb eines geschlossenen Polygons Anfangs- und Endpunkte einer Linie keine herausragende Stellung besitzen, werden diese in gleicher Weise behandelt, wobei die zugelassenen Richtungen der Variation in diesen Punkten vom Brechungswinkel abhängen. Somit wird die *Zufallsvariable Linie* im eigentlichen Sinne durch die Gesamtheit aller Linienpunkte charakterisiert. Als zugehörige Verteilung wird die Normalverteilung der Linie angenommen.

4 Simulation und Testdatensätze

4.1 Simulationen und Monte-Carlo-Methode

Ganz allgemein versteht man unter dem Begriff der Simulation die Untersuchung eines Sachverhalts oder Systems mit Hilfe eines Ersatzsystems (Beispiel: Flugsimulation). Im mathematischen Sinne ist Simulation definiert als (Piehler, Zschiesche, 1976):

Verfahren zur Durchführung von Experimenten auf einem Digitalrechner unter Benutzung mathematischer Modelle mit dem Ziel, Aussagen über das Verhalten des Systems zu gewinnen

Simulationen können in solchen Fällen eingesetzt werden, in denen herkömmliche mathematische Verfahren wegen zu großem erforderlichen Aufwands keine vernünftigen Ergebnisse liefern. Zunächst ist ein geeignetes Ersatzsystem in Form eines mathematischen Modells aufzustellen. Mit Hilfe der Ergebnisse, die eine Simulation liefert, wird versucht, Aussagen über das reale System zu machen. Ein wesentlicher Nachteil beim Einsatz von Simulationen ist durch den hohen Rechenaufwand gegeben. Daher sollte stets für leistungsfähige Hardware gesorgt werden.

Eine spezielle Unterart stellt die *Monte-Carlo-Methode* (z.B. Sobol, 1985) dar. Sie basiert auf der Generierung von großen Mengen an Zufallswerten (Realisierungen der Zufallsvariablen), die als Stichprobe dienen. Das formulierte Modell wird auf alle Realisierung nacheinander angewendet, so daß für das Ergebnis wiederum eine Stichprobe vorliegt, die es anschließend zu analysieren gilt. Mit Hilfe der Monte-Carlo-Methode können all diejenigen Vorgänge simuliert werden, auf deren Ablauf sich zufällige Faktoren auswirken. Die Güte des abgeleiteten Ergebnisses ist in erster Linie vom Umfang der Stichprobe abhängig. Je mehr Zufallswerte prozessiert werden, desto zuverlässiger wird die Güte ausfallen. Eine Steigerung allerdings wird mit einer Zunahme des Rechenaufwandes erkauft.

4.2 Testdatensätze

Im Rahmen der Arbeit wird die Monte-Carlo-Methode genutzt, um die komplexe und individuelle Fortpflanzung der geometrischen Datengenauigkeit bei der Flächenverschneidung zu untersuchen. Die Methode ist einzusetzen, da als Modell für die geometrische Genauigkeit ein stochastischer Ansatz formuliert wurde, der die Randlinien der Flächen als Zufallsvariablen betrachtet.

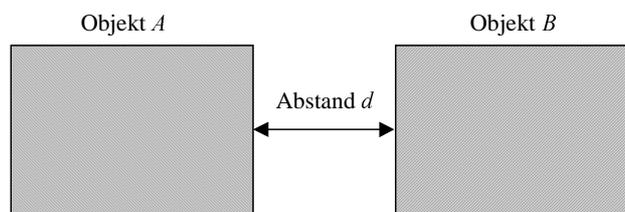


Abb. 2: Versuchsaufbau

Zur Durchführung der Simulationen ist zunächst eine große Anzahl an Realisierungen der zu untersuchenden Objekte in zufälliger Weise zu erzeugen. Der Versuchsaufbau (Abb. 2) beschränkt sich auf zwei einfache rechteckige Objekte *A* und *B*, die beide gleiche Form und gleiche Größe aufweisen und sich in einem bestimmten Abstand *d* zueinander befinden. Die dargestellte Objektgeometrie (Abb. 2) definiert die mittlere Geometrie der Linien. Alle Linien besitzen eine bestimmte Standardabweichung σ_{L_i} . Es werden verschiedene Datensätze erzeugt, die jeweils Realisierungen von beiden Objekten in gleicher Anzahl enthalten. Ein Beispieldatensatz ist in Abb. 3 zu sehen.

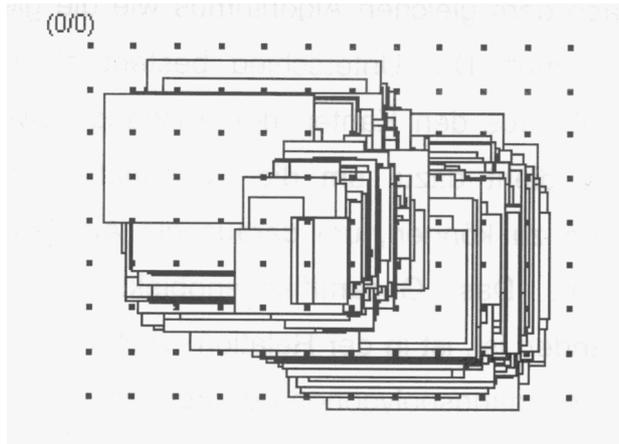


Abb. 3: Beispiel eines simulierten Datensatzes mit je 1000 Realisierungen zweier Objekte

Die Erzeugung der Realisierungen geschieht unabhängig für beide Objekte. Die verschiedenen Datensätze lassen sich in zwei Sequenzen anordnen. In der ersten Sequenz (S1) erhalten alle Linien die identische Standardabweichung $\sigma_{L_i} = \sigma_L = \sigma = 10$. Die einzelnen Datensätze der Sequenz unterscheiden sich nur durch den unterschiedlichen Abstand d der beiden Objekte. Die gültigen Abstände sind der Tab. 1 zu entnehmen. Negative Abstände entsprechen Überlagerungen der Objekte. In Datensatz 4 berühren sich die Objekte und in Datensatz 8 sind sie identisch.

Datensatz	1	2	3	4	5	6	7	8
Abstand	4σ	2σ	1σ	0	-1σ	-2σ	-4σ	-8σ

Tab. 1: Gewählte Abstände der Objekte A und B in den Datensätzen

Die zweite Sequenz (S2) unterscheidet sich von der ersten lediglich dadurch, daß für das Objekt B eine geringere Standardabweichung für die Linien von $\sigma_{L_i} = \sigma_L = 3$ angenommen wurde. Die Abstände dagegen sind identisch zu S1. Für jeden Datensatz wurden jeweils $n_A = n_B = n = 1000$ Realisierungen der Objekte erzeugt. Dieser Stichprobenumfang ist als guter Kompromiß zwischen Güte und Rechenaufwand anzusehen.

Die Simulation besteht nun daraus, daß jeweils die in einem Datensatz enthaltenen Objektrealisierungen paarweise miteinander verschnitten werden, so daß ein neuer Datensatz mit Realisierungen des Schnittobjektes entsteht, der die Grundlage für die Auswertungen und Ableitungen der Eigenschaften bildet. Die Simulation wird für alle Datensätze der beiden Sequenzen wiederholt.

5 Ergebnisse der Simulationen

Wie bereits in Abschnitt 1 erwähnt, konzentrierten sich die Untersuchungen auf ausgewählte Aufgaben, die von speziellem Interesse sind. Die einzelnen Aufgaben und die dabei erzielten Ergebnisse werden nachfolgend beschrieben und diskutiert. Die meisten Erläuterungen werden an den Daten der ersten Sequenz gegeben. Jedoch sind identischen Untersuchungen auch an den Daten der zweiten Sequenz ausgeführt worden, ohne daß darauf im speziellen eingegangen wird.

5.1 Wahrscheinlichkeit des Schnitts

Nicht alle Realisierungen der Objekte A und B in einem Datensatz schneiden sich, so daß nur eine gewisse Anzahl von Realisierungen des Schnittobjektes erzeugt wird. Abb. 4 zeigt die Häufigkeit des Schnitts für die verschiedenen Datensätze. Die Häufigkeit nimmt mit Zunahme des Abstandes ab. Während bei identischen Objekten alle Realisierungen einen Schnitt bilden, sind bei größeren Abständen (hier: ab $d = 4\sigma$) keine Schnitte mehr möglich. Im Falle von sich berührenden Objekten ($d=0$) liegt die Schnitthäufigkeit ungefähr bei der Hälfte der Realisierungsanzahl.

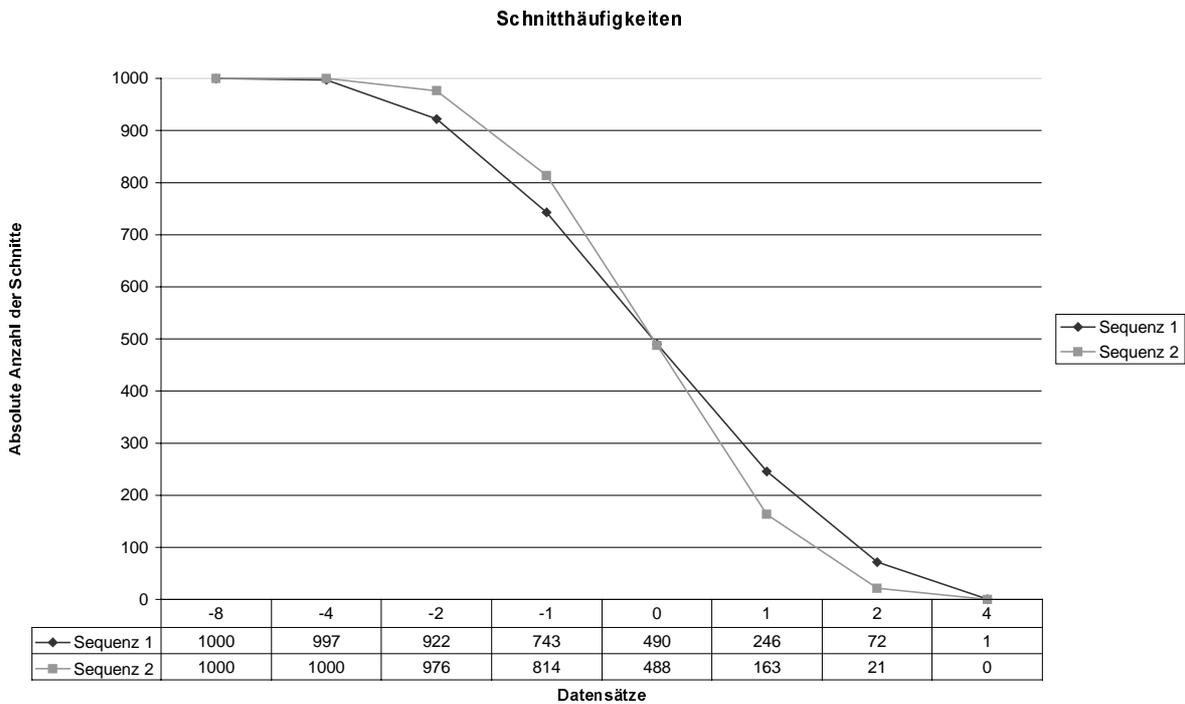


Abb. 4: Schnitthäufigkeiten

Aus diesen Werten läßt sich feststellen, daß die Wahrscheinlichkeit des Schnitts erwartungsgemäß vom Abstand d der Objekte abhängig ist. Da der Abstand in den Versuchen immer ein Vielfaches der Liniengenauigkeit beträgt, ist zusätzlich eine Abhängigkeit von der individuellen Liniengenauigkeit gegeben.

Der Abstand d kann wiederum als Zufallsvariable aufgefaßt werden, da er entsprechend der Liniengenauigkeiten variiert. Die Varianz des Abstandes σ_d^2 ergibt sich direkt aus den Liniengenauigkeiten der Objekte. Für den vorliegenden Versuchsaufbau gilt:

$$\sigma_d^2 = \sigma_{L_A}^2 + \sigma_{L_B}^2$$

Bei konstanter Liniengenauigkeit σ für beide Objekte ergibt sich dann:

$$\sigma_d^2 = \sqrt{2} \sigma$$

Aus der Annahme der Normalverteilung des Abstandes lassen sich daraus mit Hilfe der bekannten Verteilungsfunktion Wahrscheinlichkeiten für einen Schnitt p_{Schnitt} bestimmen:

$$p_{\text{Schnitt}} = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma_d} e^{-\frac{1(t-d)}{2\sigma_d^2}} dt$$

Wichtig ist die Feststellung, daß nur für $d < 0$ eine Überlagerung der Objekte (und auch der Realisierungen) und damit ein Schnitt vorliegt. Sie bestimmt die obere Integrationsgrenze der obigen Formel. Tab. 2 zeigt für die Sequenz 1 eine Gegenüberstellung der theoretisch aus der abgeleiteten Formel erzielten Werte für die Wahrscheinlichkeit des Schnitts mit den empirisch ermittelten Werten der relativen Häufigkeiten, die aus der Simulation gewonnen wurden. Die Werte stimmen gut überein, so daß die Formel durch die Simulationsergebnisse bestätigt werden kann.

Datensatz	1	2	3	4	5	6	7	8
Abstand	4σ	2σ	1σ	0	-1σ	-2σ	-4σ	-8σ
Theoretischer Wert	0.002	0.079	0.242	0.5	0.758	0.921	0.998	1.0
Relative Häufigkeit	0.001	0.072	0.246	0.490	0.743	0.922	0.997	1.0

Tab. 2: Wahrscheinlichkeit des Schnitts – Theoretischer Wert und relative Häufigkeit

5.2 Wahrscheinlichkeit der Objektzugehörigkeit

Die Objektzugehörigkeit stellt für beliebige Punkte fest, ob sie sich jeweils innerhalb oder außerhalb des Objektes befinden. Vereinigt man alle Punkte innerhalb eines Objektes zu einer Menge, dann wird dadurch die Objektausdehnung und damit die Geometrie des Objektes repräsentiert. Die Festlegung wird üblicherweise mit Hilfe des bekannten Punkt-im-Polygon-Tests getroffen. Ohne Beachtung der geometrischen Genauigkeit liefert ein solcher Test ein eindeutiges Ergebnis, wenn der Rand zum Inneren hinzuzählt. Die Eindeutigkeit geht verloren, wenn man die Objektlinien als stochastisch annimmt. Für diesen Fall ist die Objektzugehörigkeit durch einen Wahrscheinlichkeitswert auszudrücken.

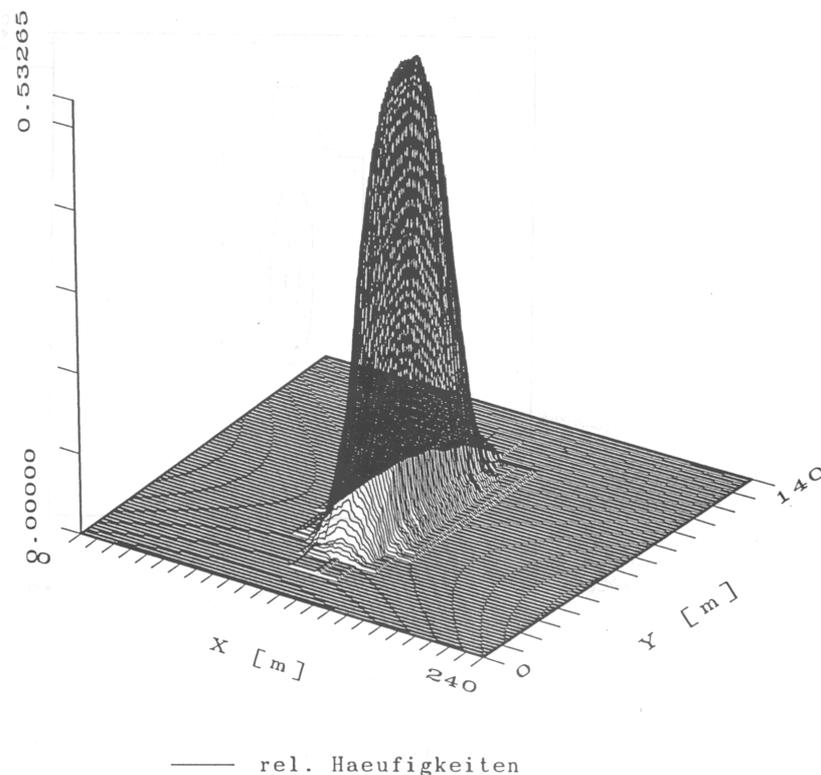


Abb. 5: Relative Häufigkeiten der Zugehörigkeiten zum Schnittobjekt

Zum Zwecke der Auswertung wird den simulierten Schnittobjekten eines Datensatzes ein gleichmäßiges Punktraster überlagert, für dessen Rasterpunkte jeweils die Häufigkeit der Ob-

jektzugehörigkeit zu den Realisierungen der Schnittobjekte bestimmt wird. Die relative Häufigkeit gemessen an der jeweiligen Anzahl der Schnittobjekte kann direkt als Wahrscheinlichkeit der Objektzugehörigkeit aufgefaßt werden. Abb. 5 veranschaulicht die relative Häufigkeit für den Datensatz 4 ($d=0$) der ersten Sequenz.

Der maximale Wert tritt in der Mitte der Schnittfläche auf, welches auch für die anderen Datensätze der ersten Sequenz gilt. Für die zweite Sequenz ist diese Position hin zum Objekt mit der geringeren Linienvarianz verschoben. Die relativen Häufigkeiten für diesen ausgewählten Punkt sind in der Abb. 6 zusammengestellt. Zusätzlich darin enthalten sind die Werte für Punkte auf der rechten Randlinie des Objektes A und für Punkte auf der linken Randlinie des Objektes B . Der maximale Wert nimmt mit Zunahme des Abstandes kontinuierlich ab. Somit ist die Wahrscheinlichkeit der Objektzugehörigkeit abhängig vom Abstand d der Objekte, von der Genauigkeit der Linien und von der Lage des Punktes relativ zu den beiden Objekten.

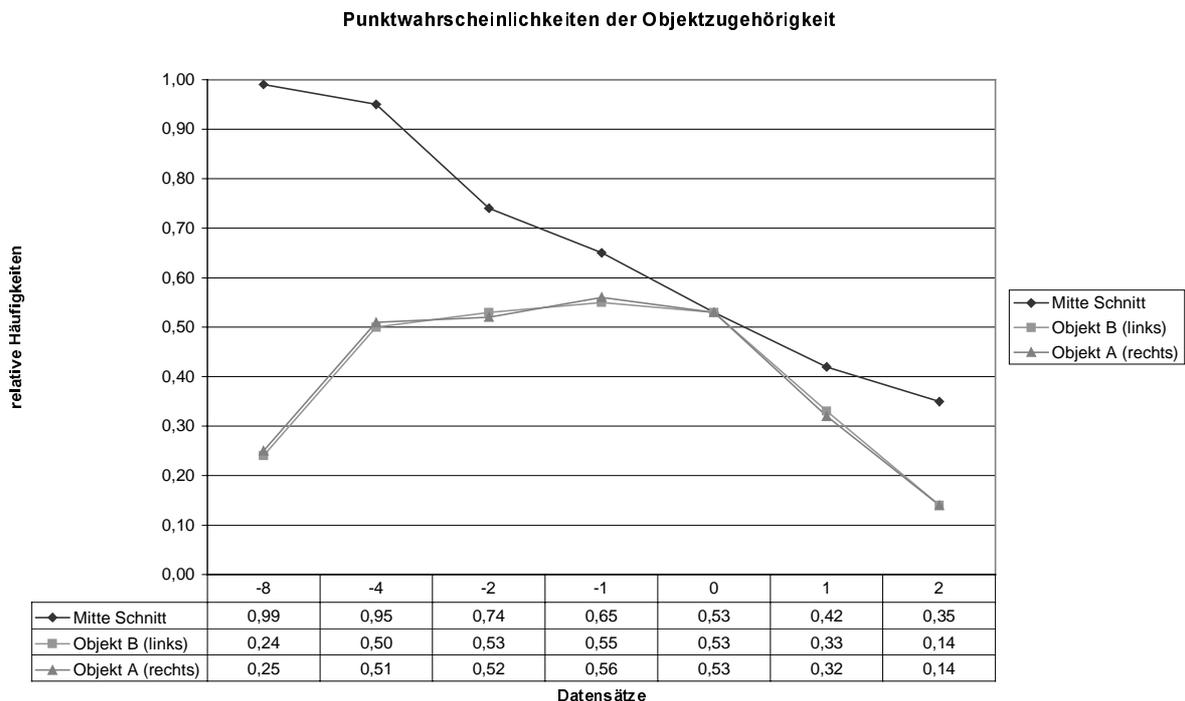


Abb. 6: Relative Häufigkeiten der Punktzugehörigkeiten zum Schnitt für ausgewählte Punkte

Betrachtet man die Flächenverschneidung im mengentheoretischen Sinn als *logisches UND*, dann kann die Wahrscheinlichkeit der Objektzugehörigkeit eines Punktes P zum Schnittobjekt S aus den Wahrscheinlichkeiten der Objektzugehörigkeiten der Eingangsobjekte A und B zu:

$$p_S = p(P \in S) = p((P \in A) \cap (P \in B)) = p(P \in A) \cdot p(P \in B) = p_A \cdot p_B$$

berechnet werden. Es gilt die Voraussetzung, daß die Zugehörigkeit des Punktes zu A und zu B unabhängig voneinander möglich ist. Z.B. ergeben sich für die Mitte des Schnitts im Datensatz 4 ($d=0$, Objekte berühren sich) $p_A = p_B = 0.5$ und damit $p_S = 0.25$. Ein Vergleich mit dem in der Simulation ermittelten Wert ($p_S = 0.53$) offenbart eine große Differenz. Begründen läßt sie sich durch die Bildung der relativen Häufigkeit als Quotient von absoluter Häufigkeit und Anzahl der existierenden Schnitte. Die relative Häufigkeit entspricht damit einer bedingten Wahrscheinlichkeit, mit der Bedingung, daß ein Schnittobjekt existiert. Aus Anwendersicht sind genau solche bedingten Wahrscheinlichkeiten erwünscht, da die Flächenverschneidung nur unter der Annahme eines existierenden Schnittes sinnvoll auszuführen ist. Die bedingte Form der obigen Formel ergibt sich zu:

$$p'_S = p(P \in S | Schnitt) = p((P \in A) \cap (P \in B) | Schnitt) = \frac{p((P \in S) \cap Schnitt)}{p(Schnitt)}$$

und damit zu:

$$p'_S = \frac{P_A \cdot P_B}{P_{Schnitt}}$$

Die Richtigkeit der Formel kann anhand der Simulationsergebnisse überprüft werden. Tab. 3 stellt die Ergebnisse der Berechnungen und die relativen Häufigkeiten am Punkt des größten Wertes (nur für Sequenz 1) zusammen. Somit ist bei der Bestimmung der Wahrscheinlichkeit der Punktzugehörigkeiten als zusätzlicher Parameter die Schnittwahrscheinlichkeit zu berücksichtigen.

Datensatz	1	2	3	4	5	6	7	8
Abstand	4σ	2σ	1σ	0	-1σ	-2σ	-4σ	-8σ
Theoretischer Wert	0.0	0.31	0.39	0.51	0.62	0.72	0.94	1.0
Relative Häufigkeit	0.0	0.34	0.42	0.53	0.64	0.74	0.94	0.99

Tab. 3: Maximale Wahrscheinlichkeiten der Zugehörigkeit zum Objekt

5.3 Wahrscheinlichste Geometrie des Schnittobjektes und Verschiebungsbetrag

Als *wahrscheinlichste Geometrie* des Schnittobjektes wird die Fläche bezeichnet, die durch die mittlere Lage der Randlinien gebildet wird. Sie wird als wahrscheinlichste Geometrie bezeichnet, da dort die Randlinien theoretisch am häufigsten auftreten. Im Falle der zugrundegelegten Normalverteilung gleicht sie der *mittleren Geometrie*, die sich aus dem Mittelwert der Linienpositionen definiert. Die wahrscheinlichste Geometrie ist für einen Nutzer von besonderem Interesse, da diese in der Regel in einem Datenbestand abgelegt sein sollte. Im Gegensatz dazu wird als *traditionelle Geometrie* die Darstellung bezeichnet, die sich ohne Berücksichtigung der geometrischen Genauigkeit ergeben würde. Traditionelle und wahrscheinlichste Geometrie sind für die Eingangsobjekte des Versuchs identisch.

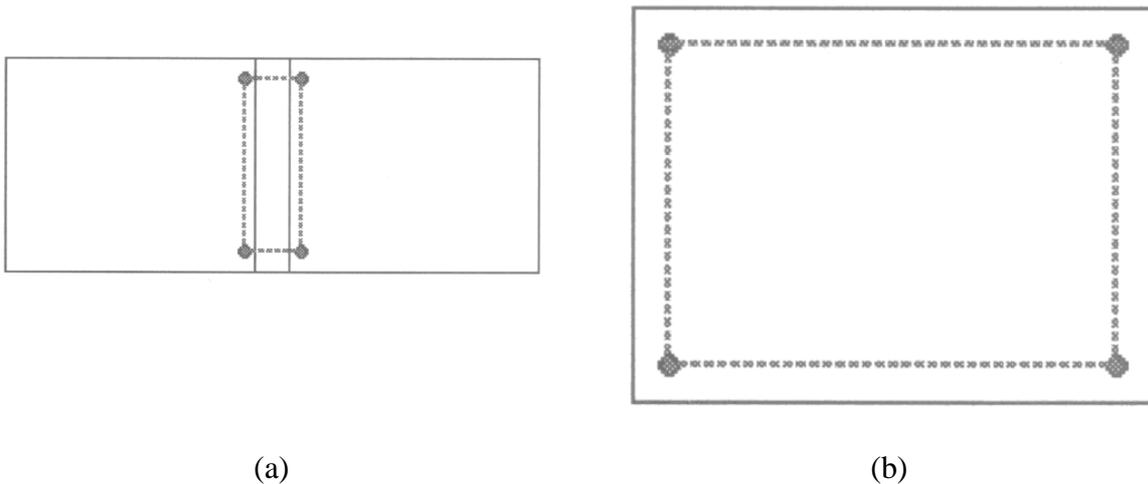


Abb. 7: Mittlere Geometrie (gestrichelt) für $d = -1\sigma$ (a) und $d = -8\sigma$ (b)

Die Bildung der traditionellen Geometrie eines Schnittobjektes (es wird eine traditionelle Flächenverschneidung mit den mittleren Eingangsobjekten durchgeführt) ist nur beim Vorhan-

densein einer Überlappung ($d < 0$) möglich. Im Falle der Datensätze 1-3 kann keine traditionelle Geometrie erzeugt werden und für den Datensatz 4 ergibt sich lediglich eine Schnittlinie. Dennoch werden während der Simulation auch in diesen Fällen Schnittobjekte erzeugt. Somit ist jeweils die Schätzung einer mittleren Geometrie möglich, indem die korrespondierenden Punkte aller erzeugter Realisierungen gemittelt werden. Abb. 7 zeigt diese Flächen als Überlagerungen zu den Eingangsobjekten für die Abstände $d = -1\sigma$ bzw. $d = -8\sigma$ (identische Objekte). Die mittlere Geometrie unterscheidet sich deutlich vom traditionell zu erwartenden Ergebnis. Differenzwerte zwischen mittlerer und traditioneller Geometrie sind für die linke und rechte Schnittlinie in der Abb. 8 aufgezeichnet. Negative Werte deuten auf eine kleinere Schnittgeometrie, positive Werte auf eine größere Schnittgeometrie als die traditionelle Geometrie hin. Es ist festzustellen, daß bei einer Überlagerung der beiden Objektflächen von ca. 50 % keine Differenzen auftreten. Mit Zu- oder Abnahme der Überlagerung werden die Differenzen immer größer. Bei Zunahme ist die mittlere Geometrie deutlich kleiner als die traditionelle Geometrie. Dagegen wird die mittlere Geometrie größer, wenn die Überlagerung abnimmt. Dieser Trend kann man sich auf die Fälle fortgesetzt vorstellen, zu denen keine traditionelle Schnittgeometrie mehr existiert. Über gleichartige Beobachtungen wird auch in Glemser (1996) berichtet.

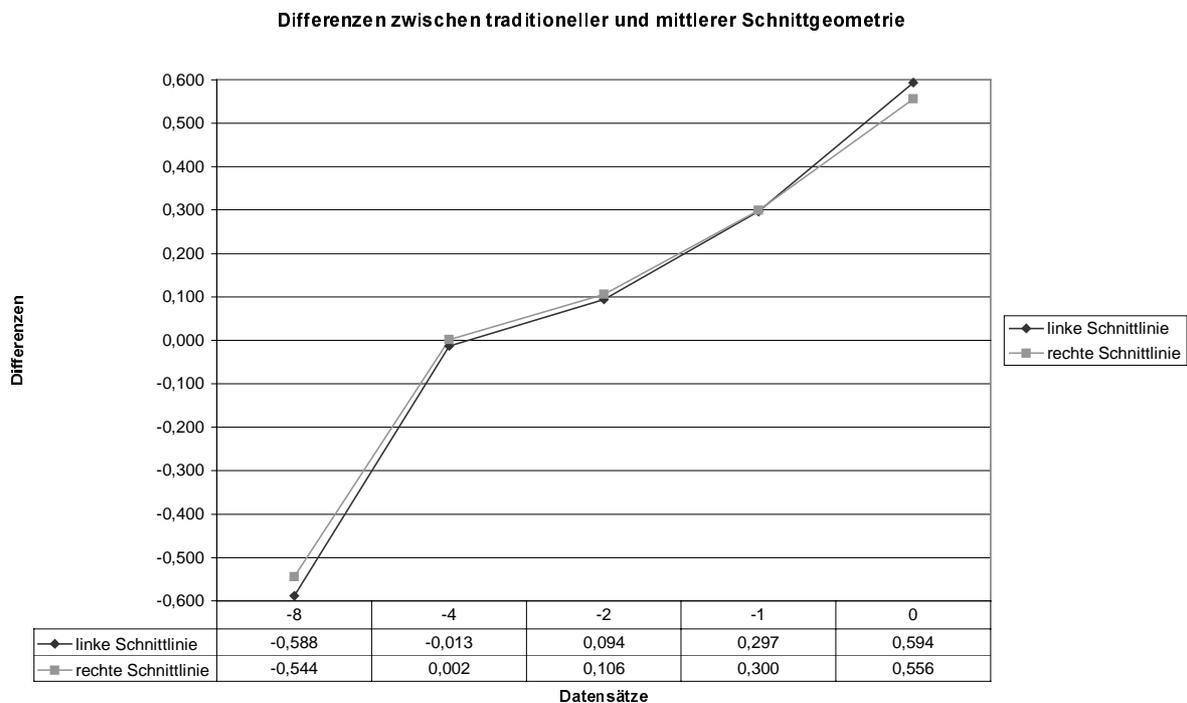


Abb. 8: Differenzen zwischen traditioneller und mittlerer Schnittgeometrie (mit Faktor σ)

Bei einer Überlagerung von 50 % sind beliebige Randlinien der Eingangsobjekte maximal voneinander entfernt. Bei Zu- bzw. Abnahme der Überlagerung nähern sich die Linien an. Eine Annäherung bedeutet demnach eine Verschiebung der mittleren Geometrie. Der minimale Abstand zwischen den Randlinien ist somit der entscheidende Parameter zur Bestimmung des Verschiebungsbetrages. Der Verschiebungsbetrag ist individuell an jedem Linienpunkt der traditionellen Geometrie zu bestimmen. Der minimale Abstand zwischen den Eingangsobjekten soll mit d_i bezeichnet werden und ist objektspezifisch zu messen. Der vergleichbare Abstand im Schnittobjekt d_s ist eine Zufallsvariable, die von der geometrischen

Genauigkeit der betreffenden Randlinien abhängt. Sie hat die Varianz $\sigma_{d_s}^2 = \sigma_{L_A}^2 + \sigma_{L_B}^2$. Ihr Mittelwert \bar{d}_s ergibt sich zu:

$$\bar{d}_s = \int_{-\infty}^{\infty} t f(d_s) dt$$

Unter der Bedingung $d_s > 0$ (nur dieser Wert ist sinnvoll) ist der Mittelwert mit einem Normierungsfaktor N zu normieren:

$$\bar{d}_s = \frac{1}{N} \int_0^{\infty} t f(d_s) dt$$

mit:

$$N = \int_0^{\infty} f(d_s) dt$$

Der gesuchte Verschiebungsbetrag ergibt sich letztlich zu:

$$\Delta = \frac{d_t - \bar{d}_s}{2}$$

Er ist nur für den Fall von Überlagerungen der Eingangsobjekte zu berechnen, da sonst keine traditionelle Schnittgeometrie existiert. Tab. 4 zeigt den Vergleich zwischen den aus der Simulation gewonnenen Daten und den theoretisch durch die Formel ermittelten Werten. Der Vergleich verifiziert die gefundene Berechnungsformel.

Datensatz	4	5	6	7	8
Abstand	0	-1σ	-2σ	-4σ	-8σ
Theoretischer Wert	0.564σ	0.289σ	0.113σ	0.0σ	-0.564σ
Simulationsergebnis	0.575σ	0.299σ	0.10σ	0.006σ	-0.567σ

Tab. 4: Verschiebungsbetrag – theoretischer Wert und aus Simulation

6 Zusammenfassung und Ausblick

Die Eigenschaften der Fehlerfortpflanzung bei der Flächenverschneidung wurden mit Hilfe einer Monte-Carlo-Simulation untersucht. Im einzelnen wurde festgestellt:

1. Die Wahrscheinlichkeit $p_{Schnitt}$, daß ein Schnitt vorliegt, hängt neben der Liniengenauigkeit entscheidend vom Abstand der Objekte ab.
2. Die Wahrscheinlichkeit p_s der Objektzugehörigkeit eines beliebigen Punktes wird von der Liniengenauigkeit, den Zugehörigkeiten zu den Eingangsobjekten und von der Schnittwahrscheinlichkeit beeinflusst.
3. Die wahrscheinlichste Geometrie stimmt nicht mit der traditionell ermittelten Schnittgeometrie, sofern diese überhaupt existiert, überein. Der Verschiebungsbetrag Δ hängt von den Liniengenauigkeiten und dem minimalen Abstand der beiden Objekte ab und ergibt einen individuellen Wert für alle Linienpunkte der traditionellen Geometrie.

Die durchgeführten Simulationen umfassen weitere Testreihen, zusätzlich zu den hier diskutierten, die unterschiedliche Liniengenauigkeiten und Überlappungsgrade abdecken (Henne-

berg, 1997). Sie sind bisher noch nicht vollständig ausgewertet, so daß weitere Erkenntnisse hinzukommen können. Übergeordnetes Ziel sollte jedoch sein, aus den gefundenen Eigenschaften ein Verfahren zu entwickeln, das die Fehlerfortpflanzung ohne aufwendige Simulationen erledigen kann.

7 Literatur

Bill, R. 1996: Grundlagen der Geo-Informationssysteme. Band 2: Analysen, Anwendungen und neue Entwicklungen. Wichmann, Heidelberg.

Caspary, W. (1992a): Qualitätsmerkmale von Geo-Daten. Zeitschrift für Vermessungswesen, Vol. 117, No. 7, 360-367, 1992.

Caspary, W. (1992b): Genauigkeit als Qualitätsmerkmal digitaler Datenbestände. In: Grünreich, D., Buziek, G. (Hrsg.) (1992), Gewinnung von Basisdaten für Geo-Informationssysteme. DVW-Schriftenreihe, Heft 4, 157-166.

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 der geometrischen Datenqualität in GIS-Funktionen. In: Proceedings des Workshops: Datenqualität und Metainformation in Geo-Informationssystemen, Universität Rostock.

Henneberg, C. (1997): Fortpflanzung der geometrischen Genauigkeit von Objekten bei der Flächenverschnidung. Diplomarbeit am Institut für Photogrammetrie, Universität Stuttgart (unveröffentlicht).

Höpke, W. (1980): Fehlerlehre und Ausgleichsrechnung. De Gruyter Verlag, Berlin.

Piehler, J, Zschiesche, H.-U. (1976): Simulationsmethoden. Mathematik für Ingenieure, Band 20. Teubner Verlag, Leipzig.

Sobol, I.M. (1985): Die Monte-Carlo-Methode. Deutsche Taschenbücher, Band 41, Verlag Harri Deutsch, Thun und Frankfurt am Main.