

Untersuchungen
zur geometrischen Datenkonsistenz
von ATKIS-Bayern

Bei der Fakultät
für Bauingenieur- und Vermessungswesen
der Universität Stuttgart
zur Erlangung der Würde eines Doktors-Ingenieurs (Dr.-Ing.)
eingereichte Dissertation

vorgelegt von

Dipl.-Ing. Silke Rossipal

aus Fürstenfeldbruck

München 1998

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München

Adresse der Deutschen Geodätischen Kommission:

Deutsche Geodätische Kommission

Marstallplatz 8 • D – 80 539 München

Telefon (089) 23 031 113 • Telefax (089) 23 031 – 100/ – 240

Telex 5 213 550 dgfi d • E-mail hornik@dgfi.badw-muenchen.de

Prüfungskommission:

Hauptberichter: Prof. Dr.-Ing.habil. Dieter Fritsch

Mitberichter: Prof. Dr.-Ing. Josef Frankenberger

Tag der Einreichung: 24.11.1997

Tag der mündlichen Prüfung: 16.01.1998

© 1998 Deutsche Geodätische Kommission, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen.

Inhaltsverzeichnis

Zusammenfassung	5
1 Einleitung	7
1.1 Grundlagen der Datenmodellierung	9
1.1.1 Konzeptionelles Modell	10
1.1.2 Logisches Modell	13
1.2 Einführende Betrachtungen zu ATKIS-Bayern	14
1.3 Grundsätzliche Betrachtungen zur geometrischen Genauigkeit von Geodaten	14
1.3.1 Inkonsistenzen bei der Definition geometrischer Grundprimitive	15
1.3.2 Erfassungsgrundlage mit nicht aktuellem Datenbestand	16
1.3.3 Lageverschiebungen aufgrund verschiedener Einpassungen (fehlerhafte Transformationsmodelle)	17
1.3.4 Isolierte (ikonisierte) Objektbereichserfassung	18
1.3.5 Generalisierungsfehler in der Erfassungsgrundlage	19
1.3.6 Was ist Redundanz bzw. was heißt Redundanzfreiheit?	20
1.4 Inhalt und Zielsetzung der Arbeit	21
2 ATKIS und ATKIS-Bayern	21
2.1 ATKIS, wie es 1989 als Standard der BRD eingeführt wurde	23
2.2 ATKIS in Bayern	28
2.3 Unterschiede zwischen ATKIS und ATKIS-Bayern - Vergangenheit und Gegenwart	31
2.3.1 Lagegenauigkeit	32
2.3.2 Erfassungsgrundlage	32
2.3.3 Datenmodelle	32
2.4 Neueste Tendenzen bei ATKIS	34
2.5 Klassifikation und Wertung von geometrischen Dateninkonsistenzen	36
3 Modelltheoretische Behandlung von geometrischen Dateninkonsistenzen	38
3.1 Grundlagen und Voraussetzungen	38
3.1.1 Transformationen	38
3.1.2 Generalisierungsmethoden für Linien	40
3.1.3 Grundsätzliche Überlegungen zur Modellierung der Genauigkeit	46
3.2 Detaillierte Beschreibung des hier angewendeten Verfahrens	49
3.2.1 Modellierung der Genauigkeit	49
3.2.1.1 Vergleich zwischen ATKIS-Bayern und DFK (innerorts)	51
3.2.1.2 Vergleich zwischen ATKIS-Bayern und Orthophoto	51
3.2.2 Inkonsistenzen innerhalb des Objektbereichs Wald aufsuchen, selektieren und bereinigen	53
3.2.2.1 Besonderheiten des Objektbereichs Wald	53
3.2.2.2 Bestimmung der Drehrichtung eines geschlossenen Polygons	53
3.2.2.3 Waldpolygone am Blattschnitt der TK 25	54
3.2.2.4 Sackgassen in Waldpolygonen - Einschnitte beseitigen	54
3.2.2.5 Durch Straßen getrennte Waldpolygone zusammenführen	58
3.2.2.6 Bildung neuer Inselfpolygone	62
3.2.3 Inkonsistenzen zwischen den Objektbereichen Wald und Siedlung aufsuchen, selektieren und bereinigen	63
3.2.3.1 Besonderheiten des Objektbereichs Siedlung	63
3.2.3.2 Zusammentreffen von Wald- und Siedlungsgebieten	63
3.2.3.3 Weitere Möglichkeiten	64
3.3 Bewertung der vorgestellten Verfahren	64

4	Exemplarische Modellrechnungen	64
4.1	Testgebiete - die Auswahl des Testgebiets Althegegnberg	64
4.2	Untersuchungsergebnisse	65
4.2.1	Ergebnisse der Genauigkeitsuntersuchungen	65
4.2.1.1	Vergleich zwischen der DFK und ATKIS-Bayern	65
4.2.1.2	Vergleich zwischen Orthophotos und ATKIS-Bayern	66
4.2.1.3	Wertung	66
4.2.2	Ergebnis der Transformationen innerorts im Objektbereich Verkehr	68
4.2.3	Weitere Auffälligkeiten im Testgebiet	69
4.3	Vorstellung und Bewertung der Testergebnisse	70
4.3.1	Der Testdatensatz vor und nach der Datenbereinigung	70
4.3.2	Zusammenfassende Analyse der Testergebnisse	70
5	Bedeutung für ATKIS-Bayern	72
6	Ausblick	73
7	Entstehung der Arbeit, Dank	77
8	Literatur	78
9	Bildnachweis	80

Anhang 81

A	Datenbereitstellung - Datenimport	81
B	Mathematische Grundlagen von fehlerhaften Geometriedaten - grundsätzliche fehlertheoretische Modelle	86
B.1	Fehlerfortpflanzung	86
B.2	Fuzzy-Datenmodelle	88
C	Testgebiet	95
C.1	Testgebiet vor der Datenbereinigung	95
C.2	Testgebiet nach der Datenbereinigung	95
	Lebenslauf	98

Zusammenfassung

Das Amtliche Topographisch-Kartographische Informationssystem (ATKIS) wird bundesweit seit 1989 von den Vermessungsverwaltungen der Länder aufgebaut. Damit werden interessenneutral Geobasisdaten für staatliche, öffentliche und private Interessenten bereitgestellt. In Bayern wurde bereits 1986 mit dem Aufbau eines Basisinformationssystems begonnen, so daß sich die ursprüngliche Erfassungsmethode von der späteren bundesweiten Vorgabe unterscheidet. Das bayerische System wird als ATKIS-Bayern bezeichnet.

Diese Unterschiede zwischen ATKIS und ATKIS-Bayern werden dargestellt, und vor allem unter dem Aspekt der geometrischen Genauigkeit beleuchtet und bewertet. Dazu werden die möglichen Erfassungsfehler und die daraus resultierenden Dateninkonsistenzen klassifiziert und es wird insbesondere auf die bayerischen Gegebenheiten eingegangen. Da in ATKIS-Bayern im Gegensatz zur der objektweisen Erfassung in ATKIS die Objekte der realen Welt in einer Ebenenstruktur abgelegt wurden, kommt es bei der Überlagerung dieser Ebenen zu Redundanzen und geometrischen Dateninkonsistenzen im Gesamtdatenbestand. Im Sinne einer bundeseinheitlichen Gesamtlösung werden derzeit die bayerischen Daten ebenfalls in die Objektstruktur von ATKIS umgesetzt. Daher gilt es die Redundanzen und Inkonsistenzen möglichst automatisiert zu beseitigen, um daraus möglich schnell die gewünschte Objektstruktur ableiten zu können. Derzeit erfolgt diese Umsetzung jedoch durch manuelle Nachbearbeitung und ist damit sehr zeit- und kostenintensiv. In dieser Arbeit werden Möglichkeiten zur halbautomatischen Beseitigung der Redundanzen aufgezeigt.

Zur Entwicklung derartiger Algorithmen wurde zuerst die geometrische Genauigkeit der ATKIS-Bayern-Daten modelliert und zwar getrennt nach Objektbereichen. Die Algorithmen wurden so entwickelt, daß sie auf jedem gängigen GIS-Produkt bzw. CAD-System nachgebildet werden können. In einem Testgebiet wurde eine Datenbereinigung durchgeführt.

Zuletzt werden weitere Modellierungsmöglichkeiten, z. B. auch die Zuhilfenahme der Fuzzy-Logik, für das vorgestellte Verfahren aufgezeigt und weitere mögliche Einsatzgebiete vorgestellt.

1 Einleitung

UMWELT - ein Wort, das heute in aller Munde ist. Gerade in den letzten Jahren hat das Thema Umwelt mehr und mehr an Bedeutung gewonnen. Es wird im Zusammenhang mit vielen Aspekten verwendet, z.B. im sparsamen Einsatz von Rohstoff- und Energievorkommen, der Reinhaltung von Luft und Wasser, als Lebensraum für Mensch und Tier sowie der Schutz von Grund und Boden als nicht vermehrbares Gut. Diese immer wichtiger werdenden Maximen sind mit unserem Streben nach wirtschaftlichem Wachstum, einem immer höherem Freizeitangebot, Forderung jedes Einzelnen nach größerer Unabhängigkeit und Mobilität, steigendem Lebensstandard und immer höherer Produktivität bei geringerem Arbeitseinsatz oft kaum vereinbar. Um all diese Forderungen in einem ökonomisch und ökologisch vertretbaren Rahmen sinnvoll miteinander zu verbinden, benötigt man umfassende und zuverlässige Informationen über unsere Umwelt - unseren Lebensraum. Diese müssen schnell verfügbar und analysierbar sein. Außerdem müssen sie schnell und einfach, sowohl in digitaler als auch in analoger Form visualisierbar sein, so daß sie auch dem Laien verständlich und plausibel gemacht werden können. Dabei spielt der Zeitfaktor in unserer schnellebigen Gesellschaft eine wesentliche Rolle. Informationen, die erst mühsam aus den verschiedensten Quellen zusammengetragen werden müssen und nicht, oder nur zum Teil, in digitaler Form vorliegen, sind dafür wenig geeignet; doch in vielen Fällen bietet sich derzeit noch keine Alternative. Hier gilt es, möglichst schnell Abhilfe zu schaffen, da die mangelnde Verfügbarkeit der Daten zu Fehlentscheidungen bei Projekten und ihrer Planung führen kann, was wiederum negative Auswirkungen auf unsere Umwelt hat. Dies ist ein nicht tragbarer Zustand, da gerade in den Industrieländern die Natur ohnehin stark belastet wird, so daß derartige Fehlentscheidungen ein nicht zu vertretendes Risiko darstellen.

Um dies zu verhindern, ist eine Sensibilisierung aller Betroffenen - Planer, Bürger und Industrie - erforderlich. Dies allein ist aber noch nicht ausreichend. Mindestens genauso wichtig ist die kooperative, fachübergreifende Zusammenarbeit. Zusammenarbeit im Bereich einzelner Verwaltungen und Organisationen, der Industrie und der Forschungseinrichtungen der Universitäten im Bereich der Geowissenschaften auch über die Staatsgrenzen hinweg ist unerlässlich. Anders sind interdisziplinäre Untersuchungen, wie z. B. Umweltverträglichkeitsprüfungen, um die Fachdaten mehrerer Dienststellen gemeinsam auszuwerten, zu analysieren und in anschaulicher, verständlicher Form zu präsentieren, nicht in angemessener Zeit und wirtschaftlich durchführbar. Der Faktor Zeit spielt dabei eine wesentliche Rolle, da nur so auf veränderte Vorgaben in angemessener Zeit reagiert werden kann.

Um Fachdaten verschiedener Stellen miteinander zu kombinieren, ist es wirtschaftlich und zeitsparend, eine gemeinsame Datenbasis zu verwenden. So kann sichergestellt werden, daß alle Partner mit den gleichen Basisinformationen arbeiten, die dann auch leichter zusammengeführt werden können.

Wie muß demnach eine Datenbasis konzipiert sein, in die auch Fachdaten verschiedener Herkunft integriert werden können, und die eine derartige Fülle an Daten verwalten, in angemessener Zeit geeignet analysieren und graphisch präsentieren kann?

Zweifelsohne können diese Aufgaben zukünftig nur mehr durch den Einsatz digitaler Daten und graphischer Datenverarbeitung mit moderner Computertechnologie in angemessener Zeit effektiv, überzeugend und aussagekräftig erledigt sowie anschaulich und übersichtlich präsentiert werden. Hierfür eignen sich Geo-Informationssysteme (GIS). In der Literatur finden sich die unterschiedlichsten Definitionen für Geo-Informationssysteme. Wesentlich jedoch sind die Bestandteile eines Geo-Informationssystems; es setzt sich aus vier Komponenten zusammen, die sich in ihrer Lebensdauer wesentlich unterscheiden:

- | | | | |
|--------------------|--------------|--------------|------------------|
| - der Hardware, | Lebensdauer: | 3-5 Jahre | |
| - der Software, | Lebensdauer: | 7-10 Jahre | |
| - den Geodaten, | Lebensdauer: | 30-100 Jahre | |
| - den Anwendungen, | Zyklus: | 10-90 Jahre | [Fritsch, 1993]. |

Die Kosten verhalten sich überschlägig zu: **Hardware : Software : Geodaten = 1 : 10 : 100**

Die wertvollste und langlebigste Komponente (abgesehen vom Anwender) eines Geo-Informationssystems sind somit die Geodaten. Mit ihrer Qualität steht und fällt das ganze System.

Geodaten werden schon seit dem 19. Jahrhundert systematisch durch Geometer erfaßt. Die erste systematische parzellenscharfe Erfassung diente der gerechten Besteuerung von Grund und Boden (Grundsteuerkataster). Beschreibt man diesen Vorgang mit modernem Wortschatz, so könnte dies lauten:

Es wurden die Geo-Basisdaten (Besitzgrenzen) unter Beachtung der Wirtschaftlichkeit in ausreichendem Maßstab (1:5000) erfaßt, mit den geeigneten Fachdaten Wirtschaftsart (z.B. Acker, Wiese, Hofraum, Gebäude, Wald, Sumpf

usw.) angereichert, anschließend durch ein passendes Verfahren taxiert (Schlüssel für die Besteuerung) und abschließend ausgewertet (Steuerbescheid an den Eigentümer verschickt).

Seither ist der Vermessungsingenieur (Geometer, Geodät) mit dem Aufbau und der Fortführung von Geodaten beschäftigt. Der gesetzliche Auftrag dazu ist bundesweit den Vermessungsverwaltungen der einzelnen Länder erteilt und z. B. in Bayern im Vermessungs- und Katastergesetz festgelegt. Mit Einführung des Bürgerlichen Gesetzbuchs (BGB) im Jahre 1900 wurde das Grundbuch für den rechtlichen Nachweis des Eigentums an Grund und Boden eingerichtet; das Liegenschaftskataster ist amtliches Verzeichnis des Grundbuchs.

Auch andere Behörden und Institutionen führen raumbezogene Daten, die oft ähnlich einem Kataster aufgebaut sind (z. B. Waldschadenskartierungen, Biotopkartierungen usw.). Aber auch ein Hotelführer oder eine Karte mit mautpflichtigen Straßen und Tunnels enthält Geodaten bzw. Geoinformationen. Hier müssen jedoch die Basis- und Fachdaten voneinander unterschieden werden. Lediglich die Erfassung und Fortführung der Basisdaten liegt bei den Vermessungsverwaltungen der einzelnen Bundesländer.

Die Erfassung und Fortführung der Geodaten muß, um den heutigen Ansprüchen gerecht zu werden, digital erfolgen. Blickt man ein wenig in die geodätische Geschichte zurück, dann zeigt sich, daß die Geodäten sich schon immer moderne Technik zu Nutze machten. So wurde der 1910 in Berlin geborene Conrad Zuse, der 1934 mit der Entwicklung einer programmgesteuerten Rechenmaschine begann und 1941 das erste programmgesteuerte Rechenggerät der Welt vorstellte, bei der Konzeption dieser wesentlich von dem Vermessungsingenieur Seifert unterstützt, der am Geodätischen Institut der Technischen Hochschule München beschäftigt war. Auch in dem klassischen geodätischen Aufgabenbereich, dem Nivellement und der topographischen Geländeaufnahme, wird die Datenverarbeitung sowohl bei der Erfassung als auch der Weiterverarbeitung der Rohdaten schon lange eingesetzt. In der topographischen Kartographie, deren analoges topographisches Kartenwerk den Ansprüchen des 20. Jahrhunderts nicht mehr genügt, war die Datenverarbeitung dagegen bisher praktisch kaum vertreten. Eine Veränderung ergab sich erst durch die Forderung nach dem Aufbau digitaler Basisdaten für Geo-Informationssysteme. Seitdem dürften wohl die größten Investitionskosten der Landesvermessungsämter in diesem Bereich liegen, der damit stark an Bedeutung gewonnen hat.

Bei der Planung für die Überführung der analogen in digitale Daten müssen folgende Punkte beachtet werden:

- Grundlage muß ein einheitliches räumliches Bezugssystem sein, in dem die topographischen Objektinformationen in einer exakten Geometrie digital abgespeichert werden, und dann als Basisdaten für externe GIS (Fachinformationssysteme) zur Verfügung stehen;
- eine rationelle, möglichst digitale Herstellung der topographischen Landeskartenwerke soll ermöglicht werden;
- die logische Datenstruktur muß so aufgebaut sein, daß eine Verknüpfung mit fachspezifischen Datenmodellen ohne Informationsverlust mit möglichst geringem Aufwand durchführbar ist.

Mit diesen Prämissen sollte möglichst schnell ein Konzept zur Schaffung, Bereitstellung und Fortführung eines digitalen Datenbestandes ausgearbeitet werden. Dieser Basisdatenbestand sollte von Landesvermessungsbehörden erfaßt und geführt werden. Daraus entwickelte sich das Jahrhundertprojekt ATKIS, das 1989 durch die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) ausgearbeitet und ins Leben gerufen wurde. ATKIS ist die Abkürzung für Amtliches Topographisch-Kartographisches Informationssystem und soll digitale Basisdaten im mittel- und kleinmaßstäbigen Bereich zur Verfügung stellen. Hier galt es schnell zu handeln, nachdem bei mehreren Institutionen im regionalen Bereich bereits ähnliche Absichten zu erkennen waren, oder diese gar schon begonnen hatten, eine eigene digitale Datenbasis aufzubauen. Hier sind z.B. das Statistische Bodeninformationssystem STABIS des Statistischen Bundesamtes oder das Maßstaborientierte Einheitliche Raumbezugssystem für Kommunale Informationssysteme MERKIS des Deutschen Städtetages, welche dieselben Ziele wie ATKIS, jedoch im großmaßstäbigen Stadtplanbereich, verfolgen, zu erwähnen. Viele weitere Systeme im Bereich des Umweltschutzes, der Raumordnung und im Zusammenhang mit Verkehrsleitsystemen, um nur einige zu nennen, werden derzeit mit großem finanziellen und technischen Aufwand aufgebaut, wobei die Frage der Fortführung oft vernachlässigt wird.

Eine Mehrfacherfassung und -fortführung ist auch aus wirtschaftlicher Sicht nicht sinnvoll, zumal die qualitativ hochwertigsten Daten, aufgrund des gesetzlichen Auftrags, die Topographie der Erdoberfläche zu erfassen, zu dokumentieren und zu aktualisieren, bei den Vermessungsverwaltungen der Länder liegen. Außerdem ist hier die Fortführung gewährleistet. Die Vermessungsverwaltungen der Bundesländer haben daraufhin Ende der achtziger Jahre beschlossen, das digitale Amtliche Topographisch-Kartographische Informationssystem ATKIS bei den Landesvermessungsämtern aufzubauen und zu entwickeln. ATKIS soll alle wesentlichen topographischen Informationen objektweise erfassen und digital abspeichern. Im Gegensatz zu den topographischen Landeskartenwerken, in denen wesentliche topographische Informationen in kartographischen Signaturen verschlüsselt sind und damit nicht originär zur Verfügung stehen, ist hier ein direkter Zugriff auf jedes einzelne Objekt möglich. ATKIS kann und will nur Basisdaten zur Verfügung stellen, die dem Benutzer eine einfache und zuverlässige Bereitstellung von aktuellen, authentischen topographischen Daten

gewährleistet. Beim Aufbau der Basisdaten soll jedoch besonders auf die Bedürfnisse von Verwaltung, Rechtsverkehr, Wirtschaft, Verkehr, Landesplanung, Bauleitplanung, Bodenordnung, Verteidigung und Forschung Rücksicht genommen werden. Um Systemunabhängigkeit zu erreichen, sollen die Daten dann über die **Einheitliche Datenbankschnittstelle (EDBS)** an die Nutzer abgegeben werden können, die hierauf dann optional ihre Fachdaten aufsetzen. Dem gesetzlichen Auftrag entsprechend, wird der Datenbestand als staatliche Dienstleistung interessenneutral geführt und auf dem aktuellen Stand gehalten. Die Basisdaten werden staatlichen, öffentlichen und privaten Interessenten gegen eine Gebühr zu Verfügung gestellt.

1.1 Grundlagen der Datenmodellierung

*Ein GIS-Datenmodell ist das Ergebnis eines **Modellierungsprozesses**, bei dem der betrachtete **Ausschnitt der realen Welt** unter Einsatz bestimmter Techniken und Methoden auf systematische Art und Weise **strukturiert und beschrieben** wird. Das Ziel ist dabei, ein für die jeweilige **Anwendung** geeignetes **Abbild der realen Welt** zu erhalten [Kaltenbach, 1995].*

Datenmodelle im obigem Sinn sind solche, die am Computer implementiert werden sollen. Zur Implementierung von Datenmodellen stehen verschiedene Techniken und Methoden zur Verfügung. Dies sind: Entity-Relationship-Technik, Ebenentechnik, die Objektklassentechnik, CAD-Datenstrukturen, topologische Datenstrukturen und hierarchische, netzwerkartige, relationale oder objektorientierte Datenbankmodelle. Diese Aufzählung zeigt nur die Vielzahl der Möglichkeiten, ist aber nicht abschließend.

Die Art der Modellierung mit dem zugrundeliegenden Datenmodell ist entscheidend für die spätere Nutzbarkeit und die Leistungsfähigkeit des Geo-Informationssystems. Denn bei der Modellierung wird festgelegt, welche Objekte der realen Welt das GIS enthält, welche Eigenschaften der Objekte erfasst werden und welche Beziehungen zwischen den Objekten untereinander bestehen.

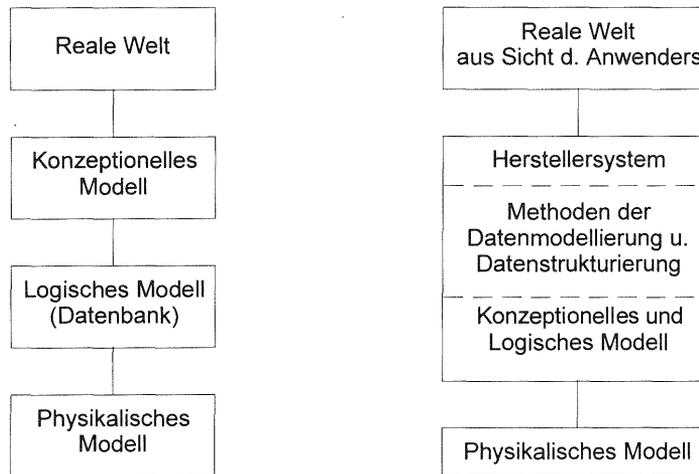


Abbildung 1.1:
Theoretische Sicht der
Modellierungsstufen

Praktische Sicht der
Modellierungsstufen

Die Modellierung ist der Prozeß, der beschreibt, wie ein bestimmter Ausschnitt der realen Welt zu einem digitalen Abbild im Computer gelangt. Dieser Prozeß teilt sich in vier Phasen. Ausgangspunkt ist die reale Welt, die aber immer durch die Brille der jeweiligen Anwendung betrachtet wird (z. B. Kataster, Forst, Energieversorgung). In einem konzeptionellen Modell wird festgelegt, **welche** Objekte, mit **welchen** charakteristischen **Merkmale**n, mit **welchen** Beziehungen untereinander über **welche** Modellierungstechnik in das GIS abgebildet werden. Die Daten müssen nach bestimmten Kriterien geordnet, strukturiert und klassifiziert werden. Dies sollte nach Möglichkeit mit standardisierten Modellierungstechniken geschehen; dann ist, zumindest theoretisch, das Modell unabhängig von einer speziellen Software. Im nächsten Schritt werden die strukturierten Daten in die logischen Strukturen der Datenbank umgesetzt. Derzeit kommen hier vorwiegend relationale Datenbanken mit einem entsprechenden Datenbankmanagementsystem (RDBMS) zum Einsatz. Der letzte Schritt ist die Abbildung in das physikalische Modell, also die Speicherung der Daten.

In der Praxis sind diese vier Stufen nicht streng voneinander zu trennen und systemabhängig. Denn jeder Anwender - und in diesem Fall ist auch ein Landesvermessungsamt ein Anwender - entscheidet sich für ein Herstellersystem, welches sowohl Methoden zur Strukturierung der Daten als auch eine Datenbank zur Verfügung stellt. Damit ist eine Trennung zwischen konzeptionellem und logischem Modell nicht mehr möglich, denn die Funktionen des Herstellersystems, auch die graphischen, bestimmen die konzeptionelle Datenmodellierung. Bietet das System lediglich die Ebenentechnik nicht aber eine Objektklassentechnik an, so ist bereits das konzeptionelle Modell nicht mehr systemunabhängig.

1.1.1 Konzeptionelles Modell

Ausgangspunkt für das konzeptionelle Modell sind die Objekte der realen Welt. Aufgrund seiner Erfahrungen nimmt der Mensch diese Objekte unterschiedlich wahr. Fragt man beispielsweise ein Kind, einen Kunsthistoriker und einen Besucher, aus welcher Kirche sie soeben herauskommen, so erhält man sicherlich die unterschiedlichsten Antworten.

Kind:	die Kirche ist groß und dunkel
Kunsthistoriker:	ein gotisches Bauwerk mit einem Haupt- und einem Querschiff.....
Besucher:	eine katholische Kirche mit einer wunderbar klingenden Orgel

Dieses kleine Beispiel soll zeigen, daß jeder Betrachter seine eigene Sichtweise von der realen Welt hat. Genauso ist das bei verschiedenen Anwendern von Geodaten; jeder betrachtet die zu erfassenden Objekte durch „seine Anwenderbrille“. Für jemand, der im Bereich Liegenschaftskataster tätig ist, sind relevante Objekte Flurstücksgrenzen, Gebäude und Katasterfestpunkte usw., während einen Umweltschützer die Belastung des Bodens und der Gewässer mit Schadstoffen usw. interessiert.

Zu Beginn der Konzeptionsphase ist zu entscheiden,

- welche Objekte für die Erfassung relevant sind,
- welche Merkmale und Eigenschaften dieser Objekte berücksichtigt werden (Größe, Nutzung, Koordinate),
- welche Beziehungen zwischen den Objekten bestehen (z. B. eine Straße führt über einen Bach),
- über die systematische Strukturierung der Daten (z. B. zur Klasse der Verkehrswege gehören Eisenbahnen, Straßen oder auch Luftfahrtstraßen).

Außerdem wird im konzeptionellen Modell festgelegt, ob und wie die **dritte Dimension** im GIS berücksichtigt wird.

Ein weiterer wichtiger Aspekt ist die Date **Zeit**. Hier stellt sich die Frage, welche Zeit gegebenenfalls zu berücksichtigen ist. Dafür kommen in Frage:

- die Zeit, zu der das Objekt entstanden ist (z. B. Baujahr einer Straße),
- der Zeitpunkt, zu dem es erstmals im GIS erfaßt wurde,
- der Zeitpunkt, an dem es zuletzt in der Datenbank geändert wurde.

Auch eine konsequente Historienführung ist gegebenenfalls zu berücksichtigen. Diese ermöglicht das „**Leben**“ eines Objekts zu verfolgen, wann es erstmals in Erscheinung getreten ist, geändert und ggf. gelöscht wurde. Auch ist zu klären, ob ein einmal gelöscht Objekt wieder erscheinen kann oder ob ein Löschvorgang nur einmal zulässig ist.

Des weiteren muß entschieden werden, welche Arten von geometrischen Daten zum Einsatz kommen - Rasterdaten oder Vektordaten oder beide.

Ein GIS-Objekt setzt sich i. d. R. aus Geometrie- und Sachdaten (auch als Attribute oder thematische Daten bezeichnet) zusammen. Aufgrund dieser Unterteilung unterscheidet man auch bei der Modellierung zwischen thematischem und geometrisch-topologischem Modellieren.

Thematisches Modellieren:

Unter thematischem Modellieren versteht man die Gliederung der realen Welt nach ihrer Thematik, wobei dies immer von der Anwendung abhängig ist. Dafür werden im allgemeinen zwei verschiedene Verfahren angewendet.

1. Das **Ebenenprinzip**: Die verschiedenen thematischen Daten werden streng durch die Abspeicherung in verschiedenen Ebenen voneinander getrennt. Werden alle Ebenen übereinander gelegt, so erhält man den Gesamtdatenbestand. Beim Ebenenprinzip ist das Folienprinzip der analogen Karten in die digitale Welt übertragen. Dieses relativ einfache Prinzip ist in vielen GIS realisiert.

2. Das **Objektklassenprinzip**: Hier erfolgt die Anordnung der verschiedenen Thematiken hierarchisch. Die Daten werden durch die Bildung von Klassen strukturiert, indem gleichartige Objekte zu Klassen zusammengefaßt werden. Klassen, die zum Teil gemeinsame Eigenschaften haben, werden zu einer Oberklasse (Hyperklasse) zusammengefaßt, z. B. die Klassen Autobahn, Bundesstraße, Kreisstraße usw. zu der Oberklasse Straße. Die Anzahl der Oberklassen ist beliebig, sollte aber überschaubar bleiben. Nicht jede Objekteigenschaft ist Grund eine eigene Klasse zu bilden. Objekteigenschaften wie Breite, Anzahl der Fahrstreifen usw. werden als Sachdaten gespeichert.

Auch die Sachdaten können auf unterschiedliche Art und Weise modelliert werden. Hierfür bieten sich im wesentlichen auch zwei Möglichkeiten an:

1. Die Objekteigenschaften werden als Attribute bei den Geometrieelementen abgespeichert, indem an das Geometrieelement (Punkt, Linie oder Fläche) ein Parameter (Deskriptor) angehängt wird, der die Eigenschaften dieses Objekts beschreibt. Diese Form der Modellierung ist wenig flexibel und nur dann sinnvoll, wenn nicht zu viele Sachdaten vorhanden sind.

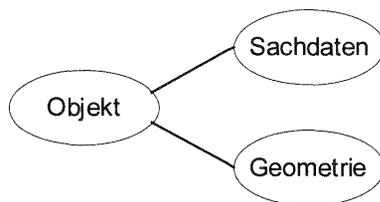
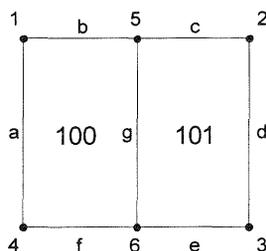


Abbildung 1.2: Objekteigenschaften

2. Die Sachdaten werden in einem relationalen Schema in einer relationalen Datenbank gespeichert und über Zeiger mit den Geometriedaten verbunden. Dies setzt aber voraus, daß das GIS die Anbindung einer relationalen Datenbank unterstützt. In modernen relationalen Datenbanken können sowohl Geometrie- als auch Sachdaten integriert gespeichert werden.



Objekt	Fläche	Kante	von	nach	li. Fläche	re. Fläche
1	100	a	4	1	0	100
		b	1	5	0	100
		g	5	6	101	100
		f	6	4	0	100
2	101	g	6	5	100	101
		c	5	2	0	101
		d	2	3	0	101
		e	3	6	0	101
		e	3	6	0	101

Flurbuch

Fläche	Lage	Nutzung	Größe	LB_NR
100	Bahnhofstr.	Bauplatz	375	300
101	Bahnhofstr.	Bauplatz	376	420
0	Außenbereich	Unland	---	---

Liegenschaftsbuch

LB_NR	Name	Geburtsdatum	Wohnort	Grundbuch
300	Test, Fritz	24.12.1925	Augsburg	20-333
420	Versuch, Hans	01.01.1950	München	20-444

Abbildung 1.3: Relationales Datenmodell, Geometriedaten (oben) und Sachdaten (unten)

Geometrisch-topologisches Modellieren:

Beim geometrischen Modellieren wird festgelegt, wie die geometrische Speicherung der räumlichen Objekte erfolgt. Dabei stehen meist die geometrischen Grundelemente Punkt, Linie und Fläche zur Verfügung. Die Koordinaten der Geometrieelemente spiegeln den Raumbezug wider. Es gibt die verschiedensten Möglichkeiten die Geometrie der raumbezogenen Objekte zu speichern, grob unterschieden in analytische und approximative Verfahren. Im GIS-Bereich wird meistens das Verfahren der Randdarstellung benützt. Bei der Randdarstellung wird das räumliche Objekt durch seine Begrenzungselemente (Flächen, Linien und Punkte) beschrieben, die entweder durch analytische Funktionen oder durch Interpolation und Approximation bestimmt sein können.

Die Topologie beschreibt die nichtmetrischen Beziehungen zwischen den Objekten. Den geometrischen Grundelementen Fläche, Linie und Punkt sind die topologischen Beziehungen Masche, Kante und Knoten zugeordnet. Wird eine Fläche abgespeichert, dann muß angegeben werden, welche Linien und Punkte zu dieser Fläche gehören. Der Grad der topologischen Beschreibung ist in den verschiedenen GIS-Produkten sehr unterschiedlich.

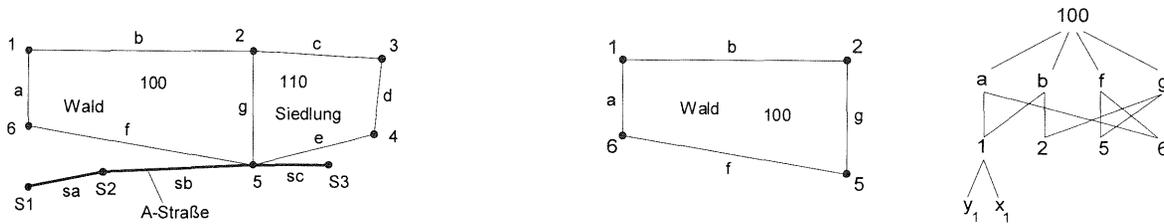


Abbildung 1.4:

a) 2D-Darstellung der Realszene

b) Topologische Zerlegung d. Fläche 100

Die alphanumerische Randdarstellung dieser Szene lautet:

Punktliste

Pkt.Nr.	y	x
1	y_1	x_1
2	y_2	x_2
...
S1	y_{S1}	x_{S1}
...

Linienliste

Linien-Nr.	Anf.-Pkt.	End-Pkt.	li. Fläche	re. Fläche
a	6	1	0	100
b	1	2	0	100
...
sa	S1	S2	---	---
...

Datenstrukturen:

Datenstrukturen sind das Ergebnis der geometrischen und topologischen Modellierung. Man unterscheidet zwischen Vektor- und Rasterdaten, wobei im folgenden nur Vektordaten betrachtet werden. Diese werden differenziert in unstrukturierte und topologisch strukturierte Daten.

Unstrukturierte Daten sind z. B. Koordinatendateien, die graphisch dargestellt nur Punktwolken sind, ohne Verbindungsinformationen, welcher Punkt zu welchem Objekt gehört. Damit ist jedem (punktförmigen) Objekt genau ein Koordinatenpaar zugeordnet. Mit Hilfe sogenannter Spaghetti-Strukturen können zwar zusätzlich Linien- und Flächenobjekte gebildet werden. Dabei werden beispielsweise die zu einer Linie gehörenden Punkte einfach hintereinander in einer Liste abgespeichert. Gehört ein Punkt zu mehreren Linien, so ist er in mehreren Listen redundant enthalten. Zwar weisen Spaghetti-Daten eine gewisse Strukturierung auf, dennoch werden diese den unstrukturierten Daten zugeordnet.

Für topologisch stukturierte Daten gibt es eine Vielzahl von verschiedenen Strukturen, die hier nicht im einzelnen aufgeführt werden. Eine gute Übersicht findet sich in [Findeisen, 1990]. Topologische Datenstrukturen sind Merkmal moderner Geo-Informationssysteme.

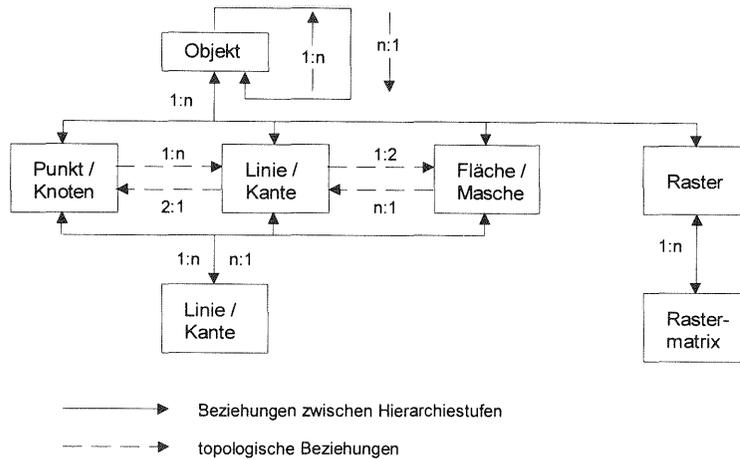


Abbildung 1.5: ATKIS-DLM-Datenmodell

1.1.2 Logisches Modell

In diesem Schritt wird das konzeptionelle Modell in einer Datenbank abgebildet. Die Abbildung in ein reines Dateisystem ist nicht mehr zeitgemäß und wird hier nicht weiter betrachtet. Gebräuchliche Modelle haben hierarchische, netzwerkartige oder relationale Struktur. In modernen GIS-Produkten werden relationale Datenbanken verwendet. Damit stehen auch alle Funktionalitäten der Datenbank zur Verfügung, wie ein Datenbankmanagementsystem, welches Mehrfachzugriffe regelt. Außerdem gewährleistet das Transaktionskonzept, daß der Datenbestand sich immer in widerspruchsfreiem Zustand befindet. Dies garantiert einen hohen Sicherheitsstandard. Fehler in der Datenbank werden durch Kontrollsysteme erkannt und behoben. Die Standardabfragesprache SQL (structured query language) bietet eine einfache Möglichkeit auf die Sachdaten zuzugreifen. Die Abfrage der Geometriedaten ist meist nur über komplexe Strukturen möglich. Auf das physikalische Modell wird hier nicht weiter eingegangen, da dies vor allem von der verwendeten Hardware abhängig ist.

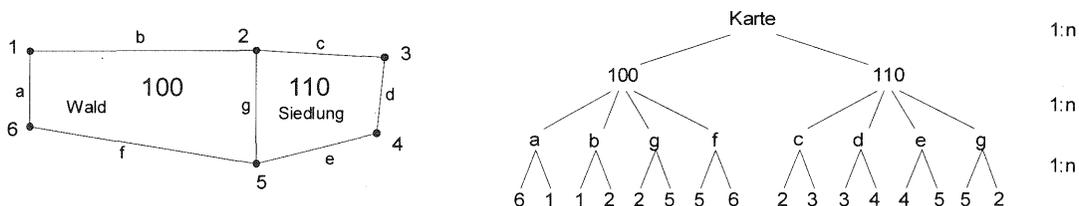


Abbildung 1.6: Logisches Datenmodell

1.2 Einführende Betrachtungen zu ATKIS-Bayern

Auf Wunsch des Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen begann das Bayerische Landesvermessungsamt 1986 mit der Digitalisierung der politischen Grenzen (Landes-, Bezirks-, Landkreis-, und Gemeindegrenzen) aus der Topographischen Karte 1:25000 (TK 25). Außerdem wurden die Gebietsnamen und die Gemeindegrenzzahlen zugefügt. Dies war der Beginn für den Aufbau eines Geographischen Informationssystems in der Bayerischen Vermessungsverwaltung - für das **Geographische Grundinformationssystem (GEOGIS)** war der Grundstein gelegt. Entsprechend dem Stand der damaligen Technik wurde das Ebenenprinzip zur Strukturierung der Daten eingesetzt. Es folgte die Erfassung der Gemarkungsgrenzen sowie die des Straßennetzes und der Gewässer, wobei zunächst nur die übergeordneten Straßen erfaßt wurden, des weiteren die Siedlungsgrenzen und die Waldgrenzen. Auch diese wurden aus der TK 25 erfaßt. Bundesweit lief erst 1989 das **Amtliches Topographisch-Kartographische Informationssystem (ATKIS)** an. Im Gegensatz zu GEOGIS, das seit 1995 als **ATKIS-Bayern** bezeichnet wird, sind die Daten im ATKIS-Modell nach dem Objektklassenprinzip strukturiert und werden in den meisten Ländern vorwiegend aus Orthophotos und der Deutschen Grundkarte (DGK) 1:5000 erfaßt.

Durch den Ministerratsbeschluß vom 18.09.1990 und die Gemeinsame Bekanntmachung der Bayerischen Staatskanzlei und der Bayerischen Staatsministerien zum Aufbau raumbezogener Informationssysteme vom 07.01.1992 ist das

Bayerische Landesvermessungsamt verpflichtet, das Geographische Grundinformationssystem GEOGIS als **Basisinformationssystem** mit hoher Priorität aufzubauen. Die wesentlichen Basisdaten sollten innerhalb von fünf Jahren, also bis Ende 1996, erfaßt sein. Dieses Ziel wurde erreicht.

Ähnliche Beschlüsse liegen in fast allen Bundesländern vor. Alle Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland sehen ihre Aufgaben darin, die raumbezogenen Basisdaten zur Verfügung zu stellen. Dies hat zum einen den Sinn, daß eine Mehrfacherfassung durch verschiedene Stellen vermieden wird, zum anderen ist die Fortführung der Daten und eine einheitliche Datenqualität gewährleistet. Mit Basisdaten werden vorwiegend geometrische Daten zu Verfügung gestellt, während in den Fachanwendungen (Fachinformationssysteme) diesen Basisdaten zusätzliche Geometrie-, aber vor allem auch Sachdaten angefügt werden. Der Datenbestand eines Fachinformationssystems eignet sich i. d. R. für eine Vielzahl von spezifischen Analysen und Auswertungen, während Basisdatenbestände naturgemäß hierfür kaum Möglichkeiten bieten.

1.3 Grundsätzliche Betrachtungen zur geometrischen Genauigkeit von Geodaten

Genauigkeit ist der Grad der Annäherung an einen **idealen** oder **erforderlichen** Wert [Mayers großes Taschenlexikon, 1995].

Der **ideale** Wert - Referenz - ist in diesem Fall die exakte Koordinate, in einem raum/zeit-stabilen geodätischen Bezugssystem. Daten mit hoher Genauigkeit und amtlichem Charakter liegen nur im Liegenschaftskataster vor, nicht aber für die Objektbereiche eines topographie-orientierten Geo-Informationssystems. Denn eine Waldgrenze muß nicht zwangsweise mit einer Flur- bzw. Grundstücksgrenze zusammenfallen. Daten mit der im Zentimeterbereich vorliegenden Genauigkeit werden allerdings z. B. zur Führung von Leitungs- oder Kanalkatastern benötigt.

Der **erforderliche** Wert hängt also entscheidend von den Bedürfnissen und Anforderungen der Nutzer ab. Diese können sehr unterschiedlich sein. An eine Biotopkartierung werden sicherlich höhere Erwartungen bezüglich der Genauigkeit gestellt als an eine Karte, in der die mautpflichtigen Straßen und Tunnels verzeichnet sind.

Da mit ATKIS bzw. ATKIS-Bayern **Geobasisdaten** bereitgestellt werden, müssen die Belange möglichst vieler Nutzer berücksichtigt werden. Daten mit hoher Genauigkeit zu erfassen und EDV-gerecht zu dokumentieren braucht viel Zeit und ist somit sehr kostenintensiv. Um einen möglichst großen Nutzerkreis anzusprechen, sind an die Basisdaten jedoch relativ hohe Genauigkeitsansprüche zu stellen, die aber auf der anderen Seite mit einem wirtschaftlich vertretbaren Aufwand in einem angemessenen Zeitraum flächendeckend auf den Markt zu bringen sind. Nicht außer acht gelassen werden darf der Aufwand für die Fortführung des Datenbestands; auch dieser steigt mit der Genauigkeit. Als Zielvorgabe für ATKIS ergab sich damit eine **erforderliche** geometrische Genauigkeit von 3-5m. Ob und wie dies im Einklang mit den obigen Forderungen steht, wird an dieser Stelle nicht diskutiert.

Spricht man von der geometrischen Genauigkeit bei Geodaten, so ist zu prüfen, ob die erforderliche Genauigkeit eingehalten ist. Ist dies nicht der Fall, dann ist diese als „ungenau“ einzustufen, also mit einem Fehler behaftet.

Wie kommt es zu diesen Fehlern, welche Ursachen haben sie und wie können sie gegebenenfalls vermieden oder behoben werden? Im folgenden wird versucht, zunächst die Ursachen und Auswirkungen der verschiedenen Erfassungsfehler zu klassifizieren und zu quantifizieren. Dabei wird noch nicht speziell auf die Situation für ATKIS-Bayern eingegangen.

Erfassungsfehler haben verschiedene Ursachen und weisen die unterschiedlichsten Wirkungen auf den Datenbestand auf.

1.3.1 Inkonsistenzen bei der Definition geometrischer Grundprimitive

Die geometrischen Grundprimitive bei planaren Geometrien sind Punkt, Linie und Fläche. Hierbei wird nicht die Definition der Grundprimitive an sich in Frage gestellt. Gemeint sind vielmehr Inkonsistenzen bei der Datenerfassung, wenn ein Objekt der realen Welt durch ein unzulängliches Grundprimitiv beschrieben wird. In der Terminologie des Objektklassenprinzips werden die Grundprimitive auch als Objekttyp bezeichnet. Inkonsistenzen zwischen zwei gleichartigen Grundprimitiven (z. B. Punkt - Punkt) sind nur zwischen verschiedenen Objektarten möglich, während Inkonsistenzen zwischen zwei unterschiedlichen Grundprimitiven (z. B. Punkt - Linie) auch innerhalb der gleichen Objektart möglich sind. Dies wird anschließend noch an Beispielen verdeutlicht. Die Definition für Objektart und Objektbereich findet sich im Abschnitt 2.1. Es werden im folgenden sechs Fälle von Inkonsistenzen unterschieden, die angeführten Beispiele beziehen sich auf den ATKIS-Objektartenkatalog.

Punkt-Punkt-Inkonsistenz:

Ein und derselbe Punkt hat eine verschiedenartige oder evtl. gar keine Bedeutung. Beispielsweise ein Lagefestpunkt (Objektbereich *Festpunkte*), der als Kilometerstein (Objektbereich *Verkehr*) abgespeichert wird, oder im Extremfall ein Lagefestpunkt, der gar nicht mehr existiert. Dabei ist zu unterscheiden, wann der Fehler der falschen Zuweisung geschehen ist. Denn nicht als Punkt-Punkt-Inkonsistenz sind Fehler zu bezeichnen, die bereits in der Erfassungsgrundlage (z. B. der TK 25) vorhanden sind. In diesem Fall rührt die Inkonsistenz aus einer falschen Darstellung in der Erfassungsgrundlage. Praktisch hat dies zwar für den Datenbestand die gleiche Auswirkung, die Ursache aber ist eine andere.

Punkt-Linien-Inkonsistenz:

Hier sind vielfältige Variationen denkbar. In ATKIS gibt es nur einige wenige Objektarten, die je nach Größe oder Ausprägung punkt- bzw. linienförmig modelliert erfaßt werden. Eine Furt (Objektbereich *Verkehr*) wird bei linienförmig modellierten Gewässern punktförmig, bei flächenförmig modellierten Gewässern linienförmig erfaßt. Eine Verwechslung beim Erfassen führt hier zu einer Punkt-Linien-Inkonsistenz. Eine wesentlich größere Zahl von Objektarten kann je nach Ausprägung punkt-, linien- oder flächenförmig sein. Dies gilt z. B. für einen Wasserfall; dabei ist die Abstufung abhängig von seiner Größe, Höhe und Breite. Auch Verwechslungen, die nicht nur innerhalb der Objektart stattfinden, sind möglich. Statt einer Schleuse, die je nach Größe punkt-, linien- oder flächenförmig zu modellieren ist, wird eine Talsperre erfaßt. Diese kann nur linien- oder flächenförmig sein. Handelt es sich um eine kleine Schleuse mit einer Länge kleiner 12 m, so ist diese punktförmig zu generieren, die Talsperre ist bei einer Länge von weniger als 12 m jedoch linienförmig darzustellen. Die Verwechslung zwischen Schleuse und Talsperre der gleichen Größe führt ebenso zu einer Punkt-Linien-Inkonsistenz. Fehler dieser Art sind in vielen Variationen möglich, auch über Objektbereiche hinweg. Hier sind vielfältige Möglichkeiten für Dateninkonsistenzen geboten. In einigen Fällen, wie z. B. bei der Furt, sind diese relativ leicht zu verifizieren, wenn die Generierung direkt durch die Modellierung des Grundelements bedingt ist. Geht man davon aus, daß in diesem Fall das Gewässer korrekt erfaßt wurde, so ist ein falsch gewähltes Grundprimitiv bei der Furt augenscheinlich. In anderen Fällen, wo die Ausprägung von der Größe abhängig ist, sind Irrtümer kaum mehr verifizierbar.

Punkt-Flächen-Inkonsistenz:

Auch hier gibt es wieder einige wenige Objektarten, die je nach Ausprägung punkt- bzw. flächenförmig modelliert erfaßt werden. Die Bahnhofsanlage wird, wenn keine weiteren Gleise von der Strecke abzweigen und die gesamte Anlage nicht größer als 0,5 ha ist, punktförmig erfaßt. Andernfalls ist nur eine flächenförmige Erfassung zulässig. Eine Verwechslung führt hier zu einer Punkt-Flächen-Inkonsistenz. Sind im Bahnhofsbereich weitere Gleise erfaßt und ist dieser punktförmig modelliert, so ist dieser Irrtum augenscheinlich, ansonsten jedoch kaum mehr aufzudecken. Punkt-Flächen-Inkonsistenzen über den Bereich einer Objektart hinaus sind vielfältig denkbar, vor allem bei „totalen“ Fehlerfassungen. Im Extremfall könnte statt einem Kilometerstein (Objektbereich *Verkehr*) ein flächenförmig modellierter Flugplatz erfaßt worden sein. Solche Irrtümer können zwar vorkommen, erscheinen in diesem Zusammenhang aber eher konstruierte Fälle zu sein. Generell sind solche „totalen“ Fehlerfassungen meist nur durch Kenntnis des jeweiligen Gebietes in der Natur aufzudecken und zu beheben.

Linien-Linien-Inkonsistenz:

Linien-Linien-Inkonsistenzen sind in vielen Variationen denkbar, sowohl objektbereichsübergreifend als auch innerhalb eines Objektbereichs. Ein Bach wird als Graben erfaßt. Beide Objektarten sind bei einer Breite unter 12 m linienförmig zu erfassen. Inkonsistenzen über den Objektbereich hinweg sind ebenso vielfältig möglich, z.B. ein Bach wird als Weg erfaßt.

Linien-Flächen-Inkonsistenz:

Einige wenige Objektarten, vor allem in den Objektbereichen *Verkehr* und *Gewässer*, sind abhängig von ihrer Ausprägung und Größe linien- bzw. flächenförmig zu modellieren. Hierunter fällt die häufig vorkommende Objektart Wasserlauf (Strom, Fluß, Bach), die bis zu einer Breite von 12 m linienförmig, darüber flächenförmig zu generieren ist. Ein Erfassungsfehler hier ist demzufolge schwer aufzuspüren. Häufiger auftreten wird die Linien-Flächen-Inkonsistenz zwischen verschiedenen Objektarten. Hier kann wieder das oben schon einmal erwähnte Beispiel mit Schleuse und Talsperre verwendet werden. In diesem Fall ist die Schleuse bei einer Länge über 12 m und einer Breite unter 12 m linienförmig zu erfassen, während die Talsperre bei einer Länge über 12 m, unabhängig von der Breite, generell flächenförmig erfaßt wird. Auch hier sind mit anderen Objektarten vielfältige Variationen denkbar.

Flächen-Flächen-Inkonsistenz:

Flächen-Flächen-Inkonsistenzen sind vor allem in den Objektbereichen *Vegetation*, *Siedlung* und *Gebiete* zu suchen, da hier fast alle Objektarten flächenförmig zu erfassen sind. Einige Beispiele: Ackerland wird mit Grünland verwechselt, Wald mit Gehölz, ein Heizwerk mit einem Kraftwerk, eine Grünanlage mit einem Friedhof, um nur einige Möglichkeiten zu nennen. Da eine große Zahl an Objektarten flächenförmig zu erfassen ist, wird die Flächen-Flächen-Inkonsistenz wesentlich öfter auftreten als die oben genannten. Die Verifizierung dieser Inkonsistenzen wird damit um so schwieriger.

1.3.2 Erfassungsgrundlage mit nicht aktuellem Datenbestand

Ist die Erfassungsgrundlage nicht auf dem aktuellen Stand, so ergeben sich mitunter nicht nur geometrische, sondern auch topologische und thematische Fehler. Eine besondere Schwierigkeit hierbei ist, daß eine einzige Erfassungsgrundlage alleine in den meisten Fällen nicht ausreichend ist, um einen fehlerfreien Datenbestand zu erhalten. Ist die Erfassungsgrundlage eine TK 25, so muß man wissen, daß diese je nach Bedarf, jedoch durchschnittlich nur alle acht bis zwölf Jahre aktualisiert wird. Sicherlich bedeutet es einen immensen Aufwand, die Kartenwerke immer aktuell zu halten, vor allem für die flächenmäßig großen Bundesländer wie Bayern, Baden-Württemberg und Niedersachsen. Aber es gibt auch noch andere Erfassungsgrundlagen, z. B. das Orthophoto (Maßstab 1:10 000 oder 1:5000) und die Flurkarte (hier: Maßstab 1:5000) bzw. die Deutsche Grundkarte (Maßstab 1:5000). Diese sind aber meist ebenfalls nicht flächendeckend aktuell vorhanden. Doch ist der Informationsgehalt einer bayerischen Flurkarte bzw. der Digitalen Flurkarte (DFK) für die Anforderungen der Basisdaten nicht ausreichend, da sie zwar die Flurstücksgrenzen und -nummern sowie die Gebäude enthalten, jedoch meist nur unzureichend aktuell die Nutzungsartengrenzen. Anders verhält es sich mit der Deutschen Grundkarte (1:5000), die auch eine Vielzahl an topographischen Informationen enthält. Diese wird aber in Bayern nicht geführt und steht auch in den anderen Bundesländern nur teilweise zur Verfügung. Hier zeigt sich, wie schwierig es ist, eine aktuelle Erfassungsgrundlage zu finden. In den meisten Fällen kann dies nur durch Kombination mehrerer Medien erreicht werden. Dies soll kurz am Beispiel der Siedlungsflächen erläutert werden.

Verwendet man die TK 25, so muß dabei bedacht werden, daß bereits erhebliche Generalisierungseffekte auftreten, worauf unten noch näher eingegangen wird, und die Daten eventuell schon über zehn Jahre alt sind. Dies führt dazu, daß die Siedlungsgrenzen teilweise so stark von den derzeitigen abweichen, daß dies nicht mehr mit „geometrischer Genauigkeit“ zu beschreiben ist, sondern nur mehr als falsch bezeichnet werden kann. Verwendet man ein Orthophoto neueren Datums, so sind die Siedlungsflächen sicherlich so gut erfaßt, daß sich dies wieder mit dem Wort „geometrischer Genauigkeit“ beschreiben läßt. Letzte Sicherheit kann man in diesem Fall allerdings nur über die aktuelle Flurkarte erhalten, vor allem in Gemeinden im Einzugsbereich von großen Ballungsräumen mit reger Siedlungsentwicklung. Ist ein Baugebiet neu ausgewiesen, so werden die Flurstücke bereits parzelliert und abgemarkt, was in der Flurkarte an Form und Größe ohne Schwierigkeit erkennbar ist, nicht aber im Orthophoto und schon gar nicht in der TK 25. Diese Flächen sind aber eindeutig als Siedlungsgebiete in den Geobasisdaten aufzunehmen.

Ähnliche Beispiele lassen sich viele finden. Wie aber wirkt sich dies auf die geometrische Genauigkeit aus?

Das genannte Phänomen soll hier anhand von Situationsbildern noch einmal verdeutlicht und erläutert werden.

Die geometrischen Informationen sind:

- Zumindest in Teilbereichen völlig falsch, d.h. ein Acker ist inzwischen bebaut, aber nicht als Siedlungsfläche erfaßt; dem Ortsbereich fehlt somit ein ganzes Gebiet.
- In der Gesamtheit geometrisch richtig erfaßt, in kleinen Bereichen aber korrekturbedürftig, z. B. die Bebauung am Ortsrand ist erweitert worden.

Erwähnt sei, daß eine Erfassung aus nicht aktuellem Kartenmaterial nicht nur zu erheblichen geometrischen, sondern auch zu topologischen und thematischen Fehlern führt.



Abbildung 1.7: Siedlungsgebiet, Überlagerung von DFK, Orthophoto und ATKIS-Bayern

1.3.3 Lageverschiebungen aufgrund verschiedener Einpassungen (fehlerhafte Transformationsmodelle)

Die Geodaten werden derzeit in das Gauß-Krüger-Koordinatensystem als einheitliches Landesbezugssystem eingepaßt. Dies geschieht bei jedem Erfassungsvorgang über vier Paßpunkte, in der Regel den Blattecken. Für die Einpassung werden gängige Transformationsmodelle wie die Helmert- oder Affintransformation verwendet. Trotz eng vorgegebener Fehlerschranken, die der Bearbeiter auch nicht umgehen kann, ist jede Einpassung individuell anders. Dies führt dann zu Verschiebungen und Verdrehungen der einzelnen Einpassungen gegeneinander. Jedes Blatt wird auf diese Weise zum Teil mehrfach eingepaßt, da nicht die gesamte Erfassung von einem Bearbeiter in einem Arbeitsgang durchgeführt werden kann.

Augenscheinlich wird dies, wenn man die erfaßten Objektbereiche mit der Digitalen Flurkarte (DFK) überlagert. Die DFK ist hier als Referenz angesehen. Am einfachsten sind diese Fehler im Objektbereich *Verkehr* zu verifizieren, da die Straßen auch in der Flurkarte leicht und eindeutig zu erkennen sind. Topologie und Thematik sind von diesem Erfassungsfehler nicht betroffen.

Sicherlich läßt sich durch die Wahl geeigneter Transformationsmodelle der Fehler minimieren. Dort wo der Fehler bereits aufgetreten ist, gestaltet sich eine nachträgliche Transformation mit einem geeigneten Modell aber schwierig, da keine Paßpunkte vorhanden sind. Geeignete Paßpunkte wären nur solche, die bei der Digitalisierung des entsprechenden Objektbereichs im gleichen Atemzug, also der gleichen Einpassung entstanden sind. Hierfür würden sich die Trigonometrischen Punkte eignen, die auch in den Flurkarten und TK 25 kartiert, jedoch in ATKIS-Bayern nicht erfaßt sind. Außerdem sind die Trigonometrischen Punkte nicht gleichmäßig über die Kartenblätter verteilt, so daß von qualifizierter Seite eine sinnvolle Auswahl vor der Erfassung getroffen werden müßte. Ferner muß sichergestellt werden, daß bei der Erfassung aller Objektbereiche die gleichen Trigonometrischen Punkte zur Einpassung verwendet werden.

Die Paßpunkte bei den Testdaten für diese Untersuchung wurden aus den Informationen der DFK gewonnen. Dazu wurden an eindeutigen Straßenknoten Paßpunkte durch Verschneidung der sich kreuzenden Straßenzüge konstruiert. Diese Paßpunkte waren gleichmäßig über das Kartenblatt verteilt. Diese Methode ist aber auch nicht durchgreifend, da derzeit erst für einen geringen Teil von Bayern die DFK vorhanden sind. Das Ergebnis der Transformation zeigt jedoch, daß dadurch die Lagegeometrie verbessert werden kann.

1.3.4 Isolierte (ikonisierte) Objektbereichserfassung

Findet die thematische Modellierung nach dem Ebenenprinzip statt, so erfolgt die Trennung von Geometriedaten verschiedener thematischer Bedeutung durch Abspeicherung in separate Gebietsdatenbanken (GDB), die wiederum mehrere Ebenen haben. Die Gesamtdarstellung erhält man durch Überlagerung aller Ebenen aller Gebietsdatenbanken. Werden die verschiedenen Ebenen getrennt voneinander und nach verschiedenen Methoden bzw. aus unterschiedlichen Datenquellen erfaßt, so kommt es bei der Überlagerung zu Verschneidungen (Inkonsistenzen) zwischen den Objektbereichen, die nicht der Wirklichkeit entsprechen und auch in der Erfassungsgrundlage nicht vorhanden waren.

Als Erfassungsmöglichkeiten kommen in Betracht:

1. Die älteste und verbreitetste Methode ist die **manuelle Digitalisierung** aus einer analogen, graphischen Vorlage mit Hilfe eines Digitalisierisches. Diese Vorgehensweise ist sehr zeit- und kostenintensiv und für den Bearbeiter sehr unbequem, hat aber den Vorteil, daß der Operateur am graphisch-interaktiven Arbeitsplatz strukturierte und mit Attributen versehene Vektordaten erzeugen kann. Aufgrund der direkten Objektbildung und Attributierung sind fast keine interaktiven Nachbearbeitungen erforderlich. Diese Methode eignet sich daher insbesondere für die Erfassung komplexer Karteninhalte mit unregelmäßiger Geometrie, viel Symbolik und unterschiedlichen Objektarten.
2. Ein für den Bearbeiter ergonomisch angenehmeres und für die Augen weniger anstrengendes Verfahren ist die **Bildschirmdigitalisierung**. Dazu wird die analoge Vorlage im Rasterscanner gescannt, anschließend georeferenziert, am Bildschirm dargestellt und dort interaktiv digitalisiert. Dabei können alle gängigen Methoden der Bildschirmdarstellung - zoomen, darstellen eines beliebigen Ausschnittes usw. - genutzt werden. Dieses Verfahren setzt jedoch ein hybrid arbeitendes Erfassungssystem voraus. Im übrigen weist das Verfahren die gleichen Vor- und Nachteile wie das unter 1. genannte auf.
3. Bei der **halbautomatischen Digitalisierung** wird die Kartenvorlage gescannt, georeferenziert und auf dem Bildschirm als Rasterbild dargestellt. Anschließend markiert der Operateur auf dem Bildschirm mit dem Mausfadenzkreuz den Linienbeginn. Diese Linie wird dann automatisch im Rasterbild bis zum nächsten Knoten verfolgt, und der Operateur markiert dort den weiteren Linienverlauf. Hinsichtlich der Erfassungsgenauigkeit ist die halbautomatische Digitalisierung der manuellen vorzuziehen, da persönliche Fehler des Bearbeiters wegfallen. Das System kann aber nur dann sinnvoll eingesetzt werden, wenn die Linien sich nicht zu oft schneiden, nicht häufig unterbrochen sind und nicht zu dicht beinander liegen. Einige Systeme bieten inzwischen auch die Möglichkeit, gestrichelte Linien zu verfolgen sowie wahlweise, sofern ein Objekt durch zwei Linien begrenzt wird (z. B. die rechte und linke Uferlinie eines Flusses), die Mittellinie oder beide Begrenzungslinien zu digitalisieren. Nicht sinnvoll erweist sich dieses Verfahren bei komplexem Inhalt, bei der Erkennung von Texten und Zahlen in Karten. Die Zeitersparnis gegenüber der manuellen Digitalisierung ist um so größer, je einfacher der Karteninhalt ist. Besonders geeignet ist das Verfahren z. B. für Isolinienkarten (Höhenlinien), die keine Knoten aufweisen. Wie auch bei der manuellen Digitalisierung können auf diese Weise strukturierte und attributierte Daten erfaßt werden.
4. Bei der **automatischen Digitalisierung** wird die analoge Karte gescannt und in ein digitales Bild umgewandelt. Anschließend wird das Rasterbild mit Verfahren der Raster-Vektor-Transformation in eine Datei von unstrukturierter Vektordaten konvertiert. Durch entsprechende Mustererkennungssoftware erfolgt dann eine automatische Strukturierung der Geometrie. Dies heißt, daß den Linien Bedeutungen zugewiesen werden. Nach Ablauf der Mustererkennung kontrolliert der Operateur die Ergebnisse und korrigiert erforderlichenfalls den Datenbestand. Der Vorteil der automatischen Digitalisierung liegt im geringeren Zeitaufwand und vor allem der hohen Genauigkeit von modernen, hochauflösenden Scannern. Leider gibt es derzeit noch keine ausgereifte Software zur kartographischen Mustererkennung, insbesondere nicht für topographische Karten. So eignet sich dieses Verfahren derzeit nur für Vorlagen, in denen aufgrund der Daten keine Knoten vorkommen können und die zudem nur Objekte einer einzigen Objektart enthalten. Insbesondere geeignet sind einfarbige Kartenvorlagen. Dies gilt auch für die beiden zuvor genannten Verfahren.

Bei der halbautomatischen und automatischen Generalisierung ist immer eine Nachbearbeitung erforderlich, zum einen bezüglich logischer Fehler, die z. B. durch Schmutzpartikel verursacht werden, zum anderen bezüglich der Geometrie. Bei der Geometrie sind einerseits Ungenauigkeiten der Hardware (hier des Scanners), z. B. Fehler in der Kreuzungsgeometrie, Achsunterbrechungen, Schleifen oder ähnliches, zu verbessern, andererseits müssen die Rasterdaten georeferenziert, d. h. der Raumbezug hergestellt werden. Dies geschieht durch eine geeignete Transformation, i. d. R. über die Blatteckenwerte der analogen Karte.

Jede dieser Erfassungsmöglichkeiten hat ihre Vor- und Nachteile und ist mit systemspezifischen Fehlern behaftet. Kombiniert man die verschiedenen Erfassungsmöglichkeiten in Verbindung mit dem Ebenenprinzip, so treten aufgrund der unterschiedlichen Genauigkeit der einzelnen Methoden Verschneidungen der Objektbereiche auf. Zudem treten geometrische Inkonsistenzen auf, da die einzelnen Objektbereiche bei der Erfassung nichts voneinander wissen. So wird

eine Straße, die auf der einen Seite an den Wald grenzt und an der anderen Seite an ein Siedlungsgebiet, dreimal unabhängig voneinander erfaßt. Dies führt zu Redundanzen in den Geodaten, die wiederum zu Dateninkonsistenzen führen. Die Konturen haben zwar visuelle Ähnlichkeit (nicht im mathematischen Sinn), jedoch ist eine Beseitigung dieser Inkonsistenzen nicht ohne weiteres möglich.

Folgendes Beispiel zur Verdeutlichung:

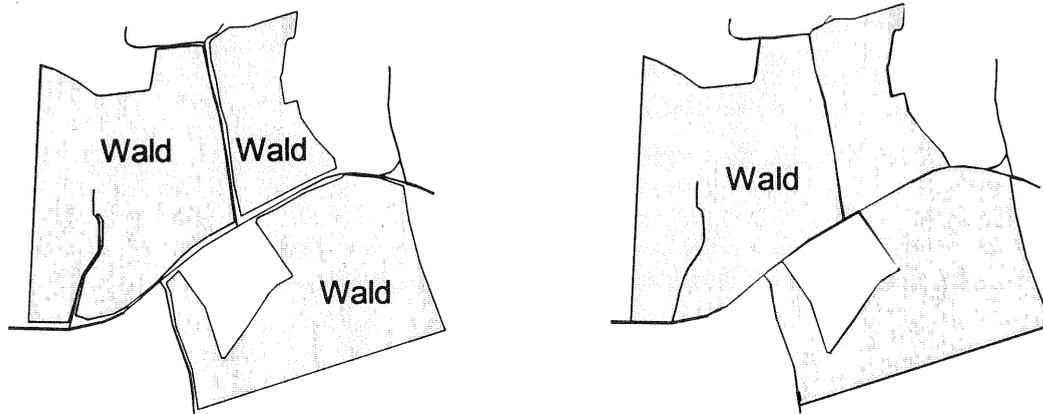


Abbildung 1.8: Waldpolygone

a) Ergebnis einer ikonisierten Objektbereichserfassung

b) tatsächliche Situation, Ergebnis einer nicht ikonisierten Objektbereichserfassung

1.3.5 Generalisierungsfehler in der Erfassungsgrundlage

Jede Kartendarstellung ist mit Generalisierungseffekten behaftet und zwar um so mehr, je kleiner der Maßstab ist. Selbst in den Flurkarten 1:5000, wo man meist noch nicht von Generalisierung spricht, tritt dieser Effekt auf. Denn Objekte, die kleiner als eine mindest darstellbare Größe von ca. 0,2 - 0,3 mm (1 - 1,5 m in der Natur) sind, werden nicht dargestellt, also generalisiert. Auch beim Orthophoto kann man im weitesten Sinne des Wortes von Generalisierung sprechen, da aufgrund der Körnung des Photopapiers bzw. aufgrund des Auflösungsvermögens des Auges die Detailerfassung nur eingeschränkt möglich ist. Noch stärker tritt dieser Effekt bereits dann beim digitalen Orthophoto auf, wenn die Größe eines Pixels in der Natur die Auflösungsschärfe wiedergibt. Bei einem Orthophoto im Maßstab 1:10000 hat ein Pixel in der Natur eine Größe von etwa 0,8 m x 0,8 m. Auch diese kommt einer Generalisierung gleich.

Echte Generalisierungsfehler:

Falsche Informationen aufgrund von Generalisierungsfehlern sind schwer zu definieren, da bei der Generalisierung dem Bearbeiter gewisse "künstlerische Freiheit" zugestanden werden muß. Eindeutig als Fehler zu beurteilen ist aber, wenn Straßenverbindungen in der TK 25 vorhanden sind, die in der Natur nicht existieren. Bei der Generalisierung dürfen lediglich unter bestimmten Voraussetzungen Informationen entfallen oder verdrängt, nicht aber neue hinzugefügt werden. Eine Erfassungsgrundlage ist damit um so besser geeignet, je größer der Maßstab ist. Aus dieser Sicht bietet hier die Flurkarte in Verbindung mit dem Orthophoto die besten Informationen. So können z. B. auch Wege oder Wasserläufe, die keine eigenen Flurstücke bilden, lokalisiert und korrekt erfaßt werden.

Die Auswirkungen von Generalisierungsfehlern in der Erfassungsgrundlage sind gravierend, da solche Fehler in den Geodaten kaum mehr aufzudecken sind, und diese sich auf Geometrie, Topologie und Thematik auswirken. Die größten Generalisierungseffekte treten im Ortsbereich auf, da hier aufgrund der großen Informationsdichte eine größere Notwendigkeit zur räumlichen Aggregation besteht.

Zu beheben sind Erfassungsfehler dieser Art nur durch manuelle Nachbearbeitung (neu digitalisieren). Doch wie können diese Fehler möglichst einfach aufgedeckt werden? Dazu bietet sich der Vergleich mit der Flurkarte oder einem digitalen Orthophoto an. Dies ist jedoch mit einem erheblichem Arbeitsaufwand verbunden. Zwar liegt die DFK in Bayern noch nicht flächendeckend vor, aber gerade für die Siedlungsgebiete wird mit erhöhter Kraft daran gearbeitet. Doch stellt sich dann die Frage, warum man nicht von vornherein die qualitativ bessere Erfassungsgrundlage verwendet.

Erfassungsfehler aufgrund der Generalisierung:

Hiermit sind geometrische Lageungenauigkeiten bzw. Informationsverluste gemeint, die mit der Verdrängung und der Vereinfachung bei der Generalisierung zusammenhängen. Verläuft ein Bach neben einer (übergeordneten) Straße, so wird seine geometrische Lage durch die übernatürlich breite Darstellung der Straße (Minstdimensionen für die Kartendarstellung) in der TK 25 beeinflusst. Im Vordergrund steht bei der Kartenherstellung der inhaltliche Zusammenhang und der Bezug zur Umgebung. Zudem müssen zwei Linien in der Kartendarstellung einen gewissen Mindestabstand voneinander haben, damit sie vom Betrachter als getrennt wahrgenommen werden können. So kommt es schnell zu größeren Lageverschiebungen, die z.B. mit einer angestrebten Lagegenauigkeit (3 - 5 Meter) für ATKIS nicht vereinbar sind. Diese geometrische Ungenauigkeit muß nicht unbedingt Auswirkungen auf die Topologie haben, vor allem nicht solange jeder Objektbereich getrennt erfaßt und vorgehalten wird. Beim Zusammenspielen mit anderen Objektbereichen kommt es jedoch zu den oben genannten Verschneidungsproblemen.

Auch in diesem Fall ist dies nur durch Nacherfassung aus einer qualitativ höherwertigen Erfassungsgrundlage zu beheben. Die Vorgehensweise ist die gleiche wie oben. Zu berücksichtigen ist hierbei, daß eventuell Informationen aufgrund der Generalisierung nicht in der geforderten Detailliertheit erfaßt werden konnten. Wird der Siedlungsbereich z. B. aus einer TK 25 ermittelt, so kann nur der Umriß des Ortsgebietes erfaßt werden, allenfalls noch die Lage der Kirche oder einer Kapelle, nicht aber eine Unterscheidung zwischen Wohn-, Misch- oder Industriegebieten.

1.3.6 Was ist Redundanz bzw. was heißt Redundanzfreiheit?

Die Redundanzen sind in ATKIS-Bayern infolge der ikonisierten Datenerfassung nach dem Ebenenprinzip sehr hoch und stellen somit die häufigste Ursache für Dateninkonsistenzen dar. Ein übergeordnetes Ziel muß daher darin bestehen, geeignete Methoden zu entwickeln, ihre Möglichkeiten und Grenzen aufzuzeigen.

Dabei steht natürlich stets im Vordergrund eine einheitliche, redundanzfreie Geometrie zu erhalten. Die Semantik spielt hierbei eine untergeordnete Rolle.

Was heißt Redundanzfreiheit?

Unter Redundanzfreiheit soll hier nicht verstanden werden, daß jeder Punkt (Linie evtl. auch Fläche) nur ein einziges Mal im Datenbestand vorkommen darf. Hierzu ein einfaches Beispiel:

Das nordwestliche Eck einer Garage liegt direkt an der Flurstücksgrenze. Die Koordinate dieses Punktes wird nur ein einziges Mal in der Datenbank vorgehalten. Ist die Garage aus dem Datenbestand zu löschen, weil sie abgerissen worden ist, so kann dies nicht ohne weiteres erfolgen, da die Information für die Flurstücksgrenze damit auch verloren ginge. Einfacher ist es, und auch nach Meinung der Autorin in der modernen objektorientierten Sicht vertretbar, diesen Punkt „redundant“ zweimal in der Datenbank vorzuhalten, einmal beim Objekt *Garage* und ein zweites Mal beim Objekt *Flurstücksgrenze*. Damit sind beide Objekte voneinander unabhängig und für mein Verständnis kann in diesem Fall **nicht** von Redundanz gesprochen werden. Ein solcher Fall ist in der folgenden Abbildung 1.10a dargestellt. Demzufolge kann in diesem Sinn auch ein Gebäude, das direkt auf die Grenze gebaut wurde, wie im rechten Teil der Abbildung sichtbar, unabhängig von der Grenzlinie dargestellt werden. Die Grenzlinie wird in diesem Fall nur durch die beiden äußeren Punkte repräsentiert. Für den Fall, daß jeder Stützpunkt nur ein einziges Mal im Datenbestand vorhanden ist, müßte konsequenter Weise dann auch die Flurstücksgrenze auf dieser Seite über alle vier Stützpunkte verlaufen. Bei einem Abriß des Gebäudes wäre zu entscheiden, ob sie erhalten bleiben oder nicht. Die Punkte repräsentieren auf diese Weise nicht mehr ihr Objekt. Diese Vorgehensweise scheint nicht sinnvoll.

Unter Redundanz wird daher in dieser Arbeit verstanden, wenn auf Grund der isolierten Datenerfassung eine inhomogene Geometrie vorzufinden ist. Auch hier ein Beispiel: Direkt an das Waldgebiet angrenzend befindet sich ein Wohngebiet. Im Datenbestand finden sich an der Grenze der beiden Objektbereiche hierfür zwei voneinander abweichende Umfangslinien mit verschiedenen Stützpunkten. Wäre der Datenbestand redundanzfrei, hätten beide Umfangslinien an dieser Stelle den gleichen Verlauf. Anderenfalls gibt es bei Verschneidungsoperationen, die von Anwendern von GIS-Daten häufig eingesetzt werden, viele kleine, nicht sinnvolle Polygone, sogenannte Sliver-Polygone.

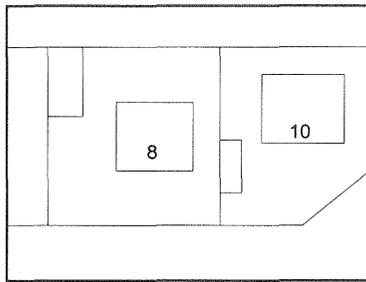
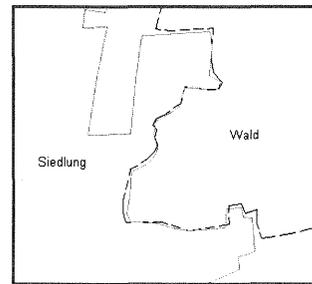


Abbildung 1.10:
a) Redundanzfreier Datenbestand



b) nicht redundanzfreier Datenbestand

1.4 Zielsetzung und Inhalt der Arbeit

Da diese Arbeit mit der Unterstützung der Bayerischen Vermessungsverwaltung entstanden ist, sollen hier vorwiegend die bayerische Situation bei der Bereitstellung von Geobasisdaten behandelt werden. Da Bayern bereits drei Jahre früher als die übrigen Bundesländer mit der ATKIS-Datenerfassung begonnen hat (siehe Kapitel 1.2), erklärt sich die etwas andere Art der Erfassung und Datenhaltung.

Inhalt der Arbeit wird deshalb sein:

- Die Grundlagen der Datenmodellierung und die Entstehung des bayerischen Modells der Basisdatenbereitstellung zu erläutern.
- Einige allgemeingültige und grundlegende Betrachtungen zu geometrischen Genauigkeiten und Inkonsistenzen der Daten vorzunehmen. Vertieft werden sollen dann aber nur die Aspekte, die speziell durch das bayerische Modell der Datenerfassung entstanden sind. Nicht weiter eingegangen wird in dieser Arbeit auf Dateninkonsistenzen, die sich weitestgehend unabhängig vom Modell erklären lassen. Konkret sind damit falsche Erfassungen gemeint, die durch menschliche Fehler des Bearbeiters entstehen und meist nur durch Zufall entdeckt und behoben werden können.
- Die Unterschiede zwischen dem bundeseinheitlichen und dem bayerischen Modell herauszuarbeiten, einander gegenüberzustellen und kritisch zu würdigen. Da das bayerische Modell die Daten nicht redundanzfrei vorhält, wird ein wesentliches Ziel dieser Arbeit sein, nach Möglichkeiten zu suchen, diesen Datenbestand möglichst automatisiert redundanzfrei zu machen.
- Einige notwendige Voruntersuchungen vorzunehmen, um die Genauigkeit der vorhandenen Geobasisdaten zu modellieren und anschließend diese, mit der Absicht einen redundanzfreien Datenbestand zu erzeugen, zu bewerten.

Ziel dieser Arbeit ist daher, die erarbeitete Ergebnisse kritisch zu würdigen und weitere Einsatzbereiche der verwendeten Algorithmen aufzuzeigen.

Bisher wurden keine Untersuchungen der ATKIS-Bayern-Daten in dieser oder ähnlicher Art und Weise vorgenommen. Diese Arbeit konzentriert sich ausschließlich auf die Besonderheiten des bayerischen Datenbestandes und wird erstmalig eine unabhängige Beurteilung der Datenqualität von ATKIS-Bayern ermöglichen. Das heißt nicht, daß bis dato keine Genauigkeitsuntersuchungen stattgefunden haben, sondern vielmehr, daß diese ausschließlich behördenintern durchgeführt wurden.

2 ATKIS und ATKIS-Bayern

Wie bereits erwähnt, werden für viele raumbezogene Anwendungen digitale Geodaten benötigt. Bereits jetzt sind mehrere staatliche und private Einrichtungen damit beschäftigt, diese Daten zu erfassen, mit zum Teil ganz unterschiedlichen Aufgabenschwerpunkten. Diese Mehrfacherfassung ist aus volkswirtschaftlicher Sicht nicht vertretbar. Auch wird in den meisten Fällen das Problem der Fortführung völlig außer acht gelassen, da zunächst alle verfügbare Energie in die Ersterfassung gesteckt wird. Die Fortführung ist aber mindestens so wichtig wie die Ersterfassung, da nicht aktuell gehaltene Daten binnen kürzester Zeit totes Kapital darstellen. Abgesehen davon sind qualitativ hoch-

wertigste Daten bei der Vermessungsverwaltung vorhanden, aufgrund des gesetzlichen Auftrags, die Topographie der Erdoberfläche zu erfassen, zu dokumentieren und zu aktualisieren. Die Vermessungsverwaltungen der Bundesländer haben die Notwendigkeit, qualitativ hochwertige Basisdaten aus einer Hand für die Allgemeinheit zur Verfügung zu stellen, sehr früh erkannt. Auch das Problem der Fortführung (Aktualisierung) der Geobasisdaten ist damit für den externen Anwender gelöst, denn er kann immer aktuelle Daten von einer zentralen Stelle in jedem Bundesland erhalten. Voraussetzung ist jedoch, daß das Datenmodell des Anwenders dies auch zuläßt. Wie bereits einleitend erwähnt, wurde bereits 1988 beschlossen, bei den Landesvermessungsämtern das digitale Amtliche Topographisch-Kartographische Informationssystem ATKIS zu entwickeln und aufzubauen.

„Interdisziplinäre Untersuchungen, wie z. B. Umweltverträglichkeitsprüfungen, die auf die Datenlieferung vieler Dienststellen angewiesen sind, sind häufig ineffektiv und wenig aussagekräftig, weil die gelieferten Basis- und Fachinformationen nicht immer integriert und zu einer fachübergreifenden Analyse und Wertung zusammengeführt werden können“ [Kophstahl, 1994]. Dem entgegenzusteuern ist das Hauptziel von ATKIS als raumbezogenes Basisinformationssystem. So sind Fachinformationssysteme, die auf ATKIS-Daten basieren, geometrisch untereinander kompatibel, sofern die Fachdaten dies auch sind. Dies bedeutet, daß bei Verschnidungen keine Genauigkeitsverluste entstehen. Da die Daten von einem Anbieter bereitgestellt werden, weisen sie außerdem einen einheitlichen Raumbezug auf. In der ATKIS-Gesamtdokumentation werden folgende Ziele genannt:

- Den Benutzern soll eine einfache und zuverlässige Beschaffung digitaler topographischer Daten ermöglicht werden, mit denen sie ihre Fachdaten verknüpfen können.
- Es soll sichergestellt werden, daß authentische und aktuelle topographische Informationen eingesetzt werden.
- Es soll die Wirtschaftlichkeit der Informationsgewinnung, -verarbeitung und -ausgabe gesteigert werden.

Durch die Erfassung von ATKIS-Daten durch die Landesvermessungsämter wird ein bundeseinheitlicher Datenbestand interessenneutral verwaltet. Die Erfassung und Abgabe von ATKIS-Daten wird als staatliche Dienstleistung verstanden [AdV, 1988].

2.1 ATKIS, wie es 1989 als Standard der BRD eingeführt wurde

Die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) hat 1989 beschlossen, bundesweit das Amtliche Topographisch-Kartographische Informationssystem (ATKIS) aufzubauen. In ATKIS sollen die topographischen Informationen als digitale Daten vorgehalten werden und damit in zeitgemäßer Form die analogen Kartenwerke ablösen bzw. ergänzen. In den analogen Kartenwerken können komplizierte topographische Zusammenhänge zwar leicht verständlich und auch für größere Gebiete übersichtlich dargestellt werden, stehen aber nicht multifunktional zur Verfügung. Eine rechnergestützte Weiterverarbeitung war bis dato nicht gegeben und es gab keine Möglichkeit die unterschiedlichen Objektarten, wie Straßen, Gewässer, Siedlungsgebiete, Naturschutzgebiete usw. unabhängig voneinander darzustellen. Dadurch wird die Übersichtlichkeit behindert und somit die bedarfsorientierte Auswahl erschwert.

Somit stellt ATKIS primär eine digitale Alternative zu den bestehenden topographischen Landeskartenwerken dar.

ATKIS enthält generell zwei verschiedene Datenmodelle:

- das **D**igitale **L**andschafts**m**odell (DLM) und
- das **D**igitale **K**artographische **M**odell (DKM).

Dabei sollten aus einem DLM mehrere DKM abgeleitet werden. Im folgenden werden die beiden Modelle beschrieben, wie sie ursprünglich geplant waren. Es wird aber schon jetzt darauf hingewiesen, daß aufgrund gravierender Probleme bei der Umsetzung des DLM in ein DKM, das DKM auf der 9. AdV-Sitzung im Frühsommer 1995 für nicht praktikabel erklärt wurde, und derzeit nicht weiter verfolgt wird.

Eine Analysemöglichkeit innerhalb des DLM ist nicht vorgesehen. Eine Datenselektion nach geometrisch-topologischen und thematischen Aspekten wird über die Einheitliche Datenbankschnittstelle (EDBS) realisiert.

Das DLM:

Im Digitalen Landschaftsmodell wird die dreidimensionale Struktur der Erdoberfläche durch digitale topographische Daten beschrieben. Dabei wird die Landschaft in topographische Objekte gegliedert, die mit Namen, Lage und Eigenschaften versehen werden. Der Inhalt des DLM ist im Objektschlüsselkatalog (ATKIS-OK) niedergelegt. Es sind drei Aggregationsstufen für jeweils verschiedene Maßstabbereiche vorgesehen. Innerhalb einer Aggregationsstufe ist das

DLM quasi maßstabsfrei und bedingt auch abbildungsunabhängig. Je kleiner der Maßstabsbereich, um so niedriger die Informationsdichte innerhalb dieses DLM. Je nach Aggregationsstufe des DLM unterscheiden sich auch die Quellen, aus denen es gespeist wird. Folgende Aggregationsstufen sind vorgesehen:

- DLM 25 1: 10000 bis 1: 50000
- DLM 200 1: 100000 bis 1: 250000
- DLM 1000 1: 500000 bis 1: 2000000

Das DLM ist im Sinne der Modelltheorie der modernen Kartographie [Findeisen, 1990] ein Primärmodell, das durch topographische Geländeaufnahme entsteht. Dabei wird das Original (die Natur) strukturiert und vereinfacht dargestellt (generalisiert). Der Raumbezug ist damit im Primärmodell bereits enthalten, nicht aber die graphische Visualisierung.

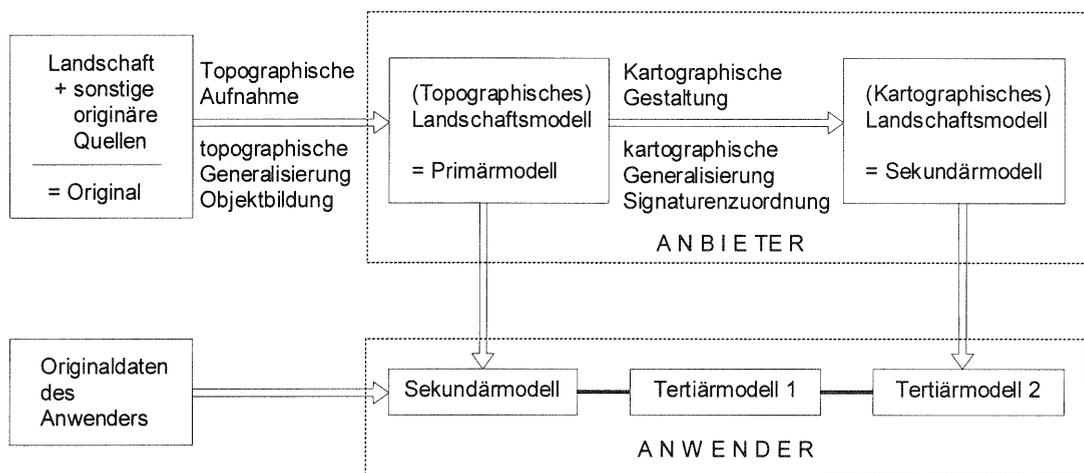


Abbildung 2.1: Kartographische Modelltheorie [Findeisen, 1990]

Die Auffassung, daß es sich bei einem DLM um ein Primärmodell handelt, kann diskutiert werden, da unter primären Erfassungsmethoden solche verstanden werden, bei denen die Daten direkt am Objekt oder aus dessen Abbild gewonnen werden [Bill, Fritsch, 1991]. Dazu zählen die Methoden der

- Vermessung mit ihren Fachbereichen
 - Erdmessung: für die Festlegung der fundamentalen Bezugssysteme
 - Landesvermessung: für die Bestimmung des Grundlagennetzes
 - Detailvermessung: für die Verdichtung des Grundlagennetzes durch Einzelpunkte
- Photogrammetrie, z.B.
 - die Stereoauswertung, mit der auch das digitale Geländemodell entsteht, und
- Fernerkundung, die mit Multispektralaufnahmen die Informationen für die Interpretation und Klassifizierung von GIS-Objekten (Attributbildung) liefert.

Folgt man nun dieser Theorie, so handelt es sich bei der Erfassung der GIS-Daten aus analogen Daten bereits um eine sekundäre Erfassungsmethode, da diese bei einem Produkt beginnt (z.B. TK 25 oder der Deutschen Grundkarte), das nur noch die Informationen enthält, die für dessen Erfassung von Bedeutung waren. Damit hat also bereits eine Selektion der Originaldaten stattgefunden. Die Genauigkeit der Daten ist geringer, meist kann dafür kein Wert angegeben werden und die Datenherkunft ist oft unklar. Die geringere Genauigkeit hat verschiedene Ursachen, die oft auch zusammenspielen:

- Die Daten wurden zu unterschiedlichen Zeiten, mit den zum jeweiligen Zeitpunkt verfügbaren Methoden erfaßt; die Methode ist im nachhinein nicht mehr erkennbar und damit kann keine Genauigkeit angegeben werden.

- Je nach Kartenmaßstab haben bereits Generalisierungsprozesse stattgefunden, die für die GIS-Datenerfassung entweder gar nicht oder unter anderen Gesichtspunkten erfolgt wären.
- Es liegen Fehler bei der Erfassung der Primärdaten vor, z.B. Interpretationsfehler, Anwendung ungeeigneter Erfassungsmethoden.

Beachtet man all diese Einflüsse, so sollte man das DLM nur dann in der Primärebene ansiedeln, wenn seine Daten wirklich durch eine Neuerfassung entstanden sind. Dies ist aus Kostengründen meist nicht der Fall und trifft auch für ATKIS und ATKIS-Bayern nicht zu. Der Pfeil in Abbildung 2.1 vom analogen Kartographischen Modell zum DLM weist zwar auf diese Unstimmigkeit hin, der Gesamteindruck dieser Theorie ist aber dennoch zumindest verwirrend.

Die Semantik wird in ATKIS im **Objektschlüsselkatalog** (ATKIS-OK) beschrieben. Er legt den Inhalt des DLM fest, indem er die Landschaft in Objekte strukturiert und diese klassifiziert. Die Objekte werden nach übergeordneten fachlichen Gesichtspunkten gebildet und können eine Identifikationsnummer sowie einen Namen zugewiesen bekommen. Der **Objekttyp** beschreibt die Ausdehnung des Objekts, nämlich punkt-, linien- oder flächenförmig. Zur Bildung der Objekte werden markante topographische Unterschiede herangezogen. Eine **Objektart** ist eine zusammenfassende Bezeichnung für gleichartige Objekte. Sie werden zur Klassifizierung benötigt. **Objektteile** sind "Teile eines Objekts", die durch einheitliche Attribute und topologische Beziehungen gekennzeichnet sind. Für ihre Bildung sind topographisch untergeordnete Merkmale und/oder semantische Merkmale ausschlaggebend. Ändert sich nur eines dieser Merkmale, muß zwangsläufig ein neuer Objektteil gebildet werden. Jeder Objektteil ist über eine Referenz eindeutig einem Objekt zugeordnet. Mit den **Attributen** werden qualitative und quantitative Eigenschaften eines Objekts oder eines Objektteils beschrieben. Attribute können sowohl den Objekten als auch den Objektteilen zugeordnet werden. Alle Attribute, die dem Objekt zugeordnet sind, beziehen sich automatisch auch auf alle Objektteile. Damit dürfen auf Objektebene nur Attribute zugewiesen werden, die für die Objektbildung notwendig sind, alle übrigen Attribute werden erst bei den Objektteilebenen belegt. Die Anwender können den ATKIS-OK um eigene, fachspezifische Objektarten und Attribute ergänzen. Jedes DLM-Objekt erhält eine Objektcoordinate. Mehrere DLM-Objekte können über Referenzen zu **komplexen DLM-Objekten** zusammengefaßt werden. Dies ist aber nur in zwei Fällen vorgesehen, für die Straße (komplex) und den Schienenverkehr (komplex). Diese sollen nur dann gebildet werden, wenn die den Bahn- bzw. Straßenkörper repräsentierenden Geometrieelemente nicht mit diesem identisch sind, oder diese zwei oder mehr Bahnstrecken bzw. Fahrbahnen tragen. Komplexe Objekte können zwar eigene Attribute und Namen, jedoch keine eigene Objektgeometrie haben. Die verschiedenen Objektarten werden zu **Objektgruppen** zusammengefaßt, die einer groben Systematisierung der Landschaft entsprechen. Die höchste Stufe in der Begriffsorganisation sind die **Objektbereiche**. Jedes Objekt gehört zu einem der sieben Objektbereiche:

- | | | | |
|----------------|----------------|--------------|-----------|
| 1. Festpunkte, | 3. Verkehr, | 5. Gewässer, | 7. Relief |
| 2. Siedlung, | 4. Vegetation, | 6. Gebiete, | |

Die zweidimensionalen Objektbereiche 1. bis 6., die sämtliche den Grundriß wiedergebenden Informationen enthalten, werden auch als **Digitales Situationsmodell** (DSM) bezeichnet. Für die dreidimensionale Beschreibung des Geländes (Objektbereich 7) wird meist der Begriff **Digitales Geländemodell** (DGM) verwendet. Dies ist jedoch nicht ganz richtig, da im Relief lediglich Höheninformationen vorgehalten werden, ohne Bezug zum Grundriß, so daß man hier besser von einem **Digitalen Höhenmodell** (DHM) sprechen sollte. Das heißt nicht, daß ein Höhenwert ohne einen Lagebezug gespeichert wird; dies würde keinen Sinn ergeben, sondern vielmehr daß eine vollständige Modellierung der Höhe z. B. auch Gebäudehöhen, Wuchshöhen bei Objekten aus dem Objektbereich Vegetation usw. beinhaltet. Erst beide Modelle zusammen beschreiben ein **Digitales Geländemodell** (DSM + DHM = DGM), das alle topographischen Informationen dreidimensional enthält [Dorrer, 1975]. Da in den meisten Fällen diese Trennung nicht so streng vorgenommen wird, wird das eigentliche DHM im folgenden, wie meist üblich, als DGM bezeichnet.

Jede ATKIS-Objektart hat eine eigene vierstellige Nummer (Objektidentifikator) im ATKIS-OK, durch die die Verknüpfung zwischen der Geometrie und der Thematik hergestellt wird. Die erste Stelle gibt den Objektbereich, die zweite Stelle die Objektgruppe und die beiden letzten Stellen die Objektart an.

Um möglichst bald ein flächendeckendes DLM bereitstellen zu können, wird von den Landesvermessungsbehörden zunächst ein reduziertes DSM mit den Objektbereichen Verkehr, Gewässer und Vegetation aufgebaut, das sogenannte DLM 25/1. Das Relief wird zunächst als unabhängiges DGM geführt und soll erst später mit dem DSM verknüpft werden.

<i>Beispiel zum ATKIS-OK:</i>	Objektbereich:	Verkehr
	Objektgruppe:	Straßenverkehr
	Objektidentifikator:	3101
	Objektart:	Straße
	Attribute:	z.B. Nutzung privat öffentlich militärisch
	Objekttyp:	linienförmig
	Objektteil:	Brücke
	Attribut:	z.B. Oberflächenmaterial Pflaster Beton Bitumen, Asphalt sonstiges

Zur geometrischen Beschreibung der Lage, der Form und des Raumbezugs dienen das Gauß-Krüger-Meridianstreifen-system (GK-System) und die NN-Höhen. Diese Geometrielemente sind ebenfalls jedem Objektteil durch Referenzen zugeordnet. Um eine geometrische Redundanzfreiheit zu erreichen, kann jedes DLM-Vektorelement zu mehreren Objektteilen gehören. Die Lagegenauigkeit der Geometriedaten soll für alle wesentlichen Objekte (z.B. Straßen, Bahnlinien...) und einzelne ausgewählte Punkte $\pm 3\text{m}$ betragen und für alle übrigen Objekte mindestens im Bereich der TK 50 (± 10 bis 30m) liegen.

Da beim DLM eine permanente Fortführung vorausgesetzt wird, ist kein Historiennachweis enthalten, lediglich das Erfassungsdatum wird beim Objekt bzw. Objektteil abgespeichert.

Die Datenbasis wird, weil je nach Anwendung und Anforderung nicht immer alle möglichen Spezifikationen nötig sind, in den drei Aggregationsstufen DLM 25, DLM 200 und DLM 1000 aufgebaut. Da im DLM noch keine Umsetzung in kartographische Zeichen stattfindet, beziehen sich die Lage- und Höhengenaugkeit des DLM nicht auf eine graphische Abbildung, sondern auf die digitalen Daten im DSM und DGM. Für das DLM 200 wird eine Lagegenauigkeit von $\pm 30\text{m}$ für Knoten- und Linienelemente sowie für alle übrigen Elemente die der TÜK 200 (50-200m) angegeben. Das DLM 1000 soll die Genauigkeit der IWK von 200-500m erreichen.

Betrachtet man die Struktur des DLM-Datenmodells, das auch als Objektklassenprinzip bezeichnet wird, noch einmal eingehender, so ergeben sich bei einer strengen hierarchischen Vorgehensweise innerhalb der Objektklassen (Objektgruppen bzw. -arten) lauter 1:n-Beziehungen zwischen den einzelnen thematischen Mengen (Abbildung 2.2). Eine Objektklasse verzweigt sich also in n Einzelobjekte, die sich wiederum in n Objektteile untergliedern. Bewegt man sich innerhalb dieses Modells in Aufwärtsrichtung, dann entspricht dies einem Generalisierungsschritt der thematischen Objektbeschreibung; die Bewegung in Abwärtsrichtung einem Spezialisierungsschritt, der auch mit einer Ausweitung der Attributstruktur verbunden ist [Molenaar, 1993]. Genau genommen gibt es bei ATKIS gar keine echten Objektklassen. Diese werden nur implizit durch die vierstellige Zahlenfolge des Objektidentifikator wiedergegeben. Dies ist die einfachste Art der semantischen Modellierung, die jedoch, da keine m:n-Beziehungen zugelassen sind, schnell zu Schwierigkeiten führen kann, denn ein Objekt darf immer nur zu einer Klasse gehören. Dieses Problem läßt sich am Beispiel des Grenzbaums darstellen, der zwei Objektbereichen zugeordnet werden kann:

- den Festpunkten, die in die Objektarten Lage-, Höhe- und Schwerefestpunkte unterteilt werden. Der Lagefestpunkt gliedert sich wiederum in die Objektteile Bodenpunkt oder Hochpunkt.
- der Vegetation, die in die Objektgruppen Vegetationsflächen und Bäume/Büsche unterteilt ist. Die Bäume/Büsche werden in den Objektarten Baum, Baumgruppe und Hecke differenziert.

Eine Doppelkodierung sieht ATKIS aber nicht vor, d.h. entweder ist das Objektteil ein Bodenpunkt oder ein Baum.

Betrachtet man unter diesem Aspekt Abbildung 2.2, so wird klar, daß der Schritt der Erfassung des DLM über den Objektartenkatalog nicht ohne Probleme erfolgen wird. Eine redundanzfreie Speicherung der Geometrie, wie sie von ATKIS propagiert wird, ist mit der strengen hierarchischen Struktur, die nur 1:n-Beziehungen zuläßt, ohne Informationsverlust nicht möglich. Vorteilhaft bei diesem Modell ist lediglich der schnelle Datenzugriff, da die Wege innerhalb des Modell fest definiert sind.

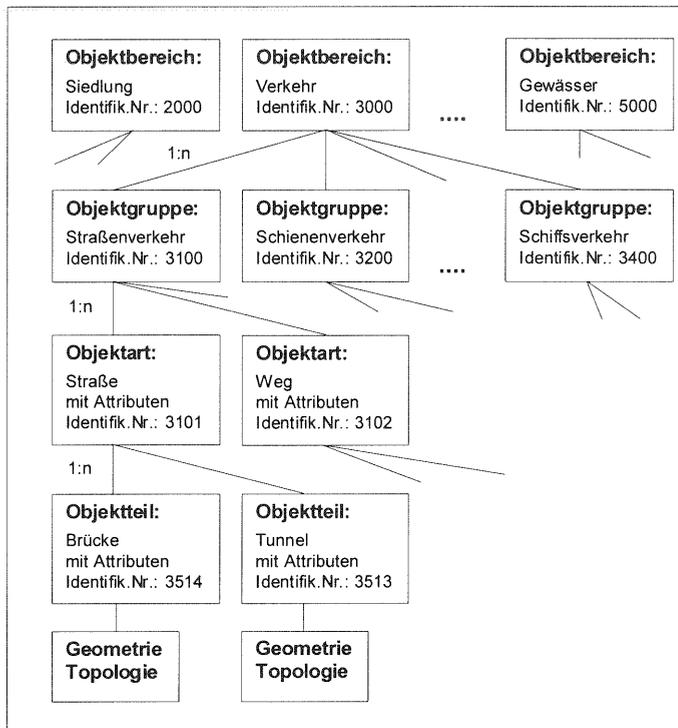


Abbildung 2.2: Objektklassenprinzip in ATKIS

Um diese Probleme in Griff zu bekommen, müssen $m:n$ -Beziehungen ermöglicht werden, d.h. es müssen Beziehungen zwischen m Objekten vom Typ 1 und n Objekten vom Typ 2 im Sinne des Entity-Relationship Modells zugelassen werden. Dazu werden Entitätsmengen gebildet, die durch die speziellen Eigenschaften ihrer Attribute gekennzeichnet sind. Ihre Elemente sind die Entitäten, die durch ein oder mehrere Attributwerte charakterisiert werden. Eine Entität kann zwar noch in Merkmale anderer Art, nicht aber in solche gleicher Art unterteilt werden; z.B. die Entität "Straße" kann nicht mehr in andere Entitäten der Art "Straße", aber in andersartige Entitäten, wie "Fahrbahnbreite" oder "Gehweg", differenziert werden. Eine Entität entspricht bei GIS-Anwendungen i. d. R. einem Objekt. Verbindungen entstehen durch Relationen zwischen den Entitäten, zwischen den Attributen und zwischen Entitäten und Attributen. Das hierarchische Modell ist bei der Bildung dieser Relationen also eingeschränkt. Verzweigungen sind nur von oben nach unten vorgesehen, Kombinationen zwischen den verschiedenen Objekten jedoch nicht. Es können damit keine Mischobjekte gebildet werden. Ein Mischobjekt könnte z.B. eine "Stadt" sein, die sich aus den Wohngebieten, den Industriegebieten und den Versorgungseinrichtungen zusammensetzt. Das Wohngebiet seinerseits ist auch ein Mischobjekt, das aus den Häusern, Straßen und Spielplätzen besteht. Kennzeichnend für Mischobjekte, die nicht unbedingt voneinander unabhängig sein müssen, ist die $m:1$ -Beziehung. Bei ihrer Bildung werden alle Attribute der Einzelobjekte aggregiert (vererbt). All diese Möglichkeiten sind beim hierarchischen Modell nicht gegeben, wodurch es in seiner Funktionalität eingeschränkt ist.

Das DKM:

Das **Digitale Kartographische Modell** ist die zweidimensionale, optisch aufgearbeitete Umsetzung der gespeicherten topographischen Informationen in kartographische Signaturen, die je nach Maßstab und Zeichenschlüssel variieren. Das DKM enthält die kartographische Gestaltung des Primärmodells. Dabei finden wieder Generalisierungs- und Strukturierungsprozesse statt. Die geometrische Dimension der Signaturen ist maßstabsbezogen und abhängig von der gewählten Abbildung. Eine Übersicht über den ursprünglich geplanten Informationsfluß in ATKIS zeigt Abbildung 2.3. Dieses logische und einfach aussehende Schema scheint auf den ersten Blick leicht realisierbar zu sein. Auf die einzelnen Bestandteile, Abschnitte und Probleme wird im weiteren in diesem Kapitel eingegangen.

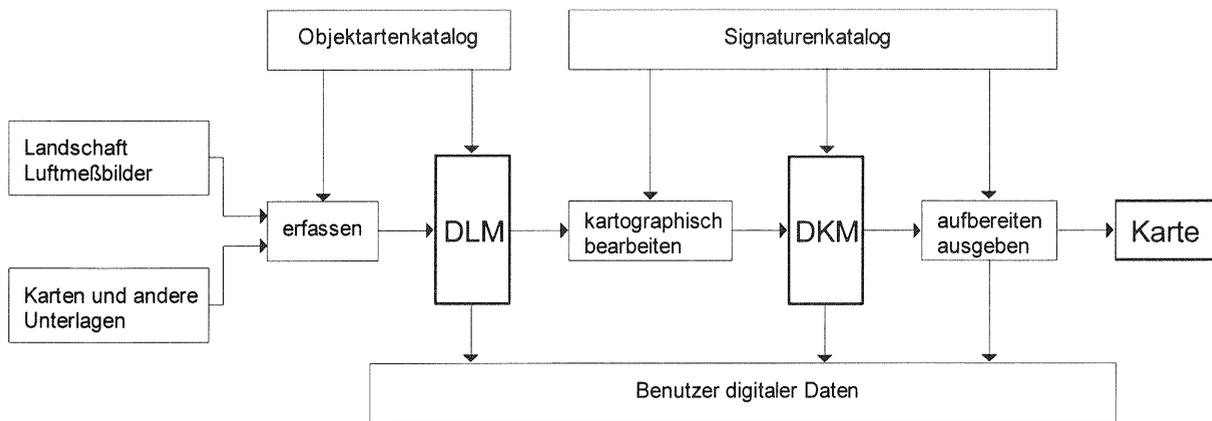


Abbildung 2.3: Ursprünglich geplanter Informationsfluß in ATKIS [AdV, 1988]

Für diese Visualisierung müssen die signaturunabhängig gespeicherten DLM-Landschaftsobjektarten in entsprechende DKM-Kartenobjektarten umgesetzt werden. Die Grundlage dafür sind die Eigennamen, Objektarten und Attribute der DLM-Objekte. Die festen Regeln für diese Umbildung mit adäquaten Signaturen sind im Signaturenkatalog (ATKIS-SK) festgelegt, also die Art der graphischen Darstellung wie Strichstärke, Mindestabstände, Auswahl- und Zusammenfassungskriterien, Verdrängungs- und Darstellungsprioritäten, Signaturverwendungsregeln, Beschriftung und dergleichen. Dabei wird deutlich, daß mit der Ableitung des DKM ein Generalisierungsprozeß verbunden ist. Somit ist das DKM maßstabsbezogen. Sollen DKM unterschiedlicher Maßstäbe aus einem DLM abgeleitet werden, so sind dafür jeweils separate Signaturenkataloge erforderlich. In der Regel wird einem DLM-Objekt ein DKM-Objekt zugeordnet; es besteht also eine 1:1 Beziehung (z.B. ein Haus im DLM entspricht einem Haus im DKM). In Sonderfällen, die einer kartographischen Generalisierung entsprechen, ist auch eine n:m Beziehung möglich, wobei $n > m$ und $m > 1$ (z.B. fünf Häusern im DLM entsprechen drei Häuser im DKM). Wie im DLM gibt es auch im DKM Objektteile, jedoch keine Attribute. Zwingend notwendige Objektteile eines DKM-Objektes sind die, welche die Kartengeometrie beschreiben. Dabei kann es zwischen DLM- und DKM-Objektteilen auch zu einer n:1 Beziehung kommen, wenn beispielsweise DLM-Attribute nicht zur Generierung der DKM-Objektteile herangezogen werden. DKM-Objektteile können auch Textinformationen sein.

Damit ist der ATKIS-Signaturenkatalog seiner Funktion nach mit den Musterblättern für die analogen Landeskartenwerke vergleichbar. Die Erzeugung des DKM läuft wegen der notwendigen Generalisierung interaktiv ab und ist damit maßgeblich von der Hard- und Softwareausstattung abhängig.

Entsprechend der Identifikationsnummer bei den DLM-Objektarten erhält jedes Kartenobjekt eine dreistellige Kartenobjektartnummer, deren erste Stelle identisch mit der Kennzahl des DLM-Objektbereichs ist. Dies kann z.B. wie folgt aussehen:

DLM-Objektart 3101:	"Straße" mit den
Attributen:	"vier und mehr Fahrspuren" und "getrennte Richtungsfahrbahnen" ergibt die
Kartenobjektart 301:	"Schnellverkehrsstraße" mit der
Verdrängungspriorität:	"8" (höchste Priorität) dem
Kartenobjekttyp:	"linienförmig" der
Darstellung:	"schwarz mit orange" und der
Darstellungspriorität:	"7".

Nicht beachtet ist hier, daß es drei Kategorien für Schnellverkehrsstraßen gibt, nämlich:

- Schnellverkehrsstraße für den Fernverkehr, Kartenobjektart: 301,
- Schnellverkehrsstraße für den Regionalverkehr, Kartenobjektart: 302 und die
- Schnellverkehrsstraße für den Nahverkehr, Kartenobjektart: 303.

So einfach wie es auf den ersten Blick hier scheint, ist die Ableitung des DKM jedoch nicht. Bisher gibt es trotz intensiver Anstrengungen keine Programmsystem, die erfolgversprechend automatisch eine kartographische Generalisierung ausführen können. Die bisherigen Prozesse laufen interaktiv ab, wobei einfache, immer wiederkehrende Situationen bereits mit Erfolg bewältigt werden können. Bei komplexeren Generalisierungsproblemen (z.B. Häuser und Verkehrswege) gibt es Algorithmen, die die Situation automatisch erkennen, welche anschließend interaktiv bearbeitet werden muß. Da bei ATKIS derzeit primär das DLM, speziell das DLM 25/1 aufgebaut wird, dessen Inhalt noch nicht

dem der analogen TK 25 gleich kommt, also für die Ableitung einer digitalen TK 25 nicht ausreicht, finden diese Probleme momentan noch wenig Beachtung.

Sicherlich kommt der analogen Ausgabe des DKM nicht mehr die gleiche hohe Bedeutung zu wie einer analogen TK 25. Das DLM mit seiner bedeutend höheren Genauigkeit und den differenzierteren Informationen ist als Arbeitsgrundlage besser geeignet. Damit ergeben sich auch für die Kartengestaltung neue Gesichtspunkte in Richtung leichter Lesbarkeit. Dazu sollen die Mindestdimensionen der Signaturen erhöht, Farben verstärkt eingesetzt und nur noch eine gut lesbare Kartenschrift (z.B. Univers) in mehreren Farben verwendet werden. Zusätzlich sollen wichtige Straßennamen aufgenommen werden. Damit wird die Verdrängung zwangsläufig größer, die geometrische Genauigkeit sinkt etwas und der Generalisierungsgrad steigt. Dies scheint der Autorin nicht sinnvoll. Diese Veränderungen in der Ausgestaltung ergeben ein völlig neues Kartenbild, welches in einigen Punkten alten kartographischen Grundsätzen widerspricht, bei denen die Karte als künstlerisches Produkt betrachtet wurde. Aus diesem Grund wird Bayern die Mindestdimensionen nicht erhöhen und keine Straßennamen aufnehmen. Ob dieses Festhalten an alten Maximen in einer modernen, computerbeherrschten Welt haltbar ist, kann bezweifelt werden, da der Einzug der digitalen Technik auch in den Bereich der Landkarten bereits stattgefunden hat. Hier kann man sogar so weit gehen und fragen, ob es überhaupt sinnvoll ist, ein DKM abzuleiten, oder ob die analoge Karte nicht in einigen Jahren so an Bedeutung verlieren wird, daß sie unter dem Aspekt "Historie" ganz aus dem Bereich der Vermessungsverwaltungen verschwinden wird.

Da die Geodaten die wertvollste und langlebigste Komponente eines GIS darstellen und die Bereitstellung mit enormem finanziellen und personellen Aufwand verbunden ist, hat man sich bei ATKIS für einen stufenweisen und bedarfsorientierten Aufbau entschlossen. Dabei stand zunächst die Verwirklichung der drei Aggregationsstufen des DLM im Vordergrund. Der Realisierung des DKM wurde zunächst kaum Beachtung geschenkt. Derzeit wird das DLM 25 bei den Landesvermessungsämtern, das DLM 200 und das DLM 1000 beim Institut für Angewandte Geodäsie (IfAG) in Frankfurt aufgebaut. Die Erfassung des DLM 25 in seiner ersten Ausbaustufe, dem DLM 25/1, die bis Ende 1995 bundesweit abgeschlossen sein sollte, erfolgt in den einzelnen Bundesländern unterschiedlich. Während manche Länder einzelne Objektbereiche flächendeckend für das ganze Bundesland erfassen, arbeiten andere Länder blattweise. Beide Varianten haben Vor- und Nachteile. Bei der flächendeckenden Erfassung können wichtige Objektbereiche vorrangig behandelt und so möglichst schnell an die Anwender abgegeben werden. Dagegen können bei der blattweisen Kompletterfassung Ballungsräume von besonderem Interesse bevorzugt erfaßt werden.

2.2 ATKIS in Bayern

Wie in Kapitel 1.2 erwähnt, hat Bayern bereits Mitte 1986 mit der Datenerfassung zum Aufbau einer digitalen topographischen Datenbasis begonnen [Roder, Wimmer, 1989]. Da Bayern mit 70.553 km² flächenmäßig das größte Bundesland der Bundesrepublik Deutschland (248.717 km²) darstellt, ist der Arbeitsaufwand für die Bereitstellung dieser Daten im Vergleich zu den anderen Bundesländern entsprechend hoch. So sah sich die Bayerische Vermessungsverwaltung Mitte der achtziger Jahre veranlaßt, sofort, noch vor einem endgültigen Abschluß der Verhandlungen der AdV über die Konzeption von ATKIS mit dem Aufbau einer derartigen Datenbasis zu beginnen. Bayern entwickelte ein auf seine Bedürfnisse zugeschnittenes Konzept, das **Geographische Grundinformationssystem** GEOGIS, das die Vorgaben von ATKIS, wie es die AdV später beschlossen hat, jedoch weitgehend berücksichtigte. GEOGIS war auf eine schnell realisierbare und überschaubare Lösung abgestimmt. Als Erfassungsgrundlage dient die analoge Topographische Karte 1:25 000 (TK 25). Wichtig war bei dem GEOGIS-Konzept, daß es auch schon die Abgabe von Teilergebnissen an die Nutzer vorsah. In seiner Endstufe sollte GEOGIS aus der digitalen topographischen Karte 1:25 000 und dem digitalen Geländemodell (DGM) bestehen, welches wie bei ATKIS separat vorgehalten wird.

Mit diesem Konzept wollte Bayern möglichst schnell dem gesetzlichen Auftrag nachkommen, die Daten der Vermessungsverwaltung dem Anwender in einer zeitgemäßen Form anzubieten. Bei der Erfassung von GEOGIS wurde Bayern immer als Gesamtheit betrachtet, d.h. nur in Ausnahmefällen sollten Teilbereiche gesondert behandelt werden. In einer ersten Stufe wurden nur wenige Objektarten erfaßt, die dafür aber in relativ kurzer Zeit bayernweit zur Verfügung standen. Dadurch war es möglich, den Nutzern auch Zwischenergebnisse zur Verfügung zu stellen. Diese Vorgehensweise hat sich bewährt und wird beibehalten. Mitte 1995 wurde die Bezeichnung GEOGIS durch ATKIS-Bayern ersetzt. Bereits hier sei erwähnt, daß Bayern sich inzwischen dem ATKIS-Konzept weitgehend angeschlossen hat, wenn auch die Realisierung noch einige Zeit in Anspruch nehmen wird. Im folgenden Text wird die Bezeichnung GEOGIS dort verwendet, wo noch deutliche Unterschiede zum ATKIS-Konzept zu erkennen sind. Auf wesentliche Unterschiede wird im folgenden noch näher eingegangen.

Art der Datenerfassung:

Die Objekte der realen Welt werden in GEOGIS **nicht** objektweise erfasst, sondern in einer Ebenenstruktur abgelegt. Der Inhalt der topographischen Datenbasis wird in Schichten (Ebenen oder Folien) strukturiert. Dies ist vergleichbar mit der Aufteilung des Kartengrundrisses einer analogen Topographischen Karte in die verschiedenen Farbfolien beim Druck. Innerhalb dieses Ebenenprinzips herrscht jedoch keine Hierarchie wie bei der objektweisen Erfassung. Alle Ebenen sind gleichberechtigt. Die Gliederung erfolgt entsprechend der Objektbereiche des ATKIS-OK. Jeder Objektbereich wird in einer eigenen Geographischen Datenbasis (GDB) abgespeichert, die wiederum in bis zu 19 Schichten unterteilt ist. Die Daten können somit nur nach Schichten selektiert werden, nicht aber nach anderen Kriterien. Die Inhalte der einzelnen Objektbereiche (entsprechen den GDB) werden nach verschiedenen Methoden erfasst.

Die Objektbereiche *Siedlung*, *Verkehr* und *Gebiete* werden über manuelle Digitalisierung, teils vom Bayerischen Landesvermessungsamt (BLVA), teils durch Vergabe an private Ingenieurbüros, erfasst. Die Erfassung erfolgt nicht direkt aus der TK 25, sondern aus den entsprechenden Blaupausen. Die schwarze Druckvorlage der TK 25 ist in vier Blaupausen im Maßstab 1:10000 aufgeteilt. Dies sind Vergrößerungen der 1:25000-er Folie. Maßgebend für den Inhalt waren die Objekte, die auch in der TK 25 enthalten waren, da diese zunächst ausschließliche Erfassungsgrundlage darstellte.

Der Objektbereich *Gewässer* wurde durch halbautomatische Digitalisierung erfasst. Die blaue Druckvorlage der TK 25 wurde dazu gescannt, anschließend georeferenziert und mittels halbautomatischer Linienverfolgung unter Berücksichtigung der Gewässerordnung digitalisiert. In einem weiteren Arbeitsschritt wurde die Fließrichtung der Gewässer erfasst. Anzumerken ist, daß verrohrte Gewässer nicht lagerichtig erfasst sind, sondern nur die Verbindung zwischen den Gewässern widerspiegeln.

Der Objektbereich *Vegetation* beinhaltet in einer ersten Stufe nur die Waldgrenzen. Dazu wurde die grüne Folie der TK 25-Druckvorlage gescannt, georeferenziert und anschließend eine automatische Digitalisierung durchgeführt. Dies bietet sich an, da die grüne Folie als Begrenzungslinien nur die Waldflächen enthält. Damit war der Objektbereich *Vegetation* der einzige, der nicht in weitere Schichten unterteilt war. Anzumerken ist hier, daß eine Waldfläche, die durch eine Straße oder ein Gewässer durchschnitten wird, bei dieser Methode in zwei getrennte Waldflächen zerlegt wird. Diese haben einen Abstand voneinander, der dem der graphischen Darstellungsbreite der Straße bzw. des Gewässers in der TK 25 entspricht (vgl. Abbildung 1.8). Diese Besonderheit wird später noch eine wichtige Rolle bei der Datenbereinigung spielen.

Ein gravierender Nachteil dieser Erfassungsart ist die redundante Abspeicherung der Geometrie, die zudem nicht zueinander paßt, da alle Objektbereiche unabhängig voneinander erfasst wurden. Dies soll die folgende Abbildung verdeutlichen:

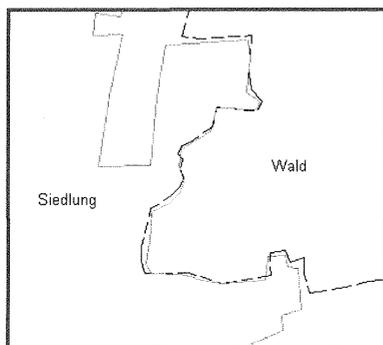


Abbildung 2.4: Redundante Daten in ATKIS-Bayern

Dennoch ist GEOGIS keine reine Digitalisierung der TK 25. Den Daten werden sehr wohl zusätzliche Informationen beigelegt, wie z. B. die Ordnung der Gewässer, die Textdeskriptoren bei den Straßen usw..

Für die weitere Vorgehensweise in Richtung ATKIS wird hier ein erheblicher Zeitaufwand für die Nachbearbeitung erforderlich, da eine Zusammenführung der getrennten GDB nicht ohne manuelle Eingriffe möglich ist. Diese Eingriffe müssen interaktiv erfolgen und sind schwer zu automatisieren. Einen Versuch hierzu zu unternehmen, ist mit ein Ziel dieser Arbeit. Eine Genauigkeitssteigerung kann bei diesem Verfahren ohne Zusatzinformationen aber nicht erwartet werden.

Maßgebend für den Inhalt der einzelnen GDB waren zunächst nur die Objekte, die auch in der TK 25 enthalten waren. Inzwischen werden diese nach den Maßgaben des ATKIS-OK für das DLM 25/1 erweitert, welches derzeit jedoch noch nicht flächendeckend vorliegt. Bei der zeitlichen Reihenfolge der Erfassung wird nach Möglichkeit auf die Bedürfnisse der Nutzer Rücksicht genommen. Aus all diesen Gründen wurde GEOGIS als Vorstufe zu ATKIS angesehen.

Die Entscheidung Bayerns, sich mittelfristig doch dem Konzept von ATKIS anzuschließen, wurde durch den Geodaten-Bedarf Bundeswehr maßgeblich beeinflusst. Diese forderte bundesweit eine einheitliche Datenstruktur und die Datenabgabe über die Einheitliche Datenbankschnittstelle (EDBS). Diese wurde als Schnittstelle für ATKIS von der AdV festgelegt. Die EDBS wurde in den anderen Bundesländern bereits für die Automatisierte Liegenschaftskarte (ALK) eingesetzt und für ATKIS ausgebaut und erweitert - an Stelle der ALK steht in Bayern die Digitale Flurkarte (DFK). Die Bundeswehr hatte für die EDBS ihrerseits extra ein eigenes Umsetzungsprogramm in Auftrag gegeben. Anfangs wurden die Daten nur über die SICAD-SQD-Schnittstelle bzw. über die Quasistandardschnittstelle DXF abgegeben. Das DXF-Format wird als Quasistandard angesehen und von vielen gängigen Softwareprodukten unterstützt. Es stammt von der Firma Autodesk, die das weitverbreitete CAD-System AutoCad vertreibt. Das DXF-Format ändert sich damit von Version zu Version geringfügig. Dies erschwert den Datenaustausch über diese Schnittstelle enorm, außerdem können keine Objekte gebildet und keine Sachdaten übergeben werden.

Die angestrebte Lagegenauigkeit von $\pm 3\text{m}$ gemäß ATKIS-Forderung wird für Bayern mittelfristig noch als unrealistisch angesehen, da die Datenerfassung weitgehend durch Digitalisierung aus der TK 25 erfolgt und nicht, wie für ATKIS vorgesehen, aus Orthophotos im Maßstab 1:10000 oder der Deutschen Grundkarte 1:5000. Die Deutsche Grundkarte gibt es in Bayern aus historischen Gründen nicht. An ihrer Stelle stehen die Bayerischen Flurkarten in Kombination mit der Höhenflurkarte. Diese wurden zwar ursprünglich ebenfalls im Maßstab 1:5000, und nur in den Realteilungsgebieten Unterfrankens in 1:2500 erstellt. Mit dem zunehmenden Wachstum der Siedlungsgebiete war dieser Maßstab jedoch nicht mehr ausreichend. Es wurde der Maßstab 1:1000 eingeführt; in diesen Bereichen werden die Flurkarten in 1:5000 nicht mehr fortgeführt. Damit verfügt Bayern über kein laufend aktualisiertes, flächendeckendes Kartenwerk im Maßstab 1:5000, das für die Erfassung von ATKIS primär herangezogen werden könnte. Die Genauigkeit von $\pm 3\text{m}$ wird im Zuge der Aktualisierung angestrebt.

Ein weiterer wesentlicher Unterschied ist, daß ursprünglich bei GEOGIS keine Objektbildung durchgeführt wurde. Lediglich bei der Erfassung des untergeordneten Straßennetzes ist ein 58 Zeichen fassender Textdescriptor für jeden Linienzug gebildet worden, in dem die ATKIS-Attributwerte angegeben werden. Dieser Textdescriptor ist als fest formatierter String definiert und hat beispielsweise beim Straßennetz folgenden Aufbau:

Stelle	Bezeichnung	Bedeutung
1-4	OAR	Objektartennummer nach ATKIS-OK
6-23	NAME	Name des Objekts aufgeteilt in:
6-9		Blattnummer der TK 25
10		Kennzeichen der Erfassungsgrundlage: G: gesamte TK 25 A, B, C, D für die entsprechende Blaupause
11-23		laufende Nummer des Objektes als Ersatz für Straßen- oder Wegebezeichnung (für spätere Objektbildung)
25-28	att1	Attributwert für die Widmung (WDM)
30-33	att2	Attributwert für die Funktion (FKT)
35-38	att3	Attributwert für die Internationale Bedeutung (IBD)
40-43	att4	Attributwert für die Fahrbahnbreite (BRF)
45-48	att5	Attributwert für die Breite des Verkehrsweges (BRV)
50-53	att6	Attributwert für die Fahrstreifenzahl (FSZ)
55-58	att7	Attributwert für den Zustand (ZUS)

Der Textdescriptor soll später für die Objektbildung nach ATKIS verwendet werden. In den vor 1993 erfaßten Bereichen muß auch dieser Textdescriptor manuell, über Nachbearbeitung, eingefügt werden.

Eine weitere Besonderheit ist der Code, der jedem erfaßten Objekt angehängt ist. Dieser enthält verschlüsselte Angaben für die spätere kartographische Wiedergabe. Dies ist bei der bayerischen Methode der Datenerfassung relativ einfach möglich, da die Basisdaten aus der TK 25 stammen und damit bereits in generalisierter Form digital vorliegen. Es ist daher keine automatische Generalisierung notwendig. Es war nämlich von Anfang an erklärtes Ziel, daß GEOGIS bzw. ATKIS-Bayern auch einen wesentlichen rationalisierenden Effekt für die Verwaltung selbst hat. Die TK 25 soll zum frühestmöglichen Zeitpunkt aus den ATKIS-Bayern-Daten ableitbar und damit leichter und schneller herstellbar sein.

Dieser Code ist ein fünfelementiger String, der z.B. für das Straßennetz wie folgt definiert ist:

Stelle	Bedeutung
1	steuert die Darstellungspriorität für die Kartenausgabe bei Überführungen Ziffer 0: dieses Element liegt „unter“ einem anderen Element und wird von diesem überführt Ziffer 1: dieses Element liegt „über“ einem anderen Element; es überführt dieses Element
2-4	beinhaltet die Kartenobjektartennummer nach den vorläufigen Richtlinien des neuen Musterblattes der TK 25; sie ordnet dem jeweiligen Element die jeweilige Signatur zu
5	dient als Platzhalter für eine mögliche Erweiterung der bisherigen Kartenobjektartennummer (Standard ist Ziffer 0)

Derzeit werden am BLVA die bereits erfaßten Daten in die von ATKIS vorgesehene objektweise Datenstruktur umgesetzt. Dies ist mit einem erheblichen Arbeitsaufwand verbunden. Von den 613 bayerischen TK 25-Blättern sind derzeit 15% bereits in die ATKIS-Datenstruktur umgesetzt. Der Abschluß der Arbeiten ist für 1999 geplant.

Objektbereich Relief - das Digitale Geländemodell

Da im Rahmen dieser Arbeit die geometrische Genauigkeit lediglich der zweidimensional vorgehaltenen Daten untersucht wird, sei der Objektbereich Relief hier nur der Vollständigkeit halber erwähnt. Es werden ohne Wertung die beim BLVA vorhandenen Daten aufgeführt.

Bereits Mitte der 80er Jahre wurde in Bayern mit dem Aufbau eines landesweiten Digitalen Geländemodells begonnen, das zunächst nur für den internen Bedarf zur Bereitstellung von Orthophotos benötigt wurde. Dieses ist seit einigen Jahren bereits komplett verfügbar und weist eine Höhengenaugigkeit von $\pm 2-3$ Metern auf. Dies entspricht der Höhengenaugigkeit der TK 25 und wird daher als DLM 25 bezeichnet. Lediglich in Waldbereichen können größere Ungenauigkeiten auftreten.

Die Ausgangsdaten für das DLM 25 wurden über photogrammetrische Profilmessung (Abstand ca. 50 Meter) aus Luftbildern mit dem Bildmaßstab 1:23 000 gewonnen. Im Hochgebirge wurden die Höhenlinien über Digitalisierung der bereits vorhandenen Aufnahmeblätter gewonnen. Die erfaßten Primärdaten wurden dann in ein einheitliches DGM-Gitter mit Maschenweiten von 50m, 100m und 200m umgerechnet und werden so den Nutzern zur Verfügung gestellt.

Außerdem wird ein DGM 5 bedarfsweise aufgebaut. Dieses hat die Genauigkeit der bayerischen Höhenflurkarte 1:5000 (HFK 5), die in freiem Gelände $\leq 0,5$ Meter beträgt. Der überwiegende Teil der Primärdaten für das DGM 5 stammt aus photogrammetrischen Gittermessungen aus Luftbildern im Bildmaßstab 1:15 000. Die Punktabstände der Primärdaten betragen maximal 25m. In geeigneten Fällen werden auch Digitalisierungen aus der HFK 5 durchgeführt, soweit diese die geforderte Genauigkeit aufweist. Außerdem beinhaltet das DGM 5 zusätzliche Informationen im Vektorformat über morphologische Strukturen wie Geländekanten, Gerippllinien und ausgewählte Einzelpunkte. Das DGM 5 wird wahlweise als geprüfter Primärdatenbestand abgegeben oder in einem regelmäßig abgeleiteten Geländegitter mit 10 m bzw. 20 m Gitterweite. Auf Wunsch werden hier auch die Strukturinformationen mit abgegeben. Nimmt der Nutzer eine gewisse Wartezeit in Kauf, so werden die Daten auf Wunsch auch originär erzeugt.

2.3 Unterschiede zwischen ATKIS und ATKIS-Bayern - Vergangenheit und Gegenwart

Die bisher teilweise schon erwähnten Unterschiede zwischen ATKIS und ATKIS-Bayern sollen im folgenden noch einmal kurz und übersichtlich zusammengestellt werden.

Eigenschaft	ATKIS	ATKIS-Bayern
Lagegenauigkeit	<ul style="list-style-type: none"> • $\pm 3-5$ Meter 	<ul style="list-style-type: none"> • wie in der TK 25, d. h. $\pm 10-20$ Meter
Erfassungsgrundlage	<ul style="list-style-type: none"> • Orthophoto, • Deutsche Grundkarte 1:5 000, • Luftbilder • in den neuen Bundesländern aus der TK 10 • weitere Quellen, die die geforderte Geometriegenauigkeit erfüllen 	<ul style="list-style-type: none"> • vorwiegend aus der TK 25 • attributive Informationen auch über: <ul style="list-style-type: none"> – Luftbilder – Stadtpläne – Karten der Straßenbauämter – militärische Straßen- und Brückenkarten – Streckenpläne der Bahn – amtliche Verzeichnisse
Datenmodell	<ul style="list-style-type: none"> • Trennung von DLM und DKM • objektorientierte Datenstruktur 	<ul style="list-style-type: none"> • keine Trennung von DLM und DKM (aufgrund der Erfassungsgrundlage nicht notwendig) • ebenenorientierte Datenstruktur

2.3.1 Lagegenauigkeit

Bei ATKIS ist bundesweit eine Lagegenauigkeit von ± 3 m angestrebt. Die Lagegenauigkeit ist von der zur Verfügung stehenden Digitalisiervorlage, der Erfassungsgenauigkeit, aber auch von der Schärfe der Objektbegrenzung abhängig. Die höchsten Anforderungen werden im Zusammenhang mit Fahrzeugnavigation an das Straßennetz gestellt. Hier beträgt die geforderte Lagegenauigkeit innerorts 2-3 m, außerhalb der Ortschaften 10-20 m. Da in Bayern aus der TK 25 erfaßt wird, kann nur eine Genauigkeit von 10-20 m erwartet werden. In einzelnen Fällen können sogar Abweichungen bis zu 40 m auftreten.

Die Gebietsgrenzen werden in der TK 25 kartographisch an das Verkehrsnetz, das bei der Generalisierung höchste Priorität hat, angepaßt. Damit ist die Lagegenauigkeit der Gebietsgrenzen oft sehr gering, wenn auch in der TK 25 streng darauf geachtet wird, daß die relative Lage der Gebietsgrenzen zu Straßen und Flüssen richtig wiedergegeben ist. Damit ist dennoch gewährleistet, daß die einzelnen Objektarten den Gemeinden, Gebieten usw. richtig zugeordnet werden können, wenn auch die absolute Lage ungenau ist.

An den Grenzen zu den benachbarten Bundesländern werden die ATKIS-Datenbestände der dortigen Landesvermessungsbehörden übernommen. Diese werden kartographisch für die bayerischen Verhältnisse nicht nachbearbeitet. Es sei nur erwähnt, daß es beim Austausch der Daten zwischen den einzelnen Bundesländern aufgrund der verschiedenen Systeme zu Problemen kommt. Durch den Datenaustausch ist gewährleistet, daß landesweit widerspruchsfreie Geometrieelemente vorhanden sind. Die Datenerfassung endet normalerweise grundsätzlich an der Grenze des jeweiligen Bundeslandes. Verläuft die Landesgrenze aber in einem flächenförmig modellierten Gewässer (z. B. Main oder Salzach), so ist es allein schon aus fachlichen Gesichtspunkten notwendig, die gegenüberliegende Uferseite mitzuerfassen. Dadurch sind Dateninkonsistenzen nicht zu vermeiden.

Eine Besonderheit muß bei den Gewässern noch erwähnt werden. Verrohrte Gewässer sind in der Regel nicht lagerichtig erfaßt, da aus wasserwirtschaftlicher Sicht nur interessant ist, welche Verbindungen zwischen den einzelnen Gewässern bestehen, nicht aber wo diese exakt liegen. Diese Besonderheit erfolgte in Absprache mit dem Bayerischen Landesamt für Wasserwirtschaft. Sie muß dem Nutzer über Metadaten ausdrücklich mitgeteilt werden.

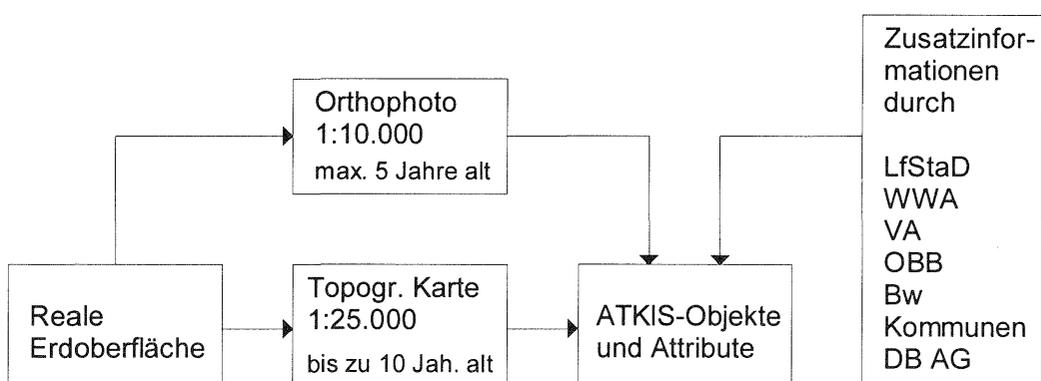
2.3.2 Erfassungsgrundlage

Ein wesentlicher Unterschied ist die Erfassungsgrundlage und die daraus resultierende Genauigkeit der Daten.

1992 stand in Bayern im Rahmen der Umstellung auf die bundeseinheitliche Datenstruktur noch einmal die Frage im Raum, auch bei der Erfassungsgrundlage auf andere Medien als die TK 25 umzusteigen, um eine Genauigkeitssteigerung zu erreichen. Zur Diskussion standen eine komplett photogrammetrische Auswertung im Bildmaßstab 1:23 000. Dieser Bildmaßstab steht durch die fünfjährige turnusmäßige bayernweite Befliegung flächendeckend zur Verfügung. Dazu wären ergänzende terrestrische Aufnahmen erforderlich gewesen. Diese Variante schied wegen des hohen Zeitaufwands und der damit verbundenen Kosten aus. Bei der Datenerfassung aus Orthophotos 1:10 000 wäre ein sehr hoher Zeitaufwand für die Beschaffung notwendiger Zusatzinformationen zur sachgerechten Objektbildung not-

wendig gewesen. Somit wurde auch diese Möglichkeit nicht als praktikabel eingestuft. So blieb als einzig gangbarer Weg nach wie vor die Erfassung aus der TK 25. Mit ein wesentlicher Grund für diese Entscheidung war zum einen, daß in der TK 25 die Erdoberfläche schon weitgehend strukturiert dargestellt ist. Zum anderen sollen die geographischen Basisdaten nicht nur möglichst schnell einem breiten Nutzerkreis zur Verfügung gestellt werden können, sondern auch einen Eigennutzen zur rationellen TK 25-Herstellung haben. Erklärtes bayerisches Ziel ist es, die Herstellung der TK 25 zu beschleunigen und zu vereinfachen. Um dies rationell durchzuführen, werden kartographisch generalisierte Vektordaten benötigt, da die automatische Generalisierung bzw. die Ableitung eines DKM aus dem DLM bis heute nicht realisiert ist. Diese Voraussetzung ist bei der Erfassung aus der TK 25 gegeben, da hinterher wieder eine Karte im gleichen Maßstab ausgegeben werden soll. Wo die TK 25 als Erfassungsgrundlage für die Objekt- bzw. Objektteilbildung für das DLM 25/1 nicht ausreichend ist, werden Karten, Orthophotos und andere Unterlagen hinzugezogen. Dies ist z. B. bei der Unterscheidung zwischen Wohnbau-, Industrie- und Gewerbeflächen unbedingt erforderlich. Von externen Stellen werden Unterlagen des Landesamtes für Statistik und Datenverarbeitung, der Wasserwirtschaftsämter, der Vermessungsämter, der Obersten Baubehörde und der Bundeswehr sowie der Deutschen Bahn AG verwendet. Mit großen Kommunen (Großstädte) werden bereits jetzt direkt Daten ausgetauscht. Die weiteren über 2000 bayerischen Kommunen sind hierfür technisch noch nicht eingerichtet.

Diese pragmatische Vorgehensweise, bei der ohne großen Aufwand eine Ableitung einer analogen kartographischen Wiedergabe möglich ist, hat sich bewährt.



LfStaD: Bayerisches Landesamt für Statistik und Datenverarbeitung
 WWA: Wasserwirtschaftsamt
 VA: Vermessungsamt
 OBB: Bayerische Oberste Baubehörde
 Kommunen: nur bei großen Kommunen
 DB AG: Deutsche Bahn AG

Abbildung 2.5: Erfassungsgrundlagen für das DLM 25/1 bei ATKIS-Bayern [Schellein, 1997]

2.3.3 Datenmodell

Das getrennte Datenmodell von DLM und DKM, so wie es 1989 von der AdV bundesweit eingeführt wurde, kam in Bayern nie zum Einsatz. Ein wesentlicher Kritikpunkt war, daß man keine Möglichkeit sah mittelfristig eine Lösung für die automatische Generalisierung zu finden, so daß kein Nutzen für die Verwaltung nach innen gegeben war. Denn aus einem DLM mit einer Genauigkeit von ± 3 m kann ohne automationsgestützte Generalisierung das DKM nicht wirtschaftlich abgeleitet werden. Dies war aber eine wesentliche Forderung, da ein derartiger Personalaufwand aus wirtschaftlichen Gründen nur dann vertretbar ist, wenn auch internes Rationalisierungspotential erwartet werden kann. Mit der bayerischen Vorgehensweise, aus der TK 25 zu erfassen, wird das Problem der DKM-Ableitung geschickt umgangen, da das DLM gleichzeitig das DKM beinhaltet. Diese beiden, im ATKIS-Konzept getrennten Modelle, werden somit zu einem zusammengeführt. Die Idee, aus einem DLM mehrere verschiedene DKM ableiten zu können, mußte dafür zurückgestellt werden, doch wurde dafür kurzfristig auch kein Bedarf gesehen.

Das ATKIS-Datenmodell in der 1989 festgeschriebenen Form wurde Mitte 1995 von der AdV aufgegeben, weil sich die Ableitung des DKM aus dem DLM als derzeit nicht lösbares Problem darstellte. Das neue Modell wurde auf der ATKIS-Symposium 1996 in Koblenz nur kurz vorgestellt und sieht wie folgt aus:

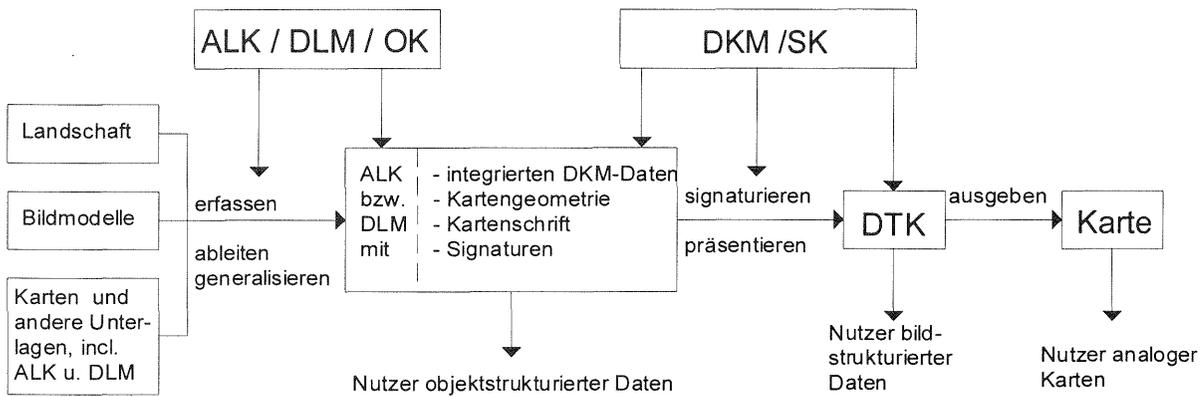


Abbildung 2.6: Neues ATKIS-Modell, Koblenz Oktober 1996

Ein weiteres Ziel der Bayerischen Vermessungsverwaltung war es, einige wesentliche Objektarten innerhalb kürzester Zeit flächendeckend für die Nutzer zur Verfügung zu stellen. Dies ist ein prinzipielles Ziel der Kundenorientierung und für die Bayerische Vermessungsverwaltung von besonderer Bedeutung. Außerdem stellte sich heraus, daß die bereits 1986 eingeführte Ebenenstruktur für die meisten Nutzer ausreichend war, so daß zunächst an diesem Konzept festgehalten werden konnte. Eine Bestätigung dieser Ansicht zeigt sich darin, daß die meisten Nutzer die Quasistandardschnittstelle DXF für die Datenabgabe verwenden, obwohl diese eine Objektbildung nicht zuläßt.

2.4 Neueste Tendenzen bei ATKIS

Bei der Weiterentwicklung des ursprünglichen Datenkonzepts von ATKIS ergeben sich derzeit sechs Themenschwerpunkte:

- Ausbau von ATKIS als Gesamtkonzept aller grund- und bodenbezogenen Basisdaten
- zukünftige Behandlung kartographischer Informationen
- Klärung der Frage, welche Basisdaten primär benötigt werden und welche Priorität sich daher für die Bereitstellung ergibt
- Entwicklung eines Konzeptes zur Präsentation und Generalisierung der Daten
- Aufgabe des vektororientierten DKM als zweites Datenmodell
- Berücksichtigung nationaler und internationaler Normungsaktivitäten für den Datenaustausch

Im folgenden sollen die einzelnen geplanten Vorhaben kurz erläutert und gewertet werden.

ATKIS als Gesamtkonzept aller grund- und bodenbezogenen Basisdaten

Aus wirtschaftlichen Gesichtspunkten ist es nicht sinnvoll, grund- und bodenbezogene Daten innerhalb einer Verwaltung redundant in verschiedenen Datenbeständen vorzuhalten. Dies bezieht sich auf die Daten, die im Liegenschaftskataster in Form der Automatisierten Liegenschaftskarte (ALK) und des Automatisierten Liegenschaftsbuch (ALB) vorgehalten werden. In Bayern ist dies etwas anders strukturiert. Die ALK heißt hier aus der Tradition der Flurkarten heraus Digitale Flurkarte (DFK) und das ALB ist hier schon zu einem Automatisierten Grund- und Liegenschaftsbuch erweitert worden (AGLB), da alle Grundbuchämter bzw. Vermessungsämter ihre originär gewonnenen Daten über Modem austauschen. Doppelarbeit wird dadurch vermieden. DFK und AGLB werden unter dem Oberbegriff Grundstücks- und Bodeninformationssystem GRUBIS zusammengefaßt. Diese Struktur wird weiter ausgebaut und unter dem Arbeitstitel AGLB 95 unter übergeordneten Gesichtspunkten völlig neu, objektweise strukturiert. Darauf soll hier jedoch nicht weiter eingegangen werden.

Was GRUBIS und ATKIS-Bayern betrifft, haben zwar die Daten dieser beiden Geo-Informationssysteme unterschiedliche rechtliche Bedeutung, doch ist hier ein einheitliches Datenmodell, das Redundanzfreiheit gewährleistet und auch vom Anwender gefordert wird, mittelfristig sicher der einzig richtige Weg.

Um dies zu realisieren, sind noch wesentliche Vorarbeiten erforderlich. So sind die Modellunterschiede zwischen DFK (ALK) / AGLB (ALB) / ATKIS-Bayern (ATKIS) herauszuarbeiten. Außerdem sind auch wesentliche Fragen zu klären,

wie z. B. in einem gemeinsamen Datenmodell auch nicht-raumbezogene Informationen, wie z. B. Eigentümer geführt werden. Das „Objekt“ Eigentümer hat keine Geometrie und kann nicht auf einen bestimmten Ort festgelegt werden. Nicht zuletzt muß geklärt werden, ob überhaupt eine objektweise Datenmodellierung sinnvoll eingesetzt werden kann. Das Bestreben der AdV, die Datenmodelle für ATKIS, ALB und ALK zusammenzuführen, wird Bayern nachhaltig unterstützen.

Wie werden zukünftig kartographische Informationen behandelt?

Das DKM wird aufgrund der erheblichen Umsetzungsprobleme nicht mehr in der Form verfolgt, wie sie 1989 beschlossen wurde. Vielmehr gilt es, die kartographischen Daten und auch andere Informationen benutzerfreundlich zu integrieren und die Konzeption des Datenmodells dahingehend zu modifizieren.

Es werden derzeit bundesweit drei Modelle für die Integration der Kartengeometrie in ATKIS diskutiert.

Dabei stellt sich die Frage, ob

1. wie in Bayern bereits geschehen, DLM-Daten bereits generalisiert sein dürfen,
2. damit verbunden auch ein Qualitätsvermerk an den Daten angebracht werden sollte, oder
3. auch eine Historienführung erforderlich ist.

Derzeit wird lediglich das Entstehungsdatum als Attribut an das Objekt mitgeführt. Das Problem der Zeitdokumentation gestaltet sich mannigfaltig, spielt aber eine immer größere Rolle im Zusammenhang mit GIS-Produkten. Eine wesentliche Frage ist, welche Zeit bzw. Zeiten verwaltet werden sollen. In Frage kommen hier mehrere Möglichkeiten, wobei evtl. nicht alle Varianten auch für alle Objektarten zutreffend sein müssen:

- **Entstehungszeit**, d.h. der Zeitpunkt, zu dem das Objekt in der realen Welt entstanden ist, z. B. für eine Straße der Zeitpunkt, bei der sie für den Straßenverkehr freigegeben worden ist bzw. evtl. schon wenn mit dem Bau begonnen worden ist
- **Erhebungszeit**, d.h. der Zeitpunkt, zu dem die Informationen zum Objekt erhoben werden
- **Datenbankzeit**, d.h. der Zeitpunkt, zu dem das Objekt in der Datenbank eingetragen wird
- **Fortführungszeit**, d.h. der Zeitpunkt, zu dem das Objekt in der Datenbank eine Änderung erfährt, also fortgeführt wird.

Ein weiterer wesentlicher Punkt ist die Betrachtung von Zeiträumen. Dazu ist es erforderlich, daß alle Veränderungen bis hin zur kompletten Löschung eines Objekts in der Datenbank abgebildet werden. Dies bedeutet, daß ein „gelöschtes“ Objekt in der Datenbank erhalten bleiben muß, aber den Status „historisch“ bekommt. Ob diese aufwendige Dokumentation des „Lebenszyklusses“ eines Objekts notwendig ist, bedarf der Klärung.

Welche Daten werden primär benötigt?

Wie sich zeigt, ist die Erfassung des DLM 25 sehr aufwendig und für viele Anwender zu detailliert, so daß sich die Frage stellt, ob die Einführung eines informationsreduzierten DLM 50 sinnvoll ist. Auch die Weiterentwicklung im Bereich der Rasterdaten muß berücksichtigt werden. Die Integration von Rasterdaten ist im ursprünglichen ATKIS-Konzept nicht vorgesehen, die Notwendigkeit war 1989 vielleicht auch noch nicht in diesem Maße erkennbar. Diese Erweiterung des ATKIS-Datenmodells wird unter dem Stichwort **Digitale Topographische Karte (DTK)** eingeführt werden und könnte somit das DKM ersetzen, das so nicht mehr an den Anwender abgegeben, wohl aber intern bei den Landesvermessungsämtern zur TK-Herstellung benötigt wird. Damit hat der Anwender auch nicht mehr das Problem zwei separate Datenmodelle vorhalten zu müssen, sondern die Trennung erfolgt dann über die Datenart.

Nationale und internationale Normungsaktivitäten

Wichtig erscheint es hier vor allem, alle nationalen, aber auch internationalen Normungsaktivitäten zu berücksichtigen, wobei „international“ nicht am europäischen Kontinent enden darf. Ohne eine einheitliche Datenschnittstelle, die den Austausch zwischen allen Anwendern ermöglicht, werden immer Kosten für Doppelerfassung entstehen, weil vorhandene Daten nicht in spezielle Anwendersysteme transferiert werden können. Auch muß bedacht werden, ob nicht auch an Anwender, die nur einen reduzierten Informationsgehalt benötigen, z. B. Daten auch ohne Attribute, abgegeben werden können.

Derartige Normungsrichtlinien finden sich sowohl in der DIN (Deutsche Industrie-Norm) als auch in ISO-Normen (International Organization for Standardization). Dem Europäischen Normungsausschuss (CEN) wurde die Aufgabe

übertragen, das Europäische Austauschformat (ETF) als allgemeines Austauschformat für raumbezogene Daten in Europa festzulegen.

2.5 Klassifikation und Wertung von geometrischen Dateninkonsistenzen

Im Kapitel 1.3 wurden eine Vielzahl an möglichen Ursachen für geometrische Dateninkonsistenzen systematisch aufgezeigt. In diesem Abschnitt wird eine Wertung in Bezug auf ATKIS und ATKIS-Bayern vorgenommen. Diese wird nicht, wie vielleicht zu erwarten ist, vorwiegend zwischen semantischen, topologischen und geometrischen Auswirkungen unterscheiden, sondern eine Einstufung in der Dringlichkeit für die Beseitigung bzw. Verringerung dieser geben.

Kapitel	Titel	Auswirkungen	Abhilfe	Priorität
1.3.1	Inkonsistenzen bei der Definition der geom. Grundprimitive	Sind schwer aufzudecken und beruhen meist auf Fehlern, die dem Bearbeiter bei der Erfassung unterlaufen sind. Derartige Fehler dürfen sich nicht häufen, sind aber in jedem Datenbestand, ob analog oder digital, vorhanden.	Penible Revision, anderfalls kaum eine Chance dies zu bemerken, da hierfür oft auch Ortskenntnis erforderlich ist.	5
1.3.2	Erfassungsgrundlage mit nicht aktuellem Datenbestand	Dabei entstehen zwar Inkonsistenzen, die für den einzelnen Nutzer evtl. schwerwiegende Konsequenzen haben können. Der Datenbestand ist veraltet. Ist der Fehler entdeckt, dann ist dieses Problem das gleiche, das bei der jeder Fortführung des Datenbestands auftritt.	Methode entwickeln, die eine halbautomatische Fortführung erlaubt. Ein Hinweis, wie diese erfolgen kann, wird in dieser Arbeit gegeben.	2
1.3.3	Lageverschiebungen aufgrund verschiedener Einpassungen	Führen zu wesentlichen geometrischen Ungenauigkeiten der Objekte. Durch die ikonisierte Datenerfassung ist damit auch die relative Lage der Objekte zueinander stark gefährdet.	Versuchen geeignete Transformationen nachträglich durchzuführen, wobei die Bereitstellung von geeigneten Paßpunkten die wesentliche Schwierigkeit ist.	4
1.3.4	Isolierte Objektbereichserfassung	Führt zu stark redundanten Datenbeständen, da die gleiche geometrische Information bei verschiedenen Objekten in verschiedenen GDB abgespeichert ist. Dies hat unmittelbare Auswirkung auf die Qualität und die Fortführbarkeit der Daten.	Wesentliches Ziel dieser Arbeit ist es, eine einfache und möglichst automatisch ablaufende Methode zu entwickeln. Ein möglicher Ansatz wird in Kapitel 3 vorgestellt.	1
1.3.5	Generalierungsfehler in der Erfassungsgrundlage	So der Fehler in der Erfassungsgrundlage oder aufgrund logische Unrichtigkeit entdeckt wird, gilt hier das gleiche wie für 1.3.2.	siehe 1.3.2	2

Anmerkungen zu dieser Tabelle:

Priorität ist hier die Dringlichkeit, mit der derartige Inkonsistenzen in einem Datenbestand behoben werden sollten. Je kleiner die Zahl, um so dringender. Da die Stufe 2 zweimal vergeben wurde, fehlt Stufe 3.

3 Modelltheoretische Behandlung von geometrischen Dateninkonsistenzen

Die modelltheoretische Behandlung der geometrischen Dateninkonsistenzen gliedert sich in mehrere Schritte. Im Vordergrund stehen hier die Dateninkonsistenzen, die aufgrund der isolierten Objektbereichserfassung entstanden sind.

3.1 Grundlagen und Voraussetzungen

In diesem ersten Abschnitt werden einige grundsätzliche Techniken vorgestellt, die in den, im Kapitel 3.2 vorgestellten und für die Situation Bayern entwickelten Verfahren, zum Einsatz kommen werden. Dabei handelt es sich vielfach um gängige in der Vermessung und Kartographie bekannte Verfahren.

3.1.1 Transformationen

Auffallend sind systematische Verschiebungen einzelner Bereiche.

Wie bereits erwähnt, wurden beim Bayerischen Landesvermessungsamt die Objektbereiche *Verkehr*, *Siedlung* und *Gebiete* durch manuelle Digitalisierung aus den vier Blaupausen der jeweiligen TK 25 erfaßt. Die Einpassung der Blaupausen zur Digitalisierung erfolgt über die Blattecken. Da insbesondere beim Straßennetz, welches außerdem in zwei Stufen - die übergeordneten und die nachgeordneten Straßen - erfaßt wurde, sehr viele Informationen zu digitalisieren waren, mußten für jedes Blatt mehrere Einpassungen mittels Transformation vorgenommen werden. Diese Einpassungen sind alle unabhängig voneinander, d. h. die Verschiebungsvektoren, der Maßstab und die Drehung weichen geringfügig, jedoch trotzdem merklich, voneinander ab, obwohl die Einpassung für den Bearbeiter nur dann erfolgreich abgeschlossen ist, wenn festgelegte Genauigkeitskriterien erfüllt sind. Dies erklärt die blockweisen Verschiebungen.

Bei der Überlagerung mit einer Flurkarte, wie in der folgenden Abbildung geschehen, wird dies deutlich sichtbar, und zwar durch die blockweisen Verschiebungen der einzelnen Abschnitte gegeneinander. Nicht erkennbar sind, zumindest in dem für diese Untersuchung vorliegenden Datenbestand, Maßstabsverzerrungen und Rotationen. Deutlich wird dieser Effekt vor allem beim Straßennetz. Da die Gebietsgrenzen bei der Generalisierung von untergeordneter Bedeutung sind, werden diese in der Blaupause nicht immer lagerichtig dargestellt. Die Verdrängung bewirkt, daß die Objekte nur relativ zueinander richtig wiedergegeben sind, so daß derartige Effekte aus diesem Grund schon nicht eindeutig ausgemacht werden können. Dies gilt auch für die Siedlungsgrenzen, da diese in der Flurkarte nicht eindeutig zu erkennen und auch in der TK 25 nur mit geringer Genauigkeit erfaßt sind.

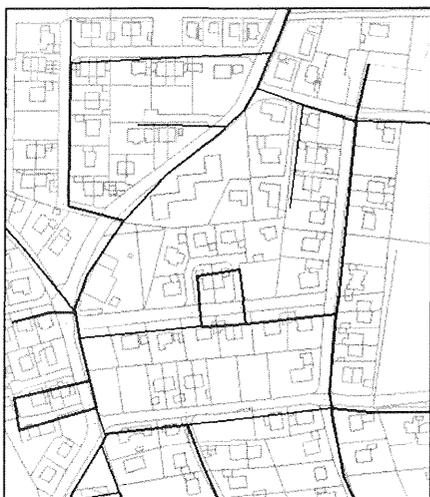


Abbildung 3.1: Überlagerung der DFK mit ATKIS-Bayern-Daten

Auch die Einpassung (Georeferenzierung) der Gewässerfolien, die zunächst gescannt werden, erfolgt über Transformation. Im Gegensatz zur manuellen Digitalisierung wird hier die Georeferenzierung jedoch nur ein einziges Mal

vorgenommen. Sie ist für die gesamte anschließende halbautomatische Digitalisierung einheitlich. Das gleiche gilt für die Erfassung des Objektbereichs *Vegetation* nach bayerischer Methode.

Bei der Georeferenzierung spielen die zur Einpassung verwendeten Paßpunkte eine wesentliche Rolle.

In den für den Objektbereich Verkehr verwendeten Blaupausen sind die Linienstärken im Vergleich zur TK 25 um den Faktor 2,5 verstärkt dargestellt. Dies gilt auch für die Paßkreuze, die für die Einpassung verwendet werden. Je dicker eine Linie ist, um so schwerer wird es für den Bearbeiter die Mitte der Linie abzuschätzen (vgl. Abbildung 3.2). Dies führt zu Ungenauigkeiten, die sich dann ggf. bei der nachfolgenden Transformation bemerkbar machen.

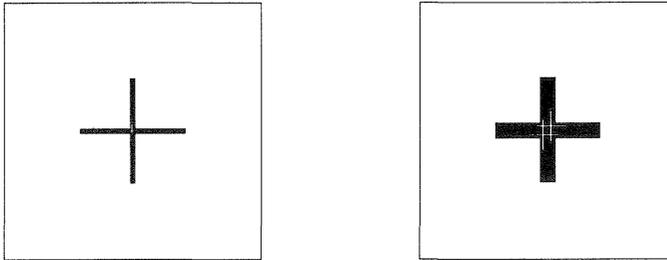


Abbildung 3.2: Optischer Eindruck bei der Digitalisierung

Die einzige Möglichkeit, diese Verschiebungen zu minimieren, ist die Durchführung einer erneuten Transformation. Diese Möglichkeit wird aber ausschließlich für den Objektbereich *Verkehr* in Betracht gezogen. Dieser liegt in der Erfassungsgrundlage weitgehend ohne Generalisierung vor, da das Verkehrsnetz bei der Generalisierung in der TK 25 nur durch Gewässer und die Lage der Trigonometrischen Punkte verdrängt wird. Berücksichtigt man dies jedoch bei der Punktauswahl und meidet solche Bereiche, so kann man beim Verkehrsnetz von bestmöglicher Lagegenauigkeit ausgehen. Voraussetzung hierfür ist jedoch, daß geeignete Paßpunkte für die Transformation zur Verfügung stehen. Als Referenzdatenbestand kommen daher nur Punkte in Frage, die als fehlerfrei zu betrachten sind, da sonst kein wesentlicher Genauigkeitsgewinn durch die Transformation zu erwarten ist. Diese Bedingung kann nur mit identischen Punkten, die aus der DFK gewonnen werden, erfüllt werden. Von Seiten der ATKIS-Bayern-Daten kommen als identische Punkte nur die Knoten des Straßennetzes in Frage. Diese sind in der DFK, die ausschließlich die Flurstücksgrenzen beinhaltet, nicht direkt enthalten. Daher bliebe allenfalls das aufwendige Verfahren, diese visuell in der DFK zu bestimmen (zu koordinieren). Da die DFK jedoch nicht flächendeckend vorliegt und auch die visuelle Bestimmung der Paßpunkte nicht als ganz fehlerfrei angesehen werden kann, gibt es praktisch keine Möglichkeit flächendeckend mit vertretbarem Aufwand die aus den verschiedenen Einpassungen herrührenden geometrischen Inkonsistenzen zu bereinigen. Eine weitere Schwierigkeit ist die Identifizierung der verschiedenen zusammengehörenden Blöcke der einzelnen Einpassungen, um sie dann gemeinsam zu behandeln. Im Testgebiet wird eine derartige Transformation versuchsweise durchgeführt, wobei klar ist, daß ein derartiges Verfahren nicht für einen bayernweiten Datenbestand realisierbar ist. Die Ergebnisse werden im Kapitel 4.2.2 vorgestellt. Wie die Paßpunktbestimmung visuell aus der DFK erfolgt, zeigt Abbildung 3.3.



Abbildung 3.3: Paßpunktbestimmung aus der DFK

3.1.2 Selektion redundanter Datensätze

Jede GBD für sich ist ein homogener Datenbestand; ATKIS-Bayern besteht aus mehreren GDB, die zusammengespielt (superpositioniert) mehrfach redundant sind.

Der nächste, wesentlich aufwendigere und schwieriger zu modellierende Schritt betrifft das Auffinden der Dateninkonsistenzen, die durch die ikonisierte Objektbereichserfassung (siehe Kapitel 1.3.4) zustande gekommen sind. Um diese Dateninkonsistenzen zu selektieren, müssen die geometrischen Informationen der einzelnen, getrennt voneinander erfaßten Ebenen, auf das notwendige Minimum, das Skelett, reduziert werden. Dieser Schritt entspricht einer Generalisierung.

Für die Generalisierung kommen in diesem Fall keine punktförmigen Objekte in Betracht, da die Generalisierung hier nicht dazu dienen soll, den Datenbestand so zu reduzieren, daß eine lesbare Karte erzeugt werden kann. Vielmehr sind alle erfaßten Objekte beizubehalten, also auch alle punktförmigen. Lediglich die Darstellung der linien- bzw. flächenhaften Objekte ist in der Wiedergabe zu vereinfachen. Für diese Generalisierung können linien- und flächenförmige Objekte in gleicher Weise behandelt werden, wenn bei der Linienvereinfachung durch Glättung und Ausdünnung beachtet wird, daß bei flächenförmig generierten Objekten die Linienpolygone wieder geschlossen sind. Ein derartiges Verfahren ist nach [Hake, 1994] eine Objektgeneralisierung, und zwar im Sinne einer Modellgeneralisierung, da hier ein neues Objektmodell geringerer Auflösung abgeleitet wird. Dabei gilt es die Linienpunkte beizubehalten, die die charakteristische Form des Objekts beschreiben. In diesem Kapitel sollen zunächst allgemein gängige Generalisierungsverfahren vorgestellt werden.

Dabei soll hier nicht das gesamte Spektrum der Möglichkeiten der Generalisierung beleuchtet werden, da dies ein eigenes Forschungsgebiet darstellt. Hier sollen lediglich einfache geometrische Verfahren kurz erläutert werden, die z. B. die Stützpunkte einer Polygonlinie auf ein verträgliches Minimum reduzieren und skelettieren, so daß aber dennoch alle wesentlichen Informationen erhalten bleiben. Nicht eingegangen wird auf wissensbasierte Systeme der Generalisierung, bei denen auch die Beziehungen zwischen mehreren Objekten eine wesentliche Rolle spielen oder auf notwendige Übertreibungen, damit der Karteninhalt lesbar wird usw..

Es wurden in der Vergangenheit eine Vielzahl von Verfahren zur Generalisierung von linearen Objekten, deren Geometrie in Vektorform vorliegt, entwickelt. Eine Auswahl der am häufigsten eingesetzten Verfahren wird im folgenden vorgestellt:

Die Verfahren der Liniengeneralisierung lassen sich je nach der zugrundeliegenden Methode in drei Gruppen unterteilen:

- geometrisches Vereinfachen der Linie, indem nach bestimmten Kriterien Stützpunkte ausgelassen werden
- die Linie durch mathematische Funktionen, wie z. B. Splines approximieren
- Ersetzen von kartographisch unbedeutenden Linienteilen eines Objektes durch weniger und einfachere Linienstücke. Dieses Verfahren entspricht der manuellen Methode, wenn z. B. mehrere Serpentinien einer Straße durch eine einzige wiedergegeben werden. An die Automation werden hier höchste Anforderungen gestellt.

Die hier aufgeführten Verfahren gehören der ersten Gruppe, der geometrischen Vereinfachung, an. Lineare Objekte sind in Vektorform durch die Abfolge ihrer Stützpunkte gegeben (Polylinien). Dabei werden diejenigen Punkte einer Linie selektiert, die, ohne daß die Form der Linie eine wesentliche Veränderung erfährt, gelöscht werden können. Es bieten sich die verschiedensten Möglichkeiten an, um die Formtreue der Linie sicherzustellen, wie z. B. Schrankenwerte für einen maximalen Abstand, eine maximale Winkeländerung usw.. Die Verfahren lassen sich nach der Art der verwendeten Schrankenwerte und Anzahl der einbezogenen Nachbarpunkte klassifizieren.

Punktverfahren:

Punktverfahren sind die einfachsten Methoden zur Reduzierung der Datenmengen. Sie benötigen keine Schwellenwerte, da die Punktauswahl völlig unabhängig von den benachbarten Punkten erfolgt. Damit bleibt die Liniengeometrie gänzlich unberücksichtigt. Eine Möglichkeit ist das *n-ter Punktverfahren*, bei dem nur jeder n-te Punkt der Stützpunktfolge erhalten bleibt, alle dazwischenliegenden jedoch entfallen. Dieses Verfahren ist sehr einfach, da keine Berechnungen erforderlich sind. Der Reduktionsgrad ist über die Größe *n* vorher schon festgelegt. Diese Methode ist gänzlich ungeeignet für Linien, die Kleinformen wiedergeben, da diese bei einer derartigen Ausdünnung nicht beibehalten werden können. Andererseits bleiben bei geradlinigen Linienverläufen zu viele Restpunkte übrig, so daß das Verhältnis zwischen Genauigkeitsverlust und Anzahl der Restpunkte in beiden Fällen nicht sehr befriedigend ist. Akzeptable Ergebnisse sind bei diesem Verfahren nur zu erwarten, wenn die ursprüngliche Linie eine extrem hohe Punktdichte aufweist und wenige Kleinformen enthalten sind. Die Punktauswahl kann auch per Zufallsgenerator vorgenommen werden, was aber keine wesentliche Auswirkung auf das Ergebnis hat.

Algorithmisierung des n-ter Punkt-Verfahrens:

```

N=3
ANZ=Anzahl der Stützpunkte d. ursprünglichen Linie
KOORD(ANZ)= Koordinatenpaar d. Stützpunktes
I=1
while I < ANZ
  {
    KOORD_N=KOORD(I)
    I=I+N
    KOORD_NEU=append KOORD_NEU KOORD_N
  }
KOORD_NEU=append KOORD_NEU KOORD(ANZ)

```

Lokale Verfahren:

Bei lokalen Verfahren werden nur die unmittelbar benachbarten Punkte eines Stützpunktes berücksichtigt. Damit wird zwar die lokale Geometrie beachtet, die Gesamtcharakteristik einer Linie bleibt aber auch hier ohne Einfluß.

Beim *Distanzverfahren* werden alle Punkte der Stützpunktmenge entfernt, die innerhalb einer bestimmten Distanz vom aktuellen Punkt liegen. Um aufwendige Berechnungen einzusparen, kann statt einem Kreis auch ein Quadrat verwendet werden, da dann die Distanzberechnung durch einen Koordinatenvergleich ersetzt werden kann. Bei diesem Verfahren ergibt sich eine gleichmäßigere Punktdichte, die Liniencharakteristik bleibt aber nahezu ohne Einfluß.

Beim *Pfeilhöhenverfahren* werden immer drei aufeinanderfolgende Punkte betrachtet, aus denen ein Dreieck gebildet wird. Vom mittleren Punkt wird dann der Lotabstand - die Pfeilhöhe - bestimmt und gegen einen Schrankenwert verglichen. Ist die Pfeilhöhe kleiner als der Toleranzwert, wird der mittlere Stützpunkt entfernt und ein neues Dreieck, beginnend beim dritten Stützpunkt gebildet. Ist die Pfeilhöhe größer als der Toleranzwert, wird der mittlere Stützpunkt auch beibehalten und von ihm beginnend ein neues Dreieck gebildet. Dadurch wird mehr Rücksicht auf die Liniencharakteristik genommen, denn die Pfeilhöhen berücksichtigen die Linienform. Der Nachteil ist, daß maximal jeder zweite Stützpunkt entfernt wird. Damit ist bei einem relativ geraden Verlauf der Linie der Reduktionsgrad zu gering. In solchen Fällen muß die Prozedur mehrmals hintereinander ausgeführt werden.

Ähnlich wie das Pfeilhöhenverfahren funktioniert auch das *Krümmungsverfahren*, nur, daß als Auswahlkriterium die Winkeländerung im mittleren Punkt des gebildeten Dreiecks als Ersatz für die Krümmung verwendet wird. Ansonsten laufen die beiden Verfahren identisch ab. Auch beim Krümmungsverfahren wird damit maximal jeder zweite Punkt gelöscht. Um bei relativ geradlinigem Linienverlauf eine höhere Reduktion zu erhalten, müssen beide Verfahren mehrmals hintereinander ausgeführt werden. Auch Kombinationen des Pfeilhöhen- mit dem Krümmungsverfahren sind denkbar. Ein Beispiel beschreibt G.F. Jenks (beschrieben in R.B. McMaster), bei dem zwei Distanzen und eine Krümmung verwendet werden.

Algorithmisierung des Disatanzverfahrens:

```

D=10 ; z. B. Distanzradius = 10
ANZ=Anzahl der Stützpunkte d. ursprünglichen Linie
KOORD(ANZ)= Koordinatenpaar d. Stützpunktes
KOORD_NEU=KOORD(1)
I=1
while I < ANZ
  {
    KREIS (KOORD(I), D)

```

```

while KOORD(I) ∈ Kreisfläche
{
  if (I < (ANZ-1))
  then I=I+1
  else break
}
; Abbruch für while-Schleife

KOORD_NEU=append KOORD_NEU KOORD(I)
}
KOORD_NEU=append KOORD_NEU KOORD(ANZ)

```

Algorithmisierung für das Pfeilhöhenverfahren:

```

PH=10 ; z. B. Pfeilhöhe = 10
ANZ=Anzahl der Stützpunkte d. ursprünglichen Linie
KOORD(ANZ)= Koordinatenpaar d. Stützpunktes
KOORD_NEU=KOORD(1)
I=1
while I ≤ (ANZ-2)
{
  LINIE=(KOORD(I); KOORD(I+2))
  ABSTAND=(KOORD(I+1), LINIE)

  if (ABSTAND < PH)
  then I=I+2
  else I=I+1

  KOORD_NEU=append KOORD_NEU KOORD(I)
}
if I = (ANZ-1)
then KOORD_NEU=append KOORD_NEU KOORD(ANZ)

```

Sektionale Verfahren:

Hier ist das von T. Lang vorgestellte Verfahren zu nennen (beschrieben in [Douglas, Peucker, 1973]), bei dem der Einfluß der Liniencharakteristik auf das Generalisierungsergebnis wesentlich größer ist, als bei den bisher beschriebenen Methoden. Der Grundgedanke hierbei ist, die Linie durch geeignete Sehnenstücke zu ersetzen, wobei die dazwischenliegenden Stützpunkte innerhalb eines Toleranzstreifens um das Sehnenstück liegen müssen. Ausgehend vom Startpunkt (*anchor*) werden zu den nachfolgenden Stützpunkten (*floater*) Sehnen gezogen, und zwar so lange, bis der Abstand eines Punktes zur Sehne zwischen anchor und floater größer als ein Schrankenwert ist. Ist dies der Fall, werden die anderen dazwischenliegenden Punkte entfernt, der Punkt vor dem floater zum neuen anchor und der Prozeß wiederholt. Der Reduktionsgrad ist hier sowohl abhängig vom Grenzwert für den maximal zulässigen Abstand als auch von der Linienform.

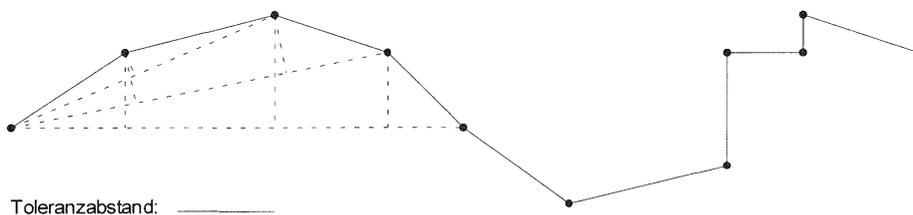


Abbildung: 3.4: Verfahren nach T. Lang [Heubach, 1992]

Algorithmisierung des Verfahrens nach T. Lang:

```

T=10 ; z. B. Toleranzabstand = 10
ANZ=Anzahl der Stützpunkte d. ursprünglichen Linie
KOORD(ANZ)= Koordinatenpaar d. Stützpunktes
KOORD_NEU=KOORD(1)

A=1
F=A+2
N=A+1

```

```

ANCHOR=KOORD(A)
FLOATER=KOORD(F)

while F ≤ ANZ
{
  LINIE=(ANCHOR; FLOATER)

  while N < F
  {
    ABSTAND=(KOORD(N), LINIE)
    if ABSTAND < T
    then N=N+1
    else N=F+100 ; beendet while-Schleife
  }

  if N = F
  then F=F+1;
     N=A+1
     FLOATER=KOORD(F)

  if N = (F+100)
  then KOORD_NEU=append KOORD_NEU KOORD(F-101)
     A=(F-101)
     ANCHOR=KOORD(A)
     F=A+3
     FLOATER=KOORD(F)
  }
KOORD_NEU=append KOORD_NEU KOORD(ANZ)

```

Globale Verfahren:

Das globale Verfahren nach [Douglas, Peucker, 1973] hat den gleichen Grundgedanken wie das Lang-Verfahren. Ein gekrümmtes Liniestück kann durch die verbindende Sehne ersetzt werden, sofern kein dazwischenliegender Stützpunkt einen größeren Abstand als den Schrankenwert von der Sehne hat. Ist der Abstand größer als der Schrankenwert, muß ein weiterer Punkt eingeschaltet werden. Dazu wird der Stützpunkt mit dem größten Abstand von der Sehne gewählt, da dieser am wahrscheinlichsten einen Knickpunkt der Linie darstellt. Das Verfahren arbeitet vom Großen ins Kleine, der Linienanfangspunkt ist *anchor*, der Linienendpunkt ist *floate*r. Nun werden die Abstände aller Linienpunkte berechnet, und der Punkt mit dem größten Abstand ermittelt. Ist sein Abstand größer als der Schrankenwert, so wird dieser Punkt der neue *floate*r. Das Verfahren wird nun mit der neuen Sehne wiederholt und zwar so lange, bis der größte Abstand von der Sehne kleiner ist als der Schrankenwert. Dann wird der *floate*r zum *anchor* und der Linienendpunkt wieder *floate*r. Das Verfahren wiederholt sich und die neue Linie wird durch die ermittelten *anchor*-Punkte gebildet. Bei diesem Verfahren wird die Liniencharakteristik vollständig berücksichtigt, da bei der Auswahl der beizubehaltenden Punkte alle relevanten Stützpunkte in die Entscheidung miteinbezogen werden.

Eine rechentechnisch günstigere und damit schnellere Variante speichert alle auftretenden *floate*r (also Punkte, deren Abstand von der Sehne durch Linienanfangs- und Linienendpunkt größer als der Schrankenwert ist) in einer Liste. Ist ein neuer *anchor* bestimmt worden, wird der neue *floate*r von der Spitze der Liste genommen. Dadurch müssen nicht erneut alle Punkte zwischen *Floate*r und Linienendpunkt geprüft werden. Speichertechnisch ist diese Variante allerdings aufwendiger und die Anzahl der beizubehaltenden Punkt ist meist etwas höher, damit der Reduktionsgrad etwas geringer.

Algorithmisierung des Douglas-Peucker-Verfahrens, Version 2:

```

T=10 ; z. B. Toleranzabstand = 10
ANZ=Anzahl der Stützpunkte d. ursprünglichen Linie
KOORD(ANZ)= Koordinatenpaar d. Stützpunktes
KOORD_NEU=KOORD(1)
A=1
F=ANZ
N=F-1
ANCHOR=KOORD(A)
FLOATER=KOORD(F)
FLOATER_LISTE={} ; leer, aber vorbelegt

while FLOATER_LISTE
{

```

```

function (PKT_MAX)

if MAX > T
then FLOATER_LISTE append FLOATER_LISTE (FLOATER, NR_MAX)
else
  if MAX = -1
  then KOORD_NEU append Koord_NEU FLOATER
      FLOATER_LISTE nil

  if MAX = -2 or MAX > 0
  then KOORD_NEU append FLOATER_LISTE FLOATER
      ANCHOR=FLOATER
      A=F
      FLOATER_NR=firstElement FLOATER_LISTE
      FLOATER_LISTE=deleteFirstElement(FLOATER_LISTE)
      FLOATER=firstElement (FLOATER_NR)
      F=secondElement (FLOATER_NR)
      MAX=0
      NR_MAX=0

}

function PKT_MAX

if (F-A) <= 1
then
  if A = ANZ
  then MAX = -1
  else MAX = -2

LINIE=(ANCHOR, FLOATER)
A=A+1

while (F-A) > 1
{
  ABSTAND=(KOORD(A), LINIE)

  if ABSTAND > MAX
  then NR_MAX=A
      MAX=ABSTAND
  A=A+1
}

```

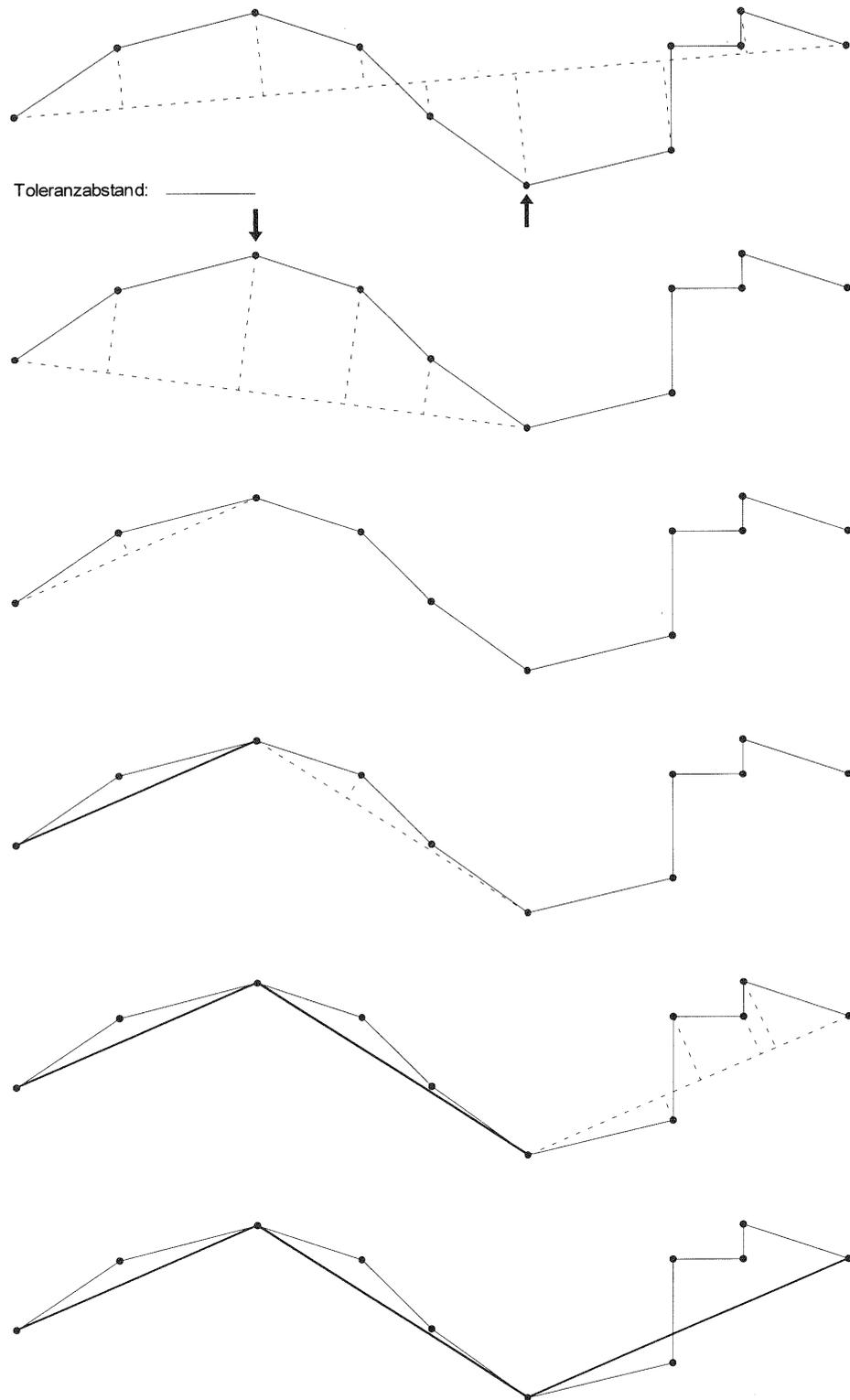


Abbildung 3.5: Douglas-Peucker-Verfahren, Version 2 [Heubach, 1992]

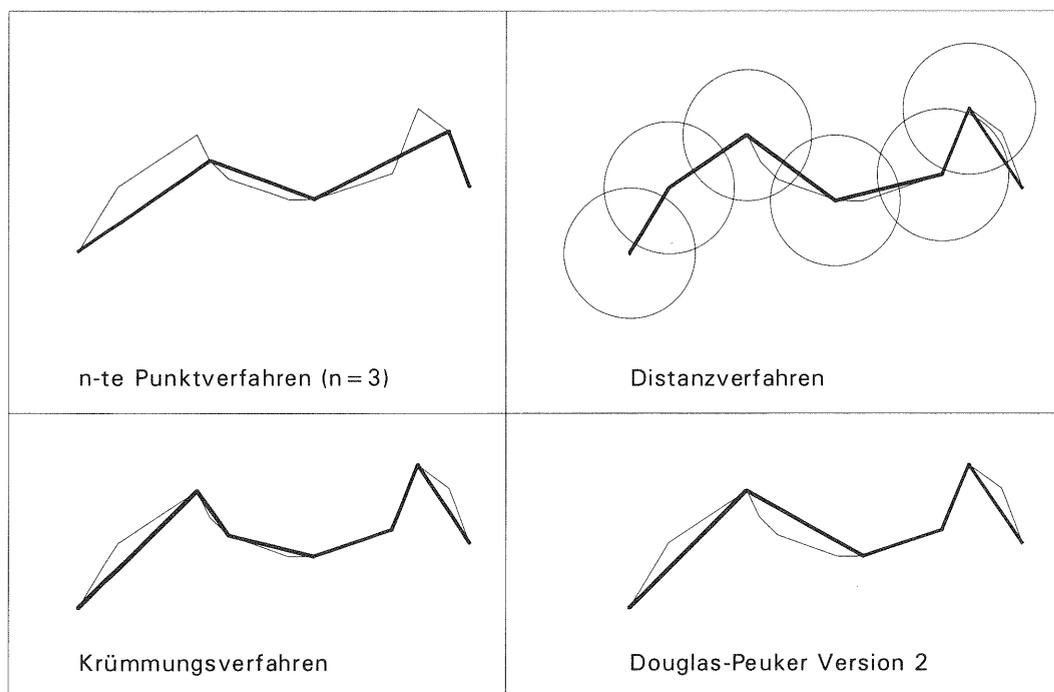


Abbildung 3.6: Visueller Vergleich der Ausdünnungsverfahren

Folgende Qualitätskriterien können allgemein für das Ergebnis einer Generalisierung herangezogen werden:

- die geometrische Genauigkeit
- die ästhetische Genauigkeit
- die topologische Richtigkeit

In diesem Fall spielt die ästhetische Genauigkeit keine Rolle, da die Generalisierung hier nur als Zwischeninformation benötigt wird, nicht aber für eine spätere graphische Darstellung.

Um diese Algorithmen und auch alle weiteren Methoden sinnvoll einsetzen zu können, muß die Genauigkeit der Geobasisdaten modelliert werden. In der geodätischen Praxis werden für solche Verfahren häufig Werte im 2σ - oder 3σ -Bereich verwendet. Damit liegen statistisch betrachtet 95,4% bzw. 99,7% aller Fälle im gewählten Bereich.

3.1.3 Grundsätzliche Überlegungen zur Modellierung der Genauigkeit

Zur Bestimmung von sinnvollen und auch statistisch haltbaren Schrankenwerten dienen die folgenden Überlegungen.

Alle Objektbereiche wurden in ATKIS-Bayern separat und nach unterschiedlichsten Methoden erfaßt. Damit ist auch die Genauigkeit für alle Bereiche zunächst getrennt zu ermitteln. Zwei wesentliche Faktoren haben Einfluß auf die Genauigkeit der Daten:

- mit welcher Genauigkeit lagen die Daten in der Erfassungsgrundlage vor?
- wie fand die Datenerfassung statt; spielen persönliche Fehler des Bearbeiters in diesem Verfahren eine Rolle?

Wie sind die Erwartungen beim ATKIS-Bayern-Datenbestand?

Die Objektbereiche *Verkehr* und *Gewässer* haben in der TK 25 höchste Priorität bei der Generalisierung, somit die höchste Genauigkeit in der Erfassungsgrundlage, gefolgt vom Objektbereich *Vegetation*. Mit geringsten Prioritäten rangieren die Bereiche *Siedlungsgrenzen* und *Gebiete*.

Da die Erfassung bei den Siedlungsgrenzen und den Verkehrswegen manuell erfolgt, sind hier noch persönliche Fehler der Bearbeiter zu erwarten. Bei der *Vegetation*, die vollautomatisch digitalisiert wird, kommen keine persönlichen Fehler vor. Bei der halbautomatischen Digitalisierung ist der Einfluß persönlicher Fehler gering, da hier lediglich Startpunkte manuell eingegeben werden und bei Lücken innerhalb der automatischen Digitalisierung nur kurze Stücke

manuell überbrückt werden müssen. Damit kann auch im Objektbereich Gewässer der persönliche Fehler vernachlässigt werden.

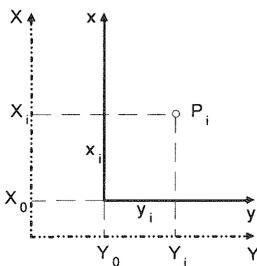
Genauigkeitsunterschiede in der Erfassungsgrundlage, der TK 25, sind nur dort vorhanden, wo mehrere Informationen in engem räumlichen Abstand zueinander dargestellt werden müssen. Doch genau diese Bereiche sind es, in denen anschließend redundante Datenbestände vorhanden sind. Zur Verdeutlichung: Grenz an den Waldrand ein Siedlungsgebiet und werden diese lediglich durch eine untergeordnete Straße voneinander getrennt, so findet für die Darstellung in der TK 25 eine Generalisierung statt. Wäre nur ein Waldgebiet darzustellen, so bräuchte keine Generalisierung für eine übersichtliche Darstellung vorgenommen werden. Damit wird deutlich, daß genau in den Bereichen, in denen redundante Datenbestände in ATKIS-Bayern vorhanden sind, auch Generalisierungseffekte eine Rolle spielen.

Aufgrund dieser Voraussetzungen ist davon auszugehen, daß die Genauigkeit des Objektbereichs Verkehr niedriger ist als die des Objektbereichs Gewässer, und zwar im wesentlichen um den Betrag des persönlichen Fehlers des Bearbeiters. Bei der Generalisierung genießen die Gewässer und Trigonometrischen Punkte gegenüber dem Verkehrsnetz zwar Vorrang, jedoch ist die Zahl der Straßen, die direkt neben einem Gewässer verlaufen gegenüber der Gesamtzahl gering und bei Trigonometrischen Punkten ist der Generalisierungseffekt auf einen sehr kleinen Bereich des gesamten Straßenverlaufs begrenzt.

Für die Modellierung der Genauigkeit werden identische Paßpunkte aus Referenzdaten bestimmt. Damit liegen sowohl im Referenzsystem als auch bei den ATKIS-Bayern-Daten zweidimensionale, kartesische Koordinaten vor. Es bietet sich für die Genauigkeitsuntersuchung eine ebene Transformation an.

Eine ebene Transformation hat maximal sechs Parameter:

- zwei Translationen eine in y- und eine in x-Richtung:

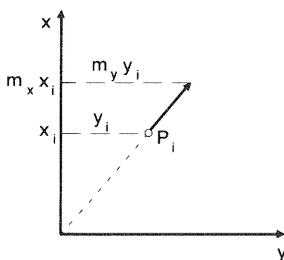


$$X_i = X_0 + x_i$$

$$Y_i = Y_0 + y_i$$

Abbildung 3.7: Translationen

- zwei Skalierungen, eine in y- und eine in x-Richtung



$$x_i \rightarrow m_x \cdot x_i$$

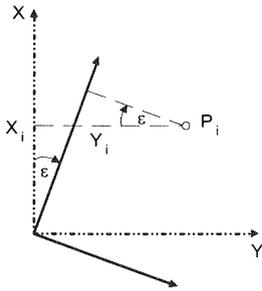
$$y_i \rightarrow m_y \cdot y_i$$

m_x = Maßstab der x-Koordinate

m_y = Maßstab der y-Koordinate

Abbildung 3.8: Maßstabsfaktoren

– eine Rotation

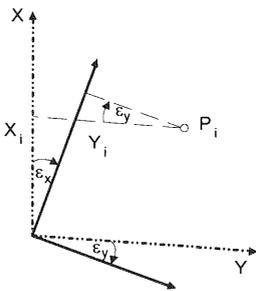


$$X_i = x_i \cdot \cos \varepsilon - y_i \sin \varepsilon$$

$$Y_i = x_i \cdot \sin \varepsilon + y_i \cos \varepsilon$$

Abbildung 3.9: Rotation

– eine Scherung



$$X_i = x_i \cdot \cos \varepsilon_x - y_i \sin \varepsilon_y$$

$$Y_i = x_i \cdot \sin \varepsilon_x + y_i \cos \varepsilon_y$$

ε_x = Winkel zwischen den x-Achsen

ε_y = Winkel zwischen den y-Achsen

Abbildung 3.10: Scherung

Die Gesamtformel lautet damit:

$$\begin{aligned} \overline{X}_i &= X_0 + \cos \varepsilon_x \cdot m_x \cdot x_i - \sin \varepsilon_y \cdot m_y \cdot y_i \\ \overline{Y}_i &= Y_0 + \sin \varepsilon_x \cdot m_x \cdot x_i + \cos \varepsilon_y \cdot m_y \cdot y_i \end{aligned}$$

Es kommen in Betracht:

- die Affintransformation mit 6 Freiheitsgraden (2 Verschiebungen, 2 Rotationen, 2 Maßstabsfaktoren) oder
- die Helmertransformtion mit 4 Freiheitsgraden (2 Verschiebungen, 1 Rotation und 1 Maßstabsfaktor).

Die **Affintransformation** enthält alle in einer ebenen Transformation möglichen Parameter. Häufig verwendet man in obiger Gesamtformel folgende Hilfsgrößen:

$$a_x = \cos \varepsilon_x m_x$$

$$a_y = \cos \varepsilon_y m_y$$

$$o_x = \sin \varepsilon_x m_x$$

$$o_y = \sin \varepsilon_y m_y$$

und damit:

$$m_x^2 = a_x^2 + o_x^2$$

$$m_y^2 = a_y^2 + o_y^2$$

$$\varepsilon_x = \arctan \left(\frac{o_x}{a_x} \right)$$

$$\varepsilon_y = \arctan \left(\frac{o_y}{a_y} \right)$$

Damit ergibt sich als Gesamtformel (Affintransformation):

$$\overline{X}_i = X_0 + a_x \cdot x_i - o_y \cdot y_i$$

$$\overline{Y}_i = Y_0 + o_x \cdot x_i + a_y \cdot y_i$$

Durch diese Ersetzung werden die nichtlinearen Winkelfunktionen ausgeschaltet. Die Gleichung für eine Ausgleichung (Parameterschätzung) wird damit linear.

Die **Helmertransformation** erhält man, indem $\varepsilon = \varepsilon_x = \varepsilon_y$ und $m = m_x = m_y$:

Damit ist:

$$\begin{aligned} a &= \cos \varepsilon \cdot m & o &= \sin \varepsilon \cdot m \\ \varepsilon &= \arctan \frac{o}{a} & m^2 &= a^2 + o^2 \end{aligned}$$

Die lineare Formel für die Helmertransformation lautet somit:

$$\begin{aligned} \overline{X}_i &= X_0 + a \cdot x_i - o \cdot y_i \\ \overline{Y}_i &= Y_0 + o \cdot x_i + a \cdot y_i \end{aligned}$$

Bei Überbestimmung ist eine Ausgleichung (Parameterschätzung) durchzuführen.

Je weniger Freiheitsgrade eine Transformation hat, um so starrer ist sie, d. h. um so weniger können Ungenauigkeiten durch die Transformationsparameter verwischt werden; es sind dennoch keine gravierenden Unterschiede zwischen beiden Transformationsarten zu erwarten (siehe Ergebnis in Kapitel 4.2.1.1). Bei diesen Transformationen geht es nicht darum, Punktfehler aufzudecken oder eine echte Transformation durchzuführen, sondern die Punktgenauigkeit zu analysieren und damit ein Maß für die mittlere Genauigkeit der Punkte zu ermitteln. Bei der Auswahl der Punkte ist darauf zu achten, daß diese möglichst gleichmäßig über das Testgebiet verteilt und eindeutig zu identifizieren sind. Als Ergebnis der jeweiligen Transformation erhält man die ausgeglichenen Parameter der jeweiligen Transformationsgleichungen und deren Varianzen. Daraus abgeleitet ergeben sich die Varianzen der Koordinatenwerte, also die „Genauigkeit der Punkte“. Die Ergebnisse der Affintransformation werden geringfügig besser ausfallen, da sie zwei Freiheitsgrade mehr besitzt.

3.2 Detaillierte Beschreibung des hier verwendeten Verfahrens

Es gilt ein Verfahren zu entwickeln, mit dem

- automatisch Redundanzen im Datenbestand aufgedeckt werden, die auf Grund der isolierten Datenerfassung nach dem Ebenenprinzip entstanden sind,
- diese zu analysieren und anschließend
- sofern möglich, die Redundanzen zu beseitigen.

Im wesentlichen gliedert sich ein solches Verfahren in vier Schritte, die anschließend detailliert beschrieben werden:

- die Genauigkeit zu modellieren,
- die Inkonsistenzen im Geodatenbestand aufzusuchen und zu selektieren,
- die Inkonsistenzen im Geodatenbestand zu bereinigen sowie
- die Kontrolle über die durchgeführten Operationen durchzuführen.

3.2.1 Modellierung der Genauigkeit

Als Maß für die Genauigkeit können verschiedene Vergleichsdaten herangezogen werden. Dafür eignet sich die Digitale Flurkarte (DFK) bzw. das digitale Orthophoto. Beide Medien liegen für Bayern noch nicht flächendeckend vor. Derzeit wird die DFK forciert in bebauten Gebieten hergestellt, während sie in den übrigen Bereichen sukzessive bei jeder Auftragsvermessung durch die Vermessungsämter erstellt wird. Flächendeckend gibt es die DFK bereits für einige Gemeinden und kleinere Städte. Daher können aus der DFK auch nur Vergleichsdaten von den Objektarten verwendet werden, die innerhalb einer Ortschaft vorkommen und in der DFK und ATKIS gleichermaßen verifizierbar sind. Die Siedlungsflächen sind in einer DFK nicht eindeutig zu erkennen, da eine Flurstücksgrenze nicht mit einer Bebauungsgrenze zusammenfallen muß. Auch eine Unterteilung zwischen Wohn- und Mischbebauung und Industrieflächen ist nicht unbedingt eindeutig zu erkennen und derzeit auch in ATKIS noch nicht erfaßt. Daher eignet sich innerorts vor allem der Objektbereich „Verkehr“, da der Verlauf von Straßen auch in der DFK eindeutig erkennbar ist. Diese Untersuchung läßt daher nur eine Aussage für einen einzigen Objektbereich zu, und dies auch nur innerorts. Dies ist deswegen gesondert zu erwähnen, da innerorts Generalisierungseffekte aufgrund der Dichte der Informationen

verstärkt auftreten. Zwar sind Straßen bei der Generalisierung mit einem hohen Gewicht versehen, doch muß im folgenden noch untersucht werden, ob das Ergebnis ohne weiteres auch auf den Außenbereich übertragen werden kann.

Das Orthophoto bietet eine ideale Grundlage für die Beurteilung der Genauigkeit, da das Orthophoto, ebenso wie das DLM in ATKIS, ein Abbild der Natur ist, nur in einer anderen Form. In einem hybriden GIS bzw. CAD-System können beide „Abbilder der Natur“ ohne großen Aufwand überlagert und identische Punkte gemessen werden. Jedoch kann das Orthophoto nicht mehr ohne weiteres als fehlerfrei betrachtet werden. Hier spielen verschiedene Faktoren eine Rolle. Zum einen hat jedes Pixel eine gewisse Größe in der Natur, vergleichbar mit der Korngröße bei herkömmlichem Filmmaterial. Dadurch wird die Auflösung eingeschränkt. Zum anderen benötigt man zur Erstellung eines Orthophotos ein Digitales Geländemodell (DGM), welches seinerseits nicht fehlerfrei ist. Besonders kritische Bereiche sind Gebäude- bzw. Waldflächen. Die Gebäudehöhe ist im DGM nicht modelliert, so daß sich hierdurch, abhängig von der Gebäudehöhe, ein Versatz zwischen Natur und Abbild im Orthophoto ergibt. Bei bewaldeten Flächen ist die Beurteilung schwierig, da hier nicht sichergestellt ist, wo die Geländehöhe für das DGM bestimmt wurde, an der Waldkrone, am Waldboden oder dazwischen. Daher muß im Waldbereich mit Verschiebungen zwischen Natur und Abbild gerechnet werden, deren Größe kaum zu beziffern ist.

Die maximale Erfassungsgenauigkeit hängt von verschiedenen Faktoren ab:

- der Zielgenauigkeit für die zu erfassende Karte:
davon wird der Operateur beeinflusst (manchmal auch unbewußt). Wenn man weiß, für welchen Verwendungszweck die Erfassung erfolgt, so wird man je nach Ziel detaillierter bzw. großzügiger erfassen. Für ATKIS-Bayern wird zwar aus der TK 25 heraus wieder für den gleichen Maßstabsbereich erfaßt, jedoch stufenweise, so daß nicht alle Informationen in einem Guß überführt werden müssen bzw. andere gar nicht enthalten sind, die aus anderen Datenquellen später einzufügen sind. Auch dieses Wissen beeinflusst den Operateur.
- der persönliche Einstellfehler:
jeder Operateur hat seine persönlichen Eigenarten, wenn er digitalisiert.
- die Reproduzierbarkeit des Digitizers.

Als Faustformel wird häufig angegeben:

$$\begin{array}{ll} \text{Zeichenfehler in der Karte:} & \sigma = 0,1 \text{ [mm]} \\ \text{Digitalisierfehler des Digitizers:} & \frac{\sigma = 0,05 \text{ [mm]} - 0,1 \text{ [mm]}}{\text{max. } 0,2 \text{ [mm]}} \end{array}$$

[Rappe 1995] gibt eine Formel für die maximal erreichbare Erfassungsgenauigkeit ds bei einer Erfassung von topographischen Karten an:

$$ds = 2 \cdot 10^{-4} \cdot m = \frac{m}{5000}$$

wobei m die Maßstabszahl der Kartenvorlage ist. Das heißt, daß bei der Erfassung der Daten aus einer TK 25 mit einer maximalen Erfassungsgenauigkeit von 5 Metern auszugehen ist.

Die oben aufgeführten Abschätzungen sind Durchschnittswerte. Da die Erfassungsmethoden sowohl für die analoge TK 25 als auch in ATKIS-Bayern für die einzelnen Objektbereiche sehr unterschiedlich sind, sind detailliertere Untersuchungen erforderlich. Diese werden getrennt nach Objektbereichen vorgenommen.

3.2.1.1 Vergleich zwischen ATKIS-Bayern und DFK (innerorts)

Für die Genauigkeitsuntersuchung innerorts wurden numerisch angesetzte Flurkarten mit den ATKIS-Daten überlagert. Die Daten aus der DFK können für diese Betrachtungen als fehlerfrei angesehen werden, sie haben Katastergenauigkeit (± 3 cm). Im folgenden Bild ist ein Ausschnitt aus dieser Überlagerung zu sehen.



Abbildung 3.11: Überlagerung DFK mit ATKIS

Wie bereits oben erwähnt, sind nur die Daten aus dem Objektbereich „Verkehr“ für diese Untersuchung geeignet. Dabei ist zu berücksichtigen, daß in der DFK nur die rechtlichen Besitzgrenzen dargestellt sind, nicht aber die dazugehörige Topographie. Im Bereich der Straßen stimmen die Besitzgrenzen mit der Topographie jedoch meistens gut überein. Visuell wurden für den Vergleich die Knotenpunkte an Straßen in der DFK bestimmt. Eine streng geometrische Konstruktion der Knotenpunkte erscheint nicht als sinnvoll und notwendig, wie im folgenden Bild zu sehen ist. Die optisch bestimmten Punkte können dennoch als fehlerfrei betrachtet werden, da zum einen keine andere genauere Grundlage als die DFK vorhanden und zum anderen eine streng geometrische Konstruktion sehr aufwendig und nicht immer eindeutig ist. Damit bringt sie keine Vorteile.

3.2.1.2 Vergleich zwischen ATKIS-Bayern und Orthophoto

Die Orthophotos wurden für die in dieser Arbeit angestellten Untersuchungen freundlicherweise vom Bayerischen Landesvermessungsamt zur Verfügung gestellt. Eine TK 25 ist aufgeteilt in neun Orthophotos, die jeweils eine Fläche von etwa von 4,5 km \times 4,0 km in der Natur abdecken. Die Größe eines Pixels und damit die Auflösung beträgt in der Natur 0,83 m. In einem ersten Schritt müssen die Orthophotos georeferenziert werden. Dies bedeutet, daß den Rasterdaten die Lagerung mitgegeben werden muß, d. h. an welchem Koordinatenwert soll es in welcher Größe positioniert werden. Diese Informationen liegen zunächst in einem separaten Datenfile vor und müssen mit den Rasterdaten verknüpft werden.

Die Methode für die Genauigkeitsuntersuchung erfolgt wie unter 3.2.1 beschrieben, ebenfalls über Helmertransformationen. Dazu werden aus den Orthophotos eindeutig verifizierbare Punkte als Referenzpunkte bestimmt. Dafür kommen in Frage:

- im Objektbereich *Vegetation* z. B. eindeutig identifizierbare Waldecken, da sonst keine weiteren Objektarten erfaßt sind.
- im Objektbereich *Gewässer* insbesondere Knotenpunkte, an denen zwei Gewässer ineinander fließen, da diese zweifelsfrei zu erkennen sind, in einigen Fällen auch Punkte, an denen Verschneidungen mit anderen Objektbereichen auftreten. Dabei muß jedoch beachtet werden, daß ausschließlich die Genauigkeit des Objektbereichs Gewässer modelliert wird und nicht auch die des anderen Objektbereichs (i. d. R. Verkehr), mit dem sich das Gewässer schneidet (Abbildung 3.13). In eindeutigen Fällen sind außerdem markante Biegungen bzw. Knickpunkte von Gewässern geeignet.
- im Objektbereich *Verkehr*, wie vorher, an Straßenknoten. Auch hier eignen sich mit Einschränkungen Verschneidungen mit anderen Objektarten (i.d.R. Straße). Nicht geeignet ist die Objektart Wald, da bei der Verschneidung mit

Straßen immer zwei Waldbegrenzungslinien in ATKIS-Bayern vorhanden sind (Abbildung 3.14). Der Schnittpunkt muß umständlich aus nicht als fehlerfrei zu betrachtenden Daten visuell bestimmt werden. Gleiches gilt auch für die Verschneidung Wald-Gewässer.

- im Objektbereich *Siedlung*, ist nicht für eine Genauigkeitsuntersuchung geeignet, da sich auch im Orthophoto die Begrenzungen für bebauete Gebiete nicht eindeutig bestimmen lassen.

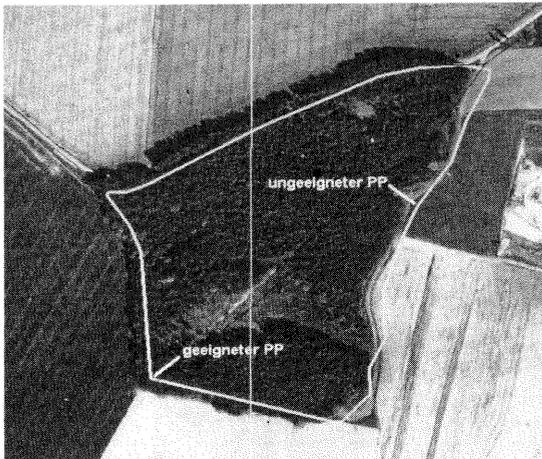


Abbildung 3.12:
geeigneter und ungeeigneter Paßpunkt im Objektbereich Vegetation

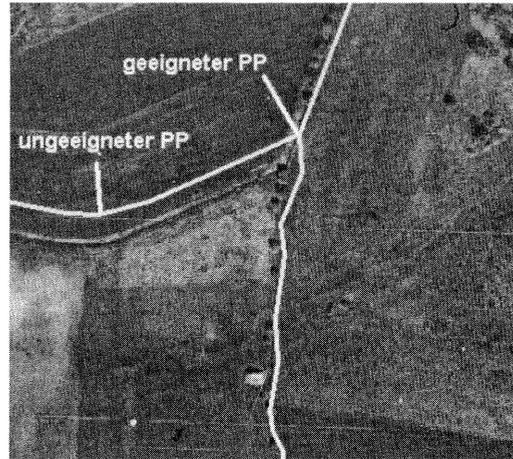


Abbildung 3.13:
geeigneter und ungeeigneter Paßpunkt im Objektbereich Gewässer

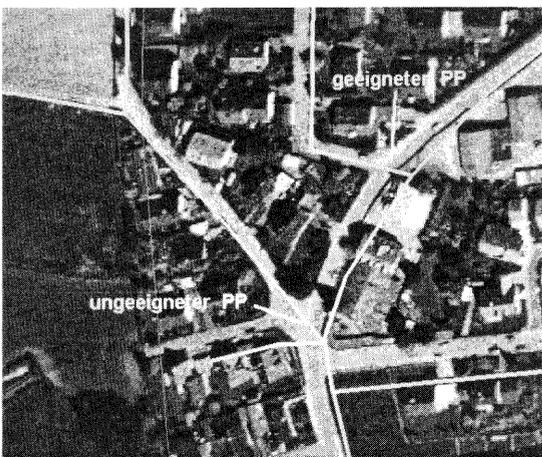


Abbildung 3.14:
geeigneter und ungeeigneter Paßpunkt im Objektbereich Verkehr

Anschließend wird die Punktgenauigkeit ebenfalls über eine Transformation bestimmt und daraus die Genauigkeit der Punkte abgeleitet, und zwar nach Objektbereichen getrennt. Die gleichmäßige Verteilung der identischen Punkte ist im Objektbereich Verkehr relativ leicht zu realisieren, da die meisten Straßenknotenpunkte im Orthophoto eindeutig zu erkennen und zu messen sind. Objekte der Objektbereiche Vegetation und Gewässer sind naturgemäß nicht gleichmäßig verteilt und vor allen Dingen oft auch nicht eindeutig im Orthophoto anmeßbar. Dennoch wurde bei der Auswahl der identischen Punkte auf eine bestmögliche Verteilung geachtet.

3.2.2 Inkonsistenzen innerhalb des Objektbereichs Wald aufsuchen, selektieren und bereinigen

Das Aufsuchen, Selektieren und Bereinigen des Geodatenbestandes wird in einem Guß behandelt, da für die verschiedenen auftretenden Fälle je ein Programm vorgestellt wird, das alle drei Aufgaben erledigt.

3.2.2.1 Besonderheiten des Objektbereichs Wald

1. Aufgrund der vollautomatischen, blattweisen Erfassung liegt dieser Objektbereich nicht blattschnittfrei vor. Dies bedeutet, wenn eine Waldfläche durch den Blattschnitt der TK 25-Blätter getrennt wird, so findet sich diese Trennlinie auch im ATKIS-Bayern - Datenbestand.
2. Die Umfangspolygone der Waldpolygone weisen unterschiedliche Drehrichtungen auf, je nach Eigenschaft des Waldpolygons. Dies wird im folgenden noch von wesentlicher Bedeutung sein.

linksdrehend sind: Waldpolygone am Blattschnitt und Inselpolygone, d. h. Polygone, die waldfreie Flächen innerhalb eines Waldpolygons darstellen.

rechtsdrehend sind: alle übrigen „normalen“ Waldpolygone ohne Besonderheiten.

3. Waldpolygone, durch die eine Straße läuft, werden in zwei Polygone getrennt (vgl. Abbildung 1.8).

3.2.2.2 Bestimmung der Drehrichtung eines geschlossenen Polygons

Werden, wie von der Polygonzugsberechnung bekannt, alle in Laufrichtung links liegenden Brechungswinkel β des geschlossenen Polygons berechnet und aufaddiert, so ergibt sich:

für ein linksdrehendes Polygon:
$$\sum_{i=1}^n \beta_i = (n - 2) \cdot 200 \text{ gon} \quad \text{und}$$

für ein rechtsdrehendes Polygon:
$$\sum_{i=1}^n \beta_i = (n + 2) \cdot 200 \text{ gon},$$

wobei n die Anzahl der Stützpunkte des Polygons ist.

Auf diese Weise werden für ein linksdrehendes Polygon die Innenwinkel und für ein rechtsdrehendes Polygon die Außenwinkel berechnet.

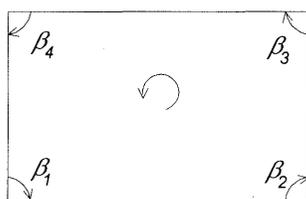
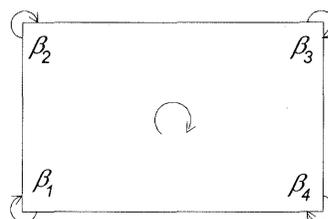


Abbildung 3.15:
a) linksdrehendes Polygon



b) rechtsdrehendes Polygon

3.2.2.3 Waldpolygone am Blattschnitt der TK 25

Da der Objektbereich, wie bereits erläutert, aus den Deckfolien der TK 25 über automatische Raster-Vektor-Konvertierung gewonnen wird, ist bei den Waldpolygonen der Blattschnitt noch erkennbar.

Die folgende Abbildung zeigt die Situation:

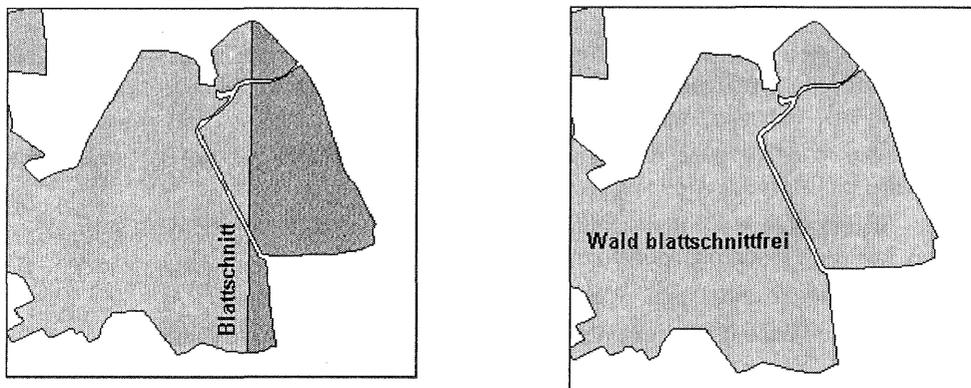


Abbildung 3.16:

a) original ATKIS-Bayern - Daten

b) bereinigte ATKIS-Bayern - Daten

Die Daten sind insofern konsistent, als die Polygonlinien der beiden benachbarten Waldpolygone entlang des Blattschnitts jeweils nur über zwei Stützpunkte repräsentiert werden. Zudem stimmen die Koordinatenwerte dieser beiden Stützpunkte in beiden Polygonen exakt überein.

Damit kann für die Vereinigung dieser beiden Waldobjekte zu einem gemeinsamen Waldobjekt eine sehr einfache Vorgehensweise angewendet werden:

1. Waldpolygon selektieren und prüfen, ob es linksdrehend ist.
2. Wenn ja, wird eine Liste der Koordinaten der Polygonstützpunkte aufgestellt. Anschließend wird um jeden Stützpunkt ein winziger Fangkreis mit nur **einem** Zentimeter Radius gelegt. Dieser minimale Fangkreis ist ausreichend, da für den Fall, daß es sich um einen Polygonstützpunkt am Blattschnitt handelt, noch ein zweiter Polygonstützpunkt auf exakt gleichem Koordinatenwert vorhanden ist. Dieser wird bereits mit einem verschwindend geringen Fangkreis zuverlässig selektiert.¹
3. Wird bei zwei aufeinanderfolgenden Stützpunkten innerhalb dieser Fangkreise das gleiche benachbarte Waldpolygon gefunden, welches ebenfalls diese beiden identischen Stützpunkte aufweist, dann wird
4. aus den beiden Stützpunktlisten ein neues nunmehr rechtsdrehendes Waldpolygon generiert.

Das Ergebnis dieser Prozedur ist rechts in obiger Abbildung dargestellt.

3.2.2.4 Sackgassen in Waldpolygonen - Einschnitte beseitigen

Da die Waldpolygone aus dem grünen Flächendecker der TK 25 generiert wurden, werden Waldflächen, in denen Straßen verlaufen, in zwei oder mehrere Waldpolygone aufgespalten. Dies geschieht durch die notwendige verbreiterte Darstellung der Straßen in der TK 25. Diese Flächen zusammenzuführen wird Thema des folgenden Kapitels sein. Doch zuvor müssen die „Sackgassen“ aus den Datensätzen beseitigt werden.

Wie kommt es zur „Sackgassenbildung“?

Die „Sackgassen“ entstehen immer dann, wenn eine untergeordnete, jedoch in der TK 25 noch doppellinig darzustellende Straße, in den Wald hinein verläuft. Damit wird an dieser Stelle das Waldpolygon aufgetrennt. Wird diese Straße im weiteren Verlauf, innerhalb des Waldpolygons, aber zu einem Feldweg, der auch in der TK 25 nur einlinig darzustellen ist, so bildet sich im Waldpolygon eine „Sackgasse“.

¹ Je nach Datenstruktur des verwendeten GIS- bzw. CAD-Produktes kann der betreffende Punkt auch direkt auf Identität überprüft werden.

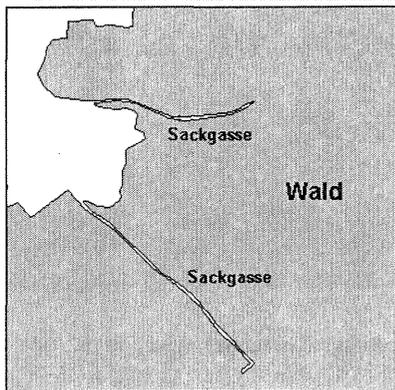
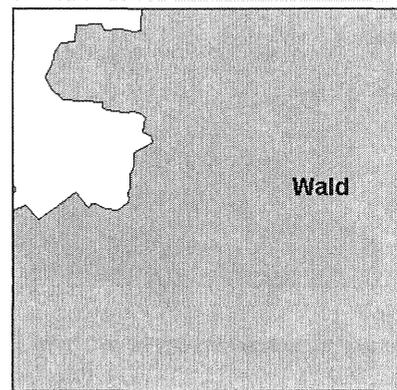


Abbildung 3.17:
a) originale ATKIS-Bayern – Daten



b) bereinigte ATKIS-Bayern - Daten

Die Vorgehensweise zur Eliminierung der Sackgassen bedarf mehrerer Schritte. Der Drehsinn des Waldpolygons spielt in diesem Fall keine Rolle, und wird daher auch nicht untersucht, da Sackgassenbildung auf alle Fälle zu beheben sind, gleichgültig, ob es sich hierbei um ein Polygon am Blattschnitt, ein Inselpolygon oder ein gewöhnliches Waldpolygon handelt.

Anwendung des Douglas-Peucker-Algorithmus zur Generalisierung:

Hiermit wird die Anzahl der Stützpunkte des Waldpolygons reduziert. Wichtig ist es, den Schrankenwert für die Generalisierung geeignet zu wählen. Dabei ist zu berücksichtigen:

1. Die Darstellungsbreite der Straße (in der TK 25), die zur Sackgassenbildung geführt hat, da dies einen wesentlichen Einfluß auf
2. das gewünschte Ergebnis nach der Generalisierung hat.

Ziel dieser Generalisierung ist es, die Sackgasse durch einerseits möglichst parallele, andererseits jedoch möglichst lange Linienstücke zu repräsentieren, die sich am Ende der Sackgasse spitzwinklig treffen, wie die folgende Abbildung zeigt.

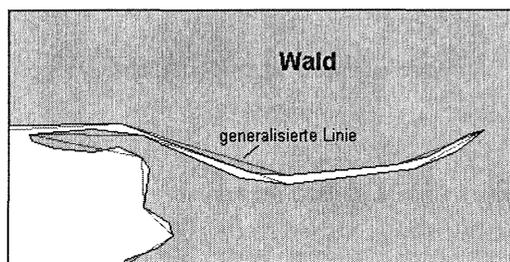
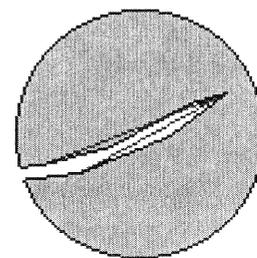


Abbildung 3.18:
a) Schrankenwert 10 m



b) Vergrößerung

Die Mindestdimension der kleinsten doppellinig darzustellenden Straße in der TK 25 ist 0,4 Millimeter; dies entspricht zehn Metern in der Natur. Alle anderen breiteren Straßen müssen nicht berücksichtigt werden, da sie praktisch nie als Sackgassen im Wald enden werden. Wird der Schrankenwert (Schwellwert) zu groß gewählt, so sind die generalisierten Linienstücke nicht mehr parallel; wird er zu klein gewählt, ist der Generalisierungseffekt minimal, die Linienstücke werden kurz und sind nicht mehr parallel zueinander.

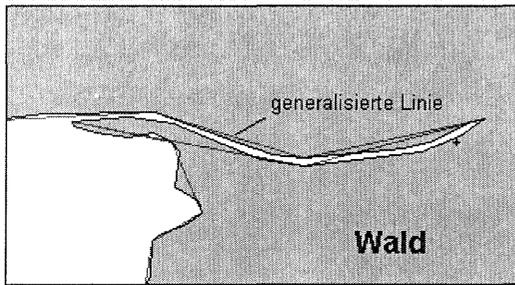
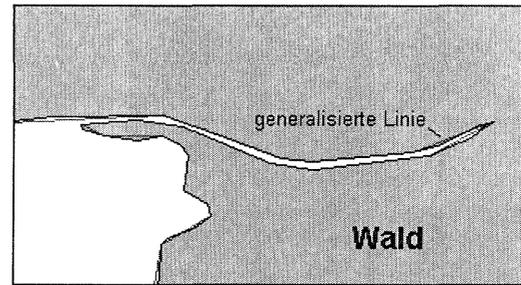


Abbildung 3.19:
a) Generalisierung mit Schrankenwert 20 m



b) Generalisierung mit Schrankenwert 5 m²

Als sinnvoller Schrankenwert hat sich die Straßenbreite von 10 m erwiesen. Damit verläuft die generalisierte Linie noch mit akzeptabler Lagerichtigkeit. Es ist immer noch eindeutig erkennbar, welches originale Linienstück dadurch wiedergegeben werden soll. Die maximale Verschiebung der generalisierten Linie beträgt damit 10 m, d. h. also maximal eine Straßenbreite. Die absolute Genauigkeit der Waldpolygone spielt für die Wahl des Schrankenwertes hier keine Rolle, da es in diesem Fall nur auf die relative Lage und die Genauigkeit in sich ankommt.

Selektieren einer Sackgasse aus der Information der generalisierten Linie:

An jedem Stützpunkt der generalisierten Linie wird geprüft, ob sich der Richtungswinkel, der aus den jeweils aufeinanderfolgenden Stützpunkten berechnet wird, sich um etwa 200 gon ändert. Ist dies der Fall, so werden die Stützpunktnummern der Punkte 1 und 3 in Laufvariablen gespeichert.

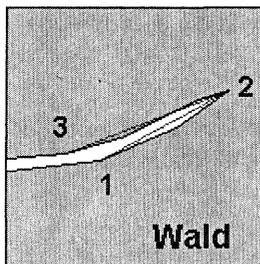
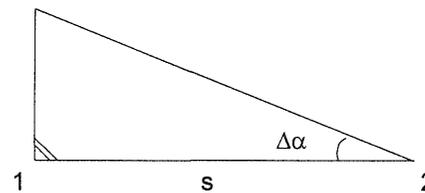


Abbildung 3.20:
a) Ende der Sackgasse



b) Entscheidungskriterium für Scheitelbildung

Was bedeutet „um etwa 200 gon“?

Als maximal zulässige Winkelabweichung $\Delta\alpha$ wird in diesem Fall sehr einfach berechnet:

$$\tan \Delta\alpha = \frac{12 \text{ [m]}}{s \text{ [m]}}$$

wobei s die Entfernung zwischen den beiden Stützpunkten ist. Dabei wird die jeweils kürzere der beiden Strecken (von 1 nach 2, bzw. von 2 nach 3) verwendet. Der Wert von 12 Metern ergibt sich wie folgt:

Die Straßenbreite beträgt durchschnittlich 10 Meter. Durch die Generalisierung ergeben sich kleinere zusätzliche Verschiebungen, so daß die generalisierten, zusammengehörigen Linienstücke auch einen etwas größeren Abstand als 10 Meter voneinander haben können. Diese Toleranz ist mit einer 20%-igen Vergrößerung auf 12 Meter aufgefangen. Zudem geht sicherheitshalber die kleinere der beiden Strecken in die Berechnung ein, auch dadurch wird die maximale Winkelabweichung etwas vergrößert.

Ist ein Stützpunkt der generalisierten Linie als Scheitelpunkt der Sackgasse identifiziert worden, werden die Stützpunktnummern der Punkte 3 und 1 gespeichert. Anschließend wird die generalisierte Linie vom Punkt 1 aus rückwärts und von Punkt 3 aus vorwärts weiter verfolgt. Auch hier werden die Richtungswinkel der weiteren Linienstücke miteinander verglichen. Lediglich bei der Berechnung der maximalen Winkelabweichung wird die Schranke von 12 Meter auf 7,5 Meter reduziert. Dies ist dadurch begründet, daß mit Ausnahme am Scheitelpunkt die zusammengehörenden, generalisierten Linienstücke weitgehend nahezu parallel verlaufen, während sie am Scheitel zusammenlaufen müssen.

² dort, wo nur eine Linie sichtbar ist, fallen generalisierte und originale Linie zusammen

Kleinere Abweichungen werden hiermit dennoch aufgefangen. Werden bei der Verfolgung der generalisierten Linie wieder zwei parallele, zusammengehörige Linienelemente gefunden, so werden jeweils die neuen Stützpunktnummern gespeichert. Dieses Verfahren funktioniert zuverlässig, da die Waldpolygone vollautomatisch digitalisiert wurden und somit keine persönlichen Fehler eines Operateurs enthalten sind. Nur so ist es möglich, daß die beiden Straßenseiten bei den Waldpolygonen mit Sackgassenbildung sehr parallel verlaufen.

Generierung des neuen Polygons ohne „Sackgasse“:

Da die gespeicherten Stützpunktnummern, die bei der Linienverfolgung ermittelt wurden, aus der generalisierten Linie hervorgegangen sind, muß jetzt bei der Generierung des neuen Polygons wieder auf die ursprüngliche Linie zurückgegangen werden. Um die neue Linie abzuleiten, wird von der kürzeren Seite das Lot auf die längere Seite gefällt, und dort der Schnittpunkt S mit der ursprünglichen Linie berechnet. Die folgende Abbildung verdeutlicht die Schnittpunktbestimmung.

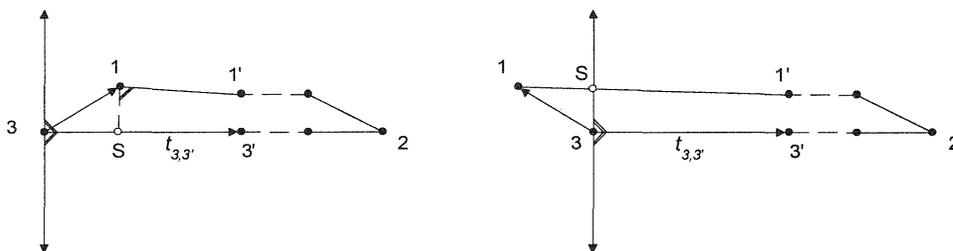


Abbildung 3.21: Richtungswinkel als Entscheidungskriterium für die Schnittpunktbestimmung

Folgende Fallunterscheidung ist zu treffen:

Der Schnittpunkt liegt auf dem Linienelement zwischen den Stützpunkten 3 und 3', wenn gilt:

$$\text{für } \left\{ (t_{3,3'} - 100) \mid 0 \leq (t_{3,3'} - 100) \leq 200 \right\} \quad \text{wenn } (t_{3,3'} - 100) \leq t_{3,1} \leq (t_{3,3'} + 100)$$

$$\text{für } \left\{ (t_{3,3'} - 100) \mid 200 < (t_{3,3'} - 100) < 400 \right\} \quad \text{wenn } t_{3,1} \mid \left\{ (t_{3,3'} - 100) < t_{3,1} < 400 \vee 0 \leq t_{3,1} \leq (t_{3,3'} - 100 \text{ gon}) \right\}$$

Der Schnittpunkt liegt auf dem Linienelement zwischen den Stützpunkten 1 und 1', wenn gilt:

$$\text{für } \left\{ (t_{3,3'} + 100) \mid 0 \leq (t_{3,3'} + 100) \leq 200 \right\} \quad \text{wenn } (t_{3,3'} + 100) \leq t_{3,1} \leq (t_{3,3'} - 100)$$

$$\text{für } \left\{ (t_{3,3'} + 100) \mid 200 < (t_{3,3'} + 100) < 400 \right\} \quad \text{wenn } t_{3,1} \mid \left\{ (t_{3,3'} + 100) < t_{3,1} < 400 \vee 0 \leq t_{3,1} \leq (t_{3,3'} + 100 \text{ gon}) \right\}$$

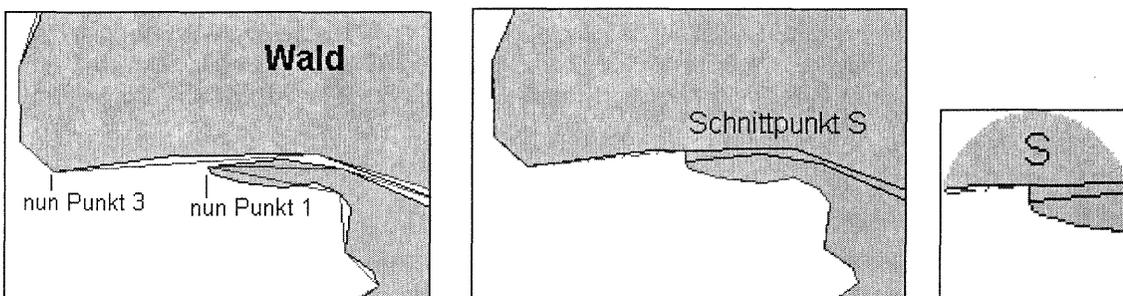


Abbildung 3.22:

a) Ende der Linienverfolgung, gespeicherte Anfangs- und Endpunkte

b) bereinigter Datenbestand, wobei das Lot auf die Originallinie gefällt wurde

c) Vergrößerung

In einigen Fällen findet bei der Generalisierung des Waldpolygons am Ende der Sackgasse keine klare Scheitelbildung statt, wie die folgende Abbildung zeigt. Diese Situation tritt dann auf, wenn kurz vor Ende der Sackgasse die Straße noch eine Biegung macht, und/oder das Ende der Sackgasse etwas ausgefranst dargestellt ist.

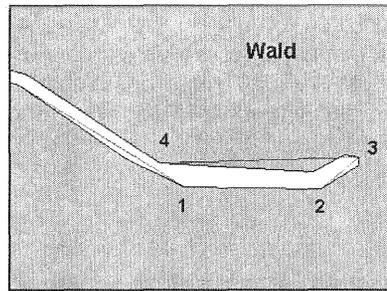


Abbildung 3.23: keine klare Scheitelbildung

Gesetzt den Fall, daß nach der ersten Methode kein Scheitel gefunden wurde, so ist ein zweiter Suchlauf zu starten, der eine Hilfslinie zwischen den Punkten 1 und 3 bildet und diese für die Scheitelbildung verwendet. Voraussetzung dafür ist, daß Punkt 2 von dieser Hilfslinie nicht weiter als 15 Meter entfernt ist. Dieser Wert von 15 Metern entspricht dem Wert einer nochmaligen Generalisierung mit den gleichen Eingangswerten wie zuvor. Ist der Abstand größer, so kann nicht mehr davon ausgegangen werden, daß es sich um ein Wegendstück handelt. Außerdem wird nur dann angenommen, daß eine Sackgassenbildung vorliegt, wenn bei der Linienverfolgung rückwärts von Punkt 1 bzw. vorwärts von Punkt 4 aus mindestens noch ein zusammengehörendes, paralleles Linienpaar gefunden wird.

3.2.2.5 Durch Straßen getrennte Waldpolygone zusammenführen

In diesem Abschnitt soll gezeigt werden, wie zwei eigentlich zusammengehörende Waldpolygone, die durch eine Straße getrennt wurden, wieder zu *einem* Waldpolygon zusammengeführt werden können, so daß ein einheitliches Waldpolygon entsteht.

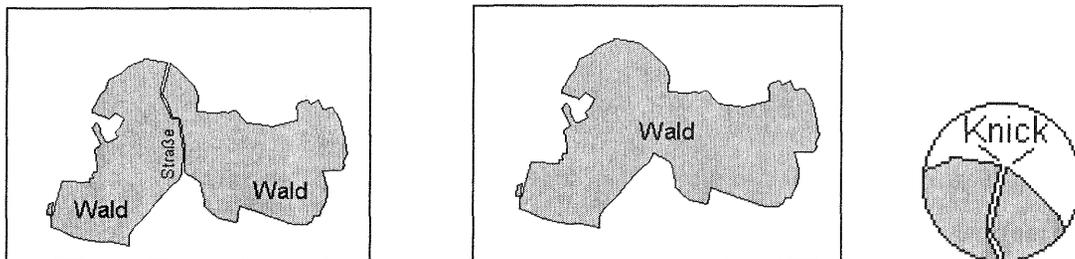


Abbildung 3.24:

a) original ATKIS-Bayern – Daten

b) bereinigte ATKIS-Bayern – Daten

c) Vergrößerung

Voraussetzung für den Einsatz dieser Prozedur ist, daß die Waldpolygone am Blattschnitt bereits bereinigt sind. Damit ist gewährleistet, daß alle Waldpolygone, die für diese Nachbearbeitung in Frage kommen, rechtsdrehend sind. Inselpolygone kommen hierfür generell nicht in Betracht. Gestartet wird die folgende Prozedur an jedem beliebigen Waldpolygon.

Anwendung des Douglas-Peucker-Algorithmus zur Generalisierung:

In dieser Prozedur wird der Douglas-Peucker-Algorithmus dazu verwendet, das Umfangspolygon der Waldfläche so zu repräsentieren, daß vor allem die wesentlichen Knickpunkte im Polygon selektiert werden können. Dies ist deshalb wichtig, da an Stellen, an denen eine Straße in ein Waldpolygon mündet, i. d. R. zumindest eines der beiden Polygone eine Richtungsänderung erfährt - einen „Knick“ hat. Der Schrankenwert ist in diesem Fall so zu wählen, daß die Kleinstruktur der Linie eliminiert wird, jedoch nur so, daß noch ausreichend viele Knickpunkte erhalten bleiben. Als sinnvoll hat sich hier der Wert von 15 Metern erwiesen.

Vorgehensweise an den Knickstellen:

Eine Knickstelle wird immer dann angenommen, wenn aufeinanderfolgende Richtungswinkel, berechnet aus immer zwei hintereinanderliegenden Stützpunkten der generalisierten Linie, sich um mehr als 20 gon ändern. Anschließend

wird um jede Knickstelle ein Fangkreis mit einem Radius von 40 Metern gelegt. Hier ist zu prüfen, ob innerhalb dieses Fangkreises sich ein weiteres, rechtsdrehendes Waldpolygon (damit wird sichergestellt, daß kein Inselpolygon gefiltert wurde) **und** außerdem ein Polygon einer Straße zu finden ist. Beide Bedingungen müssen zwingend erfüllt sein. Diese Selektion ist sehr einfach, da die ATKIS-Bayern - Daten so strukturiert sind, daß innerhalb jeder GDB (hier: Verkehr) verschiedene Ebenen vorhanden sind. Damit ist für jede Straßenkategorie eine eigene Ebene vorhanden (Bundesstraßen, Kreisstraßen, Gemeindestraßen usw.), die Selektion kann über die Ebenenstruktur erfolgen.

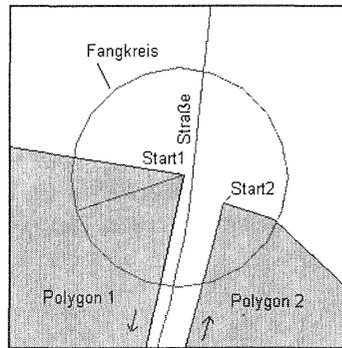


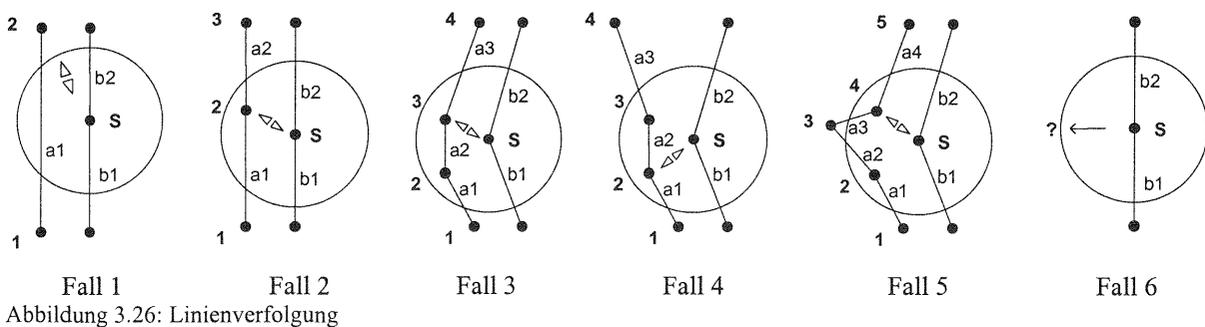
Abbildung 3.25: Fangkreis um den Knickpunkt
die Pfeile geben die Laufrichtung der Polygone an.

Anschließend ist der Stützpunkt des zweiten Polygons zu bestimmen, der am nächsten am Knickpunkt von Polygon 1 liegt - hier mit Start2 bezeichnet. Er stellt den zweiten Startpunkt für die Linienverfolgung dar. Da der Fangkreis um den Knickpunkt der generalisierten Linie gezogen wurde, ist nun noch zu überprüfen, ob noch ein anderer Punkt der Originallinie näher an Start2 liegt als der Knickpunkt. Ist dies der Fall, ist dieser Startpunkt für die Linienverfolgung im Polygon 1 - Start1, anderenfalls wird der Knickpunkt dafür verwendet.

Linienverfolgung:

Für die Linienverfolgung werden nun ausschließlich die beiden selektierten Waldpolygone verwendet. Es ist nicht möglich auch das Straßenpolygon heranzuziehen, da diese Polygone meistens sehr kurz sind, d. h. meist weniger als fünf Stützpunkte haben, die außerdem oft sehr eng beieinander liegen. Die Ursache dafür dürfte die manuelle Digitalisierung sein. Diese wirkt sich jedoch nicht besonders nachteilig aus, da dennoch bei der Linienverfolgung sichergestellt wird, daß entlang der verfolgten Linie auch ein Straßenpolygon vorhanden ist. Dies wird aufgrund der oft minimalen Länge der Straßenpolygone nicht immer das gleiche sein.

Für die Linienverfolgung werden ausschließlich die Originallinien herangezogen - nicht die generalisierten. Die Linienverfolgung wird von der Polygonlinie 2 aus gestartet, indem die Stützpunkte hier zunächst rückwärts verfolgt werden (Anschließend erfolgt im Bedarfsfall noch eine Linienverfolgung vorwärts). Um jeden weiteren Stützpunkt wird wieder ein Fangkreis gelegt, in dem sich wiederum eine Straße finden muß sowie die beiden vorher ermittelten Waldpolygone. In der folgenden Situationsdarstellung werden die Linienelemente von Polygon 1 mit a_i und die des Polygons 2 mit b_i bezeichnet.



Fall 1:

Es wird nur ein Linienelement a_1 von Polygonlinie 1 im Fangkreis gefunden. Es ist auch in den folgenden Fällen gleichgültig, ob die Linienelemente der Polygonlinie 1 ganz im Fangkreis liegen oder diesen nur kreuzen. Zu prüfen ist,

ob der Richtungswinkel von a1 mit dem von b1 oder b2 übereinstimmt. Ist dies der Fall, wird dem Punkt S der Stützpunkt der Polygonlinie 1 zugeordnet, der näher an S liegt, in obiger Abbildung also Punkt 2.

Fall 2:

Es werden zwei Linienelemente a1 und a2 im Fangkreis gefunden. In diesem Fall ist zu prüfen, ob der Richtungswinkel von a1 oder a2 dem von b1 oder b2 zuzuordnen ist. Wenn ja, wird wiederum der Stützpunkt der Polygonlinie 1 dem Punkt S zugeordnet, der am nächsten an S liegt.

Fall 3 und 4:

Die Zuordnung erfolgt wie in den vorigen beiden Fällen, nur daß hier die Elemente a1, a2 und a3 gefunden wurden. Dem Punkt S wird im Fall 3 folglich der Punkt 3 zugeordnet, da der Richtungswinkel von a3 mit dem von b2 übereinstimmt. Im Fall 4 wird Punkt 2 zugeordnet, da der Richtungswinkel von a1 dem von b1 zugeordnet werden kann. Jedoch harmonisiert der Richtungswinkel von a2 weder mit dem von b1 noch mit dem von b2.

Fall 5:

Das Verfahren bleibt das gleiche, es soll nur gezeigt werden, daß kleine Unebenheiten im Linienvverlauf keine Rolle spielen, da sie von diesem Verfahren gefiltert werden, denn a3 kann weder b1 noch b2 zugeordnet werden. Dies ist sehr wichtig, da sonst die Linienvverfolgung zu häufig und zu früh abreißt.

Fall 6:

Hier findet keine Zuordnung statt.

Größe des Fangkreises:

Die Größe des Fangkreises ist abhängig von den Mindestdimensionen der doppelartig darzustellenden Straßen in der TK 25. Diese schwanken (inklusive Strichstärken der Linien) zwischen 0,9 und 1,7 Millimetern, was umgerechnet mit dem Kartenmaßstab einer Breite zwischen 22 und 42 Metern in ATKIS-Bayern entspricht. Ohne Nachteile kann die Fangkreisgröße generell auf 45 Meter gesetzt werden - eine Unterscheidung je nach Straßentyp ist nicht notwendig, da der Fangkreis nur eine Suchhilfe darstellt. Wesentlich entscheidender ist, daß bei der Linienvverfolgung immer die gleichen benachbarten Waldpolygone im Fangkreis selektiert werden und die oben genannten Bedingungen erfüllt sind.

Wann können zwei Richtungswinkel einander zugeordnet werden?

Der Richtungswinkel berechnet sich:

$$\tan t = \frac{y_E - y_A}{x_E - x_A} \quad \text{und} \quad s^2 = (y_E - y_A)^2 + (x_E - x_A)^2$$

wobei y_A und x_A die Koordinaten des Anfangs- und y_E und x_E die des Endpunktes sind.

Daraus folgt:

$$\frac{\partial t}{\partial y_E} = \frac{x_E - x_A}{s^2}, \quad \frac{\partial t}{\partial x_E} = \frac{y_A - y_E}{s^2}$$

$$\frac{\partial t}{\partial y_A} = \frac{x_A - x_E}{s^2}, \quad \frac{\partial t}{\partial x_A} = \frac{y_E - y_A}{s^2}$$

und damit:

$$\sigma_t^2 = \left(\frac{\partial t}{\partial y_E} \right)^2 \cdot \sigma_{y_E}^2 + \left(\frac{\partial t}{\partial x_E} \right)^2 \cdot \sigma_{x_E}^2 + \left(\frac{\partial t}{\partial y_A} \right)^2 \cdot \sigma_{y_A}^2 + \left(\frac{\partial t}{\partial x_A} \right)^2 \cdot \sigma_{x_A}^2, \quad \text{mit}$$

$$\sigma_{y_A} = \sigma_{y_E} = \sigma_{x_A} = \sigma_{x_E} = \sigma_0$$

$$\sigma_t^2 = 2 \cdot \left(\frac{x_E - x_A}{s^2} \right)^2 \cdot \sigma_0^2 + 2 \cdot \left(\frac{y_E - y_A}{s^2} \right)^2 \cdot \sigma_0^2$$

$$\sigma_t^2 = 2 \cdot \frac{\sigma_0^2}{s^4} \cdot \left((x_E - x_A)^2 + (y_E - y_A)^2 \right)$$

$$\sigma_t^2 = 2 \cdot \frac{\sigma_0^2}{s^4} \cdot s^2 = 2 \cdot \frac{\sigma_0^2}{s^2}$$

damit ergibt sich die Genauigkeit des Richtungswinkels zu:

$$\sigma_t = \sqrt{2} \cdot \frac{\sigma_0}{s}$$

wobei σ_0 die Lagegenauigkeit im Objektbereich Vegetation und s die Länge des Liniensegments ist. Für s wird in jedem Fall immer die Länge des kürzesten Liniensegments verwendet, das innerhalb des aktuellen Fangkreises liegt (vgl. Abbildung 3.26, Fall 3: hier wäre es die Länge des Liniensegments a_2). Wie im Kapitel 4 noch gezeigt wird, liegt die mittlere Lagegenauigkeit im Objektbereich Vegetation bei ± 7 Metern. Dieser Wert wird für diese Untersuchung noch verdoppelt - damit sind statistisch betrachtet 95,4% aller Fälle abgedeckt - und wird hier gerundet mit 15 Metern eingeführt.

Damit ist ein sinnvolles Maß für die zulässige Abweichung zweier einander zuzuordnender Richtungswinkel ermittelt.

Das **Ende** der Linienverfolgung ist dann erreicht, wenn keine Zuordnung mehr stattfinden kann. Als Ergebnis sind die Stützpunktnummern der Anfangs- und Endpunkte der Linienverfolgung beider Polygonlinien gespeichert - und zwar die Nummern der Originallinie.

Zusammensetzen der neuen gemeinsamen Waldpolygonlinie:

Die Endpunkte, die bei der Linienverfolgung ermittelt werden, liegen nicht zwingend einander gegenüber. Es ist sinnvoll einen weiteren Punkt in die Linie einzurechnen. Dazu wird das Lot auf die andere Linie gefällt.

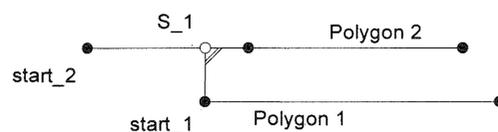


Abbildung 3.27: Schnittpunktbildung

Die Schnittpunkte sind sowohl am Beginn der Linienverfolgung (S_1) als auch am Ende (S_2) zu bilden (Abbildung 3.28).

Für die Bildung des neuen Waldpolygons sind vier Fälle zu unterscheiden, je nachdem, wo die Anfangs- und Endpunkte der ursprünglichen Polygonlinien ($A1/E1$ bzw. $A2/E2$) zu liegen kommen. Die folgende Abbildung zeigt eine schematische Darstellung der Situation, wobei die Schnittpunkte S_1 und S_2 nur symbolisch zwischen den beiden ursprünglichen Polygonen dargestellt sind. Die Pfeile geben die Laufrichtung der Polygone an.

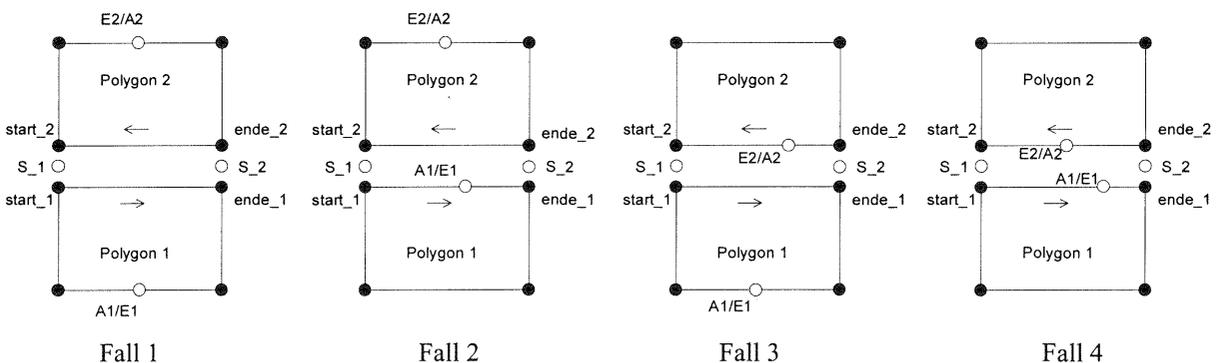


Abbildung 3.28: Bildung des neuen Polygons

Eindeutiges Entscheidungskriterium für den Zusammenbau des neuen Waldpolygons ist:

- Fall 1:** $start_1 < ende_1$ und $start_2 > ende_2$
- Fall 2:** $start_1 > ende_1$ und $start_2 > ende_2$
- Fall 3:** $start_1 < ende_1$ und $start_2 < ende_2$
- Fall 4:** $start_1 > ende_1$ und $start_2 < ende_2$

Die neuen Waldpolygone werden als rechtsdrehende Polygone gebildet. Dies ist wichtig, wenn diese Prozedur mit dem gleichen Waldpolygon mehrfach durchzuführen ist. Der Fall tritt immer dann auf, wenn ein zusammengehörendes Waldpolygon von mehreren Straßen durchzogen wird.

Anzumerken ist, daß die Lage des Straßenpolygons in diesem Algorithmus unverändert bleibt. Die Redundanzen, die ursprünglich vorhanden waren, werden allein durch die Bereinigung des Waldpolygons beseitigt.

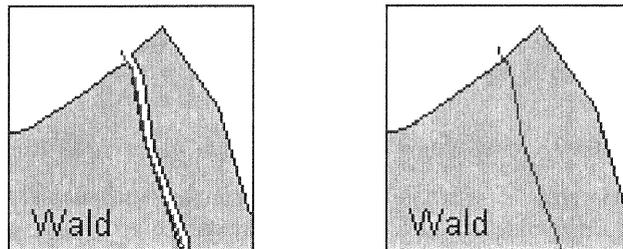


Abbildung 3.29:

a) Waldpolygon nicht bereinigt

b) Waldpolygon bereinigt

3.2.2.6 Bildung neuer Inselpolygone

Werden zwei Waldpolygone, wie im letzten Abschnitt beschrieben, vereinigt, so kommt es zu Situationen, in denen neue linksdrehende Inselpolygone zu bilden sind.

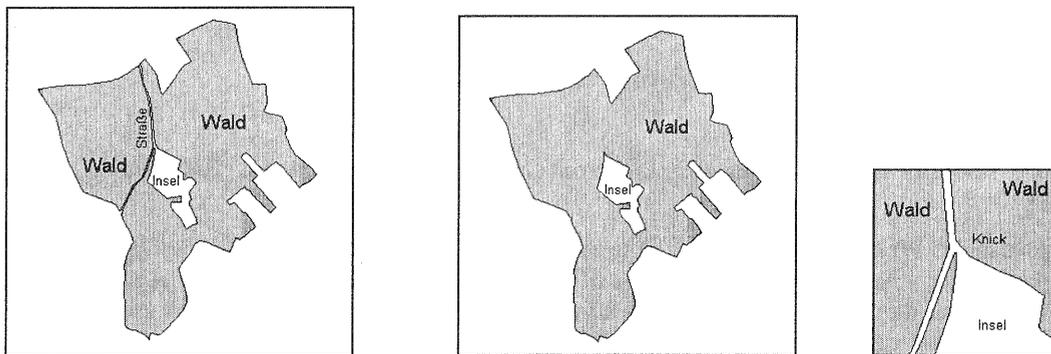


Abbildung 3.30: Waldpolygone

a) original ATKIS-Bayern – Daten

b) bereinigte ATKIS-Bayern – Daten

c) Vergrößerung

Die Vorgehensweise in diesem Programm ähnelt der vorher genannten.

1. Der Douglas-Peucker-Algorithmus wird auch hier mit den gleichen Parametern, wie im vorhergehenden Abschnitt beschrieben, verwendet, um die Originallinie auszudünnen und anschließend Knickpunkte zu selektieren. Denn auch bei der „Inseluche“ ist es immer so, daß zumindest am Anfang oder am Ende des neu zu bildenden Inselpolygons eine Knickstelle im ursprünglichen Polygon zu finden ist.
2. Um die Knickstelle wird ein Fangkreis mit einem Radius von 50 Metern gelegt und überprüft, ob ein anderes Liniensegment des gleichen Waldpolygons gefunden wird. Um sicher zu sein, daß es sich auch zumindest um eine kleine Insel handelt ist die Voraussetzung für die Inselbildung, daß zwischen dem Knickpunkt und dem Stützpunkt des anderen Linienelements mindestens fünf weitere Stützpunkte liegen. Diese bilden die Insel. Anderenfalls handelt es sich nicht um eine Inselbildung, da in einem Fangkreis von 50 Metern mit sehr großer Wahrscheinlichkeit noch ein weiteres Linienelement der gleichen Polygonlinie zu finden ist. Die Größe des Fangkreises ist wieder mit der Straßenbreite zu begründen, da auch in diesen Fällen die Ursache der Verlauf einer Straße in diesem Bereich ist.
3. Ein weiterer wesentlicher Punkt ist in diesem Fall noch zu berücksichtigen. Es handelt sich nur dann auch wirklich um eine Insel, wenn der Drehsinn des Teilpolygons zwischen „Beginn und Ende“ entgegen dem des Gesamtpolygons ist, sonst handelt es sich um eine Ausbuchtung. Dies bedeutet im Fall der Waldpolygone, daß das Inselstück linksdrehend sein muß.

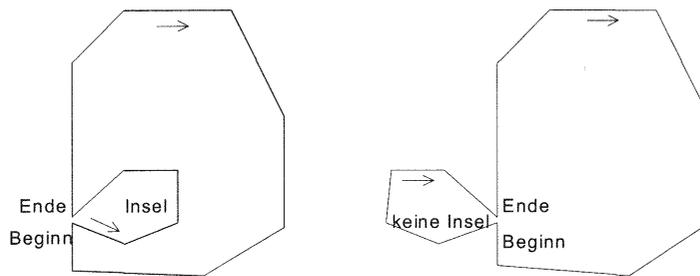


Abbildung 3.31:
a) Inselbildung

b) keine Inselbildung

4. Wurde ein Inselstück festgestellt, werden zwei getrennte neue Waldpolygone gebildet:
- das rechtsdrehende Gesamtpolygon, bei dem die Punkte zwischen Beginn und Ende ausgelassen werden und
 - das linksdrehende Inselpolygon, gebildet aus den Stützpunkten zwischen Beginn und Ende.

3.2.3 Inkonsistenzen zwischen den Objektbereichen Wald und Siedlung aufsuchen, selektieren und bereinigen

3.2.3.1 Besonderheiten des Objektbereichs Siedlung

Der Objektbereich Siedlung weist zwei Besonderheiten auf:

1. Die Umfangspolygone der Siedlungsflächen haben uneinheitlichen Drehsinn, jedoch ohne Systematik.
2. Da die Erfassung aus der TK 25 erfolgt, ist die Genauigkeit der Polygone gegenüber allen anderen untergeordnet.

3.2.3.2 Zusammentreffen von Wald- und Siedlungsgebieten

Die Grenzen von Wald- und Siedlungspolygonen treffen in Bereichen aneinander, wo das Ortsgebiet bis an den Waldrand heranreicht. Auch hier treten Redundanzen auf, die auf die getrennte Objektbereichserfassung zurückzuführen sind. Da beide Objektbereiche flächenförmig erfasst sind, müssen, im Gegensatz zu Kapitel 3.2.2.5, für die Datenbereinigung beide Objektarten (bzw. -bereiche) herangezogen werden.

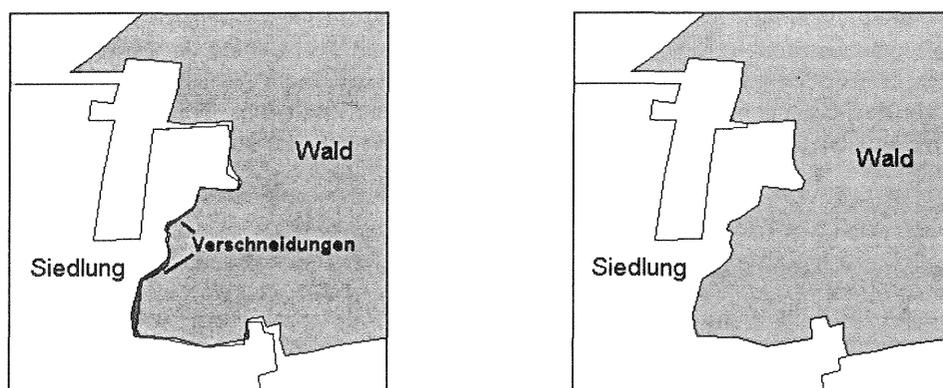


Abbildung 3.32: Zusammentreffen der Objektbereiche Siedlung und Vegetation
a) originale ATKIS-Bayern – Daten

b) bereinigte ATKIS-Bayern - Daten

Die Verfahrensweise gleicht dennoch der unter 3.2.2.5 beschriebenen. Auch hier wird ausgehend vom Waldpolygon der Douglas-Peucker-Algorithmus angewandt und in gleicher Weise die Linienverfolgung an den Knickstellen gestartet.

Der wesentliche Unterschied besteht bei der Neubildung der redundanzfreien, bereinigten Daten. Wie zuvor schon erläutert, können die Siedlungspolygone aus der TK 25 nur mit untergeordneter Genauigkeit erfasst werden. Die Kartierung des Waldes ist in der TK 25 i. d. R. relativ genau, da in freien Lagen auch kaum Verdrängungen durch Generalisierungen vorkommen. Zudem erfahren Waldflächen selten Veränderungen in der Natur bezüglich ihrer Lage, so daß

die Waldpolygone eine deutlich höhere Lagegenauigkeit aufweisen als die Siedlungspolygone. Damit kann anschließend an die Linienverfolgung zwischen dem Siedlungs- und dem Waldpolygon im redundanten Bereich die Lage des Waldpolygons auch für das Siedlungspolygon angehalten werden, ohne eine echte Genauigkeitseinbuße in Kauf nehmen zu müssen. Dies vereinfacht die Bereinigung des Datenbestands, da somit auch hier, wo zwei flächenförmig modellierte Objektbereiche aufeinander treffen, die Lage eines Objektbereichs beibehalten werden kann.

3.2.3.3 Weitere Möglichkeiten

Die bisher aufgeführten Verfahren lassen sich ohne weiteres auch auf andere Objektarten bzw. Objektbereiche übertragen. Dabei ist immer die Drehrichtung der Polygone zu berücksichtigen, und für die Generierung des neuen Polygons sind die Genauigkeiten der Objektarten zu modellieren.

Insbesondere kommt das unter 3.2.3.2 beschriebene Verfahren auch für die Bereinigung der Redundanzen zwischen flächenförmig modellierten Gewässern und dem Objektbereich Wald in Frage.

3.3 Bewertung der vorgestellten Verfahren

Allen Verfahren liegt ein der Situation angepaßter, jedoch möglichst einfacher Algorithmus zugrunde. Die Algorithmen sind so aufgebaut, daß sie in allen handelsüblichen GIS-Produkten, die eine programmierbare Schnittstelle haben, ohne große Schwierigkeiten nachgebildet werden können. Da im Rahmen dieser Arbeit ausschließlich auf die geometrischen Eigenschaften von ATKIS-Bayern eingegangen wird, ist auch der Einsatz eines programmierbaren CAD-Systems ausreichend, um die Algorithmen einzusetzen.

4 Exemplarische Modellrechnungen

In diesem Kapitel werden die zuvor in 3.2 theoretisch vorgestellten Verfahren anhand eines Datensatzes getestet, es werden Erfolge und Grenzen des Verfahrens aufgezeigt.

4.1 Testgebiete - die Auswahl des Testgebiets Althegenberg

An dieser Stelle soll nicht versäumt werden, der Bayerischen Vermessungsverwaltung für die kostenlose Bereitstellung der Daten zu danken; hier insbesondere dem Bayerischen Landesvermessungsamt, welches die Datensätze aus ATKIS-Bayern und die Digitalen Orthophotos unkompliziert und schnell zur Verfügung stellte und dem Vermessungsamt Fürstenfeldbruck, welches die DFK ebenso schnell und unkonventionell verfügbar machte.

Welche Voraussetzungen muß ein Testgebiet erfüllen, das für eine Untersuchung der geometrischen Genauigkeit eines Geobasisdatenbestandes herangezogen werden soll?

Hierzu sollen einige Mindestkriterien herangezogen werden, die der Autorin besonders wichtig erscheinen:

1. der Datensatz muß vollständig sein,
2. gewisse Ortskenntnis im Testgebiet soll vorhanden sein,
3. die Daten sollen möglichst einfach zu beschaffen sein.

Da nicht für alle Bereiche Bayerns zum Zeitpunkt der ersten Testdatenabgabe alle Objektbereiche vollständig erfaßt waren, bestand eine wichtige Bedingung für die Auswahl eines Testgebiets darin, daß in diesem bereits alle Objektbereiche von ATKIS-Bayern erfaßt waren. Die Ortskenntnis schien wichtig, um gegebenenfalls Unstimmigkeiten zwischen DFK und ATKIS-Bayern vor Ort zu klären und damit auch Aussagen über die Aktualität treffen zu können. Diese Bedingung ist mit dem Wunsch der möglichst einfachen Beschaffung gut zu vereinbaren gewesen, da Daten bei meinem heimatlichen Vermessungsamt Fürstenfeldbruck, an dem ich während meiner Referendarzeit ein dreiviertel Jahr gearbeitet habe, sehr leicht zu beschaffen waren. Zum Zeitpunkt der ersten Datenabgabe waren am Vermessungsamt Fürstenfeldbruck die Digitalen Flurkarten für die Gemeinden Eichenau und Althegenberg bereits flächendeckend vorhanden. Dabei war nicht von Bedeutung, ob diese streng numerisch angesetzt, oder über Digitalisierung und Homogenisierung entstanden waren. Beide Methoden können für den hier benötigten Verwendungszweck als fehlerfrei ange-

sehen werden. Die Daten aus ATKIS-Bayern waren ebenfalls für die beiden Gebiete vorhanden. Die Entscheidung fiel auf das Gebiet um Althegnenberg, eine ländliche Gemeinde im Nordwesten des Landkreises Fürstentfeldbruck. Folgende Gründe sprachen für diese Entscheidung:

- da die Siedlungsflächen bisher bei der Erfassung von ATKIS-Bayern nicht in Wohn-, Misch- und Industriegebiete unterteilt wurden, bietet Eichenau, eine Wohngemeinde mit geringen Industrieflächen, die noch zum Einzugsbereich von München gehört, keine Vorteile für diese Untersuchung.
- die Landschaft im ländlichen Bereich von Althegnenberg weist eine Vielzahl verschiedenster Wasserläufe auf, wie Bäche, Entwässerungskanäle, Kläranlagen, kleine Weiher usw., während das Gebiet um Eichenau zwar nicht wasserarm ist, jedoch wesentlich nur vom Wasserlauf der Amper geprägt wird.
- insgesamt ist die Landschaft um Althegnenberg ländlich geprägt - viele kleinere Ortschaften und Weiler, keine großen Gemeinden
- die Dichte der Verkehrswege war in beiden Gebieten etwa gleich
- Waldgebiete sind sowohl um Althegnenberg als auch um Eichenau vorhanden, wenn auch nur mit mäßiger Dichte, da der gesamte Landkreis Fürstentfeldbruck nur relativ wenige Waldflächen aufweist.

4.2 Untersuchungsergebnisse

4.2.1 Ergebnisse der Genauigkeitsuntersuchungen

4.2.1.1 Vergleich zwischen der DFK und ATKIS-Bayern

Es wurde für den Vergleich mit der DFK einmal eine vierparametrische Helmerttransformation und eine sechsparemetrische Affintransformation durchgeführt. Wie zu erwarten war, fielen die Verbesserungen bei der Affintransformation geringer aus, denn je mehr Parameter - bei gleicher Anzahl der identischen Punkte - verwendet werden, desto geringer die Verbesserungen.

Objektbereich Verkehr (innerorts):

Helmerttransformation:	Drehwinkel:	$\varepsilon = 0,22007 \text{ gon}$
	Maßstabsfaktor:	$m = 1,00021366$
	Translation 1. Achse:	$Y_0 = -1287,88 \text{ m}$
	Translation 2. Achse:	$X_0 = 1425,06 \text{ m}$
Standardabweichung der Restklaffen in y-Richtung:		$\sigma_y = 4,07 \text{ m}$
Standardabweichung der Restklaffen in x-Richtung:		$\sigma_x = 4,00 \text{ m}$
und damit eine mittlere Genauigkeit von:		$\sigma = 4,15 \text{ m.}$

Affintransformation:	Winkel der beiden Koordinatenachsen zueinander:	$\varepsilon = 99,5578 \text{ gon}$
	Maßstabsfaktor der 1. Achse:	$m_y = 1,0065978$
	Maßstabsfaktor der 2. Achse:	$m_x = 0,9957727$
	Translation der 1. Achse:	$Y_0 = -4154,50 \text{ m}$
	Translation der 2. Achse:	$X_0 = 106,39 \text{ m}$
Standardabweichung der Restklaffen in y-Richtung:		$\sigma_y = 3,836 \text{ m}$
Standardabweichung der Restklaffen in x-Richtung:		$\sigma_x = 3,965 \text{ m}$
und damit eine mittlere Genauigkeit von:		$\sigma = 4,01 \text{ m.}$

4.2.1.2 Vergleich zwischen Orthophotos und ATKIS-Bayern

Da wie oben gezeigt der Unterschied zwischen der Helmert- und der Affintransformation sehr gering ist, wird für diesen Vergleich nur mehr die Helmerttransformation verwendet.

Objektbereich Verkehr:	Drehwinkel:	$\varepsilon = -0,01839$ gon
	Maßstabsfaktor:	$m = 1,00066662$
	Translation 1. Achse:	$Y_0 = -1411,1$ m
	Translation 2. Achse:	$X_0 = -4838,8$ m
	Standardabweichung der Restklaffen in y-Richtung:	$\sigma_y = 9,2$ m
	Standardabweichung der Restklaffen in x-Richtung:	$\sigma_x = 6,9$ m
	und damit eine mittlere Genauigkeit von:	$\sigma = 8,2$ m

Objektbereich Vegetation:	Drehwinkel:	$\varepsilon = -0,02839$ gon
	Maßstabsfaktor:	$m = 1,00099790$
	Translation 1. Achse:	$Y_0 = -2040,7$ m
	Translation 2. Achse:	$X_0 = -7309,1$ m
	Standardabweichung der Restklaffen in y-Richtung:	$\sigma_y = 7,0$ m
	Standardabweichung der Restklaffen in x-Richtung:	$\sigma_x = 6,6$ m
	und damit eine mittlere Genauigkeit von:	$\sigma = 6,9$ m

Objektbereich Gewässer:	Drehwinkel:	$\varepsilon = -0,000015$ gon
	Maßstabsfaktor:	$m = 0,99999592$
	Translation 1. Achse:	$Y_0 = 5,0$ m
	Translation 2. Achse:	$X_0 = 34,8$ m
	Standardabweichung der Restklaffen in y-Richtung:	$\sigma_y = 6,9$ m
	Standardabweichung der Restklaffen in x-Richtung:	$\sigma_x = 6,5$ m
	und damit eine mittlere Genauigkeit von:	$\sigma = 6,8$ m

4.2.1.3 Wertung

Die Lagegenauigkeit des Objektbereichs Verkehr innerorts ist im Testdatenbestand mit $\pm 4,0$ Metern sehr hoch. Bemerkenswert ist jedoch, daß dieser doppelt so hoch ist wie außerhalb des Ortsbereichs. Zwar wurde anfangs bereits darauf hingewiesen, daß der Vergleich zwischen DFK und ATKIS-Bayern innerorts nicht ohne weiteres auf die übrigen Verkehrswege übertragen werden darf, doch der Faktor 2 wurde nicht erwartet.

Um sicher zu sein, daß die Differenzen nicht auf eine wesentlich verminderte Lagegenauigkeit der Orthophotos gegenüber der DFK zurückzuführen ist, wurde noch ein Vergleich zwischen Orthophoto und DFK für den Objektbereich Verkehr außerhalb der Siedlungsgebiete durchgeführt. Dies wurde dadurch möglich, daß das Vermessungsamt Fürstentfeldbruck seit Mitte 1997 auch über einige DFK im Außenbereich verfügt. Diese Genauigkeitsuntersuchung wurde, wie in Kapitel 4.2.1 beschrieben, durchgeführt und brachte folgendes Ergebnis:

Verkehr DFK-Orthophoto:	Drehwinkel:	$\varepsilon = -0,01077$ gon
	Maßstabsfaktor:	$m = 0,99959431$
	Translation 1. Achse:	$Y_0 = 235,6$ m
	Translation 2. Achse:	$X_0 = 66,3$ m
	Standardabweichung der Restklaffen in y-Richtung:	$\sigma_y = 0,98$ m
	Standardabweichung der Restklaffen in x-Richtung:	$\sigma_x = 0,74$ m
	und damit eine mittlere Genauigkeit von:	$\sigma = 0,9$ m

Diese Genauigkeitsuntersuchung zeigt, daß die Orthophotos eine mittlere Lagegenauigkeit von etwa einem Meter aufweisen und erwartungsgemäß damit ohne weiteres als Referenz für die ATKIS-Bayern - Daten herangezogen werden können.

Mögliche Ursachen für diese doch gravierenden Differenzen zwischen ATKIS-Bayern und DFK (außerorts) sind:

- Da im Ortsbereich die Darstellung in der TK 25 besonders schwierig ist - aufgrund der Vielzahl der darzustellenden Objekte -, wird hier besonders sorgfältig gearbeitet und die hohe Generalisierungspriorität bei den Verkehrswegen sorgt für eine besonders exakte Darstellung.
- Auch bei der manuellen Digitalisierung wird der Bearbeiter alleine durch die Vielzahl der dargestellten Objekte in der Erfassungsgrundlage im Unterbewußtsein immer wieder daran erinnert, besonders sorgfältig zu arbeiten. Hier scheint die Informationsdichte „psychologisches Moment“ zu sein.
- Der Umkehrschluß besteht darin, dort wo geringe Informationsdichte vorhanden ist, ist die absolut exakte Lage des Objekts nicht so entscheidend, es ist eindeutig und „sowieso mit keinem anderen zu verwechseln“ - der Bearbeitung wird dadurch weniger Beachtung geschenkt.

Dieses Untersuchungsergebnis ist zwar bemerkenswert, wird aber aus folgenden Gründen nicht weiter berücksichtigt werden:

- flächenmäßig sind die Straßen innerhalb geschlossener Ortschaften gegenüber den restlichen Objektarten in der Unterzahl, und wesentlich wichtiger,
- da derzeit im Ortsbereich als weitere Objekte nur Gewässer erfaßt sind, kommt es kaum zu Redundanzen, die im folgenden behoben werden müssen.

Damit ist eine separate Modellierung für innerörtliche Straßen nicht notwendig.

Auffallend ist zudem, daß die Lagegenauigkeit der Objektbereiche *Vegetation* und *Gewässer* gleich ist, während der Objektbereich *Verkehr* eine um etwa 20% niedrigere Genauigkeit aufweist. Dieser Unterschied ist nicht mit Generalisierungseffekten zu begründen, da diese immer nur dann eine Rolle spielen, wenn mehrere Objekte im engen räumlichen Abstand aufeinandertreffen. Dies trifft gerade für den Objektbereich *Verkehr* außerhalb von Ortschaften selten zu, da dessen Objekte lediglich durch Gewässer verdrängt werden, während die *Vegetation* (die Wälder) bei der Generalisierung durch die Mindestbreiten der Straßen deutlich verdrängt werden kann. Damit ist dies vorwiegend auf die manuelle Digitalisierung des Objektbereichs *Verkehr* zurückzuführen und die Differenz der Ergebnisse ist als persönlicher Fehler des Bearbeiters zu betrachten. Dieser beträgt damit etwa 1,3 Meter, umgerechnet über den Maßstab 1:10000 der Blaupause macht dies 0,13 Millimeter Ungenauigkeit bei der Erfassung.

Da die Differenz zwischen dem Objektbereich *Gewässer* und *Vegetation* mit 0,1 Meter nicht signifikant ist, kann ein wesentlicher Einfluß der Generalisierung hieraus nicht abgeleitet werden.

4.2.2 Ergebnis der Transformationen innerorts im Objektbereich Verkehr

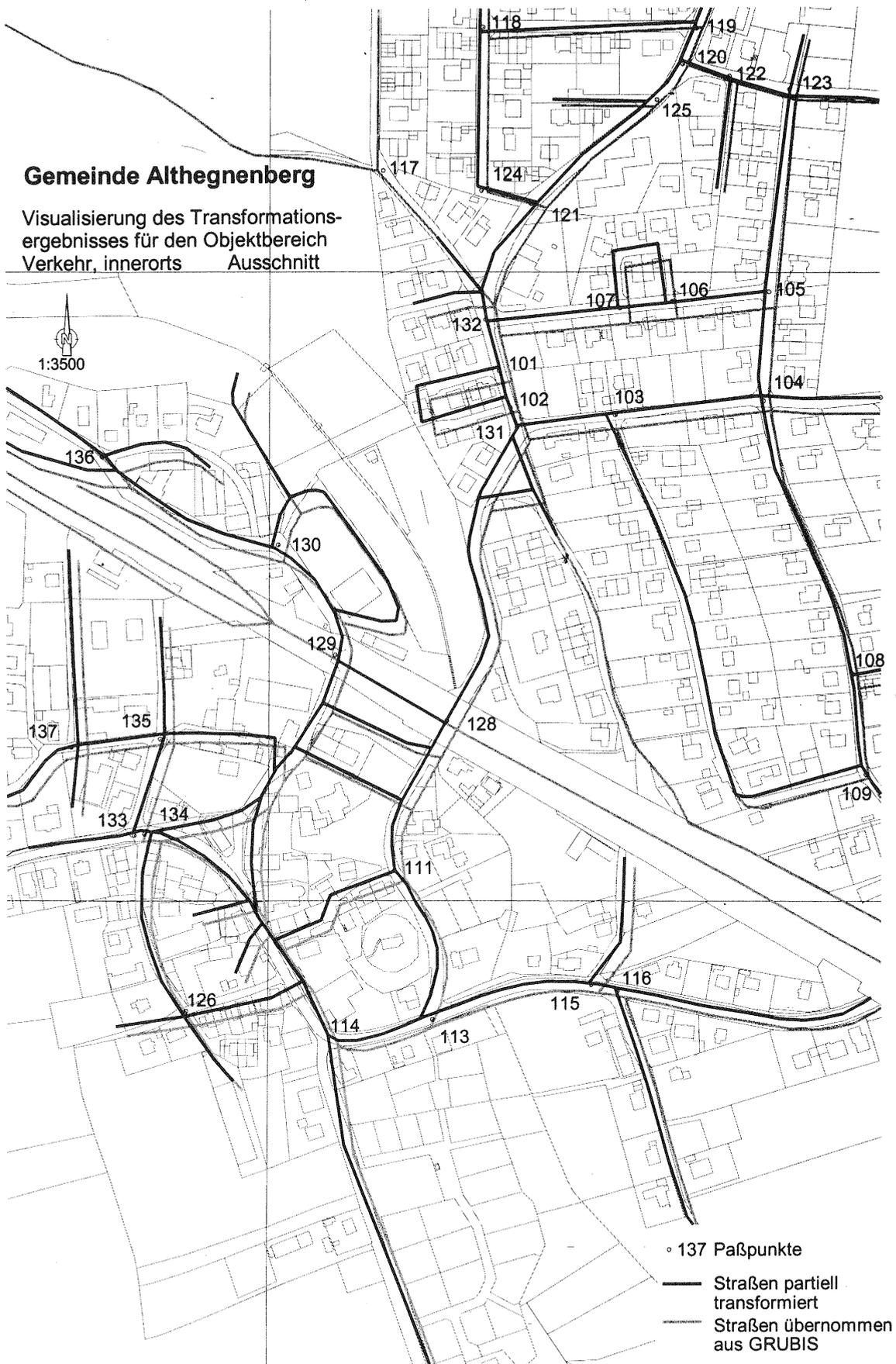


Abbildung 4.1: DFK überlagert mit ATKIS-Bayern-Daten vor und nach der Transformation

In Kapitel 3.1.1 wurden die auffälligen blockweisen Verschiebungen des Objektbereichs *Verkehr* gegenüber der DFK erwähnt. Bereits angesprochen wurde, daß eine Genauigkeitssteigerung durch Transformationen im großen Stil sicherlich nicht praktikabel ist und damit nur der Vollständigkeit halber erwähnt wird.

Die Gebiete mit einheitlichen Verschiebungsvektoren wurden über mindestens fünf visuell bestimmte Paßpunkte transformiert. Die folgenden Abbildungen visualisieren das Ergebnis in einem Teilbereich.

Anschließend wurde, um das Ergebnis der Transformationen bewerten zu können nochmals, wie im Kapitel 4.2.1.1 beschrieben, eine erneute Genauigkeitsuntersuchung mit den gleichen Paßpunkten durchgeführt.

Verkehr - DFK nach der Transformation:

Drehwinkel:	$\varepsilon = -0,07229$ gon
Maßstabsfaktor:	$m = 0,99743329$
Translation 1. Achse:	$Y_0 = 1495,3$ m
Translation 2. Achse:	$X_0 = 397,9$ m
Standardabweichung der Restklaffen in y-Richtung:	$\sigma_y = 3,18$ m
Standardabweichung der Restklaffen in x-Richtung:	$\sigma_x = 2,90$ m
und damit eine mittlere Genauigkeit von:	$\sigma = 3,0$ m

Wie das Ergebnis der zweiten Genauigkeitsuntersuchung zeigt, brachte die Transformation eine Genauigkeitssteigerung um einen Meter; dies entspricht einer Verbesserung um 25%. Vergleicht man dieses Ergebnis mit obiger Abbildung, so erscheint die erreichte Genauigkeitssteigerung relativ gering, da die Lage der transformierten Straßen visuell wesentlich verbessert erscheint.

Hauptverantwortlich hierfür ist die Lage der Paßpunkte an den Knoten der Verkehrswege, die für die Genauigkeitsuntersuchung herangezogen wurden.

Die Lagegenauigkeit des Knotenpunktes wurde durch die Transformation nicht erheblich verbessert, während der Verlauf der Straßen, die von diesem Knoten abgehen, eine deutliche Lageverbesserung erfahren. Dies ist jedoch nicht mit Zahlen belegbar, da im Verlauf des Straßen keine geeigneten Paßpunkte bestimmt werden können. Diese Einschränkung gilt natürlich für die Genauigkeitsuntersuchung im Objektbereich *Gewässer* gleichermaßen. Zudem ist die Lage der Knotenpunkte zueinander nicht immer richtig. Diese Ungenauigkeiten können jedoch über partielle Transformationen nicht behoben werden.

Folgende Abbildung verdeutlicht dies. Die relative Lage der Paßpunkte 101 und 102 zueinander paßt nicht, sie liegen zu nah beieinander, dennoch konnte durch die Transformation die absolute Lage der Straßen verbessert werden.

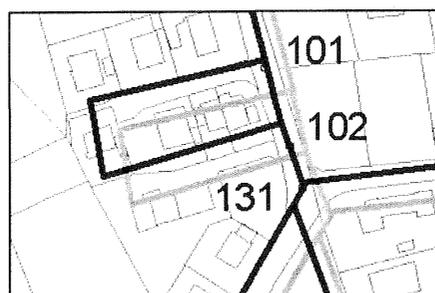


Abbildung 4.2: Lageungenauigkeiten der Paßpunkte zueinander

4.2.3 Weitere Auffälligkeiten im Testgebiet

Auch in Altheggenberg, das sowohl zum Einzugsbereich von München gehört als auch zu dem von Augsburg, war in den letzten Jahren eine deutliche Bautätigkeit zu verzeichnen. Am südlichen Ortsrand wurde 1992 ein neues Baugebiet erschlossen. Dieses war in der DFK binnen kürzester Zeit zu erkennen, da die Grundstücke entsprechend parzelliert wurden. In der TK 25, die einem Aktualisierungsrhythmus von mehreren Jahren unterliegt, war dies noch nicht enthalten. Damit fehlte es auch in ATKIS-Bayern. Dies ist ein typischer Fall, wo die Erfassungsgrundlage nicht dem aktuellen Stand der Örtlichkeit entsprach. Würde ein Orthophoto zur Datenerfassung mit herangezogen, so wäre

zumindest auf einigen Grundstücken eine Bebauung erkennbar gewesen und die Ortsgrenze hätte damit geometrisch genauer erfasst werden können.



Abbildung 4.3: Siedlungsgrenzen, Überlagerung von DFK, Orthophoto und ATKIS-Bayern

Des Weiteren fielen im Ortsbereich einige Straßen auf, die in der TK 25 als solche dargestellt und aus der photogrammetrischen Auswertung als solche erkennbar waren. Bei genauerem Betrachten und Überlagern mit der DFK stellten diese sich jedoch als private Hofeinfahrten oder andere private Wege heraus. Dieser Verdacht wurde durch eine Besichtigung in der Örtlichkeit bestätigt. Dies ist ein typischer Generalisierungsfehler in der Erfassungsgrundlage, der nur durch Ortskenntnis behoben werden kann.

Die hier aufgeführten Beispiele sind sicher keine Einzelfälle; sie treten im Datenbestand von ATKIS-Bayern immer wieder auf.

4.3 Vorstellung und Bewertung der Testergebnisse

4.3.1 Der Testdatensatz vor und nach der Datenbereinigung

Die Abbildungen finden sich in Anhang C.

4.3.2 Zusammenfassende Analyse der Testergebnisse

Die im Rahmen dieser Arbeit vorgestellten Algorithmen hatten zum Ziel, die vorhandenen Redundanzen im Datenbestand von ATKIS-Bayern zu beseitigen. Damit wird die Datenqualität für den Anwender deutlich gesteigert. Anwender ist in diesem Fall auch das Bayerische Landesvermessungsamt selbst, da als Eigennutzen von ATKIS-Bayern eine schnellere und vereinfachte Herstellung der TK 25 von Anfang an angestrebt wurde.

Nicht verbessert wird mit diesen Algorithmen die absolute Lagegenauigkeit der einzelnen Objektarten, zumal die hier zu bereinigenden Redundanzen nur bei flächenhaft zu modellierenden Objektarten auftreten. Linienhaft zu modellierende Objektarten werden im Zusammenhang mit diesen Algorithmen nur als Auswahlkriterien verwendet, werden aber nicht in ihrer geometrischen Lage verändert.

Eine Verminderung der Lagegenauigkeit durch die Anwendung dieser Verfahren ist aus mehreren Gründen nicht zu befürchten:

- Wann immer möglich, werden die Originalkoordinaten der Daten beibehalten. Es werden bewußt keine äquivalenten geometrischen Linien berechnet. Solche Verfahren sind nur dann sinnvoll, wenn dadurch ein echter Gewinn erzielt werden kann. Dies ist hier aber nicht der Fall, da die Schranken für die Suchprozeduren an die Lagegenauigkeit des jeweiligen Objektbereichs angepaßt sind.

- Bei der Bereinigung der durch Straßen zerschnittenen Waldflächen werden die Originalkoordinaten beibehalten und lediglich am Anfangs- und Endpunkt der Linienverfolgung je ein neuer Schnittpunkt berechnet, der in den Datensatz eingefügt wird. Dieser liegt jedoch auf einer ursprünglich bereits vorhandenen Linie des Ausgangsdatsatzes und ist damit indirekt doch wieder ein Element des Ausgangsdatsatzes.

Im folgenden noch einige Detailbeispiele:

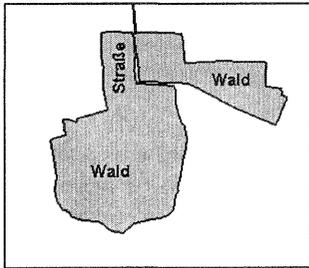


Abbildung 4.4: Datenbereinigung unterbrochen

Abbildung 4.4 zeigt zwei ursprünglich durch eine Straße getrennte Waldstücke. Die Datenbereinigung ist hier richtigerweise abgebrochen worden, da die, die Waldstücke trennende Straße, im Testdatensatz abgebrochen war.

In der folgenden Abbildung ist trotz Datenbereinigung ein kleiner Teil des Straßeneinschnittes übrig geblieben, da die beiden Straßenseiten hier nicht parallel verlaufen. In der analogen TK 25 ist an dieser Stelle des Straßenverlaufs eine Ausweichbucht dargestellt. Damit wurde der Waldecker an dieser Stelle auch etwas ausgeweitet.

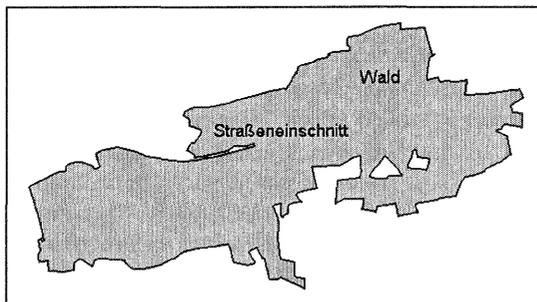


Abbildung 4.5: aufgeweiteter Straßenverlauf

Die beiden nächsten Abbildungen zeigen Beispiele, bei denen keine Linienverfolgung stattfinden konnte, da die aneinandergrenzenden Objekte sich nur auf minimalster Länge annähern.

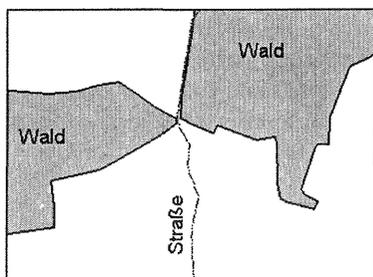
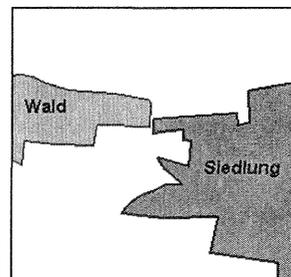


Abbildung 4.6: minimale Überlappungsbereiche
a) zwischen zwei Waldflächen



b) zwischen Wald- und Siedlungsfläche

5 Bedeutung für ATKIS-Bayern

Da die gesamte Arbeit immer mit dem Hintergrund einer praktischen Umsetzung angefertigt wurde, soll diese in diesem Kapitel noch etwas näher ausgeführt werden.

Wie können diese Algorithmen und Methoden gewinnbringend in die Praxis umgesetzt werden?

In Kapitel 2.2 wurde erwähnt, daß am Bayerischen Landesvermessungsamt derzeit die Arbeiten zur Umsetzung der ATKIS-Bayern - Daten in die objektweise Datenstruktur von ATKIS voll im Gange und sehr kosten- und zeitintensiv sind. Ein Ziel dieses Bearbeitungsvorgangs ist u. a., die Redundanzen im vorhandenen Datenbestand zu beseitigen. Dies geschieht derzeit durch zeitintensive manuelle Nachbearbeitung. Genau diese Arbeiten werden mit diesen Algorithmen wesentlich vereinfacht.

In dieser Arbeit erfolgte der Einsatz der Prozeduren halbautomatisch. Sie wurden händisch angestoßen und die Auswahl, welche Prozedur für den folgenden Arbeitsschritt gewählt wird, erfolgte visuell durch den Bearbeiter. Für den produktiven Einsatz ist es durchaus denkbar - unter bestimmten Voraussetzungen - einen vollautomatischen Batchbetrieb einzurichten. Dafür müßten folgende Voraussetzungen zwingend gegeben sein:

1. Es muß sichergestellt sein, daß die Drehrichtung der Polygone der verschiedenen Objektarten im gesamten Datenbestand einheitlich sind und die Bedingungen erfüllen, wie sie in Kapitel 3.2.2.3ff bei den einzelnen Algorithmen beschrieben wurden.
2. Die Abfolge der einzelnen Prozeduren muß aufeinander abgestimmt sein. Dies hat jedoch zur Folge, daß eine hohe Anzahl von Schritten routinemäßig immer in einer festen Reihenfolge ablaufen, auch wenn dies in vielen Fällen unnötig sein kann.

Hier ein Vorschlag für eine mögliche Abfolge der Prozeduren für einen vollautomatischen Betrieb.

Begonnen wird die Verarbeitung bei den **Waldpolygonen**:

- Im gesamten Datensatz werden die Waldpolygone am Blattschnitt vereinigt.
- Im gesamten Datensatz werden die Sackgassen in den Waldpolygonen beseitigt.
- Im gesamten Datensatz werden die Waldpolygone, die durch Straßen voneinander getrennt sind, wieder zusammengefügt. Dabei ist zu beachten, daß nach jedem erfolgreichen Lauf, also nach jedem neu zusammengeführten Polygon sofort, bevor weitere Waldpolygone zu einem zusammengeführt werden, geprüft wird, ob dadurch eine neue Sackgasse oder ein zu beseitigendes Inselfolygon entstanden ist. Die beiden zuletzt genannten Algorithmen werden somit sicher häufig vergeblich über den Datensatz geschickt. Dies ist aber unbedingt erforderlich, da sonst viele Inkonsistenzen nicht behoben werden können.
- Der Objektbereich *Vegetation* ist damit homogen; es folgt nun die Behebung der Redundanzen zwischen den Objektbereichen *Vegetation* und *Siedlung*. Dies erfordert einen einmaligen Durchlauf über alle „neuen“ Waldpolygone.
- Zuletzt erfolgt die Beseitigung der Redundanzen zwischen den Objektbereichen *Gewässer* und *Vegetation*, wobei hier in einem ersten Schritt die „Waldinseln“, die z. B. durch einen See entstanden sind und daran anschließend die Verschneidung zwischen Wäldern und flächenhaft zu modellierenden Gewässern bearbeitet werden.

Dabei stellt die heikelste Aufgabe die Bereinigung der Dateninkonsistenzen innerhalb des Objektbereichs *Vegetation* dar, da hier nach jeder erfolgreichen Datenbereinigung noch eine Überprüfung bezüglich erneuter Sackgassen- und Inselbildung anzuschließen ist. Scheitert hier einer der Nachbearbeitungsalgorithmen aufgrund knapp überschrittener Schranken, wie in Kapitel 4.3.1 gezeigt, so ist eine händische Nachbearbeitung an dieser Stelle unumgänglich.

Anwendung bei der Datenfortführung:

Ein weiterer Anwendungsbereich ist im Bereich der Datenfortführung denkbar, insbesondere im bayerischen Datenmodell. In der ersten, bereits abgeschlossenen Stufe, wurden lediglich die Objektarten erfaßt, die in der TK 25 enthalten waren. Bereits im DLM 25/1 werden eine Vielzahl weiterer Objektarten und Attribute erfaßt werden, die nur aus großmaßstäbigeren Kartenwerken bzw. aus Orthophotos ermittelt werden können, wie z. B. die Trennung zwischen Laub- und Nadelwald oder die Unterteilung der Siedlungsgebiete in Wohn- und Industrieflächen. Großmaßstäbiger Erfassungsgrundlagen bringen gleichzeitig eine Steigerung der Lagegenauigkeit mit sich. In diesem Zusammenhang bietet sich auch der Einsatz der Algorithmen, die Redundanzen zwischen zwei verschiedenen Objektarten beseitigen. Wird eine Laubwaldfläche aus einem Orthophoto erfaßt, so treten dadurch wiederum Verschneidungen mit angrenzenden Objektarten auf, die aus deren niedrigerer Lagegenauigkeit resultieren. Damit wäre eine sukzessive Genauigkeits-

steigerung im gesamten Datenbestand möglich, da jedes neue, zusätzlich erfaßte Objekt damit auch die Lage seiner benachbarten Objekte verbessert. Dieses Verfahren funktioniert ausschließlich für flächenhaft zu modellierende Objektarten.

Die in dieser Arbeit vorgestellten Verfahren nützen im wesentlichen die geometrischen Informationen für die Bereinigung der Redundanzen. Grundsätzlich gibt es zwei verschiedene Ansätze für die Vorgehensweise: Während bei der *geometrischen Integration* ausschließlich die geometrischen Eigenschaften der Objekte betrachtet werden, werden bei der *semantischen Integration* auch Sachdaten und Objektstrukturen herangezogen. Die folgende Abbildung zeigt dies in einem Beispiel.

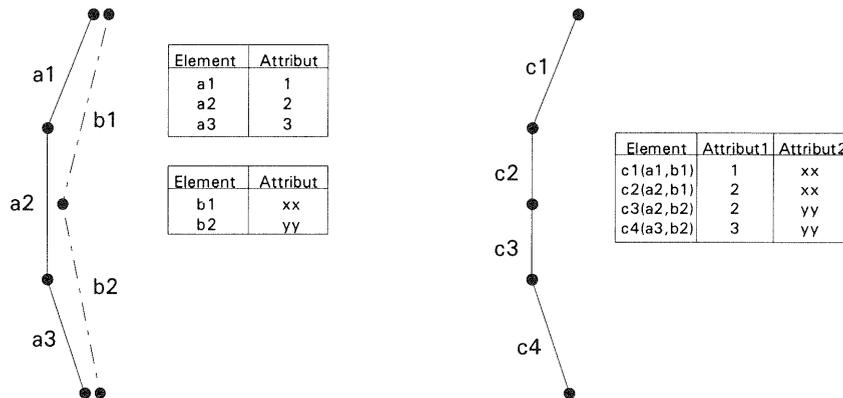


Abbildung 5.1: Semantische Integration

Ob die semantische oder die geometrische Integration sinnvoller ist, hängt vom Datenmodell ab. Die semantische Integration ist dann vorteilhaft, wenn durch sie bereits eine eindeutige Identifizierung des zweiten Objekts erfolgen kann. Auch Mischformen bei der Vorgehensweise sind möglich, z. B. wenn die Menge der in Frage kommenden Objekte durch die semantische Information bereits stark reduziert wird. Dies ist beispielsweise bei der Unterteilung der Waldflächen in Laub-, Nadel- und Mischwaldflächen der Fall. Hier kommen für den, wie oben vorgeschlagen, fortzuführenden Datensatz zur Bereinigung neu entstehender Redundanzen

1. nur Waldflächen in Frage und
2. nur solche Waldflächen, deren Erstellungsdatum älter ist als das der neu zu klassifizierenden Fläche.

Je mehr Semantik im Datensatz bereits enthalten ist, um so kleiner wird die Menge der möglichen Treffer sein. Dieser Aspekt sollte bei einer derartigen Implementierung noch eingehender betrachtet und untersucht werden.

6 Ausblick

In diesem Kapitel soll kurz eine Möglichkeit aufgezeigt werden, wie die zuvor vorgestellten Algorithmen mit Hilfe der Fuzzy-Logik modelliert werden könnten. Diese Methode ist wesentlich komplexer und nicht in allen Programmiersprachen ohne weiteres leicht abzubilden.

Alle Algorithmen, die im Rahmen dieser Arbeit entwickelt wurden, sind konventionell aufgebaut. Für Genauigkeitsabschätzungen und die daraus resultierenden Entscheidungskriterien wurde das klassische Fehlerfortpflanzungsgesetz verwendet. Damit wird mit einem festen Grenzwert die Entscheidung „ja“ oder „nein“ getroffen. Solche scharfen Grenzen sind nicht sehr realitätsbezogen und widersprechen dem menschlichem Empfinden. Die Fuzzy-Logik (unscharfe Logik) bietet hier erweiterte Möglichkeiten. Im Gegensatz zur klassischen Logik ist sie geeignet, Zusammenhänge und verbal formuliertes Wissen auf einem Computer nachzubilden. Damit ist es möglich differenziert Schlußfolgerungen zu ziehen sowie Analysen und Steuerungen komplexer Vorgänge durchzuführen. Eine kurze Zusammenfassung der wichtigsten Begriffe der Fuzzy-Logik findet sich im Anhang B.2.

Wie könnte die Fuzzy-Logik in den hier vorgestellten Algorithmen verwendet werden?

Die Linienverfolgung ist wesentlicher Bestandteil der zuvor vorgestellten Verfahren. In Abbildung 4.5 ist ein Beispiel dargestellt, in dem die Linienverfolgung abgebrochen wurde, da die Grenzwerte im Algorithmus geringfügig überschritten waren. Visuell betrachtet bzw. mit Worten beschrieben könnte eine Verbalisierung der Situation folgendermaßen lauten:

Die beiden Straßenseiten liegen zwar etwas weiter auseinander und damit ist auch die Differenz der Richtungswinkel gegebenenfalls etwas größer, aber die Straße verläuft dennoch weiter.

Diese Formulierung ist sehr vage, entspricht aber unserem menschlichen Empfinden.

Welche Kriterien können für die Fuzzifizierung herangezogen werden und welche Regeln können daraus abgeleitet werden?

Zunächst sei darauf hingewiesen, daß alle konventionellen Reglertypen auch über Fuzzy-Controller mit Fuzzy-Logik abgebildet werden können. Dies erscheint zwar zunächst unsinnig, aber es ermöglicht

1. einen universellen Reglertyp zu realisieren, der je nach Parametrisierung der Eingangswerte konventionelles oder Fuzzy-Verhalten zeigt.
2. in Fällen, in denen ein konventioneller Regler bereits im Einsatz ist, kann dieser zunächst ohne Verlust durch einen entsprechend parametrisierten Fuzzy-Controller ersetzt werden, um daraus schrittweise eine Verbesserung des Regelverhaltens durch Optimierung der Zugehörigkeitsfunktionen und der Regelbasis herbeizuführen.

Die zweite Möglichkeit scheint in diesem Zusammenhang interessant. Im folgenden soll an einem simplen Beispiel gezeigt werden, wie die Umsetzung eines konventionellen Reglertyps zu einem Fuzzy-Controller funktionieren kann.

Betrachtet man den einfachsten Fall eines Zweipunktreglers, der lediglich zwei Möglichkeiten der Stellgröße hat, nämlich u_{min} (für $x \leq 0$) und u_{max} (für $x > 0$), so ist die Kennlinie eines solchen Reglers ist sprunghaft.

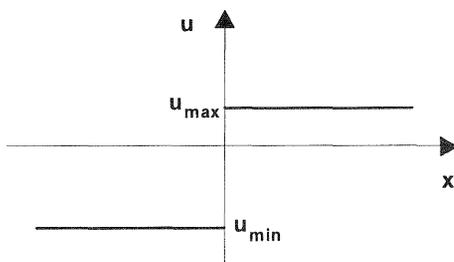


Abbildung 6.1: Kennlinie eines Zweipunktreglers

Eine derartige sprunghafte Kennlinie erhält man auch bei einem Fuzzy-Controller, indem sich die Zugehörigkeitsfunktionen für die Eingangsgröße nicht überlappen. Die linguistischen Terme für die Eingangsgrößen ($x \leq 0$ bzw. $x > 0$) lauten *negativ* und *positiv*. Als Stellgröße dienen hier die Terme *min* und *max*. Die Regelbasis lautet:

WENN $x = \textit{negativ}$ DANN $u = \textit{min}$
 WENN $x = \textit{positiv}$ DANN $u = \textit{max}$.

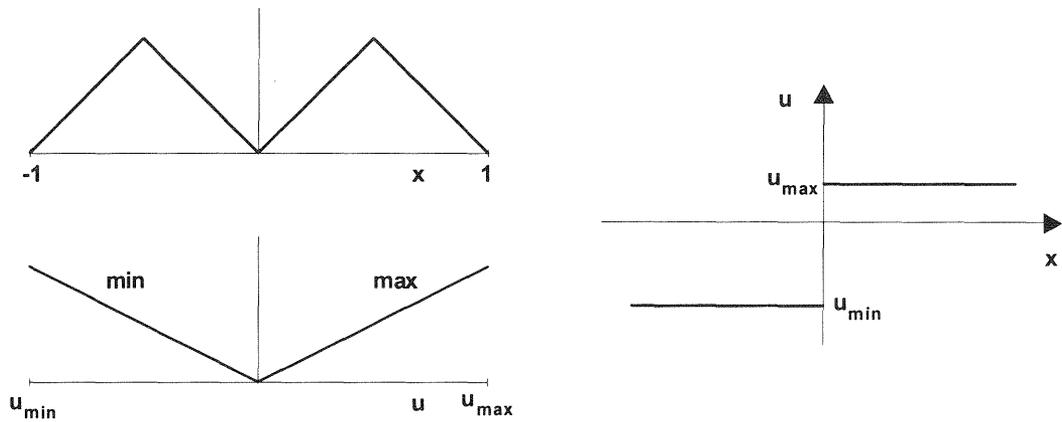


Abbildung 6.2: Zweipunktregler als Fuzzy-Controller

Dieses Beispiel ist unmittelbar einleuchtend, da für $x \leq 0$ nur die erste Regel und für $x > 0$ nur die zweite Regel aktiv ist. Die in Abbildung 6.2 dargestellte Realisierung ist nicht die einzige Möglichkeit der Realisierung. (Für die Defuzzifizierung wird hier von der Schwerpunktmethod ausgegangen). So ist für die Festlegung des Eingangsgröße-Fuzzy-Sets nur die Einflußbreite von Bedeutung nicht aber die Form. Das Fuzzy-Set der Stellgröße läßt noch größere Variationsmöglichkeiten zu, es ist lediglich an die Bedingung geknüpft, daß der linke Rand an u_{\min} und der rechte Rand an u_{\max} grenzt.

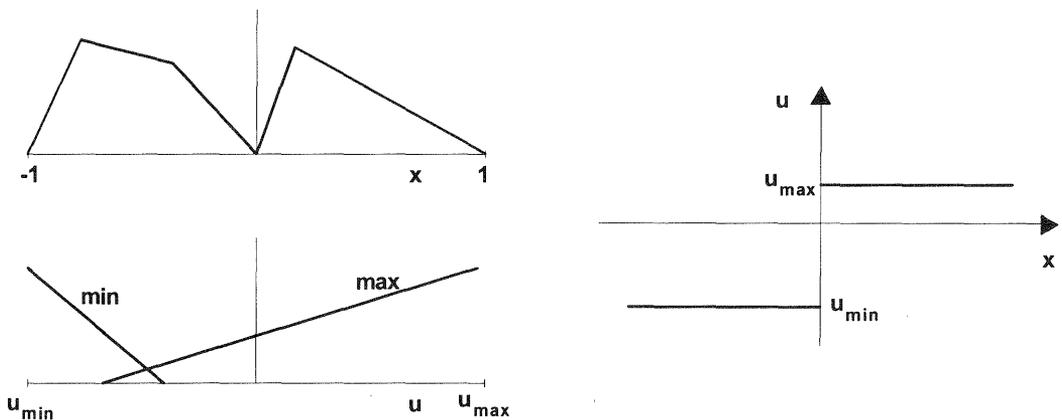


Abbildung 6.3: Alternativer Fuzzy-Controller für den Zweipunktregler

Im Fall der Linienverfolgung sind jedoch zwei Eingangsgrößen vorhanden: Zum einen der Abstand d der beiden Linien voneinander und $d\alpha$, die Winkeldifferenz zwischen Ausgangsline und verfolgter Linie. Zwei Eingangsgrößen werden mit dem MIN-Operator miteinander verbunden (siehe Anhang B.2).

Zunächst ein einfacher Vorschlag für die Fuzzifizierung der Eingangsgrößen und der Stellgröße u , die in diesem Fall entweder den Abstand oder die Winkeldifferenz benutzen:

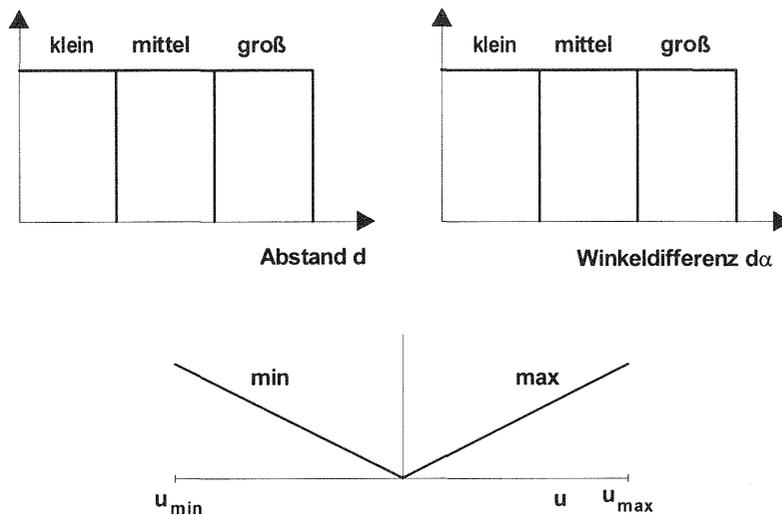


Abbildung 6.4: Vorschlag eines Fuzzy-Controller für die Linienverfolgung

Die Regelbasis könnte wie folgt lauten:

WENN $d = \text{klein}$ UND $d\alpha = \text{klein}$ DANN $u = \text{max}$
 WENN $d = \text{klein}$ UND $d\alpha = \text{mittel}$ DANN $u = \text{max}$
 WENN $d = \text{klein}$ UND $d\alpha = \text{groß}$ DANN $u = \text{min}$
 WENN $d = \text{mittel}$ UND $d\alpha = \text{klein}$ DANN $u = \text{max}$
 WENN $d = \text{mittel}$ UND $d\alpha = \text{mittel}$ DANN $u = \text{max}$
 WENN $d = \text{mittel}$ UND $d\alpha = \text{groß}$ DANN $u = \text{min}$
 WENN $d = \text{groß}$ UND $d\alpha = \text{klein}$ DANN $u = \text{min}$
 WENN $d = \text{groß}$ UND $d\alpha = \text{mittel}$ DANN $u = \text{max}$
 WENN $d = \text{groß}$ UND $d\alpha = \text{groß}$ DANN $u = \text{max}$.

Dabei bedeutet: $u = \text{min}$: Zuordnung ist nicht möglich → Linienverfolgung wird abgebrochen
 $u = \text{max}$: Zuordnung ist möglich → Linienverfolgung wird weitergeführt.

Als Kriterium für die Entscheidung, wann eine Winkeldifferenz klein, mittel oder groß ist, erscheint es sinnvoll die Grenzwerte, die sich aus dem Fehlerfortpflanzungsgesetz ergeben, mit einzubeziehen, da diese zulässigen Grenzwerte für die Winkeländerungen sich damit in Abhängigkeit von der Länge des jeweiligen Polygonstücks ändern.

Über die Form, Einflußbreite und Überlappungsbereiche soll hier bewußt keine Aussage getroffen werden; ebenso ist obige Regelbasis nur ein Vorschlag.

Die hier im Ausblick vorgestellte alternative Modellierungstechnik soll vielmehr als Denkanstoß für weitere Arbeiten dienen. Eine sensibel funktionierende Regelbasis muß schrittweise sukzessive entwickelt werden, bis die gewünschte Feinabstimmung auch zuverlässig erreicht ist.

7 Entstehung der Arbeit, Dank

Mit dem Abschluß meines Studiums an der Technischen Universität München im September 1992 stellte sich die Frage, wie mein Einstieg ins das Berufsleben als Vermessungsingenieurin aussehen sollte. Hier boten sich mir die verschiedensten Möglichkeiten an. Der meist gewählte Weg der Studienabgänger zu dieser Zeit war, aufgrund der guten konjunkturellen Lage, in einem Ingenieurbüro tätig zu werden. Die Alternative hierzu war das Referendariat für den höheren vermessungstechnischen Verwaltungsdienst, eine nochmalige 27-monatige Ausbildung, die mit der Großen Staatsprüfung endet. Den doch wesentlich weiter gesteckten Rahmen, den die Ausbildung als Vermessungsingenieur umfaßt, wie Photogrammetrie, Erdmessung und der Geodäsie wurde nur von einem kleinen Prozentsatz der Abgänger genutzt. Besonders die neuen Bereiche, die sich durch Geoinformatik und Geo-Informationssysteme erschließen, blieben fast gänzlich unbeachtet. Gerade dieser zukunftssträchtige Bereich, mit seinen neuen Aufgaben und den damit verbundenen Schwierigkeiten, schien mir ein interessanter Aufgabenbereich. Einen Beitrag hierzu leisten die Vermessungsverwaltungen der Bundesländer, die Basisdaten für Geo-Informationssysteme zur Verfügung zu stellen, da diese in analoger Form, als topographische Karten, Flurkarten und Deutsche Grundkarten usw., immer schon in ihrem Zuständigkeitsbereich geführt worden sind.

Die Möglichkeit eine Dissertation an einem Institut oder Lehrstuhl anzufertigen, erschien mir erstrebenswert. Im Januar 1993 bekam ich eine Anstellung am Institut für Photogrammetrie (IFP) der Universität Stuttgart. Am IFP werden drei Forschungsschwerpunkte betrieben - digitale Photogrammetrie, Sensorik und Geoinformatik - wobei sich meine Interessen auf den Bereich Geo-Informationssysteme richteten. Da die Vermessungsverwaltungen der Bundesländer einen wesentlichen Beitrag zur Bereitstellung von Geobasisdaten leisten, entstand daraus der Gedanke, diese Arbeit in Absprache mit dem Bayerischen Staatsministerium der Finanzen in München zu erstellen, der obersten Dienststelle der Bayerischen Vermessungsverwaltung. Damit war mir die Möglichkeit gegeben, gleichzeitig das Referendariat abzuleisten. Im gegenseitigen Einverständnis und Interesse wurde die Grundlage dieser Arbeit geschaffen. So begann ich im Mai 1993 das Referendariat, legte die Große Staatsprüfung im Juni 1995 ab und bin seit August 1995 fest in den Dienst der Bayerischen Vermessungsverwaltung übernommen. Das Referendariat gliedert sich in praktische und theoretische Ausbildungsabschnitte. Um auch während dieser Ausbildung etwas Zeit für meine Doktorarbeit zu gewinnen, stellte mich das Bayerische Staatsministerium in den praktischen Ausbildungsabschnitten immer wieder vom Dienst frei. Für dieses große Entgegenkommen und die freundliche Unterstützung möchte ich mich an dieser Stelle herzlich bedanken, vor allem bei Herrn Univ. Prof. Dr.-Ing. Josef Frankenberger. Ohne dieses hätte die Arbeit so nicht durchgeführt werden können. Besonderer Dank gebührt auch Herrn Univ. Prof. Dr.-Ing. habil. Dieter Fritsch, der dieses Vorhaben uneingeschränkt von Anfang an unterstützte.

Mein Dank gebührt aber vor allem auch meinen Familienangehörigen, die mich mit Geduld und Nachsicht während der letzten Jahre begleitet haben, um mir die Fertigstellung dieser Arbeit zu erleichtern. Die Gewährung dieses Freiraumes ermöglichte es mir die Arbeit abzuschließen.

8 Literatur

- AdV (1988): Amtliches Topographisch-Kartographisches Informationssystem (ATKIS). Landesvermessungsamt Nordrhein-Westfalen, Bonn.
- AdV-Symposium ATKIS (1996): Das Geoinformationssystem ATKIS und seine Nutzung in Wirtschaft und Verwaltung, Vorträge anlässlich des 3. AdV-Symposiums ATKIS vom 29. und 30. Oktober 1996 in Koblenz, Landesvermessungsamt Rheinland-Pfalz.
- AdV-Projekt ATKIS-Entwicklung, Projektdokumentation der Expertengruppe ATKIS-Entwicklung, Stand 30.12.96
- ALK (1986): ALK-Dokumentation 2.1: Einheitliche Datenbankschnittstelle (EDBS), AG Hannover, Niedersächsisches Landesvermessungsamt - Landesvermessung Hannover.
- Aumann, G. (1993): Aufbau qualitativ hochwertiger Digitaler Geländemodelle aus Höhenlinien, Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften, München, Reihe C, Nr. 441.
- Autodesk (1994): Handbuch für Benutzeranpassungen - AutoCAD Release 13, Autodesk Inc., veröffentlicht durch: Autodesk Development B.V., Neuchâtel (Schweiz).
- Baumann, E. (1993): Vermessungskunde, Band 2 - Punktbestimmung nach Höhe und Lage, vierte, bearbeitete und erweiterte Auflage, Dümmler Verlag, Bonn
- Bayerisches Landesvermessungsamt (1993): Informationen des Bayerischen Landesvermessungsamtes zur Abgabe von Rasterdaten, Stand 9/93.
- Bayerische Staatskanzlei und alle Staatsministerien (1992): Gemeinsame Bekanntmachung über den Aufbau raumbezogener Informationssysteme vom 07.01.92.
- Bayerisches Staatsministerium der Finanzen (1993): Richtlinien zum Datenaustausch für das amtliche Grundstücks- und Bodeninformationssystem (DatRI-GRUBIS), Bekanntmachung vom 25.06.1993.
- Bill, R., Fritsch, D. (1991): Grundlagen der Geo-Informationssysteme, Band 1, Hardware, Software und Daten, Wichmann Verlag, Karlsruhe.
- Bill, R. (1996): Grundlagen der Geo-Informationssysteme, Band 2, Analysen, Anwendungen und neue Entwicklungen, Wichmann Verlag Heidelberg.
- CARD/1 (1993): Handbuch zum CAD-System CARD/1 der Firma Basedow & Tornow Hamburg (Selbstverlag).
- Douglas, D., Peucker, D. (1973): Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. The Canadian Cartographer, Vol. 10.
- Ebner, H. (1988): Neue Entwicklungen in den Bereichen Digitales Geländemodell und Digitale Photogrammetrie, 2. Informationsveranstaltung über die graphische Datenverarbeitung der Bayer. Vermessungsverwaltung, München, Hrsg. Bayer. Landesvermessungsamt, 77-79.
- Findeisen, D. (1990): Datenstrukturen und Abfragesprachen für raumbezogene Informationen. Kirschbaum Verlag, Bonn.
- Frankenberger, J. (1987): Konzept der graphischen Datenverarbeitung in der bayerischen Vermessungsverwaltung, 1. Informationsveranstaltung über die graphische Datenverarbeitung der Bayer. Vermessungsverwaltung, München, Hrsg. Bayer. Landesvermessungsamt, 7-48.
- Frankenberger, J. (1993): Folien über Aufbau und Ausstattung der Bayerischen Vermessungsverwaltung (Handreichung).
- Franz, D. (1991): Statistik - Eine Einführung in die Wahrscheinlichkeitsrechnung, Qualitätskontrolle und Zuverlässigkeit für Techniker und Ingenieure, Hüthig Buch Verlag, Heidelberg.
- Fritsch, D. (1995): Vorlesungen: Ausgleichsrechnung und Statistik I, Institut für Photogrammetrie, Universität Stuttgart, Handreichung für Studenten.
- Fritsch, D. und Anders, K. (1996): Objektorientierte Konzepte in Geo-Informationssystemen, Geo-Informationssysteme (GIS) 9(2), 2-14.

- GEOGIS (1992): Bericht der Projektgruppe GEOGIS, interner Bericht des Bayerischen Landesvermessungsamtes München (nicht veröffentlicht).
- Göpfert, W. (1987): Raumbezogene Informationssysteme, Wichmann Verlag, Karlsruhe.
- Grünreich, D. (1993): Die Basisfunktion des topographischen Informationssystems ATKIS für Umwelt-informationssysteme, 3. Internationales Anwenderforum Duisburg 3.-4. März '93, 497-509.
- Hake, G. (1994): Kartographie, 7. Auflage, Walter de Gruyter Verlag, Berlin - New York .
- Heubach, G. (1992): Eine Untersuchung verschiedener geometrischer Generalisierungsansätze für lineare Objekte, Diplomarbeit am Institut für Photogrammetrie, Universität Stuttgart (unveröffentlicht).
- Kahlert, J. und Frank, H. (1994): Fuzzy-Logik und Fuzzy-Control, 2. Auflage, Vieweg & Sohn Verlagsgesellschaft Braunschweig.
- Kahmen, H. (1993): Vermessungskunde, 18. Auflage, Walter de Gruyter Verlag, Berlin - New York.
- Kaltenbach, H. (1995): Datenmodellierung in GIS - Stand der Entwicklung, erschienen in: Neunte Informationsveranstaltung 1995 der Bayer. Vermessungsverwaltung über die graphische Datenverarbeitung, Bayerisches Landesvermessungsamt München im Auftrag des Bayer. Staatsministeriums der Finanzen, S. 7 - 20.
- Koch, K. R. (1987): Parameterschätzung und Hypothesentests, Dümmler Verlag, Bonn.
- Kopstahl, E. (1994): Überblick über Anwendungen des Geoinformationssystems ATKIS - Datenintegrationskonzept, erschienen in: R. Harbeck, Hrsg., Das Geoinformationssystem ATKIS und seine Nutzung in der Wirtschaft und Verwaltung, Vorträge anlässlich des AdV-Symposiums ATKIS am 15. und 16. Juni 1994, Landesvermessungsamt Nordrhein-Westfalen, 33 - 46.
- Ludwig, R., Hampp, D. und Walk, R. (1996): Aktuelles zum Grundstücks- und Bodeninformationssystem (GRUBIS) in Bayern, erschienen in: Zehnte Informationsveranstaltung 1996 der Bayer. Vermessungsverwaltung über die graphische Datenverarbeitung, Bayerisches Landesvermessungsamt München im Auftrag des Bayer. Staatsministeriums der Finanzen, 81 - 94.
- Meyers großes Taschenlexikon (1995): Band 8, B.I.-Taschenbuchverlag, Mannheim.
- Molenaar, M (1993): Object hierarchies and uncertainty in GIS or why is standardisation so difficult?, Geo-Informationssysteme (GIS), 6(4), 22-28.
- Newman, W. M. /Sproull R. F. (1986): Grundzüge der interaktiven Computergrafik, McGraw-Hill Book Company GmbH Hamburg.
- Rappe, B. (1995): Erfassung und Integration von Geo-Daten aus unterschiedlichen Quellen, erschienen in: G. Buziek, Hrsg., GIS in Forschung und Praxis, Verlag Konrad Wittwer Stuttgart, 123 - 140.
- Reiß, P. (1988): Digitale Geländemodelle in der Bayerischen Landesvermessung, DVW-Mitteilungsblatt Bayern, Heft 1, 1988, 11-35.
- Reiß, P. (1990): Das digitale Geländemodell des Freistaates Bayern - Aufbau, Genauigkeit, Datenabgabe, Anwendungsmöglichkeiten, erschienen in: Vierte Informationsveranstaltung 1990 der Bayer. Vermessungsverwaltung über die graphische Datenverarbeitung, Bayer. Landesvermessungsamt München im Auftrag des Bayer. Staatsministeriums der Finanzen, 83-102.
- Reiß, P. (1996): Luftbildkarten, digitale Orthophotos und Digitales Geländemodell - Basisprodukte der Bayer. Vermessungsverwaltung für den Aufbau von Informationssystemen, erschienen in: Zehnte Informationsveranstaltung 1996 der Bayer. Vermessungsverwaltung über die graphische Datenverarbeitung, Bayerisches Landesvermessungsamt München im Auftrag des Bayer. Staatsministeriums der Finanzen, 21 - 34.
- Roder, B. (1988): Das Geographische Grundinformationssystem des Bayerischen Landesvermessungsamtes, DVW-Mitteilungsblatt Bayern, Heft 2, 1988, 159-178.
- Roder, B., Wimmer, H. (1989): Geographisches Grundinformationssystem (GEOGIS) - Basis für Umweltdokumentation und Planung in Bayern, in: Geo-Informationssysteme, Hrsg.: Schilcher M., Fritsch, D., Wichmann Verlag, Karlsruhe, 293-310.
- Rudolph, D. (1991): Die AutoCAD Programmierung, Dr. L. Rossipaul Verlagsgesellschaft m.b.H., München.
- Sachs, L. (1988): Statistische Methoden: Planung und Auswertung, 6. Auflage, Springer-Verlag Berlin.
- Schellein, H. (1995): GEOGIS 25 - Datenstruktur, Datenerfassung und Datenabgabe, erschienen in: Neunte Informationsveranstaltung 1995 der Bayer. Vermessungsverwaltung über die graphische Datenverarbeitung, Bayerisches Landesvermessungsamt München im Auftrag des Bayer. Staatsministeriums der Finanzen, 21 - 29.

- Schellein, H. (1996): Aktueller Stand der Erfassung und Abgabe der ATKIS-Daten, erschienen in: Zehnte Informationsveranstaltung 1996 der Bayer. Vermessungsverwaltung über die graphische Datenverarbeitung, Bayerisches Landesvermessungsamt München im Auftrag des Bayer. Staatsministeriums der Finanzen, 45 - 49.
- Schellein, H. (1997): Integrierte Modellierung amtlicher Basisinformationssysteme ALB/ALK/ATKIS; ATKIS - Erfahrungen in Bayern, erschienen in der Handreichung zum Münchner Fortbildungsseminar Geoinformationssysteme 1997, Geodätisches Institut der Technischen Universität München.
- Schilcher, M. (1995): Technologiewandel im Vermessungswesen durch Geoinformationssysteme, Mitteilungsblatt des DVW-Bayern 3/1995, 251 - 288.
- Schmidt, H. (1997): Was ist Genauigkeit? - Zum Einfluß systematischer Abweichungen auf Meß- und Ausgleichungsergebnisse, erschienen in: Vermessungswesen und Raumordnung Heft 59/4, 212 - 228.
- Vermessung 2000 (1993): Der Bayerische Weg im Vermessungswesen, Stellungnahme zu Privatisierungsbestrebungen vom 05.03.93, veröffentlicht in: DVW-Mitteilungsblatt des Landesvereins Bayern, Heft 2, 1993, 202-204.
- Walk, R. (1993): Händische Mitschrift aus dem Vortrag vom 18.05.93 im Ausbildungsabschnitt 1 beim Referendarsjahrgang 1993, Elektronische Datenverarbeitung bei der Vermessungsverwaltung.
- Walk, R. (1993, 2): Datenaustauschkonzept der staatlichen Vermessungsämter für das amtliche Grundstücks- und Bodeninformationssystem (GRUBIS), veröffentlicht im DVW-Mitteilungsblatt des Landesvereins Bayern, Heft 4, 1993, 405-448.
- Walter, Volker (1996): Zuordnung von raumbezogenen Daten - am Beispiel der Datenmodelle ATKIS und GDF, Dissertation am Institut für Photogrammetrie der Universität Stuttgart, Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften, München, Reihe C.
- Wimmer, H. (1993): Das Geographische Grundinformationssystem GEOGIS in Bayern, Entwicklung und Erfassungsstand.

9 Bildnachweis

Es wurden folgende Orthophotos (M = 1:10000) mit der Genehmigungsnummer 1770/97 vom 03.04.1997 im Bereich der TK 25 Nr. 7732 für die Veröffentlichung zur Verfügung gestellt:

Nummer	Bildnummer	Bildflugnummer
7732/1	093	95 106/0
7732/2	091	95 106/0
7732/3	089	95 106/0
7732/4	122	95 106/0
7732/5	124	95 106/0
7732/6	126	95 106/0
7732/7	147	95 106/0
7732/8	145	95 106/0
7732/9	143	95 106/0

Anhang

A Datenbereitstellung - Datenimport

Die Dateninformationen der Digitalen Flurkarten werden in Bayern in einem von der Bayerischen Vermessungsverwaltung entwickelten Schnittstellenformat, das in der DatRi-GRUBIS eingehend beschrieben ist, an die Nutzer abgegeben [Bayerisches Staatsministerium der Finanzen, 1993].

In diesem Datenformat wurden auch die DFK-Daten für diese Arbeit zur Verfügung gestellt. Konverter für CAD-Systeme, die dieses Format verarbeiten können, werden inzwischen von privaten Anbietern vertrieben. Zum Zeitpunkt der Datenabgabe für diese Arbeit standen derartige Umsetzer noch kaum zur Verfügung. Die DFK-Daten sollten zunächst in das CAD-System *CARD/I* eingespielt werden. Im folgenden werden die für diese Arbeit wesentlichen Bestandteile der DFK-Schnittstelle erläutert und der relativ einfache Weg der Datenkonvertierung aufgezeigt.

Da die DFK ein digitales Basiskartenwerk darstellt, sollten die Daten über eine möglichst einfache Schnittstelle bereitgestellt werden, die dem Anwender selbst die Möglichkeit gibt, aus den vorgegebenen Daten kompliziertere Strukturen, wie z.B. Objekte, zu entwickeln. Das Schnittstellenformat sollte dabei verschiedene Anforderungen erfüllen [Walk, 1993, 2]:

- Redundanzfreiheit der Daten und damit auch weitgehender Verzicht auf Objektstrukturen,
- Unabhängigkeit von firmenspezifischen Schnittstellenformaten,
- Möglichkeit zum Austausch von Differenzdaten,
- einfache Darstellung der analogen Kartenelemente in digitaler Form,
- Verzicht auf Ebenenstrukturen,
- einfache Möglichkeit zur Weiterverarbeitung der Daten.

Unterteilung des Inhalts einer Flurkarte:

- Linien und Bögen (z.B. Flurstücksgrenzen, Gebäudegrenzen)
- Texte (z.B. Straßennamen, Flurstücksnummern)
- Symbole (z.B. Grenzpunkte, Nutzungsarten)

Ausgehend von dieser Einteilung besteht das DFK-Schnittstellenformat im wesentlichen aus drei verschiedenen Datensätzen, dem Koordinatensatz, dem Liniensatz/Bogensatz und dem Textsatz/Symbolsatz, die wie folgt aufgebaut sind.

Der Koordinatensatz:

Der Koordinatensatz dient dem Austausch von Gauß-Krüger-Koordinaten der Grenz-, Gebäude- und sonstiger Punkte.

Feld	UFeld	Feldbezeichnung	von	bis	Länge	Bemerkung
1		Kennung	1	6	6	
1	1	Vorzeichen	1	1	1	+ Einfügeblock - Löschblock
1	2	Kennzahl	2	6	5	immer 9990
2		Flurkarte	8	15	8	
3		Punktnummer	17	20	4	
4		Rechtswert	22	30	9	die erste Stelle der GK-Koordinate entfällt, Angabe in cm-Einheiten
5		Hochwert	32	40	9	die erste Stelle der GK-Koordinate entfällt, Angabe in cm-Einheiten
6		Abmarkungsart	42	43	2	10 = Genzstein 11 = Grenzstein aus Kunststoff

						12 = Grenznagel
						13 = Grenzrohr
						14 = Meißelzeichen
						15 = unbehauener Feldstein
						16 = Grenzpflock
						17 = Uferlinie
						18 = Grenzpunkt, Abmarkung zurückgestellt
						19 = Gebäudeeckpunkt, Brechpunkt einer Nutzungsartengrenze, Grenzpunkt unabemarkt
						0 = Grenzpunkt, sonstige oder unbekannte Abmarkung
7	Kennung über Art und Entstehung der Koordinaten	45	45	1		4 = d-Koordinate, durch Homogenisierung 5 = d-Koordinate, durch exakten numerischen Ansatz 7 = g-Koordinate, Zuordnung bei Homogenisierung 8 = g-Koordinate, durch exakten numerischen Ansatz
8	Letzter schreibender Zugriff Satzendezeichen	47	47			Zeitpunkt des Eintrags oder der Löschung des Punktes im Koordinatenarchiv, Zeitangabe in Sekunden seit dem 01.01.1970, 00:00:00 Greenwich Mean Time

Der Liniensatz/Bogensatz:

Der Liniensatz/Bogensatz beinhaltet die Punktnummern des Bogens/der Linie und enthält keine Koordinaten. Die Grundrißlinien eines Gebäudes werden aufeinanderfolgend im Uhrzeigersinn dargestellt (einfache Objektstruktur).

Liniensatz:

Feld	UFeld	Feldbezeichnung	von	bis	Länge	Bemerkung
1		Kennung	1	6	6	6000 - 9899
1	1	Vorzeichen	1	1	1	+ Einfügeblock - Löschblock
1	2	Kennzahl	2	6	5	z. B. 6000 = Flurstücksgrenze (Gerade) z. B. 6050 = Grundriß Hauptgebäude (Gerade)
2		Flurkarte Linienanfang	8	15	8	spezielle Verschlüsselung, unterschieden nach g-Koordinaten und d-Koordinaten
3		Punktnummer Linienanfang	17	20	4	
4		Flurkarte Linienende	22	29	8	
5		Punktnummer Linienende Satzendezeichen	31	34	4	

Bogensatz:

Feld	UFeld	Feldbezeichnung	von	bis	Länge	Bemerkung
1		Kennung	1	6	6	6000 - 9899
1	1	Vorzeichen	1	1	1	+ Einfügeblock - Löschblock
1	2	Kennzahl	2	6	5	z. B. 6001 = Flurstücksgrenze (Bogen) z. B. 6051 = Grundriß Hauptgebäude (Bogen)
2		Flurkarte Bogenanfang	8	15	8	spezielle Verschlüsselung, unterschieden nach g-Koordinaten und d-Koordinaten
3		Punktnummer Bogenanfang	16	20	4	
4		Flurkarte Bogenende	22	29	8	
5		Punktnummer Bogenende	31	34	4	
6		Radius	36	41	6	
6	1	Vorzeichen	36	36	1	+ Mittelpunkt rechts vom Bogenanfangspunkt

6	2	Betrag	37	41	5	- Mittelpunkt links vom Bogenanfangspunkt in cm-Einheiten, max. Zentriwinkel 200 gon
---	---	--------	----	----	---	---

Textsatz/Symbolsatz:

Der Textsatz/Symbolsatz dient zum Datenaustausch von Texten und Symbolen. Beide haben den gleichen Aufbau; die Trennung zwischen beiden findet über die Kennzahl des Datensatzes statt.

Feld	UFeld	Feldbezeichnung	von	bis	Länge	Bemerkung
1		Kennung	1	6	6	2000 - 2999 für M = 1:1000, Kartenrand 3000 - 3999 für M = 1:5000, Kartenrand 4000 - 4999 für M = 1:1000 5000 - 5999 für M = 1:5000
1	1	Vorzeichen	1	1	1	+ Einfügeblock - Löschblock
1	2	Kennzahl des Datensatzes	2	6	5	für Symbolsatz in 1:1000 immer 4450 für Symbolsatz in 1:5000 immer 5450 für Symbolsatz für Kartenrand in 1:1000 immer 2450 für Symbolsatz für Kartenrand in 1:5000 immer 3450 Textsatz z.B. 4001 Hausnummer
2		Flurkarte	8	15	8	spezielle Verschlüsselung
3		Rechtswert der tatsächlichen Text-/Symbolposition	17	25	19	die erste Stelle der GK-Koordinate entfällt, Angabe in cm-Einheiten
4		Hochwert der tatsächlichen Text-/Symbolposition	27	35	9	die erste Stelle der GK-Koordinate entfällt, Angabe in cm-Einheiten
5		Differenz Rechtswert zum Text-/Symbolbezugspunkt	37	42	6	in cm-Einheiten
5	1	Vorzeichen	37	37	1	
5	2	Differenzbetrag	38	42	5	
6		Differenz Hochwert zum	44	49	6	in cm-Einheiten
6	1	Vorzeichen	44	44	1	
6	2	Differenzbetrag	45	49	5	
7		Richtungswinkel des Textes/Symbols	51	57	7	
8		Textinhalt/Kennzahl d. Symbols	59		256	

Zu Beginn jeder DFK-Datei steht als erste Zeile der Startsatz, der allgemeine Angaben zum Datenaustausch und zum Datenblocks enthält.

Feld	UFeld	Feldbezeichnung	von	bis	Länge	Bemerkung
1		Kennung	1	6	6	
1	1	Vorzeichen	1	1	1	
1	2	Kennzahl	2	6	5	immer 9999
2		Kennung d. abgebenden Stelle	8	12	5	Eindeutige Kennung z. B. des Vermessungsamtes
3		Kennung f. Datenaustauschverfahren	14	15	2	z. B. KG = Komplettaustausch DFK z. B. DG = Differenzaustausch DFK
4		Zeit der Datenausspielung	17	26	9	Zeitangabe in Sekunden seit dem 01.01.1970 00:00:00 Greenwich Mean Time
5		Minimaler Rechtswert	28	36	9	eines Rechteckfensters
6		Minimaler Hochwert	38	46	9	

7		Maximaler Rechtswert	48	56	9	
8		Maximaler Hochwert	58	66	9	
		Satzendezeichen				

Dem Startsatz folgen immer mindestens drei, meist vier Gebietssätze, die das aus der DFK ausgelesene Gebiet durch Angabe der Eckpunktkoordinaten beschreiben. Bei Abgabe im Flurkartenblattschnitt sind dies die vier Blattecken der Flurkarte, bei blattschnittfreier Abgabe ein beliebiges Polygon.

Feld	UFeld	Feldbezeichnung	von	bis	Länge	Bemerkung
1		Kennung	1	6	6	
1	1	Vorzeichen	1	1	1	
1	2	Kennzahl	2	6	5	immer 9998
2		Rechtswert	8	16	9	die erste Stelle der GK-Koordinate entfällt, Angabe in cm-Einheiten
3		Hochwert	18	26	9	die erste Stelle der GK-Koordinate entfällt, Angabe in cm-Einheiten
		Satzendezeichen				

Im Anschluß an die Gebietssätze folgen die Koordinatensätze, die Liniensätze/Bogensätze und dann die Textsätze/Symbolsätze. Am Ende jeder DFK-Datei steht der Schlußsatz mit der Kennung 1.

Feld	UFeld	Feldbezeichnung	von	bis	Länge	Bemerkung
1		Kennung	1	6	1	
1	1	Vorzeichen	1	1	1	
1	2	Kennzahl	2	6	5	immer 1
		Satzendezeichen				

Die Felder der verschiedenen Sätze sind jeweils durch \$-Zeichen voneinander getrennt.

Beispiel für eine DFK-Datei:

```
+ 9999$00084$KG$ 770379986$ 44965778$ 33546267$ 45012737$ 33593229
+ 9998$ 44965778$ 33546529
+ 9998$ 44966040$ 33593229
+ 9998$ 45012737$ 33592967
+ 9998$ 45012476$ 33546267
+ 9990$20010800$3626$ 45003499$ 33559477$ 0$7$ 770379986
+ 6000$20010800$3626$20010800$3670
+ 6020$20010801$ 300$20010801$ 497
+ 6400$20010831$ 10$20010831$ 152
+ 2050$20010802$ 45035956$ 33569487$-23865$+20423$100.000$35.9
+ 4001$20010801$ 44966388$ 33577459$+ 0$+ 0$338.755$32
+ 4301$20010801$ 44967599$ 33557550$+ 329$- 70$100.000$8497-2070
+ 4450$20010801$ 44966223$ 33559463$+ 0$+ 0$223.973$60
+ 1
```

Arbeitsablauf zur Umsetzung der Daten:

Zunächst wurden die Daten der DFK-Differenzschnittstelle in ein für das CAD-System CARD/1 lesbares Format gebracht. Hierbei wurden aus jeder *.DFK - Datei drei CARD/1-ASCII-Dateien erzeugt, die wie folgt aufgebaut sind.

Die Koordinatendatei *.ASC:

Die CARD/1-ASCII-Koordinatendateien dienen zum Transportieren von Vermessungspunkten zwischen Rechnern und Programmen.

Format: Ein Punktdatensatz pro Zeile

Bezeichnung	1. Spalte	Format	Bemerkung
Punktnummer	1	A12	linksbündig
Rechtswert	13	F12.4	
Hochwert	25	F12.4	
Flags	49	I5	
Herkunft	54	I2	
Punktart	56	I4	
Vermarkungsart	60	I2	
Frei	65	I3	
Bemerkung	68	A12	

Die Spalten „Höhe“ bis „Bemerkung“ sind optional.

Die Polygondatei *.POL:

Die CARD/1-ASCII-Polygondateien dienen zum Transportieren von Polygonen zwischen Rechnern und Programmen. Da die Datei nur Punktnummern enthält, kann sie nur mit der Punktdatenbank oder mit einer CARD/1-ASCII-Koordinatendatei verwendet werden.

Format: Ein Polygon besteht aus einer Reihe von Punktnummern (min. 2, max. 100). Jede Punktnummer steht in der Datei in einer extra Zeile. Leerzeilen trennen Polygone von-einander.

Bezeichnung	1. Spalte	Format	Bemerkung
Punktnummer	1	A12	linksbündig
Linienart	13	I4	
Elementtyp	17	I4	für Verbindung zum vorhergehenden Punkt 1 = Gerade 2 = Linksbogen < Halbkreis 3 = Linksbogen > Halbkreis 4 = Rechtsbogen < Halbkreis 5 = Rechtsbogen > Halbkreis 8 = Kurve 9 = Kurvenendpunkt
Elementattribute	21	I4	1 = Linienunterbrechung
Radius	25	F10.3	für Elementtyp 2.....5 ohne Vorzeichen

Die Textdatei *.TXA:

Die CARD/1-ASCII-Textdateien dienen zum Transportieren von Texten zwischen Rechnern und Programmen.

Format: Ein Textdatensatz hat zwei Zeilen

Bezeichnung	1. Spalte	Format
Y-Koordinate	1	F12.3
X-Koordinate	13	F12.3
rel. Y-Verschiebung	25	F6.2
rel. X-Verschiebung	31	F6.2
Textbreite	37	F6.2
Texthöhe	43	F6.2
Winkel	49	F7.2
Neigung	56	F7.2
Textart	63	I4
Schriftart	67	I4
Attributcode	71	I6
Text	1	A80

Die Dateien wurden mittels eines kleinen C-Programmes in die für CARD/1 lesbaren Dateien umgewandelt. Dabei war zu beachten, daß die Punktnummern nur im Zusammenhang mit der Angabe der Flurkarte eindeutig sind, so daß bei der Konvertierung für jede Flurkarte ein eigener Kennbuchstabe vor die Punktnummer gesetzt wurde, damit die Flurkarten in einem gemeinsamen Projekt bearbeitet werden können. Da die DFK-Daten ausschließlich als Referenzdatenbestand verwendet werden und keine Weiterbearbeitung in dieser Arbeit nötig war, genügte es nur zwei Schichten (Layer, Ebenen) zu generieren: die Schicht „Linien“ für die Informationen der *.POL - Dateien und die Schicht „Texte“ für die Informationen der *.TXA - Dateien.

Die Daten aus ATKIS-Bayern wurden vom Bayerischen Landesvermessungsamt im DXF-Format zur Verfügung gestellt. Zur Konvertierung von DXF-Dateien hat CARD/1 ein eigenes Modul, so daß dies keine größeren Schwierigkeiten bereitete. Es wurde für jeden Objektbereich eine eigene Schicht angelegt. Auch für die DFK wird in nächster Zukunft die DXF-Schnittstelle für den Datenaustausch mit den Nutzern zur Verfügung stehen.

Obwohl sowohl die DFK- als auch die ATKIS-Bayern-Daten in GK-Koordinaten vorliegen, können beide Datenbestände nicht ohne weiteres zusammengespielt werden, da den GK-Koordinaten der DFK die erste Stelle (1000 km-Stelle) sowohl in y- als auch in x-Richtung fehlt. Daher wurden den ATKIS-Bayern-Daten die erste Vorkommastelle gekappt und dann beide Datensätze zusammengespielt.

B Mathematische Grundlagen von fehlerhaften Geometriedaten - grundsätzliche fehlertheoretische Modelle

B.1 Fehlerfortpflanzung

Bei der Auswertung von Meßprozessen spielt die Zufallsvariable eine wichtige Rolle, da man jede Messung als Realisierung einer Zufallsvariablen auffassen kann. Die Realisierung eines derartigen Experiments kann entweder diskrete Werte (z. B. beim Würfeln = ganze Zahlen zwischen 1 und 6) oder aber, wie in der Vermessung meistens vorkommt, unendlich viele Werte annehmen. Daher unterscheidet man **diskrete** und **kontinuierliche** (stetige) Zufallsvariablen.

Definition 1: Man bezeichnet L als diskrete Zufallsvariable, wenn sie lediglich endlich viele oder abzählbar unendlich viele Werte annimmt. Sind l_1, l_2, \dots, l_i diese Werte und $f(l_1), f(l_2), \dots, f(l_i)$ ihre Wahrscheinlichkeiten, so nennt man $f(l_i)$ die Dichte oder die Verteilung von L .

Definition 2: Man bezeichnet L als stetige Zufallsvariable, wenn eine nichtnegative integrierbare Funktion $f(l)$ existiert, die für ein beliebiges $l \in \mathfrak{R}$ die Beziehung $F(l) = \int_{-\infty}^l f(t) dt$ erfüllt, wobei $F(l)$ die Verteilungsfunktion von L und t eine Integrationsvariable bedeuten. Die Funktion $f(l)$ heißt Dichte (stochastische Verteilung) oder auch einfache univariate Verteilung.

Definition 3: Als Erwartungswert (1. Moment) einer Zufallsvariablen $L: S \rightarrow \mathfrak{R}$ bezeichnet man den **Mittelwert** μ oder $E(L)$, definiert durch:

$$\mu = E(L) := \int_{-\infty}^{+\infty} l \cdot f(l) dl$$

für eine kontinuierliche Zufallsvariable, für den Fall, daß das Integral existiert.

Damit folgt für einen $n \times 1$ Zufallsvektor $\mathbf{l} = [L_1, \dots, L_n]^T$ mit der Dichte $f(l_1, \dots, l_n)$ der Erwartungswert von $L_i (i \in \{1, \dots, n\})$ durch:

$$\begin{aligned} \mu_i = E(L_i) &:= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} L_i(s) f(l_1, \dots, l_n) dl_1 \dots dl_n \\ &= \int_{-\infty}^{+\infty} l_i g(l_i) dl_i \end{aligned}$$

mit Hilfe der Randdichte $g(l_i)$ definiert, sofern das Integral existiert.

Definition 4: Momente, die auf den Zufallsvektor $\mathbf{l} - E(\mathbf{l}) := [L_i - \mu_i, \dots, L_n - \mu_n]^T$ bezogen werden, heißen **zentrale Momente**. Speziell bezeichnet man das 2. zentrale Moment der beiden Zufallsvariablen L_i und L_j als **Kovarianz**.

$$\begin{aligned} \sigma_{ij} = C(L_i, L_j) &:= E((L_i - \mu_i)(L_j - \mu_j)) = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (l_i - \mu_i)(l_j - \mu_j) g(l_i, l_j) dl_i dl_j, \end{aligned}$$

und das 2. zentrale Moment der Zufallsvariablen L_i allein als **Varianz** oder **Dispersion**:

$$\begin{aligned} \sigma^2 = V(L_i) = D(L_i) &:= E((L_i - \mu_i)^2) = \\ &= \int_{-\infty}^{+\infty} (l_i - \mu_i)^2 g(l) dl_i. \end{aligned}$$

Insgesamt faßt man das 2. zentrale Moment in der $n \times n$ **Kovarianzmatrix** oder der **Dispersionsmatrix** zusammen.

$$D(\mathbf{l}) = (\sigma_{ij}) = C(L_i, L_j) = E((\mathbf{l} - E(\mathbf{l}))(\mathbf{l} - E(\mathbf{l}))^T).$$

Satz: Ist \mathbf{l} ein $n \times 1$ Zufallsvektor mit der Kovarianzmatrix $D(\mathbf{l})$ und A eine $m \times n$ Matrix, sowie \mathbf{c} ein $m \times 1$ Vektor von Konstanten, so ergibt sich die Kovarianzmatrix des linear transformierten $m \times 1$ Zufallsvektors $\mathbf{y} := A\mathbf{l} + \mathbf{c}$ nach dem **Fehlerfortpflanzungsgesetz**:

$$D(\mathbf{y}) = D(A\mathbf{l} + \mathbf{c}) = AD(\mathbf{l})A^T$$

Bei nichtlinearen Funktionen $y: L(S) \rightarrow \mathfrak{R}$ wird diese zunächst um einen genäherten Erwartungswert $E(\mathbf{l})_0 = [E(L_1)_0, \dots, E(L_n)_0]^T$ nach Taylor entwickelt. Damit ergibt sich:

$$\begin{aligned} y(\mathbf{l}) &= y(\mathbf{l}_0) + \frac{\partial y}{\partial l_1} \Big|_{l_1=E(L_1)_0} \cdot (l_1 - E(L_1)_0) + \dots + \frac{\partial y}{\partial l_n} \Big|_{l_n=E(L_n)_0} \cdot (l_n - E(L_n)_0) \\ &= y(\mathbf{l}_0) + y' \cdot \Delta x \end{aligned}$$

wobei $y' := \frac{\partial y}{\partial a_j} \Big|_{l_j} = E(l_j)_0$ die entsprechende $1 \times j$ Zeile einer JACOBI-Matrix und $\Delta l := (l_i - E(L_i)_0)$ der zugehörig $n \times 1$ Zufallsvektor von kleinen Differenzen ist.

[Fritsch, 1995].

B.2 Fuzzy-Datenmodelle

Fuzzy-Logik bietet die Möglichkeit unscharfe Informationen zu beschreiben. Dazu sollen zunächst die Begriffe „unscharf“ und „Unschärfe“ geklärt werden. Wir sind es im Alltag gewohnt mit unscharfen Informationen umzugehen. Dabei sind drei wesentliche Arten der Unschärfe zu unterscheiden.

Stochastische Unschärfe:

Die Wahrscheinlichkeit für das Eintreten eines Ereignisses wird in Zahlen zwischen 0 und 1 ausgedrückt. Wahrscheinlichkeiten werden nach bestimmten mathematischen Verfahren berechnet oder auch geschätzt. Sie haben nicht zwingend einen Bezug zum tatsächlichen Eintreten des Ereignisses (auch nach zweimaligem Werfen einer Münze muß noch kein „Kopf“ erschienen sein). Wahrscheinlichkeiten täuschen damit oft scharfe Realitäten nur vor.

Sprachliche Unschärfe:

Sprachliche Unschärfe drückt sich z. B. in den Begriffen wie „schönes Wetter“, „große Häuser“, „instabile Währung“ usw. aus. Zwar können solche sprachlichen Unschärfen im Zusammenhang zu Begriffsdefinitionen konzentriert werden. Ein Computer jedoch kann solche Zusammenhänge nicht erkennen und damit auch nicht weiterverarbeiten.

Informale Unschärfe:

Unter informaler Unschärfe versteht man Unschärfen, die aufgrund fehlender Informationen oder mangels Wissen entstehen. Meist sind diese Informationen dann auch nicht ohne weiteres zu beschaffen evtl. auch gar nicht zu bekommen, wie z. B. Kreditwürdigkeit, Wettervorhersagen oder dergleichen.

In der Mathematik, den Natur- und Ingenieurwissenschaften waren bisher unscharfe Formulierungen unerwünscht, jedoch unvermeidbar. Es waren kaum geeignete Mittel bekannt, um diese zu verarbeiten. Daß sie aber vorhanden waren, zeigt sich im kleinen z.B. bei Grenzwertbetrachtungen in der Mathematik. Dort betrachtet man z.B. das Verhalten einer Funktion für "sehr große Zahlen". Dies zeigt, daß in bestimmten Fällen sich auch in den "exakten" Wissenschaften geeignete Hilfsmittel zur Handhabung von Unschärfen zurechtgelegt haben. Es gilt also nur noch geeignete Hilfsmittel zu finden, die solche unscharfe Betrachtungen im Bereich der Ingenieurwissenschaften auch zulassen und ihre Weiterverarbeitung erlauben. Diese Möglichkeiten bietet die Fuzzy-Logik, die erstmals bereits 1965 von L. A. Zadeh entworfen und publiziert wurde. Inzwischen wird seine Theorie erfolgreich in der ingenieurtechnischen Praxis eingesetzt.

Ein wesentliches Kennzeichen der "strengen" Logik ist das klassische Zweiwertigkeitsprinzip, d.h. eine Information gehört immer zu einem der drei folgenden Paare "wahr \leftrightarrow falsch", "ja \leftrightarrow nein" oder binär ausgedrückt, ist "1 \leftrightarrow 0". Um unscharfe Informationen in computergesteuerten Prozessen verarbeiten zu können, müssen die unscharfen Informationen in diese computergerechte Form überführt werden, sie müssen modelliert werden. Was hiermit gemeint ist, soll an einem einfachen Beispiel kurz erläutert werden.

Betrachtet wird die Menge der Zahlen, die "viel größer als 1" sind. Eine "klassische" Modellierung könnte wie folgt aussehen:

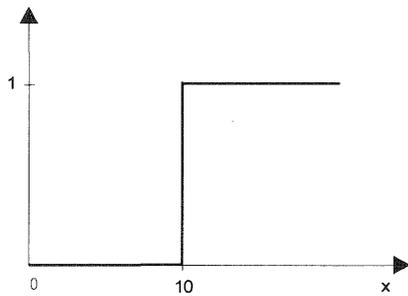


Abbildung B.1: Klassische Modellierung der Menge der Zahlen „viel größer als 1“

Folgerung: $x = 10$ ist viel größer als 1
 $x = 9.9$ ist nicht viel größer als 1

Diese „harte“ Trennung an der Stelle 10 widerspricht unserem Empfinden. Abhilfe bringt hier die Modellierung als unscharfe Menge, wo der Übergang fließend formuliert werden kann. Die könnte wie folgt geschehen:

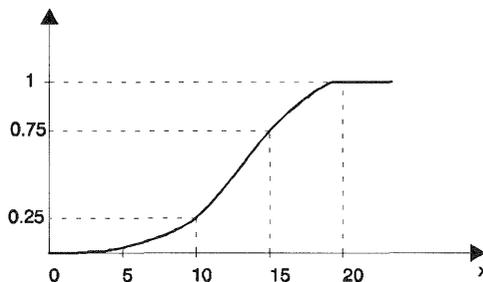


Abbildung B.2: Modellierung mit einer unscharfen Menge

Folgerung: Die Werte $x \leq 1$ erfüllen das Kriterium „viel größer als 1“ sicher nicht, entsprechend ist das Kriterium für $x \geq 20$ sicherlich erfüllt. Dazwischen existiert aber ein fließender Übergangsbereich, in dem die Werte mehr oder weniger zur Menge gehören.

Wie hier in diesem Beispiel gezeigt, sind für Fuzzy-Mengen auch Werte zwischen 0 und 1 zulässig. Ohne nähere Begründung sei hier darauf hingewiesen, daß in der Fuzzy-Theorie immer „normale“ Mengen zugrundegelegt werden. Eine Menge ist dann eine normale Menge, wenn sie mindestens einen Wert hat, für den gilt $\mu(x) = 1$.

Theoretisch ist jede Funktion, die das Intervall $[0;1]$ als Wertemenge hat, für eine Modellierung denkbar. In der Regel werden jedoch folgende Möglichkeiten der Modellierung (auch kombiniert) besonders häufig eingesetzt, da sie mathematisch leicht zu formulieren und für die meisten Anwendungen geeignet sind. Auf spezielle Variationsmöglichkeiten soll hier nicht eingegangen werden.

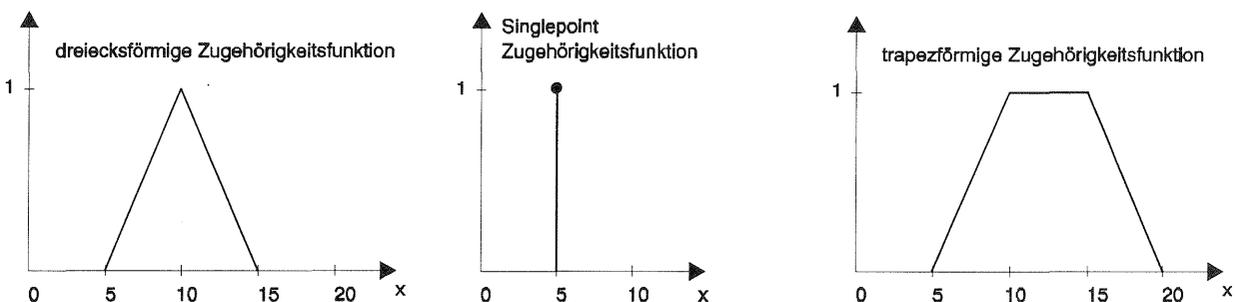


Abbildung B.3: Verschiedene Möglichkeiten für Zugehörigkeitsfunktionen

Ohne mathematische Definition werden im folgenden die Begriffe Einflußbreite und Toleranz anhand einer Zeichnung erklärt, da diese bei der Modellierung von unscharfen Mengen von Bedeutung sind.

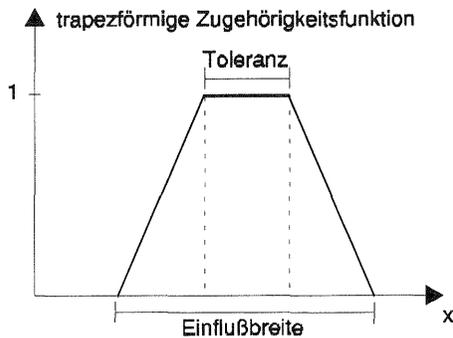


Abbildung B.4: Einflußbreite und Toleranz

An einem einfachen Beispiel soll eine fuzzytypische Modellierung gezeigt werden. Als zu modellierende Kenngröße wird in diesem Beispiel die Temperatur im Bereich von 0 - 100°C verwendet. Die typischen Bezeichnungen, um die Kenngröße Temperatur (linguistische Variable) im Alltagsbereich zu beschreiben lauten: „hoch, niedrig, lauwarm, kalt...“. Um eine Kenngröße zu modellieren, bedient man sich sogenannter linguistischer Terme, die in diesem Beispiel mit „sehr niedrig“, „niedrig“, „mittel“, „hoch“ und „sehr hoch“ gewählt und wie folgt modelliert werden:

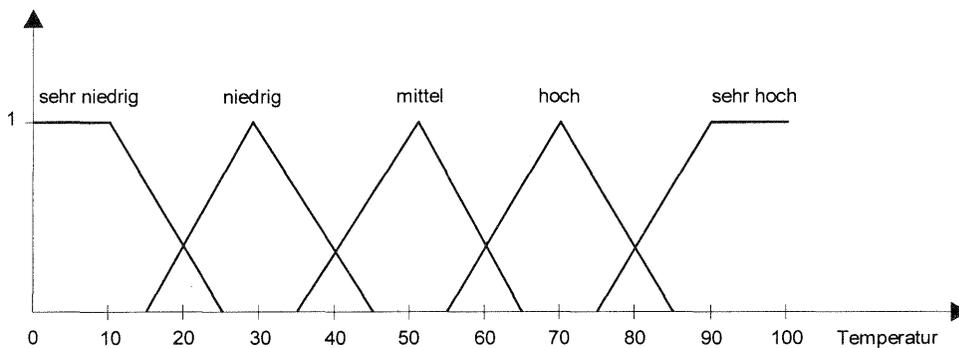


Abbildung B.5: Satz von linguistischen Termen für die linguistische Variable Temperatur

Obiges Bild zeigt eine häufig verwendete Modellierung von Fuzzy-Mengen, bei der der Rand des Wertebereiches einer linguistischen Variable durch trapezförmige Fuzzy-Mengen abgedeckt wird. Für den Zwischenbereich werden häufig symmetrische dreiecksförmige Fuzzy-Mengen verwendet. Die Anzahl der linguistischen Terme hängt vom Anwendungsfall ab, bewegt sich aber im Regelfall zwischen 2 und 7. Mit steigender Anzahl nimmt die Unschärfe ab, dies bedeutet aber für die spätere Realisierung einen erheblich größeren Aufwand. Variationsmöglichkeiten bieten sich durch die Einflußbreite und dem Überlappungsbereich der einzelnen Mengen. In Bereichen, in denen kleine Abweichungen einen großen Einfluß haben (eine hohe Auflösung erforderlich ist), wählt man Zugehörigkeitsfunktionen mit geringer Einflußbreite. Außerdem ist es nicht sinnvoll die Einflußbreite kleiner zu wählen als ein zu erwartender Meßfehler sein kann.

Der nächste Schritt ist die sogenannte **Fuzzifizierung**. Damit bezeichnet man den Übergang vom scharfen Signalwert T auf den zugehörigen Fuzzy-Signalwert T^* .

$$T^* = (\mu_{\text{sehr_niedrig}}(T), \mu_{\text{niedrig}}(T), \mu_{\text{mittel}}(T), \mu_{\text{hoch}}(T), \mu_{\text{sehr_hoch}}(T)).$$

Das fuzzifizierte Signal ist ein n -Tupel von Zugehörigkeitsgraden, wobei n die Anzahl der linguistischen Terme auf dem Signalkanal angibt.

Im Beispiel der Temperatur liefert die Fuzzifizierung folgendes Ergebnis:

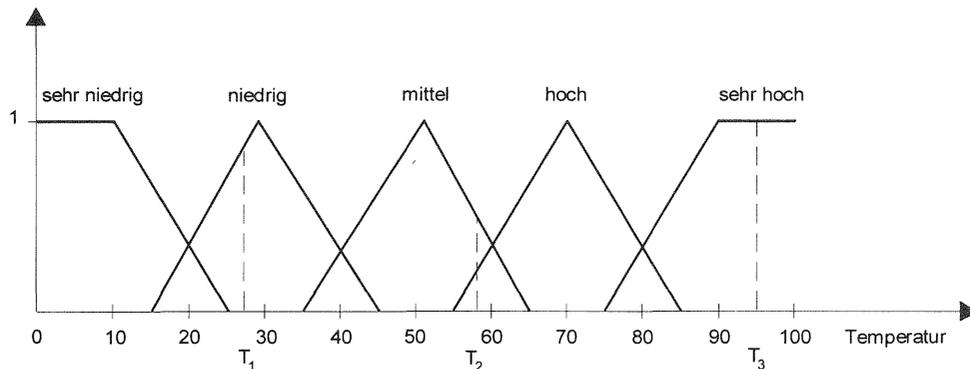


Abbildung B.6: Fuzzifizierung scharfer Temperaturwerte

$$T_1 = 28^\circ \xrightarrow{\text{Fuzzifizierung}} T_1^* = (0, 0,9, 0, 0, 0)$$

$$T_2 = 58^\circ \xrightarrow{\text{Fuzzifizierung}} T_2^* = (0, 0, 0,5, 0,2, 0)$$

$$T_3 = 95^\circ \xrightarrow{\text{Fuzzifizierung}} T_3^* = (0, 0, 0, 0, 0, 1)$$

Man ist es gewohnt, Informationen durch UND und ODER miteinander zu verknüpfen.

In der klassischen Mengenlehre wird die Menge aller Elemente, die zu einer Menge M_1 **und** einer Menge M_2 gehören als **Schnittmenge** beider Mengen bezeichnet.

$$M_1 = \{x | x \in \mathfrak{R}, 1 \leq x \leq 4\} \quad \text{und} \quad M_2 = \{x | x \in \mathfrak{R}, 2 \leq x \leq 5\}$$

$$M_1 \cap M_2 = \{x | x \in \mathfrak{R}, 2 \leq x \leq 4\}$$

Als **Vereinigungsmenge** bezeichnet man die Menge aller Elemente, die zur Menge M_1 **oder** zur Menge M_2 gehören.

$$M_1 = \{x | x \in \mathfrak{R}, 1 \leq x \leq 4\} \quad \text{und} \quad M_2 = \{x | x \in \mathfrak{R}, 2 \leq x \leq 5\}$$

$$M_1 \cup M_2 = \{x | x \in \mathfrak{R}, 1 \leq x \leq 5\}$$

Diese Definitionen können auf Fuzzy-Mengen übertragen werden.

Definition 1: Seien μ_1 und μ_2 zwei Fuzzy-Mengen auf der Grundmenge G , dann heißt:

$$\mu_1 \cap \mu_2 : G \rightarrow [0, 1] \quad \text{mit} \quad (\mu_1 \cap \mu_2)(x) := \text{MIN}(\mu_1(x), \mu_2(x))$$

der **Durchschnitt** der Fuzzy-Menge μ_1 und μ_2 . Hierfür soll im folgenden zur Modellierung die Verknüpfung

$$\mu_1 \text{ UND } \mu_2$$

herangezogen werden.

Definition 2: Seien μ_1 und μ_2 zwei Fuzzy-Mengen auf der Grundmenge G , dann heißt:

$$\mu_1 \cup \mu_2 : G \rightarrow [0, 1] \quad \text{mit} \quad (\mu_1 \cup \mu_2)(x) := \text{MAX}(\mu_1(x), \mu_2(x))$$

die **Vereinigung** der Fuzzy-Menge μ_1 und μ_2 . Hierfür soll im folgenden zur Modellierung die Verknüpfung

$$\mu_1 \text{ ODER } \mu_2$$

herangezogen werden.

Veranschaulicht werden diese Operationen an linguistischen Termen *einer* linguistischen Variablen.

Die Verknüpfung kann über UND oder ODER erfolgen. Die UND-Verknüpfung erfolgt mit dem MIN-Operator und die ODER-Verknüpfung mit dem MAX-Operator.

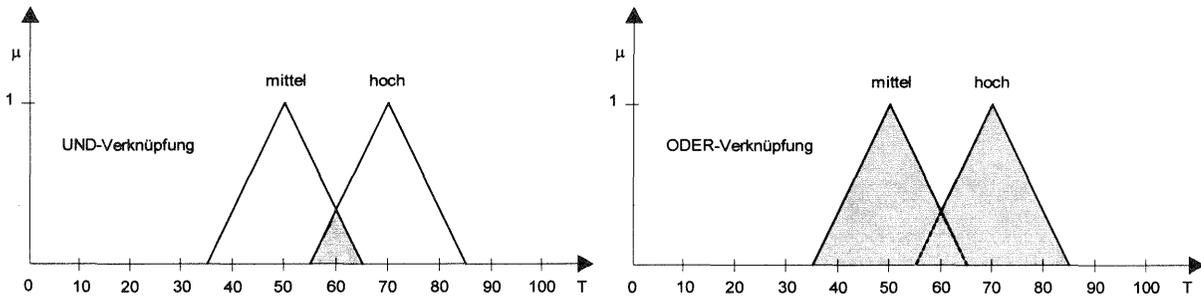


Abbildung B.7:
UND-Verknüpfung der Fuzzy-Mengen „mittel“ und „hoch“ der Temperatur T

ODER-Verknüpfung der Fuzzy-Mengen „mittel“ und „hoch“ der Temperatur T

Der nächste Schritt ist eine Verarbeitungsvorschrift in Form von WENN... DANN... - Regeln aufzustellen. Eine solche Verarbeitungsvorschrift bezeichnet man als **Fuzzy-Inferenz**, die eine Schlußfolgerung als Ergebnis hat. Für die Modellierung von solchen Regeln kann, wie bei der UND-Verknüpfung, der MIN-Operator verwendet werden.

Dies soll wieder am Beispiel der Temperatur verdeutlicht werden, die hier um die Wärmezufuhr in % erweitert wird.

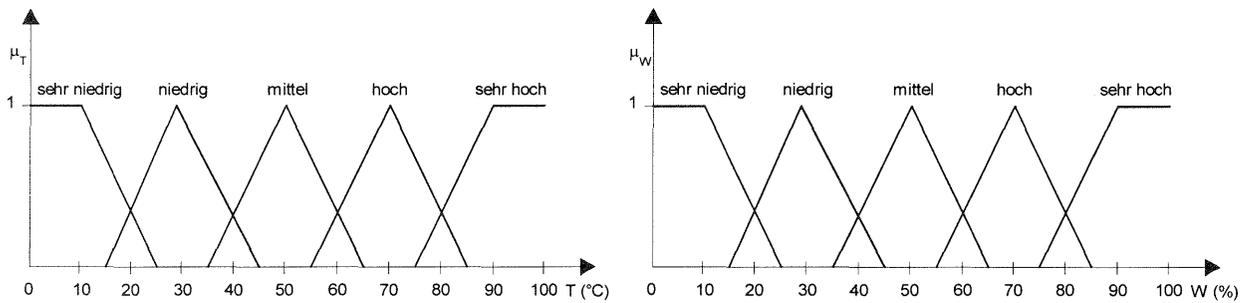
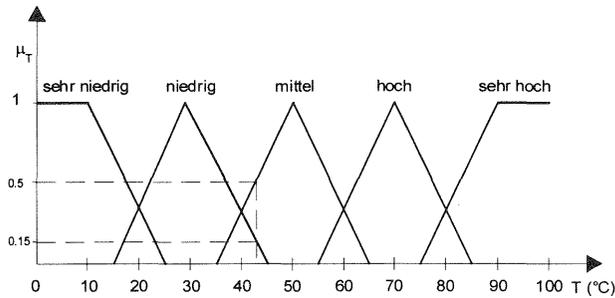


Abbildung B.8: Linguistische Terme für Temperatur T und Wärmezufuhr W

Es werden folgende fünf Regeln in einer **Regelbasis** aufgestellt:

- R_1 : WENN Temperatur T = *sehr niedrig* DANN Wärmezufuhr W = *sehr hoch*
- R_2 : WENN Temperatur T = *niedrig* DANN Wärmezufuhr W = *hoch*
- R_3 : WENN Temperatur T = *mittel* DANN Wärmezufuhr W = *mittel*
- R_4 : WENN Temperatur T = *hoch* DANN Wärmezufuhr W = *niedrig*
- R_5 : WENN Temperatur T = *sehr hoch* DANN Wärmezufuhr W = *sehr niedrig*

Nun soll für einen scharfen Temperaturwert von $T = 43 \text{ }^\circ\text{C}$ eine geeignete Wärmezufuhr ermittelt werden. Die Fuzzifizierung ergibt:



$$T^* = \begin{pmatrix} \mu_{T \text{ sehr_niedrig}}(T) \\ \mu_{T \text{ niedrig}}(T) \\ \mu_{T \text{ mittel}}(T) \\ \mu_{T \text{ hoch}}(T) \\ \mu_{T \text{ sehr_hoch}}(T) \end{pmatrix} = \begin{pmatrix} 0 \\ 0.15 \\ 0.5 \\ 0 \\ 0 \end{pmatrix}$$

Abbildung B.9: Fuzzifizierung der scharfen Eingangsgröße

Hieraus ist erkennbar, daß für den Eingangswert von $T = 43 \text{ }^\circ\text{C}$ nur die Regeln R_2 und R_3 **aktiv** sind,
 d. h. der WENN-Teil der von R_2 ist zu $\mu_{T \text{ niedrig}}(T) = 0.15$ erfüllt, der **Erfüllungsgrad** ist $H_2 = 0.15$
 der WENN-Teil der von R_3 ist zu $\mu_{T \text{ mittel}}(T) = 0.5$ erfüllt, der **Erfüllungsgrad** ist $H_3 = 0.5$.

Im nächsten Schritt müssen die einzelnen Ausgangs-Fuzzy-Mengen ermittelt werden. Jede aktive Regel liefert über das Inferenzschema eine resultierende Ausgangs-Fuzzy-Menge, indem der Erfüllungsgrad der Regel auf die jeweilige Fuzzy-Menge übertragen wird. Dazu wird das Minimum von Erfüllungsgrad und Ausgangs-Fuzzy-Menge

$$\text{MIN}(H_i, \mu_{w_i}(W))$$

gebildet, d. h. die Ausgangs-Fuzzy-Menge wird in der Höhe H_i abgeschnitten.

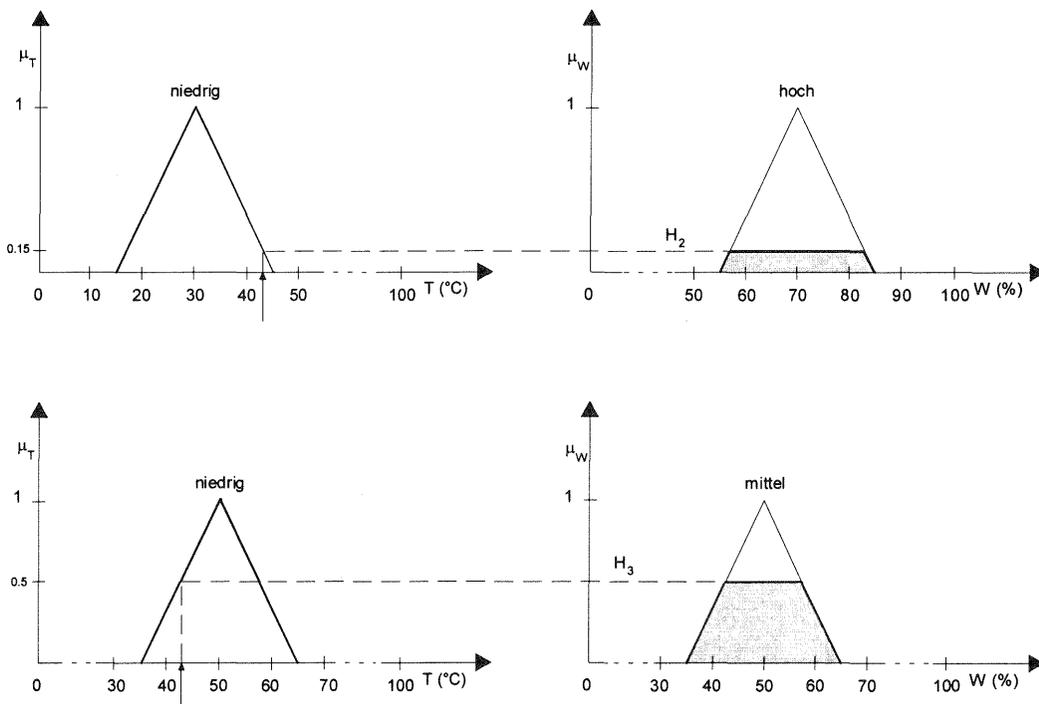


Abbildung B.10: Auswertung der Regeln R_2 (oben) und R_3 (unten)

Da die einzelnen Mengen implizit über ODER miteinander verknüpft sind, müssen die zugehörigen Ergebnis-Fuzzy-Mengen über den MAX-Operator zur resultierenden Ausgangs-Fuzzy-Menge vereinigt werden:

$$\mu_{W \text{ res}} (W) = \text{MAX}_{i=1, \dots, 5} \left(\text{MIN} (H_i, \mu_{W_i} (W)) \right).$$

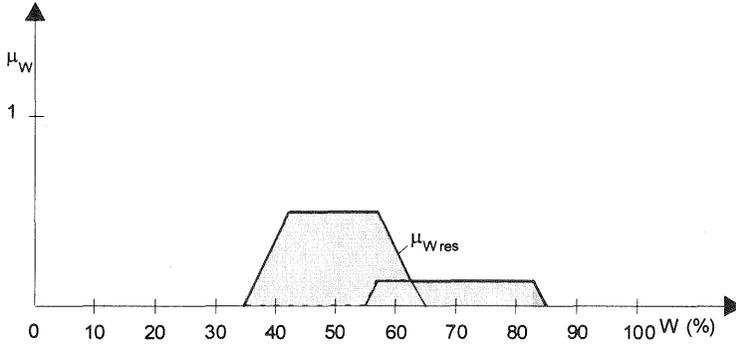


Abbildung B.11: Überlagerung der einzelnen Ausgangs-Fuzzy-Mengen über den MAX-Operator

Der letzte Schritt ist die sogenannte **Defuzzifizierung**. Hierfür wird in der Literatur eine Vielzahl von Möglichkeiten angegeben. Hier wird nur das gebräuchlichste - die **Schwerpunktmethode** („Center of Gravity“-Methode) vorgestellt. Bei der Schwerpunktmethode wird der Flächenschwerpunkt der Ausgangs-Fuzzy-Menge μ_{res} über die Ausgangsgrößen gebildet und die Abszisse als Ausgangsgröße y_{res} bestimmt. Diesem Verfahren liegt die exakte Formel

$$y_{res} = \frac{\int_0^{\infty} y \mu_{res}(y) dy}{\int_0^{\infty} \mu_{res}(y) dy}$$

zugrunde. Meist ist jedoch die Näherungsformel ausreichend, bei der die Abszissen y_i der Schwerpunkte der Ausgangsmengen aller m Regel, die dreiecks- oder trapezförmig sein sollten, in einer mit dem Erfüllungsgrad H_i gewichteten Summe berechnet werden:

$$y_{res} = \frac{\sum_{i=1}^m y_i H_i}{\sum_{i=1}^m H_i}$$

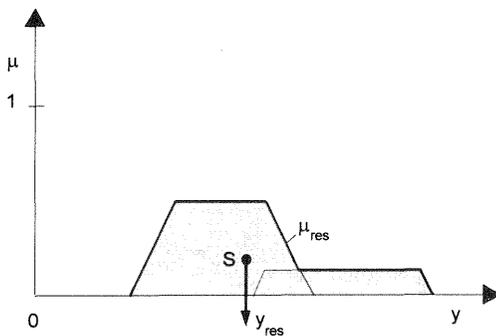
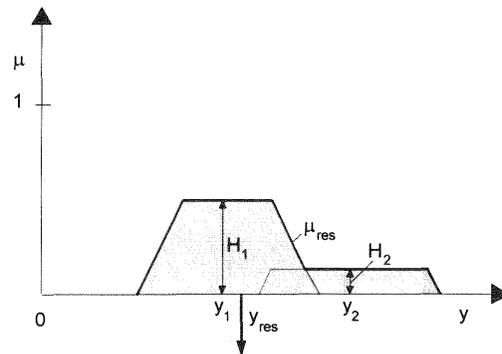


Abbildung B.12: a) Defuzzifizierung nach der Schwerpunktmethode



b) Anwendung der Näherungsformel für die Schwerpunktmethode

Die Defuzzifizierung ergibt in obigem Beispiel nach der Näherungsformel (vgl. Abbildung B.10):

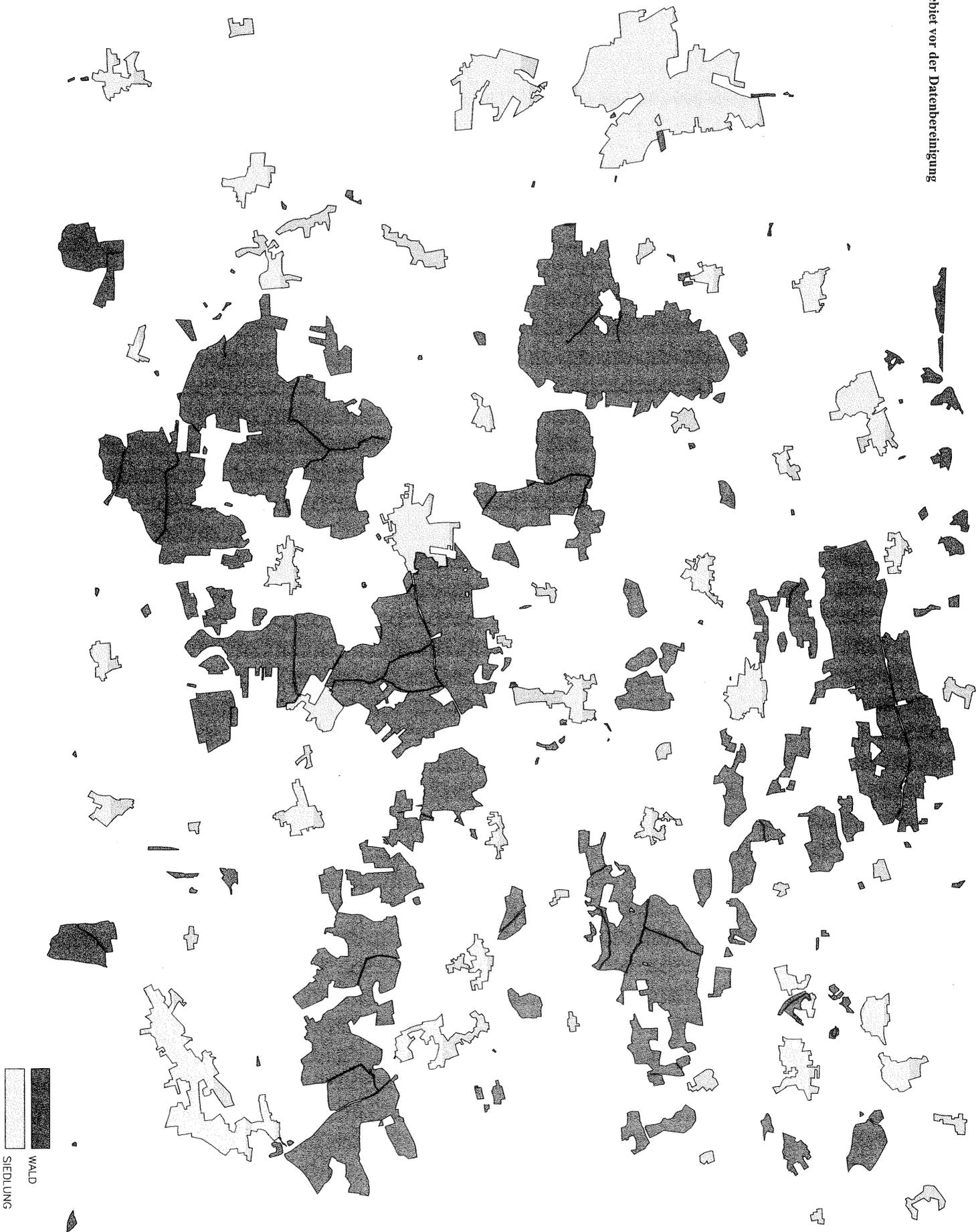
$$y_{\text{res}} = \frac{y_1 H_1 + y_2 H_2}{H_1 + H_2} = \frac{50\% \cdot 0.5 + 70\% \cdot 0.15}{0.5 + 0.15} = 54.6\% .$$

Für weitere differenziertere Betrachtungen wird auf die Literatur verwiesen [Kahlert, Frank, 1994].

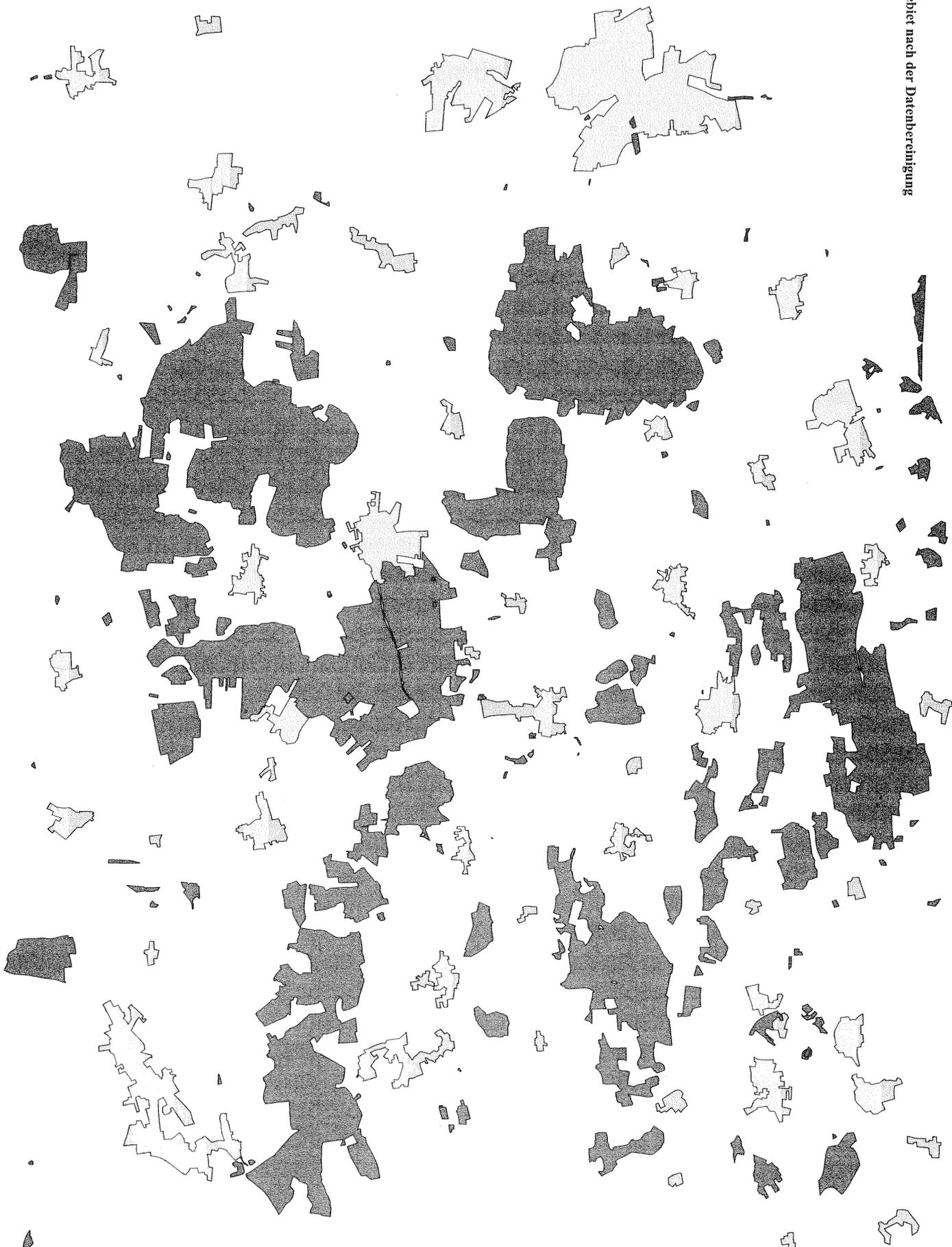
C Testgebiet

C.1 Testgebiet vor der Datenbereinigung

C.2 Testgebiet nach der Datenbereinigung



WALD
SEDLUNG



Lebenslauf

21. August 1968 geboren in München, als Tochter von Gunter und Marielene Rossipal
- 1974 - 1978 Gundschule in Emmering
- 1978 - 1988 Graf-Rasso-Gymnasium in Fürstenfeldbruck
29. Juni 1988 Abitur
- 1988 - 1992 Studium des Vermessungswesens an der Technischen Universität München
30. September 1992 Diplomhauptprüfung
- Okt. - Dez. 1992 Beschäftigung beim Ingenieurbüro G. Rossipal, Fürstenfeldbruck
- Jan. - April 1993 Mitarbeiterin am Institut für Photogrammetrie der Universität Stuttgart und Beschäftigung beim Ingenieurbüro G. Rossipal, Fürstenfeldbruck zu je 50%
- Mai 1993 - Juli 1995 Referendarin für den höheren vermessungstechnischen Verwaltungsdienst in Bayern am Bayerischen Landesvermessungsamt in München
- seit August 1995 Vermessungsrätin z. A. an der Bezirksfinanzdirektion München, Vermessungsabteilung, Referat 46 – Zentrale Dienstleistungen in der Informations- und Kommunikationstechnik