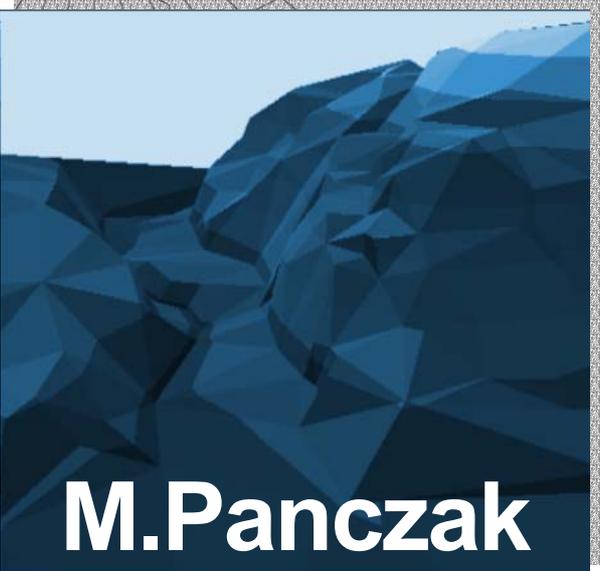


Diplomarbeit

Konzeption und prototypische Realisierung einer flexiblen, blattschnittfreien DGM Datenverwaltung sowie eines plattformunabhängigen Zugriffs auf Funktionen und abgeleitete Ergebnisse durch einen Web-Service.

The logo for FH-Karlsruhe features a dark blue background with a white wireframe mesh of a mountain range. The text 'FH-Karlsruhe' is written vertically in white on the right side.

FH-Karlsruhe

The name 'M. Panczak' is displayed in white text on a dark blue background, which is a low-poly mountain range graphic.

M. Panczak



FH Karlsruhe

UNIVERSITY OF APPLIED SCIENCES
HOCHSCHULE FÜR TECHNIK



Baden-Württemberg
Landesanstalt für Umweltschutz



©2003 Marc Panczak

Fachhochschule Karlsruhe – Hochschule für Technik

Fachbereich Geoinformationswesen
Studiengang Kartographie und Geomatik

Landesanstalt für Umweltschutz

Baden-Württemberg
Karlsruhe Referat 53
ITZ [Informationstechnisches Zentrum]

Diplomand, Autor & Herausgeber:

Marc Panczak

Betreuer (Fachhochschule):

Prof. Dr.-Ing. Gerhard Schweinfurth

Betreuer (Landesanstalt für Umweltschutz):

Biologiedirektor Manfred Müller

Betreuer:

Dipl.-Ing (FH) Wolfgang Schillinger

Alle Rechte vorbehalten. Kein Teil dieser Arbeit darf in irgendeiner Form (Druck, Fotokopie oder einem anderen Verfahren) ohne schriftliche Genehmigung des Autors reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Deckblatt:

Bilder und Gestaltung M. Panczak

Kartennetz: August Epicycloidal Projektion

Bilder: 3D Ansichten, Ausschnitte des DHM50 und 1m-DGM. [Karlsruhe-Durlach und Karlsruhe-Grötzingen]

Erklärung

Hiermit versichere ich, dass ich diese Diplomarbeit selbständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.

Karlsruhe, den

M. Panczak

Danke

An dieser Stelle möchte ich mich bei Herrn Prof. Dr.-Ing. G. Schweinfurth für die Betreuung dieser Arbeit bedanken.

Bedanken möchte ich mich auch bei Herrn Biologiedirektor M. Müller, W. Schillinger, B. Ellmenreich und allen weiteren Mitarbeitern des Referats 53 in der Landesanstalt für Umweltschutz für die fachliche und technische Betreuung.

Besonderer Dank gilt auch meiner Familie, die es mir ermöglicht hat zu studieren, außerdem Danke ich Familie Georg und J. Scheibe für ihre Hilfestellung.

Vorwort

Zum Aufbau und zur Formatierung dieser Arbeit:
Alle Überschriften und gliedernden Absätze sind, wie in der Abbildung unten gezeigt, fett formatiert und mit einer linksbündigen, durchgehenden Nummerierung versehen. Alle Überschriften finden sich im „ausführlichen Inhaltsverzeichnis“ wieder. Alle weiteren Texte sind vom linken Rand einfach eingerückt. Schlüsselbegriffe werden durch Versalien hervorgehoben.

		Erste Überschrift
10	Erste Überschrift	Konzeption und prototypische Realisierung einer flexiblen, blattschnittfreien DGM Datenverwaltung sowie eines plattformunabhängigen Zugriffs auf Funktionen und abgeleitete Ergebnisse durch einen Web-Service.
10.1	Zweite Überschrift	Konzeption und prototypische Realisierung einer flexiblen, blattschnittfreien DGM Datenverwaltung sowie eines plattformunabhängigen Zugriffs auf Funktionen und abgeleitete Ergebnisse durch einen Web-Service.
10.1.1	Dritte Überschrift	Konzeption und prototypische Realisierung einer flexiblen, blattschnittfreien DGM Datenverwaltung sowie eines plattformunabhängigen Zugriffs auf Funktionen und abgeleitete Ergebnisse durch einen Web-Service.

Inhalt [kurz]

1	Einleitung	12
2	Grundlagen	14
3	Mögliche DGM Modelle	23
4	Importmöglichkeiten nach ArcGIS	26
5	Programmierte Werkzeuge im Überblick	26
6	Möglichkeiten einer blattschnittfreien Speicherung (landesweiter Zugriff)	26
7	Optimierungen	26
8	DGM-Web-Service	26
9	Resümee	26
	Glossar & Abkürzungen	26
	Tabellen- und Abbildungsverzeichnis	26
	Quellen- und Literaturverzeichnis	26
	Verzeichnis Anhänge	26
	Anhänge ab	26

Inhalt [ausführlich]

1	Einleitung	12
2	Grundlagen	14
2.1	Definition des Begriffs Digitales Geländemodell	14
2.2	Die zur Verfügung stehenden Daten	14
2.2.1	Projektentwicklung und -ablauf	14
2.2.2	Datenerfassung mit Laserscanning	15
2.2.3	Datenverarbeitung	17
2.2.4	Datenformate	18
2.3	Arbeitsumgebung	19
2.3.1	Hardware	19
2.3.2	Software	20
3	Mögliche DGM Modelle	23
3.1	Definition von GRID und TIN	24
3.2	Gegenüberstellung von GRID und TIN	24
3.3	Raster/GRID und TIN in ArcGIS	26
4	Importmöglichkeiten nach ArcGIS	26
4.1	Konvertierung in ESRI GRID	26
4.2	Konvertierung in SDE-RASTER	26
4.3	TIN als Datei	26
4.3.1	TIN erstellen	26
4.3.2	TIN aus GRID oder Punktwolke erstellen	26
4.4	TIN in der Datenbank	26
4.5	GRID oder TIN	26
5	Programmierte Werkzeuge im Überblick	26
5.1	DGM-Tools, Import- und Verwaltungswerkzeuge	26
5.2	Mosaik-Erstellung und SDE-Loader	26
5.3	Rasterkatalog	26
5.4	Raster Clipper	26
5.5	Dummy Builder	26
5.6	Raster 2 ASCII	26

6	Möglichkeiten einer blattschnittfreien Speicherung (landesweiter Zugriff)	26
6.1	Grenze durch 32Bit Architektur	26
6.2	Lösungen für großräumige Bearbeitungen	26
6.2.1	NODATA-Bereiche	26
6.2.2	Maßstabsreduzierung	26
6.2.3	Rasterkatalog	26
6.2.4	TIN	26
6.2.5	Landesweiter Zugriff auf die Datenbank	26
6.3	Zusammenfassen der Möglichkeiten	26
7	Optimierungen	26
7.1	Optimierung von Rasterdaten in der Datenbank	26
7.1.1	Ladeprofile allgemein	26
7.1.2	Ladeprofile für DGM-Daten	26
7.2	Integer- vs. Floating Point-GRID	26
8	DGM-Web-Service	26
8.1	DGM-Web-Service mit ArcIMS	26
8.1.1	Was ist ArcIMS?	26
8.1.2	Voraussetzung für den DGM-Web-Service	26
8.1.3	Zusammenstellen des Geodatenmaterials	26
8.1.4	Funktionalität und Design	26
8.2	Web-Service durch COM-Komponente	26
8.2.1	Architektur	26
8.2.2	Realisierung	26
8.2.3	Probleme	26
9	Resümee	26
	Glossar & Abkürzungen	26
	Tabellen- und Abbildungsverzeichnis	26
	Quellen- und Literaturverzeichnis	26
	Printmedien	26
	Internet	26
	Verzeichnis Anhänge	26
	Anhänge ab	26

Das Thema dieser Diplomarbeit lautet: Konzeption und prototypische Realisierung einer flexiblen, blattschnittfreien DGM-Datenverwaltung sowie eines plattformunabhängigen Zugriffs auf Funktionen und abgeleitete Ergebnisse durch einen Web-Service.

Welche Gründe motivieren zu dieser Aufgabenstellung?

Zahlreiche hydrologische und ökologische Fragestellungen setzen eine genaue Kenntnis der Geländehöhe voraus. In Baden-Württemberg wurde das bisherige Digitale Höhenmodell (DHM, s. Glossar) mit Hilfe photogrammetrischer Verfahren erstellt (50m Auflösung). Zur Zeit (2002/03) werden mittels Laserscan-Befliegung zwei neue hochauflösende Digitale Geländemodelle (DGM s. Glossar) mit der Auflösung 1m und 5m erarbeitet.

Der Aufbau eines landesweiten Geländemodells aus Milliarden von Punkten und die Verwaltung der etwa 36 Milliarden abgeleiteter Folgepunkte setzen leistungsfähige Datenstrukturen und Zugriffsmechanismen voraus.

SCHLEYER 2001 schreibt in seinem Bericht, dass „mit der landesweit flächendeckenden Einrichtung eines neuen, hochgenauen Laserscan-DGM, die Vermessungsverwaltung Baden-Württemberg ein Projekt anpackt, das wohl als einmalig in der Bundesrepublik bezeichnet werden kann und mit dem der Vermessungsverwaltung eine Oberflächenbeschreibung des Geländes in hoher Qualität zur Verfügung stehen wird. Damit wird es für viele Anwendungen interessant. Bis in den großmaßstäbigen Bereich können z.B. Höhenlinien und Böschungen mit sehr hoher Qualität abgeleitet werden. Wissenschaftler versprechen sich sehr viel genauere Hochwasserberechnungen und Prognosen durch präzise Simulationen. Aber auch populäre Anwendungen wie virtuelle Flüge über Landschaften, beispielsweise bei der Wettervorhersage einer Nachrichtensendung, werden sehr viel realistischer sein. Wird dem DGM ein Orthophoto überlagert, stehen Perspektivansichten von hoher photorealistischer Qualität zur Verfügung. Im Bereich Mobilnetzplanung kann das Digitale Oberflächenmodell (DOM, s. Glossar) genauere Ergebnisse liefern. Die Aufzählung der Anwendungen, bis hin zu Ableitungen aus dem DGM, wie landesweite Profile, Neigungsmodelle, Neigungsisolines oder Höhengeschichtpläne, ließe sich beliebig fortsetzen [...]“

JÄGER 2000 S. 1f schreibt, „Ziel dieser Aufnahme ist es zum einen, sämtliche Reliefinformationen originär in nur einem digitalen Basisdatenbestand zu führen, der alle Anforderungen, insbesondere auch die Höhendarstellungen der verschiedenen Maßstabbereiche, zu erfüllen hat. Zum anderen können damit die bisherigen, teilweise inkonsistenten, redundant geführten und veralteten Höheninformationen in der HFK25 (Höhenflurkarte 1:2.500), der DGK 5 (Deutsche Grundkarte 1:5.000), den topographischen Karten und im DHM abgelöst werden.“

Um dieses Ziel zu erreichen, soll mit Hilfe dieser Diplomarbeit zunächst eine Konzeption für eine prototypische Realisierung einer flexiblen, blattschnittfreien DGM-Datenverwaltung geschaffen werden.

Wie kann solch eine flexible Datenverwaltung für die vorliegenden ASCII-DGM-Daten realisiert werden?

Zuerst soll inventarisiert werden, welche Mittel vorhanden sind: welche Programme werden durch die vorhandene Software vorgegeben, welche Möglichkeiten bieten sie und die vorhandene Hardware und welche weiteren Faktoren sind möglicherweise zu berücksichtigen.

Anschließend wird der Import für die Datenhaltung eine wichtige Rolle spielen. Die DGM-Daten werden nicht auf einmal für das ganze Land bereitgestellt, sondern werden sukzessive, je nach Voranschreiten der Nachbearbeitung, durch das LVA geliefert. Diese Tatsache erfordert einen einfachen und weitgehend automatisierten Weg, um die Daten zu importieren.

Die Optimierung der Datenhaltung stellt einen weiteren wichtigen Aspekt vor und während des Datenimports dar. Die Faktoren Geschwindigkeit und Speichervolumen werden dafür eine maßgebende Rolle spielen, nicht nur im Hinblick auf den Import und die Datenhaltung, sondern auch in Bezug auf die Bereitstellung und Weiterverarbeitung der Daten.

Weitere Punkte ergeben sich durch Probleme, die bei den vorher genannten Schritten auftreten werden, z.B. die Realisierung des *landesweiten* Zugriffs auf die DGM-Daten.

Zuletzt wird der plattformunabhängige Zugriff auf Funktionen und abgeleitete Ergebnisse durch einen Web-Service behandelt. Um die DGM-Daten aktuell und dynamisch einer breiten Öffentlichkeit zur Verfügung zu stellen, bietet sich der Zugriff über einen Web-Service im Intranet oder Internet an.

2 Grundlagen

Das folgende Kapitel gibt einen Überblick über digitale Geländemodelle und die zu bearbeitenden Daten des Landesvermessungsamts (LVA) Baden-Württemberg. Es werden u.a. der Entwicklungsstand der einzelnen Modelle, die wichtigsten Parameter der DGM-Daten, die Erfassung und die Schritte der Nachbearbeitung erläutert.

Daneben wird die Hard- und Software für die Bearbeitung und Datenhaltung in der Landesanstalt für Umweltschutz (LfU) Baden-Württemberg unter der Überschrift „Werkzeuge“ kurz beschrieben.

14

2.1 **Definition des Begriffs Digitales Geländemodell**

Digitale Geländemodelle sind wichtige Bausteine beim Aufbau einer umfassenden Struktur für Geodaten. Solche Daten beschreiben die reale Welt, wozu auch das Gelände gehört – die Topographie, die Berge, Täler, Geländeformen, Einschnitte, Aufschüttungen, Trassen usw. Der Begriff des Geländes kann in diesem Zusammenhang umfassender gebraucht werden als in der Umgangssprache üblich. „Jede Erscheinung, die im Wesentlichen als eine stetige Funktion des Ortes (der Lage) gewertet werden kann [...]“, kann nach BARTELME (1994, S.117f) mit einem DGM abgebildet werden.

BARTELME meint mit diesem letzten Satz, dass sich auch andere Thematiken modellieren lassen, z.B. ist es möglich Hangneigungen, Bodenerträge, Ozonwerte oder abstrakte Themen wie das Bruttosozialprodukt auf ein Gitter zu interpolieren. Was die Begriffe Gitter oder englisch „GRID“ und „interpolieren“ in diesem Zusammenhang bedeuten, wird in Kapitel 3 „GRID und TIN“ erläutert.

Die Geländeoberfläche als digitales Modell wird kurz als DGM oder im englischen als DTM – Digital Terrain Model – bezeichnet. In einigen Fällen, in denen besonders hervorgehoben werden soll, dass das Hauptaugenmerk auf Höhenwerten liegt, wird auch der Begriff DHM – digitales Höhenmodell – oder der englische Begriff DEM – Digital Elevation Model – verwendet.

Im Rahmen dieser Arbeit wird der allgemeine Begriff DGM oder DGM-Daten benutzt, da es sich eindeutig um Höhendaten handelt.

2.2 **Die zur Verfügung stehenden Daten**

Im Folgenden werden die zur Verfügung stehenden Daten erläutert. Daher wird zunächst ein Überblick über das gesamte DGM-Projekt von der Datenerfassung über die Nachbearbeitung bis zur Auslieferung durch das LVA gegeben.

2.2.1 **Projektentwicklung und -ablauf**

Bisher wurde vom LVA ein Höhenmodell mit der Rasterweite von 50m zur Verfügung gestellt. Dieses DHM50 eignet sich für Arbeiten im Übersichtsmaßstab, nicht aber für hochauflösende Modelle und Planungen. Mit einem neuen DGM

werden zukünftig flächendeckend sehr präzise Höhendaten bereitgestellt, die für zahlreiche anstehende Projekte eine optimale Datengrundlage bieten werden.

Das Fachkonzept für das DGM-Projekt wurde im Januar 1999 beschlossen. Den Werkvertrag für die Erfliegung der Daten erhielt die Firma TopScan®. Projektbeginn war im Jahr 2002. Der Abschluss der Befliegung ist voraussichtlich noch 2003. Die Fertigstellung des 1m Rasters einschließlich der Nachbearbeitung ist für das Jahr 2005 geplant.

Für das 5m Raster liegen zur Zeit keine zeitlichen Angaben vor.

Der momentane Projektstand (Juli 2003):

- beflogene Fläche: ca. 28.000km² von ca. 35.800km²

Wenn die Rohdaten den Nachbearbeitungsprozess bei TopScan® und anschließend bei dem LVA durchlaufen haben, liegen sie als bereinigte, regelmäßige 1m und 5m ASCII-DGM-Dateien vor und werden vom LVA u.a. an die LfU (Landesanstalt für Umweltschutz) ausgeliefert. Zur Zeit (Oktober 2003), liegen nach 5 Datenlieferungen knapp 3.000km² 1m und 5m-DGM-Daten vor. Einen ausführlicheren Überblick über den Erfassungs- und Nachbearbeitungsprozess geben die folgenden Abschnitte.

Ein weiterer der LfU vorliegender Datensatz enthält ein 300km² großes Testgebiet mit interpolierten Rasterdaten und je eine 1km² Beispielkachel für klassifizierte, bereinigte und unbereinigte Punktwolken für die Generierung von TIN-Datensätzen (Triangulated Irregular Network, s. Kapitel 3.1).

2.2.2 Datenerfassung mit Laserscanning

Die Befliegung der ca. 35.800km² findet in den vegetationsarmen Wintermonaten statt.

Bei der Laserscan-Messung werden von einem im Flugzeug installierten System in regelmäßigen Abständen Laserimpulse ausgesendet, die vom Erdboden und den darauf befindlichen Objekten reflektiert werden. Durch Laufzeitmessung wird die Entfernung zwischen dem Lasersensor und den Geländeoberflächen bestimmt. Dabei wird die Zeit zwischen dem ausgesendeten Impuls und der Registrierung des reflektierten Signals gemessen. Der Laserimpuls wird zudem quer zur Flugrichtung durch einen rotierenden Spiegel abgelenkt. Dadurch wird ein Geländestreifen mit einer Zickzacklinie entlang des Flugwegs abgetastet. Als Multisensorsystem basiert die Laserscanner-Messung auf folgenden Hauptkomponenten:

- Laserdistanzmesser mit Scanvorrichtung,
- Inertiales Navigationssystem (INS) und
- GPS-Empfänger (Globales Positionierungssystem) und GPS-Referenzstation am Boden. (s. Abb. 2)

Die dreidimensionalen Laserpunkte werden mit Hilfe von räumlichen Vektoren berechnet, deren Anfangspunkte durch GPS, deren Richtung durch INS und deren Länge durch Entfernungsmessung bestimmt werden.

Eine charakteristische Eigenschaft von Laserscannern ist die Unterscheidung von Mehrfachreflexionen (s. Abb. 1). Für einen einzelnen Laserimpuls können über der Vegetation mehrere Reflexionen beobachtet werden, da Teile des Lichtimpulses an der Vegetation reflektiert werden, andere bis auf den Erdboden vordringen. Mit den heutigen Laserscannern können sowohl die ersten Reflexionen (first-pulse), als auch die letzten Reflexionen (last-pulse) aufgezeichnet werden.

Mit Hilfe von Filteralgorithmen werden, basierend auf Höheninformationen benachbarter Punkte, die erfassten Punkte verschiedenen Punktgruppen zugeordnet, man spricht von Klassifizierung. Man unterscheidet die tiefsten Punkte, auch Bodenpunkte genannt und die übrigen Punkte, auch als Hochpunkte oder Vegetationspunkte bezeichnet. In der Regel repräsentieren die tiefsten Punkte aus der last-pulse-Aufzeichnung das Gelände.

Die Eckwerte der Datenaufnahme lassen sich wie folgt beschreiben:

Der mittlere Punktabstand beträgt etwa 1,5m bei einer Höhengenaugigkeit von $\pm 0,15\text{m}$ (97% aller Punkte $< 0,5\text{m}$). Die Klassifizierung nach Boden- und Hochpunkten erfolgt automatisch. Georeferenziert werden die Daten in der Universalen Transversalen Mercator Projektion (UTM), mit dem Ellipsoiden ETRS89 als Grundlage. Die Methode der Erfassung ist ausschließlich Laserscanning. Die Daten werden in $1\text{x}1\text{ km}^2$ großen Kacheln ausgeliefert.

(vgl. SCHLEYER 2001, S. 119)

SIGLE 2002 S. 5 fasst die Aufnahmemethode folgendermaßen zusammen: „Das Laserscanning liefert sehr dichte Punktwolken mit Punktabständen von ca. 0,5m bis 5m. Die Höhenmessungen haben eine hohe Genauigkeit von 0,1m bis 0,2m. Somit wird das Laserscanning vor allem beim Aufbau hochgenauer Geländemodelle eingesetzt. Die Messung erfolgt während des Überflugs vollautomatisiert.

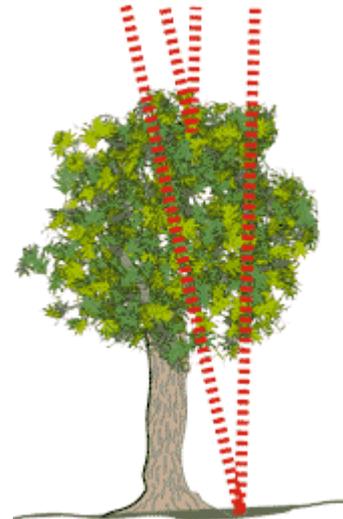


Abb. 1: Mehrfachreflexion



Abb. 2: Schema des Laserscannings

Ein großer Vorteil des Laserscanning ist die Fähigkeit, in Waldgebieten neben Punkten auf der Vegetationsoberfläche auch Punkte am Boden zu erfassen und somit flächendeckend die für ein DGM benötigten Daten zu liefern. Neben dem DGM können aus Laserscanning-Daten auch digitale Oberflächenmodelle (DOM) abgeleitet werden, die die Geländeoberfläche mit Vegetation und Bebauung beschreiben.

Ein Nachteil des Laserscanning kann sein, dass es nur begrenzt skalierbar ist. Das Verfahren liefert grundsätzlich genaue und dichte Punktwolken und ist somit für eine schnelle DGM-Erfassung für große Flächen mit eingeschränkter Genauigkeit nicht wirtschaftlich.

Ein weiteres Problem ist die fehlende absolute Selbstkontrolle. Zur Qualitätskontrolle werden externe Daten wie z.B. Orthophotos, photogrammetrische Stereomodelle oder digitale Karten benötigt.

Eine direkte Messung morphologischer Daten wie Geländekanten oder Gerippllinien ist nicht möglich. Es gibt zwar Ansätze, diese automatisch aus der 3D-Punktwolke abzuleiten, dies ist jedoch nur mit beschränkter Genauigkeit möglich.“

2.2.3 Datenverarbeitung

Die Firma TopScan[®] führt nach der Datenaufnahme eine Datenaufbereitung aus. Aus den aufgenommenen Laserpunktdaten werden verschiedene Punktgruppen gebildet, die sich mit den folgenden Schritten beschreiben lassen:

- Eine Auswahl der höchsten Punkte innerhalb eines vorgegebenen Gitters (z.B. 5m) liefert die Daten zur Ableitung eines DOM.
- Die Basis für das DGM bilden die last-pulse-Daten, die automatisch in Boden- und Hochpunkte klassifiziert werden.
- Die first-pulse-Daten werden ebenfalls nach Boden- und Hochpunkten klassifiziert. Sie werden später beim Landesvermessungsamt zunächst nicht weiterverarbeitet, sondern lediglich einer Plausibilitätsprüfung unterzogen.
- Die Daten der Reflexionsintensität werden zur Erstellung von Graustufenbildern genutzt, die u.a. bei der Nachbearbeitung der Laserscan-Daten als Interpretationshilfe dienen.

Die Auslieferung der Daten von TopScan[®] an das LVA erfolgt in ETRS89 („Earth Resources Technology Satellite“ Stand 1989, Landsat NASA) als Bezugssystem und UTM als Abbildungssystem. Die Höhenwerte beziehen sich als ellipsoidische Höhen auf GRS80 (Geodätisches Referenz System 1980).

Insgesamt sollen aus den aufgezeichneten Daten drei Modelle entstehen:

- Ein 1m-DGM (Gitterweite 1m, ohne morphologische Strukturelemente)
- Ein 5m-DGM (Gitterweite 5m, ohne morphologische Strukturelemente),
- Ein 5m-DGM (wie oben, nur mit geomorphologischen Strukturelementen).

Die Arbeitsschritte beim LVA gliedern sich in Vorverarbeitung, interaktive Nachbearbeitung mit Ableitung der morphologischen Strukturelemente und der Gitterberechnung. Es folgt eine Beschreibung der einzelnen Stufen der Verarbeitung mit ihren Unterschritten:

- Prüfung der Daten auf Vollständigkeit und Plausibilität. Dies beinhaltet die richtige Zuordnung von Kachel und Inhalt und die Lesbarkeit des Datenträgers.

Vorverarbeitung

- Transformation der Daten von ETRS89/UTM und ellipsoidischer Höhen nach Gauß-Krüger und Höhen ü. NN. Die Arbeit erfolgt mit der Software GEOMATRA.
- Kachelung der transformierten Punkte in 1km² große Kacheln, begrenzt durch die Koordinatenachsen des Gauß-Krüger Systems (s. Abb. 3). Die Bearbeitung erfolgt mit dem Programm EXTRA.
- Qualitätskontrolle durch den Vergleich mit vom LVA festgelegten Kontrollflächen und anderen bekannten Punkten. Durchgeführt wird die Kontrolle mit der Software SCOPE.

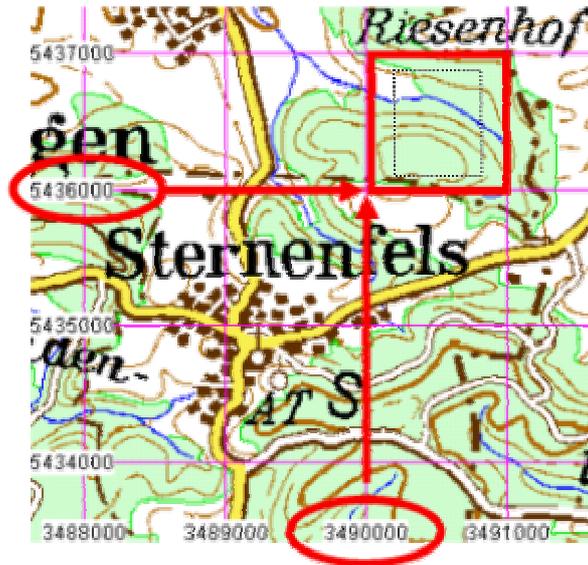


Abb. 3: Kacheleinteilung mit Gauß-Krüger

Interaktive Nachbearbeitung, Ableitung der morphologischen Strukturelemente

- Da die Algorithmen zur automatischen Klassifizierung der Laserscan-Punkte in Boden- und Hochpunkte auf den Höhendifferenzen benachbarter Punkte beruhen, sind Fehlinterpretationen unvermeidlich. Diese gilt es durch eine interaktive Nachbearbeitung der Klassifizierung zu beseitigen. Hierbei werden falsch klassifizierte Punkte unter Zuhilfenahme von Orthophotos und Gebäuden aus der Automatisierten Liegenschaftskarte (ALK) als Interpretationshilfe der richtigen Punktgruppe zugeordnet (Boden- oder Hochpunkt). Als Werkzeug dient die Software GVE.
- Aus den so bereinigten Daten werden automatisch Bruchkanten abgeleitet, die ebenfalls interaktiv nachbearbeitet werden müssen.
- Bei den Daten zur Ableitung eines DOM sind die für die gewünschte Modellbildung nicht relevanten Objekte zu eliminieren. Hierbei handelt es sich vor allem um Hochspannungsleitungen, Funk-, und Fernsehtürme oder Brücken. Diese Objekte werden in einer gesonderten Datei abgelegt, um sie für spezielle Anwendungen, z.B. für die Flugsicherung, wieder in das DOM aufnehmen zu können.

Gitterberechnung

Aus den interaktiv nachbearbeiteten Daten werden die oben genannten drei Modelle mit dem Programm SCOPE abgeleitet. (vgl. SCHLEYER 2001, S. 222f)

2.2.4 Datenformate

Die ausgelieferten Daten des LVA enthalten je nach Dateierweiterung eine der oben genannten Punktgruppen und Untergruppen, sofern es sich um Punktwolken handelt. Sie werden zur Erstellung von TIN-Datensätzen benutzt.

Den größten Teil der vorliegenden Daten stellen die ASCII DGM-Daten dar. Sie enthalten regelmäßige aus den Rohdaten interpolierte Koordinaten in 1m und 5m Rastern. Die Namenskonvention aller Dateien ist durch die Kachelung an das Gauß-Krüger-Koordinatensystem gebunden.

Die 1km²-Kacheln sind begrenzt durch ganzzahlige 1000m-Gitterlinien des Gauß-Krüger-Meridianstreifensystems. Die Dateinamen werden aus den Gauß-Krüger-Koordinaten der südwestlichen Kachelecke gebildet (s. Abb. 3).

Betrachtet man die ASCII Kacheln näher, stellt man fest, dass die tatsächliche Ausdehnung statt der erwarteten 1000 x 1000m, 1001 x 1001m beträgt. Die Gründe liegen nach Auskünften des LVAs in der Konformität zu den digitalen Orthophotos, die die gleiche Ausdehnung haben. Jedes DGM ist also um einen Meter am nördlichen und westlichen Rand breiter als erwartet. Diese Tatsache wird später bei der Programmierung berücksichtigt.

Die nachfolgende Aufzählung beschreibt alle Datenformate (*.DATEIENDUNG), die auf der CD des in Kapitel 2.2.1 genannten Testgebiets enthalten sind.

Beispiel: (s. Abb. 3)
Gauß-Krüger-Koordinaten der südwestlichen Kachelecke:

Rechtswert	Hochwert
3490000	5436000
Dateiname:	
34905436.*	

DGM-Gitter (ASCII)	Regelmäßiges Punktraster berechnet aus den nachbearbeiteten Bodenpunkten der Laserscan-Aufnahme. Gitterweite 1m und 5m (*.1m oder *.5m).
DOP (TIFF)	Digitales Orthophoto.
Punktwolke original (ASCII)	*.GRD = automatisch klassifizierte Bodenpunkte der Laserscan-Aufnahme.
	*.VEG = automatisch klassifizierte Hochpunkte der Laserscan-Aufnahme.
Punktwolke bereinigt (ASCII)	*.GDE = interaktiv nachbearbeitete Bodenpunkte.
	*.VEG = Hochpunkte. Dieser Datensatz ist nicht bereinigt und kann unplausible Werte enthalten.
	*.GEB = interpolierter Datensatz zum Schließen von Lücken im DGM verursacht durch Gebäude.
	*.BRU = Punkte auf Brücken.
	*.WAS = ebene Wasserflächen künstlich erzeugt.

2.3 Arbeitsumgebung

2.3.1 Hardware

Arbeitsrechner

Betriebssystem	Microsoft Windows 2000 Professional
Version	5.0.2195 Service Pack 3 Build 2195
Betriebssystemhersteller	Microsoft Corporation
Systemname	R53P99
Systemtyp	X86-basierter PC
Prozessor	X86 Family 15 Model 2 Stepping 4 GenuineIntel ~2017MHz (~2GHz)
BIOS-Version	Award Medallion BIOS v6.0
Gesamter realer Speicher	1,048GB
Gesamter virtueller Speicher	3,570GB

2.3.2 Software

ArcGIS (8.3 floating Licence)

ArcGIS ist der Name für eine Produktfamilie von aufeinander abgestimmten GIS (s. Glossar) Produkten der Firma ESRI.

Die ArcGIS Strategie sieht vor, dass alle Daten in Zukunft in objektrelationalen Datenbanken gespeichert werden. Dateien als Datenquellen werden weiterhin unterstützt, sie sind in der Abbildung unten jedoch nicht enthalten. Innerhalb von ArcGIS wird zwischen den Serverdiensten und den Clients unterschieden. Die Serverdienste ArcSDE (Spatial Data Engine, s. Glossar) und ArcIMS (Internet Map Server) stellen Grundfunktionalität zur Verfügung und sind zugleich die Bindeglieder zwischen den GIS-Clients und der unternehmensweiten IT-Infrastruktur. Die Clients sind funktional abgestufte Anwendungen. Hier stehen die ArcGIS

Desktop Clients ArcView, ArcEditor und ArcInfo im Vordergrund. Diese können über Erweiterungen mit zusätzlicher Funktionalität schrittweise ausgebaut werden. (vgl. www.esri.de)

ArcEditor 8.3

Für die clientseitigen Arbeiten wurde ArcEditor verwendet. ArcEditor setzt sich wie ArcView und ArcInfo aus den drei Basis-komponenten ArcToolbox, ArcMap und ArcCatalog zusammen. Der Funktionsumfang dieser Komponenten, kann mit Hilfe von Erweiterungen ausgebaut werden.

Die Schwerpunkte der ArcToolbox liegen hauptsächlich in der Bereitstellung von Werkzeugen für den Im- und Export von Daten und bei Konvertierungsaufgaben. Was der Explorer für Windows ist, ist der ArcCatalog für ArcGIS. Mit ihm lassen sich alle Geodaten auf der Festplatte und in der Datenbank verwalten. ArcMap ist das eigentliche Editier-, Analyse- und Layout- bzw. Präsentationswerkzeug.

Spatial Analyst

Mit dem Spatial Analyst werden ArcGIS Desktop Produkte im Bereich Rasterfunktionalität ergänzt. Diese Erweiterung stellt eine Palette von Funktionen zur räumlichen Modellierung und Auswertung zur Verfügung. Zellbasierte Rasterdaten können erzeugt, abgefragt, analysiert und kartografisch aufbereitet werden. Durch die Raster/Vektorkonvertierung, das Berechnen und Analysieren von Oberflächen und Funktionalitäten zur Rasterverschneidung erhält man aus bestehenden Datensätzen neue Informationen. (vgl. www.esri.de)

ESRI ++

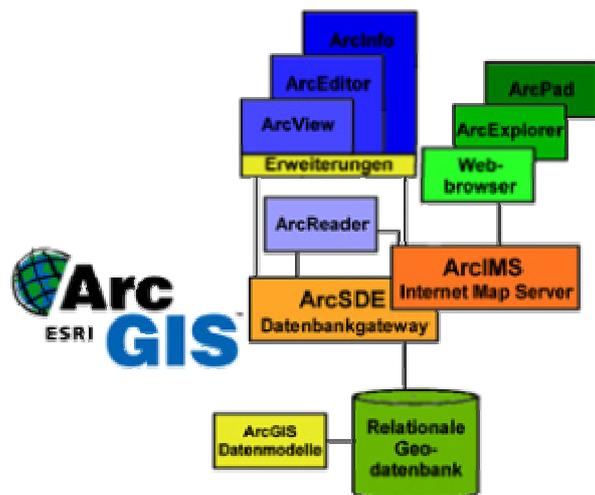


Abb. 4: ArcGIS Produktfamilie

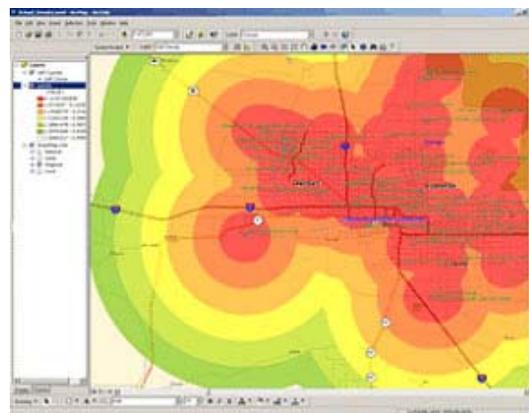


Abb. 5: Spatial Analyst Rasteroperation, Pufferung.

3D-Analyst

Die Erweiterung 3D-Analyst erlaubt es dem Benutzer, Oberflächendaten zu erzeugen, diese zu analysieren, zu modifizieren und in Echtzeit zu visualisieren. Mit ArcGIS 3D-Analyst hat man Möglichkeiten zur Erstellung, Bearbeitung und Analyse von 3D-Daten: Oberflächenberechnungen durch Interpolationsverfahren, Oberflächenanalysen (Neigung, Exposition, Schummerung usw.), Berechnung von Profilen, interaktive Abfrage von 3D-Elementen, perspektivische Darstellung von Oberflächen, Volumenberechnungen, Sichtbarkeitsanalysen, Navigation und 3D-Flüge aufzeichnen, sowie Export und Ausgabe im VRML 2.0-Format.

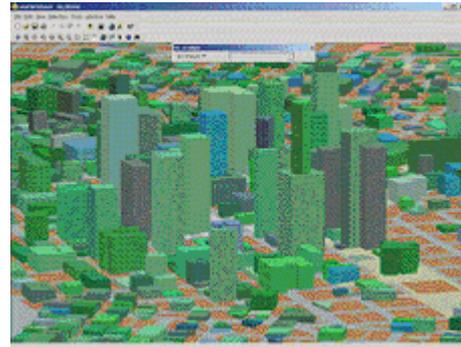


Abb. 6: ArcScene, 3D Darstellung

ArcScene als Visualisierungsapplikation von ArcGIS 3D-Analyst ermöglicht es dabei, den Blickwinkel auf die Oberflächendaten beliebig zu verändern und zu zoomen. Mit ArcGIS 3D-Analyst können GRID, TIN, CAD-Files (*.DXF, *.DWG usw.), Rasterdaten, ArcInfo Coverages und Shapedateien für dreidimensionale Darstellungen verwendet werden. (vgl. www.esri.de)

ArcSDE

ArcSDE ermöglicht den Zugriff auf raumbezogene Daten in relationalen Datenbanksystemen. ArcSDE ist ein GIS Gateway in die Datenbank.

Durch die Fähigkeit, raumbezogene Daten im Datenbanksystem zu verwalten und einer Vielzahl von Nutzern parallel zugänglich zu machen, ist ArcSDE ein zentraler Baustein in GIS Installationen mit mehreren Benutzern. Zudem vereinheitlicht ArcSDE als Schnittstelle die Zugriffe auf Datenbanksysteme. Egal, ob man mit Oracle, Microsoft SQL Server, IBM DB2 oder IBM Informix arbeitet, der Zugang zu den Geodaten in der Datenbank ist für die Anwendung immer gleich. Wenn man ein System benötigt, in dem mehrere Benutzer zeitgleich editierend zugreifen, bietet ArcSDE umfassende Zusatzfunktionalität gegenüber einem Standardone Desktop System:

- Zentrale Datenhaltung im frei wählbaren DBMS.
- Hohe Leistungsfähigkeit und Performanz bei Massendaten und/oder hohen Zugriffszahlen.
- Mehrbenutzerfähigkeit mit Versionsverwaltung und langen Transaktionen.
- Unterstützung eigener Geodatenmodelle; Aufbau derselben mittels ArcGIS Desktop oder mit Standard CASE-Tools.
- Offenheit für andere Klienten (auch CAD).
- OGC-Konformität (s. Glossar).
- Zugriff auf Geodatentabellen via SQL Statements (vgl. www.esri.de).

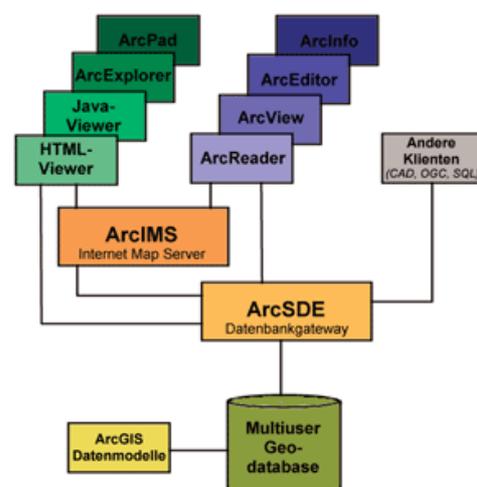


Abb. 7: ArcSDE Position in der Produktfamilie

ArcIMS

ArcIMS ist ein internetbasiertes GIS, mit dem man Geodaten in Form digitaler Karten oder interaktiver Anwendungen zentral aufbereiten kann und für andere Nutzer sowohl innerhalb einer Organisation (Intranet) als auch über das Internet zugänglich machen kann. ArcIMS verfügt sowohl über Client- als auch über Server-Technologie. Mit ArcIMS kann man Webauftritte bereichern, indem man interaktive GIS Komponenten integriert.

HTML- und Java-Viewer gehören zu ArcIMS. Zahlreiche weitere Clients, wie z.B. alle ArcGIS Desktop Produkte, ArcPad und drahtlose Geräte wie WAP-Handys oder PDAs können über vorbereitete Schnittstellen angebunden werden (vgl. www.esri.de).

Oracle 8i bzw. 9i

Im Laufe der Diplomarbeit wechselte die installierte Version der Oracle Datenbank von 8i auf 9i. Oracle 8i bzw. 9i ist ein Relationales Datenbankmanagementsystem (RDBMS). Der Oracle Server ist ein Repositorium für umfangreiche Datenmengen und gewährt dem Benutzer einen schnellen Zugriff auf diese Daten. Der Oracle Server erlaubt die gemeinsame Nutzung der Daten durch mehrere Anwendungen und Benutzer; die Information ist an einer Stelle gespeichert und wird von vielen Systemen genutzt (ABBEY 2000). Die Oracle Datenbank ist durch ArcSDE mit ArcGIS verbunden. In Ihr werden später die DGM-Daten abgelegt.



Server der Oracle Datenbank

Der Windows 2000 Server ist ein Server, der für kleine und mittlere leistungsfähige Netzwerke geeignet ist und auch eine solide Plattform für Webseiten oder E-Commerce-Lösungen darstellt (www.microsoft.de). Auf ihm ist die Oracle-Datenbank installiert, in der die DGM-Daten gespeichert werden sollen.



3 Mögliche DGM Modelle

Das Kapitel enthält eine Vor- und Gegenüberstellung der beiden Dateiformate die für die Darstellung und Speicherung der DGM-Daten in Frage kommen.

Mit der Ankündigung eines neuen und genaueren, landesweiten DGM stellt sich aus verschiedenen Gründen die Forderung nach einer flexiblen, blattschnittfreien DGM-Datenverwaltung (s. Kapitel 1, „motivierende Gründe“).

Eine höhere Auflösung ist in den meisten Fällen, so auch hier, mit einem erhöhten Speichervolumen verbunden, da wesentlich mehr Informationen und somit größere Datenmengen anfallen. Die Frage ist nun, welche DGM-Modelle stehen zur Verfügung und können sie die Anforderungen nach Flexibilität, Blattschnittfreiheit und landesweitem Zugriff erfüllen? Diese Frage soll in diesem und im nächsten Kapitel erörtert werden.

Die DGM-Daten werden vom LVA, wie bereits bei der Vorstellung der Daten oben beschrieben, in zwei Arten an die LfU ausgeliefert (Primärdaten). In beiden Fällen handelt es sich um ASCII-Dateien mit XYZ-Koordinaten (s. Abb. 8). Zum einen als regelmäßiges, interpoliertes Raster und zum anderen als klassifizierte Punktwolken.

Um diese ASCII-DGM-Daten als solche mit ArcGIS oder einer anderen Anwendung verarbeiten zu können, d.h. Analysen in Form von Verschneidungen, Berechnungen von Höhenlinien, Hangneigungen oder auch nur die Speicherung der Daten durchführen zu können, müssen die Daten in ein für das Programm verarbeitbares Modell konvertiert werden (Sekundärdaten).

Dabei stehen zwei grundlegende Modellierungsarten zur Auswahl. Zum einen das Rastermodell, wobei das Ergebnis entweder als ESRI-GRID (auf der Festplatte) oder als SDE-Raster (in der Datenbank) bezeichnet wird und zum anderen die Dreiecksvermaschung, auch Triangulation genannt, wobei TIN-Datensätze (Triangulated Irregular Network) erstellt werden.

Erst die Interpolation zu einem regelmäßigen Gitter oder die Vermaschung zu einem TIN, machen die aufgenommenen und nachbearbeiteten Primärdaten zu einem verarbeitbaren DGM für entsprechende Anwendungen.

Im Falle der ASCII-DGM-Dateien, die ein regelmäßiges Raster bilden, müsste man aufgrund der bereits interpolierten Daten von Sekundärdaten sprechen – es liegt in diesem Fall bereits ein Rastermodell vor. Folglich müsste von einer „Konvertierung“ der ASCII-DGM-Daten gesprochen werden und zwar in ein Format, das von einer entsprechenden Anwendung verarbeitet werden kann.

Im Falle der Punktwolken kann man nicht von Konvertierung sprechen, da die Rohdaten zwar nachbearbeitet wurden, diese Nachbearbeitung allerdings keine Interpolation beinhaltet. Es handelt sich somit um Primärdaten, die in ein Raster- oder Vektormodell überführt werden müssen.

Datei	Bearbeiten	Format ?
3403000.	5323000.	191.10
3403000.	5323001.	190.99
3403000.	5323002.	190.91
3403000.	5323003.	190.87
3403000.	5323004.	190.87
3403000.	5323005.	190.89
3403000.	5323006.	190.92
3403000.	5323007.	190.94
3403000.	5323008.	190.94
3403000.	5323009.	190.92
3403000.	5323010.	190.91
3403000.	5323011.	190.91
3403000.	5323012.	190.89
3403000.	5323013.	190.86
...		

Abb. 8: 1m ASCII DGM-Datei

Nachdem bereits öfters die Bezeichnungen Raster, GRID und TIN gefallen sind, sollen die Begriffe definiert und die Zusammenhänge zwischen ihnen erklärt werden.

3.1 **Definition von GRID und TIN**

Das **GRID** oder **Raster** gehört zur Familie der Rasterdaten, die entsprechend des Rasterdatenmodells zur Abbildung von Bildrastern bzw. Geodaten durch eine Matrix von Bildelementen (Pixel, Rasterzelle oder Zelle) gespeichert werden. Der Begriff **PIXEL** kommt aus dem Bereich der Bildverarbeitung. Um deutlich zu machen, dass es sich hier um Geobezüge handelt, wird der Begriff Rasterzelle verwendet.

Das Rasterdatenmodell ist ein spezielles Datenmodell zur Verwaltung von Geometriedaten, das die Rasterzelle allgemein als Träger der geometrischen

Information nutzt. Bei Geodaten kann diese Trägereigenschaft um zusätzliche Information (Attribute) oder einer Referenz auf externe Attributdaten erweitert sein.

Im Falle der DGM-Daten ist zusätzlich zur Lageinformation die Höhe als Attribut abgespeichert. Durch die regelmäßige Unterteilung des Georaums in Zeilen und Spalten benötigt das Raster lediglich die Angaben bezüglich eines Aufsatzpunktes, die Ausdehnung und die Richtung zur Georeferenzierung. Jede Rasterzelle erhält dabei einen Wert, der als thematisches Attribut aufgefasst werden kann. In diesem Fall sind das, wie bereits erwähnt, die Höhenangaben für das DGM. (vgl. FINDEISEN 1990, S.74ff)

Im weiteren Verlauf sind die Begriffe Raster, GRID und ESRI-GRID als Synonyme aufzufassen. Letzteres ist das ESRI eigene Rasterformat, das im Kapitel 3.3 näher erläutert wird.

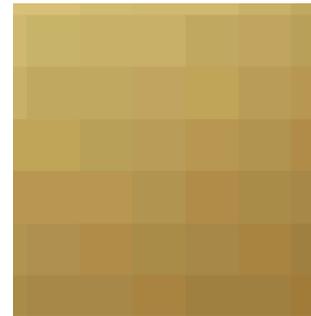


Abb. 9: GRID mit Pixelstruktur

Ein **TIN** ist eine Speicherform für dreidimensionale Oberflächen, bei der eine nicht attributierte, regelmäßige oder unregelmäßige 3D-Punktmenge über ein Dreiecksnetz vermascht ist. Das Dreiecksnetz besteht aus sich nicht überlappenden Dreiecken, die wiederum aus Kanten und Knoten bestehen. Das Haupteinsatzgebiet solcher TIN sind digitale Geländemodelle, aber auch andere Kontinua könnten auf diese Weise dargestellt werden. Ein TIN erlaubt eine direkte Speicherung primärer dreidimensionaler Messreihen ohne die Notwendigkeit einer vorherigen Interpolation, wie das bei Rastern der Fall ist, da sie hinsichtlich der Lage und damit auch der Dichte ihrer Knoten flexibel sind.

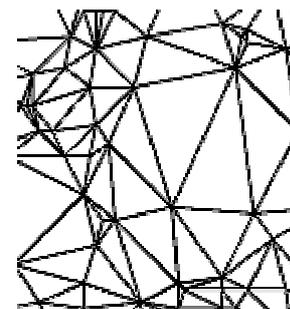


Abb. 10: TIN, 2D-Drahtmodell

3.2 **Gegenüberstellung von GRID und TIN**

Beide Modelle haben ihre speziellen Anwendungsbereiche und Vor- und Nachteile. Diese werden in der folgenden Auflistung zusammengestellt. Inwieweit die aufgelisteten allgemeinen Merkmale die Anforderung der Aufgabenstellung erfüllen können, wird anschließend erklärt.

GRID vs. TIN

Aufbau	Im Gegensatz zum TIN weist das GRID einen einfachen Aufbau aus einem regelmäßigen Raster auf, das sich mit EDV (elektronischer Datenverarbeitung) gut bearbeiten lässt. Ein TIN hingegen hat mit seinen Knoten, Kanten und Maschen einen vergleichsweise komplexen Aufbau.
Bearbeitung in Form von Analysen	Die Funktionen der Raster- und damit auch der GRID-Verarbeitung zeichnen sich nach BARTELME 2000, S.124 gegenüber Vektoroperationen (TIN) „durch ihre extreme Einfachheit aus. Ihr Nachteil besteht hauptsächlich in der großen Anzahl von Schritten, die für eine Bearbeitung eines gesamten Bildes notwendig sind“. Aber mit der Steigerung der Hardwareleistung in den letzten Jahren wird dieser Faktor immer mehr in den Hintergrund gedrängt, ist aber zur Zeit bei großen Datenmengen wie den vorliegenden keinesfalls zu verachten.
Oberflächenanpassung und Informationsverlust	In der Regel ist ein TIN anpassungsfähiger als ein GRID, da es in der Auflösung flexibel ist, d.h. dort wo das Relief eine höhere Energie aufweist werden mehr Punkte zur Vermaschung genutzt, im ebenen Gelände dagegen werden nur weniger Maschen erzeugt. Ein GRID ist mit seiner festen Rasterweite nicht flexibel, was erklärt, warum bei einer zu großen Rasterweite morphologische Kleinformen wortwörtlich durch das Raster fallen.
Speicherplatz	Die zuvor genannte Flexibilität von TIN Datensätzen hilft auch bei der Speicherung. Bei demselben Geländeausschnitt werden bei einem TIN im Vergleich zu einem GRID wesentlich weniger Informationen benötigt, um mit der gleichen Auflösung das Gelände zu beschreiben.

(vgl. WEIDENBACH 1998, S.99f)

Einige der genannten allgemeinen Merkmale lassen sich nicht auf das neue hochauflösende DGM übertragen. Beispielsweise fällt das Argument der einfachen oder komplexen Algorithmen zur Bearbeitung weg. Die gängigen Anwendungen, wie das Erzeugen von Höhenlinien, liegen bereits als Funktion im Softwarepaket ArcGIS für beide Modelle bei. Dabei kann man für einzelne manuelle Bearbeitungen die Erweiterungen wie Spatial Analyst oder 3D Analyst benutzen oder durch Programmierung mit ArcObjects (s. Glossar) eigene angepasste Werkzeuge erstellen.

Bei der Programmierung muss weiterhin die Bearbeitungszeit beachtet werden, da diese durch einen komplexen Algorithmus möglicherweise verlängert wird. Außerdem findet in ArcGIS bei verschiedenen Anwendungen auf TIN Datensätze, eine interne, temporäre Umwandlung zu einem GRID statt, was zusätzlich Zeit kostet.

Der Speicherplatz wird gerade wegen der hohen Auflösung ein Faktor im Vergleich von GRID und TIN sein.

Die Tatsache wiederum, dass je nach Rasterauflösung morphologische Formen bei einem GRID wegfallen können, greift nicht, da die DGM-Daten im Rasterformat mit ihrer 1m Auflösung den Anwendungen eine absolut hinreichende Genauigkeit bieten werden.

Welches der beiden Modelle sich letztendlich in der vorhandenen Umgebung durchsetzt, wird in den folgenden Kapiteln erarbeitet. Der praktische Umgang mit den beschriebenen Daten und Programmen wird zeigen, welches Format die Bedingungen erfüllen kann.

3.3 **Raster/GRID und TIN in ArcGIS**

Rasterdaten werden in ArcGIS als Datei (z.B. ESRI-GRID) und in der Datenbank (SDE-Raster) unterstützt. TINs hingegen werden voraussichtlich erst ab der Version 9 in der Datenbank unterstützt und können somit vorerst nur als Datei auf der Festplatte abgelegt werden.

Diese Tatsache ist für die DGM-Erstellung als TIN ein großer Nachteil im Hinblick auf die Realisierung eines blattschnittfreien landesweiten Zugriffs, wie man später sehen wird. Hinzu kommt, wie bereits oben erwähnt, dass bei verschiedenen Anwendungen von TIN-Datensätzen temporär für die Berechnungen GRIDs angelegt werden, wodurch man längere Bearbeitungszeiten erhält. Gerade bei großen Datenmengen wird sich diese Tatsache bemerkbar machen.

Doch bevor die Daten abgelegt und weiterarbeitet werden können, müssen sie zuerst in die vorhandene EDV Struktur der LfU importiert werden.

4 Importmöglichkeiten nach ArcGIS

Das folgende Kapitel beschäftigt sich mit den Importmöglichkeiten der DGM-Daten nach ArcGIS. Damit ist gemeint, wie verschiedenen Dateiformate (ESRI GRID, TIN) in ArcGIS erzeugt werden können. Dabei werden die notwendigen Schritte sowie Probleme erläutert.

4.1 Konvertierung in ESRI GRID

27

Ziel der Konvertierung ist es, die beschriebenen interpolierten DGM-Daten des LVAs in ein Raster zu überführen, das in ArcGIS bearbeitet werden kann. Als Dateiformat stehen mehrere Möglichkeiten zur Verfügung. ArcEditor unterstützt u.a. Rasterformate wie TIFF (Tag Image File Format) oder das ERDAS IMAGINE-Format, doch für die Weiterverarbeitung eignet sich am besten das ESRI eigene Rasterformat, ESRI GRID, im späteren Verlauf nur GRID genannt. Da bei allen anderen Formaten bei Analyseaufgaben, wie dem Erstellen von Hangneigungen oder Höhenlinien, intern ein temporäres GRID für die Berechnung erstellt wird. Dies stellt, wie bereits erwähnt, einen zeit- und rechenintensiven Prozess dar, je nach Aufwand und Größe der Datei. Deshalb werden ausschließlich ESRI GRIDs verwendet. Für den Aufbau eines ESRI GRIDs benötigt man den Ursprung, den die linke untere Ecke bildet, die Kantenlänge einer Rasterzelle, sowie die Anzahl der Zeilen und Spalten des Rasters; optional besteht die Möglichkeit auch Werte für nicht vorhandene Daten, so genannte NODATA-Werte, zu bestimmen. Wird dafür kein Wert gesetzt, wird ein Standardwert eingesetzt. Dieser Wert ist in der Regel der kleinste Wert des Datentyps, aus dem das Raster gebildet wird.

Mit Hilfe dieser Angaben über Ursprung, Rasterzellengröße und Zeilen-Spalten-Anzahl ist es möglich, mit nur einer Koordinate, nämlich dem Ursprung, das Raster aufzubauen. Man muss nicht für jede Rasterzelle eine XY-Koordinate abspeichern, sondern nur den Z-Wert, dies reduziert den Speicherplatz. Für eine schnelle Ansicht, z.B. in ArcMap, lassen sich optional so genannte PYRAMID LAYER erstellen. Diese kann man, wie in der Abbildung gezeigt, als auflösungsreduzierte Kopien des Original-GRIDs betrachten. Pyramid Layer haben den Vorteil, dass je nach Zoomstufe nicht alle Informationen des Originals geladen werden müssen, sondern nur der Pyramid Layer, der für den entsprechenden Maßstab optimal ist. Dabei hat man je nach Erstellungsart des GRIDs, Einfluss auf den Komprimierungsgrad und die Höhe der Pyramide (s. Abb. 11).

Mit Pyramid Layern ist es möglich, auch sehr große GRIDs schnell anzeigen zu lassen. Allerdings muss man mit etwa 10-35% höherer Speicherbelegung gegenüber dem Original-GRID rechnen. Die Pyramid Layer werden in *einer* RRD-Datei pro GRID, mit dem gleichen Namen wie das GRID selbst, abgelegt.

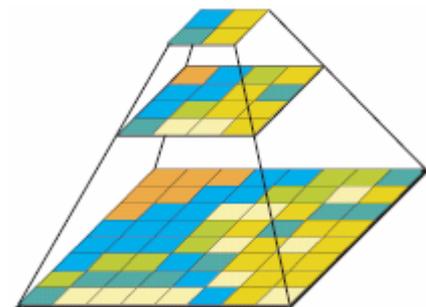


Abb. 11: Pyramid Layer

Es gibt mehrere Wege, ein GRID mit ArcGIS zu erzeugen. Bei den meisten muss viel von Hand editiert werden, darüber hinaus sind viele Zwischenschritte nötig.

Deshalb kommen diese Möglichkeiten zwar nicht für große Datenmenge in Frage, für einzelne GRIDs und für Testzwecke können sie aber gut genutzt werden.

Mit der ArcToolbox lässt sich ein weitgehend automatischer Import von ASCII Dateien nach ESRI GRID realisieren (Anhang 7), wobei dies für einzelne Dateien oder für mehrere Dateien gleichzeitig möglich ist. Die Prozessierung von mehreren Dateien auf diese Art nennt man Stapelverarbeitung (englisch: Batch-Verarbeitung).

In beiden Fällen (Einzel- oder Stapelverarbeitung) müssen die ASCII Dateien in einer bestimmten Formatierung vorliegen. Einige Headerinformationen müssen gegeben sein, und die Koordinaten müssen zeilenweise, von links oben nach rechts unten sortiert, untereinander stehen (s. Abb. 19, Kapitel 4.5.1). Beide Eigenschaften sind bei den LVA-Daten nicht gegeben und müssen deshalb manuell ergänzt bzw. geändert werden. (programmiertechnische Lösung in Kapitel 4.5.1)

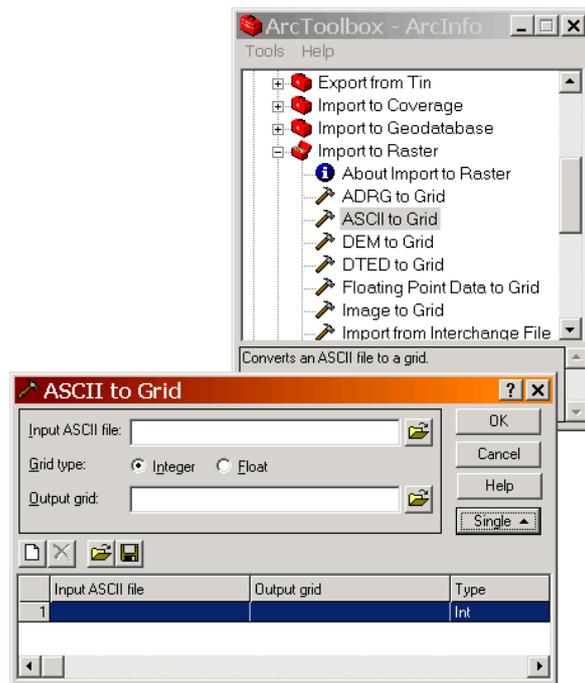


Abb. 12: ArcToolbox, ASCII to GRID

Weitere Möglichkeiten, einzelne Dateien zu importieren bietet ArcMap. Über das Laden von XY-Daten (s. Glossar) kann direkt oder indirekt über eine Shapedatei ein GRID erstellt werden. Eine Anleitung ist in Anhang 8 zu finden.

Die 1m ASCII-DGM-Dateien enthalten je Datei in der Regel 1.002.001 Koordinatentripel (1001 x 1001m) und benötigen auf der Festplatte knapp 28MB Speicher. Umgewandelt in ein ESRI GRID belegen sie ohne Pyramid Layer 3,87MB. Davon ausgenommen sind Dateien, die auf der Landesgrenze von Baden-Württemberg liegen. Diese Dateien sind nur mit Koordinaten bis zur Landesgrenze gefüllt und enthalten somit eine geringere Menge Koordinaten und folglich eine geringeres Datenvolumen.

4.2 Konvertierung in SDE-RASTER

Schaut man sich die Importwerkzeuge für die Datenbank im ArcCatalog und der ArcToolbox an, so muss für einen Datenbankimport ein Raster, ein Coverage oder eine Shapedatei vorliegen. Folglich müssen zuerst, wie im Kapitel zuvor gezeigt, ESRI GRIDs vorliegen. Diese GRIDs können auf zwei Arten in die Datenbank geladen werden.

Zum einen kann man das Administrator-Konsolenwerkzeug SDELOADER benutzen, oder man benutzt das RASTER TO GEODATABASE Werkzeug. Letzteres steht im ArcCatalog und in der ArcToolbox zur Verfügung.

Das Programm SDELOADER hat den Nachteil, dass es zwar Rasterdaten, aber nicht das ESRI GRID-Format unterstützt und somit nicht genutzt werden kann.

Wird ein GRID in die Datenbank geladen, ist es dort als SDE-Raster abgelegt. Werden im Batch-Betrieb mehr als eine Datei in die Datenbank geladen, werden alle Raster aneinandergehängt und zu einem großen MOSAIK vereint. Ein Mosaik ist ein Raster, in diesem Fall im ESRI GRID-Format, das aus mehreren kleineren, zusammenhängenden, formatgleichen Rastern gebildet wurde. Mehr zu Mosaiken wird in Kapitel 5.2 erklärt.

Auch bei den Rastern in der Datenbank lässt sich – wie bei einem GRID auf der Festplatte – festlegen, ob man Pyramiden bilden möchte oder nicht. Diese und weitere Einstellungen sind optional.

Aus verschiedenen Gründen ist es interessant zu wissen, wie die Spatial Data Engine (SDE), Rasterdaten in Oracle ablegt. Innerhalb dieser Arbeit war es u.a. wichtig zu wissen, wie viel Speicher ein Raster oder mehrere Raster in der Datenbank (SDE-Raster) im Vergleich zu einem GRID auf der Festplatte belegen. Wie Rasterdaten abgelegt werden und wie viel Festplattenplatz sie beanspruchen, beschreibt der nächste Abschnitt.

Vor der Beschreibung des Aufbaus eines SDE-Rasters, sei vorweg kurz erwähnt, wie sich ein Raster allgemein aufbaut, damit die folgenden Begriffe eindeutig verstanden werden.

Ein Raster besteht allgemein aus einem oder mehreren Rasterbändern, in der Bildverarbeitung auch als (Farb-) Kanal bekannt. Schwarzweißbilder haben wie die Höhendaten *ein* Rasterband, da nur *eine* Informationsschicht vorhanden ist. Mehrere Bänder sind beispielsweise bei Farb- oder Multispektralbildern vorhanden, die für eine Darstellung kombiniert werden. Ein Rasterband wiederum besteht aus Pixeln oder wie in der Definition für Raster genannt – Rasterzellen, welche die eigentlichen Informationen enthalten.

In der Datenbank wird ein Raster in vier Tabellen durch die

Als **Business-Tabelle** versteht ArcSDE jede in ArcSDE von einem ArcSDE Client, vom „sdetable Administrator Konsolenwerkzeug“ oder der „ArcSDE C API SE_table_create“-Funktion angelegte Tabelle.

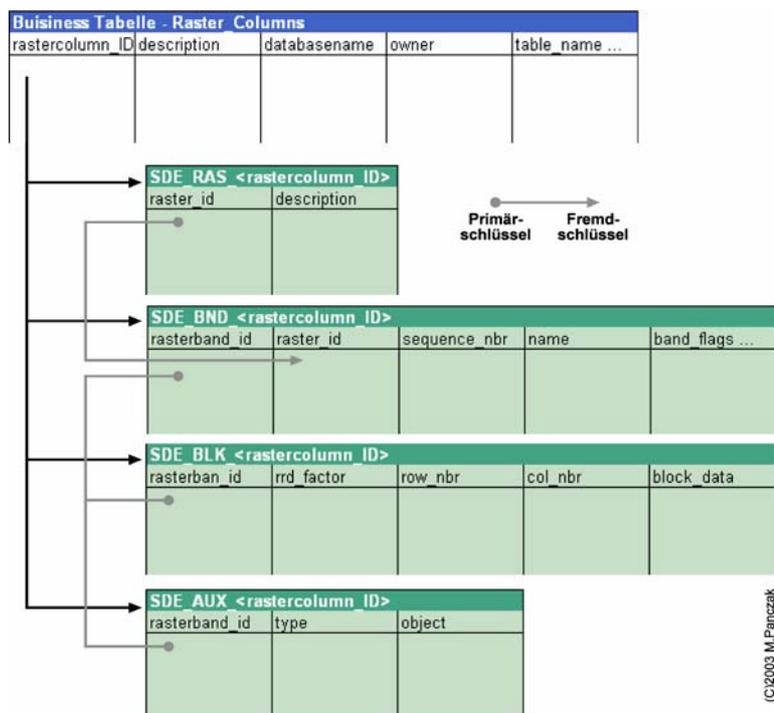


Abb. 13: Rastertabellen und ihre Beziehungen

organisiert und abgelegt:

- Die Raster-Blocks-Tabelle (SDE_BLK_<rastercolumn_ID>), speichert die eigentlichen Rasterdaten für jedes Rasterband. Die Daten sind innerhalb der

Tabelle als so genannte Kacheln (englisch: TILES) abgelegt. Schaut man sich die Tabelle in der Datenbank (Oracle) an, besteht diese Tabelle aus BLOBs (Binary Large Objects).

- Die Rasterband-Tabelle (SDE_BND_<rastercolumn_ID>) speichert Metadaten über das betreffende Band eines Rasters. Neben weiteren Informationen enthält diese Tabelle Angaben zur Größe von Kacheln und Pixeleigenschaften.
- Die Raster-Description-Tabelle (SDE_RAS_<rastercolumn_ID>) speichert die Beschreibung des Rasters unter der Spalte „raster_ID“ (s. Abb. 13). Diese Spalte wird mit dem neuen Datensatz an eine Business-Tabelle angehängt, die somit die Raster_ID beinhaltet. Durch diese Raster_ID werden die Business-Tabelle und die vier Rastertabellenarten verbunden.
- Die Tabelle SDE_AUX_<rastercolumn_ID> gehört ebenfalls zu jedem Raster und speichert zusätzliche Informationen, wie Colormap, Statistik und eine optionale Bitmaske, die für Überlappungen und zum Bilden eines Mosaiks benutzt wird.

Abbildung 13 zeigt die eben genannten Beziehungen zwischen den Tabellen.

4.3 TIN als Datei

4.3.1 TIN erstellen

TIN Modelle werden, wie in der Definition im Kapitel 3.1 beschrieben, aus „primären dreidimensionalen Messreihen ohne die Notwendigkeit einer vorherigen Interpolation“ erstellt.

Für diese Art der Erstellung bieten sich die Punktwolken des LVAs an, die für drei 1km^2 -Testgebiete (Stadt, Wald, Mischgebiet) vorliegen.

Für Testzwecke können einzelne TINs manuell folgendermaßen erstellt werden:

Die Punktwolken werden als Punktthema (XY-Layer) in ArcMap geladen, Dafür muss zuvor eine Kopfzeile mit bestimmten Informationen in die ASCII-Datei eingefügt werden.

Die einzelnen Koordinaten müssen durch Tabulatoren getrennt sein, ebenso die Kopfzeile, die als Überschrift für die jeweilige Koordinatenspalte dient, welche mnemotechnisch mit X, Y, Z bezeichnet werden.

Danach können mit dem 3D-Analyst TINs erstellt werden.

Doch es gibt auch die Möglichkeit, aus den vorliegenden ESRI GRIDS, TINs zu erstellen. Der Ablauf der Erstellung ist der gleiche, mit dem Unterschied, dass bei letzterem als Input ein Raster und kein Punktthema vorliegt. Welche Vor- und Nachteile die jeweilige Methode hat, wird in Kapitel 4.3.2 aufgezeigt.

Die einzelnen Schritte für die Erstellung im Detail, sind im Anhang 3 beschrieben.

In beiden Fällen arbeitet man beim Erstellen mit einer vertikalen Toleranz als Parameter. Was diese Toleranz bewirkt und wie sich Änderungen auswirken, zeigt der folgende Abschnitt.

Die Toleranz gibt die maximale Anzahl von Einheiten (hier Zentimeter) an, die eine TIN Oberfläche von der Zellmitte (im Falle eines Rasters) des Inputrasters abweichen darf. Ist die Abweichung größer als die angegebene Toleranz, wird ein eigener Knoten wie im Fall II (s. Abb. 14) gebildet. In Fall I liegt die Höhendifferenz innerhalb des Toleranzbereichs, d.h. im Punkt B wird kein Knoten gebildet. Infolgedessen bewirkt die Toleranz, dass ein relativ ebenes Gelände bei der Wahl einer entsprechenden Toleranz mit nur wenigen Punkten beschrieben werden kann. Eine kleine Toleranz ergibt folglich ein TIN, dass eine höhere Detailtreue im

Vergleich zur Rasteroberfläche aufzeigt. Eine große Toleranz hingegen generalisiert die Oberfläche.

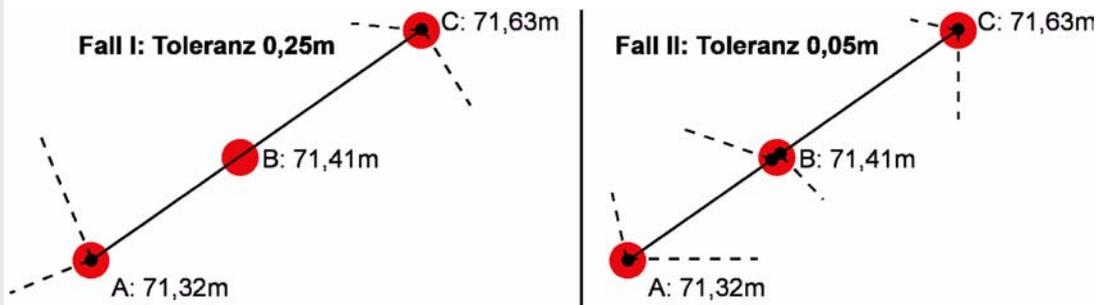


Abb. 14: vertikale Toleranz bei der TIN Erstellung

31

Würde man diese Toleranz weglassen oder die Toleranz gleich 0 setzen, was den gleichen Effekt hat, würde das Ergebnis der Vermaschung wie in den beiden folgenden Beispielen aussehen (s. Abb. 15).

Links lag als Input eine Punktwolke vor, rechts ein GRID. Bei beiden wurde die Toleranz gleich 0 gesetzt, dadurch wurde jeder Punkt vermascht. Zum einen ergibt sich ein wesentlich höheres Datenvolumen, zum anderen entspricht das nicht dem Vorhaben, eine möglichst detailgetreue Beschreibung der Oberfläche mit möglichst wenigen Punkten zu erreichen.

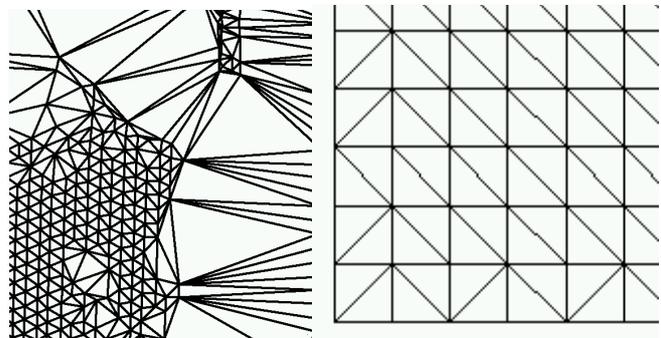


Abb. 15: TIN Erstellung ohne Toleranz

Legt man eine Toleranz bei der Erstellung fest, sieht das Ergebnis wie in Abbildung 16 aus.

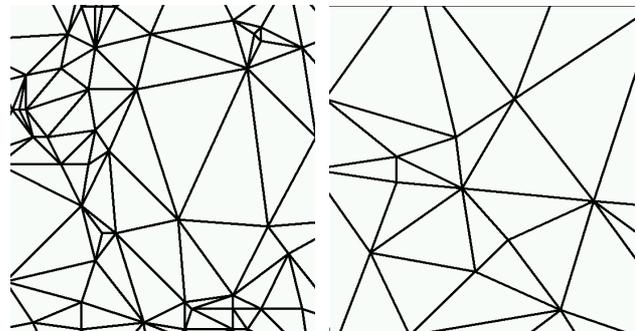


Abb. 16: TIN Erstellung mit Toleranz

Abbildung 16 zeigt links ein TIN mit einer Punktwolke als Ausgangsdatei und rechts die Erstellung mit einem GRID. Wie man bereits aus diesem

kleinen Bildausschnitt erahnen kann, besteht das TIN, das aus einem GRID erstellt wurde (rechts), aus weniger Dreiecken als das aus einer Punktwolke entstandene. In diesem Beispiel lag der Faktor etwa bei zehn; 27.622 Knoten gegenüber 2.571 Knoten. Das drückt sich auch in einer unterschiedlichen Dateigröße aus. Ein TIN aus einem GRID (0,20MB), war im Test um ein Vielfaches kleiner, als ein TIN aus einer Punktwolke (1,18MB).

Einen Vergleich der Knotenanzahl und der Größe des Datenvolumens bei den genannten Inputarten, jeweils mit und ohne Toleranz, zeigt Tabelle „Vergleichstest der Datenvolumina“ in Anhang 12.

Bei den genannten Größenunterschieden, muss man sich die Frage stellen, ob die kleineren TINs aus GRID-Dateien den geforderten Genauigkeitsansprüchen genügen. In diesem Fall würde eine Einsparung von Speicher gegenüber TINs aus

Punktwolken einen großen Vorteil bringen. Der Vorteil wäre sogar noch größer, vergleicht man die Einsparung gegenüber ESRI GRIDs (3,87MB). Dieser Sachverhalt wird im Folgenden überprüft.

4.3.2 TIN aus GRID oder Punktwolke erstellen

In diesem Abschnitt gilt es zu erörtern, ob die Genauigkeit von TINs, die aus GRIDs erstellt wurden, ausreichend im Hinblick auf einen möglichst hohen Detailgrad, im Vergleich zu TINs aus Punktwolken ist.

Die nachfolgende Abbildung (Abb. 17) zeigt die vier Möglichkeiten, die zur TIN Erstellung im Rahmen des 1m-DGMs für Baden-Württemberg in Frage kommen. Die zwei Hauptunterschiede (grün und rot), zeigen jeweils Wege mit vertikaler Toleranz (grün) und ohne Toleranz (rot). Diese beiden Hauptunterschiede unterteilen sich jeweils in eine Erstellung von TINs aus ESRI GRIDs bzw. direkt aus Punktwolken.

Wie bereits anfangs dieser Arbeit erwähnt, werden beide Formen – Punktwolke und ASCII-Raster, vom LVA an die LfU geliefert.

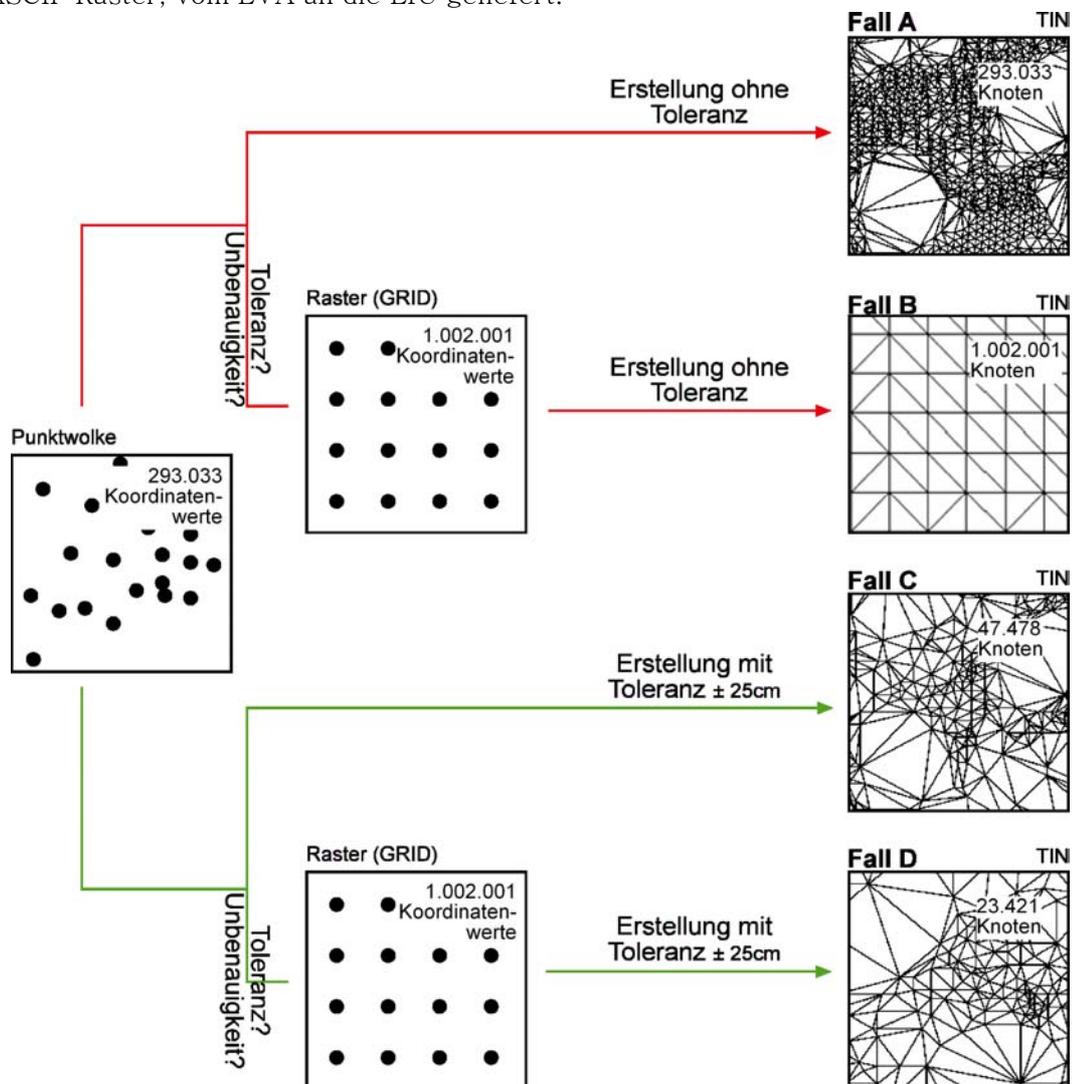


Abb. 17: TIN Erstellung aus Punktwolken oder Raster, mit und ohne Toleranz

Was man unter einer Toleranz bei der TIN Erstellung zu verstehen hat, wurde einen Abschnitt zuvor erklärt und soll nur kurz mit anderen Worten zusammengefasst werden.

Bei der Toleranz handelt es sich um einen Parameter, der einen vertikalen Grenzwert für die Knotenbildung zwischen benachbarten Höhenwerten bei der Vermaschung angibt. Liegt ein benachbarter Höhenwert außerhalb dieser Toleranz, wird er bei der Triangulation berücksichtigt, andernfalls wird er weggelassen (s. Abb. 14: vertikale Toleranz bei der TIN Erstellung).

Die größte Knotenzahl und infolgedessen der höchste Detailgrad aller Fälle A bis D (Abb. 17) ist gegeben, wenn das TIN direkt aus einer Punktwolke ohne die Angabe einer Toleranz erzeugt wird. Wie man an der Anzahl der Knoten im Fall A im Vergleich zur Anzahl von Koordinatenwerten der Punktwolke sehen kann, werden alle gemessenen Werte der Punktwolke verarbeitet. Dieser Vorgang ist somit umkehrbar. Es gibt weder einen Informationsverlust noch eine Verzerrung von Werten.

Im LVA werden die Rasterdateien durch Interpolation aus den Punktwolken-Dateien *.GDE (Ground), *.WAS (künstlich erzeugten Gewässeroberflächen), *.LUE (Lücken) und *.GEB (Gebäudegrundflächen) erstellt.

Betrachtet man den Fall B (Abb. 17), so wird deutlich, dass durch Weglassen des Toleranzwertes jeder Koordinatenwert berücksichtigt und trianguliert wird. Dadurch entsteht das regelmäßige Muster, das in Fall B zu sehen ist.

Da das Inputraster von Fall B aus einer Interpolation der Punktwolke entstanden ist, und das TIN in Fall A eine „1 zu 1“ Wiedergabe der Punktwolke darstellt, hat das TIN in Fall B zwar mehr Knoten, aber keinen höheren Informationsgehalt gegenüber Fall A.

Zusammengefasst bedeutet das: da Fall A bei gleichem Informationsgehalt weniger Knoten aufweist und deshalb auch weniger Dateivolumen besitzt, kann der Fall B bei dieser Fragestellung vernachlässigt werden.

Der Fall C beinhaltet ebenfalls ein TIN, das aus der gleichen Punktwolke wie Fall A erstellt wurde, mit dem Unterschied, dass eine Toleranz von $\pm 25\text{cm}$ angewandt wurde. Durch diese Toleranz werden benachbarte Höhenwerte, die innerhalb des Grenzwertes liegen, vom Erstellungsprozess ausgeklammert – es wird generalisiert. Das Ergebnis ist in diesem Fall ein TIN, das etwa ein Sechstel der Knoten von Fall A enthält. Mit der Reduzierung der Knoten vermindert sich auch das Speichervolumen.

Bei ersten empirischen Vergleichen von Folgeprodukten (Höhenlinien, Schumierungen) der genannten TINs fielen keinerlei gravierende Unterschiede auf.

Bleiben Fall C und D (Abb. 17), die jeweils ein TIN darstellen, das mit einer Toleranz erstellt wurde.

In beiden Fällen beträgt die Toleranz $\pm 25\text{cm}$, dabei weisen beide Fälle große Unterschiede in der Anzahl der Knoten auf. Die Reduzierung der Knoten in den vorliegenden Testgebieten (Mischgebiet, Stadt, Wald), reichte von 50–10% des Ausgangswerts.

Wie erklärt sich dieser Unterschied? Da in beiden Fällen (C und D) die gleiche Toleranz angewandt wurde, muss die Ursache für den Unterschied an anderer Stelle gesucht werden. Nimmt man Abbildung 17 zur Hilfe bleibt das Raster bzw. dessen Entstehung als Erklärung.

Bei der Interpolation, also der Erstellung des Rasters aus der Punktwolke, wird die Oberfläche geglättet und somit generalisiert. Das bedeutet:

- An Stellen mit einer hohen Punktdichte in einer Punktwolke werden Koordinatenwerte zu neuen Werten zusammengefasst.
- Dafür werden im Gegenzug an Stellen, an denen wenige oder gar keine Punkte vorhanden sind, auf Grund des Bestrebens ein Raster zu bilden, interpolierte Werte eingefügt.

Die letzten Punkte fallen bei der TIN Erstellung durch die Verwendung der Toleranz meist weg, da sie ein relativ stetiges Gelände mit Unebenheiten innerhalb der Toleranz beschreiben.

Durch die Glättung/Generalisierung, werden bereits bei der Erstellung des Rasters mehrere Punkte zu einem interpolierten Punkt zusammengefasst und die Punktzahl dadurch reduziert.

Insgesamt sind folglich bei der TIN-Erstellung in Fall D weniger relevante Koordinatenpunkte bei der Erzeugung beteiligt, als das im Fall C der Fall ist.

Dies erklärt, warum ein TIN aus einem Raster, das unter Berücksichtigung einer Toleranz erstellt wurde, kleiner ist als ein TIN, das direkt aus einer Punktwolke (ebenfalls mit Toleranz) erstellt wurde, obwohl das Raster mehr Punkte enthält als die Punktwolke.

Bleibt noch die Frage offen, wie genau bzw. wie ungenau die TINs aus Raster werden. Denn möglicherweise ist die Genauigkeit der kleineren TINs (kleiner aus Sicht des Datenvolumens) ausreichend für einige Anwendungen. Diese Anwendungen könnten u.a. von den schnelleren Bearbeitungszeiten profitieren. Aber auch die gesamte Datenhaltung würde sich aufgrund der reduzierten Datenmenge vereinfachen.

Fassen wir die Parameter der Erstellungsprozesse einmal zusammen:

Bei der TIN Erstellung direkt aus Punktwolken und der TIN Erstellung aus Raster legt man eine Toleranz fest – diese ist bekannt (hier $\pm 25\text{cm}$). Dass dabei ein Unterschied zwischen den Ergebnissen der beiden Erstellungsarten besteht, wurde im vorigen Abschnitt beschrieben. Als Ursache wurde die Erstellung des Rasters mit der dazugehörigen Interpolation angegeben.

Nach Angaben des LVAs wird zur Interpolation der Raster die Interpolationsmethode CURVE (Glockenkurve) des Programms SCOPE eingesetzt. Auf diese Interpolationsmethode wird ein Filter mit einer Toleranz von $\pm 20\text{cm}$ gesetzt. Je nach Höhe des Filterbetrags erfolgt eine schwächere oder stärkere Glättung des Geländes.

Bei der genannten Toleranz handelt es sich allerdings nicht um eine absolute Toleranz, sondern um einen *angestrebten mittleren Filterbetrag* von $\pm 20\text{cm}$. Das bedeutet, dass Punkte über 20cm liegen können. Wird der Betrag um das Fünffache überschritten (1m), schaltet das Programm von der Interpolationsmethode CURVE auf die Interpolationsmethode LINEAR um.

Für den Fall D kann dadurch, dass es sich bei der Rastererstellung um einen angestrebten und keinen absoluten Filterbetrag handelt, nur schwierig – wenn überhaupt – eine Genauigkeit angegeben werden. Ein Vergleich der Fälle C und D in Zahlen ist daher nur sehr wage, wenn überhaupt möglich.

Tatsache ist, dass definitiv eine Reduzierung der Knoten und somit des Detailgrads in Fall D gegenüber Falls C erfolgt. Da momentan keine konkreten Genauigkeitsansprüche für Anwendungen vorliegen, ist eine Aussage, welche Erstellungsmethode zu empfehlen ist, schwierig. Aus Sicht eines hohen Genauigkeitsanspruchs ist Fall A am besten geeignet, allerdings spricht die große Datenmenge dagegen. Der umgekehrte Fall, kleine Datenmenge bei niedrigem Genauigkeitsanspruch, ist in

Fall D vorhanden. Einen guten Kompromiss stellt Fall C (TIN aus Punktwolke mit Toleranz) wegen einer effektiven Datenhaltung und einer zügigen Verarbeitung dar. Durch entsprechende Wahl der Toleranz kann man hier die Genauigkeit variieren.

4.4 **TIN in der Datenbank**

Die Möglichkeit, TINs mit Hilfe der Spatial Data Engine in einer Datenbank zu speichern und zu verwalten, besteht zu diesem Zeitpunkt (2003) noch nicht. Voraussichtlich wird dies ab der nächsten Version von ArcGIS (ArcGIS 9.0) möglich sein.

4.5 **GRID oder TIN**

Fasst man das letzte Kapitel zusammen, werden GRIDs sowohl dateibasiert als auch in der Datenbank unterstützt. TINs hingegen werden nur auf der Festplatte unterstützt.

Ein weiteres Argument für das ESRI GRID-Format gleich zu Beginn des Kapitels 4.1 war die temporäre Umwandlung von Rasterformaten Dritter (TIFF, ERDAS Imagine) und TIN-Datensätze in ein ESRI GRID, um Folgeoperationen, wie die Berechnung von Höhenlinien, durchführen zu können. Außerdem verfügen ESRI GRIDs über Pyramid Layer, was ebenfalls in Kapitel 4.1 dargestellt wurde. Diese ermöglichen GRIDs, dass sie auch bei einer großen Datenmenge schnell am Monitor aufgebaut werden.

TINs hingegen besitzen, je nach Wahl der Toleranz bei der Erstellung, von vorneherein ein kleineres Datenvolumen.

Ein entscheidendes Argument wird im Kapitel 6 für das ESRI GRID-Format geliefert. Dafür ist es allerdings notwendig, dass eine möglichst große Anzahl von DGM-GRIDs vorliegt. Dazu müssen zuerst Werkzeuge vorhanden sein, die Konvertierungs- und Importaufgaben weitgehend automatisch für die DGM-Daten erledigen.

5 Programmierete Werkzeuge im Überblick

[Graphische Übersicht s. Anhang 13: Programmierete Werkzeuge]

Beim Import von Raster bzw. GRID und TIN wurden einige Wege beschrieben, wie man mit ArcGIS die betreffenden Formate erzeugt. Doch alle vorgestellten Möglichkeiten bieten keinen direkten Weg, die LVA ASCII-Dateien zu importieren, entweder sind zu viele manuelle Schritte notwendig oder es können nur einzelne Dateien importiert werden, in den meisten Fällen treffen beide Unzulänglichkeiten zu.

36

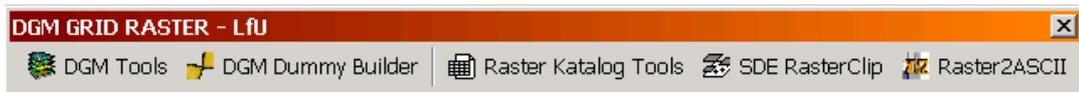


Abb. 18: Werkzeugleiste DGM-GRID-RASTER in ArcMap

Aufgrund der enormen Anzahl der erwarteten ASCII Dateien (36.000) und des damit verbundenen Datenvolumens, das es zu verarbeiten gilt, sowie der portionenweisen Anlieferung der ASCII Dateien ist eine einfache flexible und vor allem weitgehend automatische Verarbeitung der Daten nötig. Zu diesem Zweck wurden verschiedene Werkzeuge mit Hilfe der VB- und VBA-Schnittstelle (Visual Basic [for Applications]) in ARCOBJECTS programmiert. Diese Werkzeuge stellen Funktionen für den Daten Im- und Export der Rasterdaten bereit, außerdem bieten weitere Werkzeuge typische Funktionalitäten im Zusammenhang mit Rastern. Dazu gehört das Erzeugen von Mosaiken oder Rasterkatalogen, sowie Werkzeuge für die Weiterverarbeitung.

ArcObjects ist die Basis von ArcGIS, es ist eine Sammlung von Softwarekomponenten mit GIS Funktionalität und programmierbaren Schnittstellen. Die ArcObjects Technologie beruht auf dem COM Protokoll (s. Glossar). Anpassungen, wie in diesem Fall das Programmieren eigener Werkzeuge oder anderer Funktionalitäten, werden mit dem vorhandenen VBA-Editor oder einer anderen COM-kompatiblen Programmiersprache wie VB, Visual C++ oder Delphi, programmiert.

Eine graphische Übersicht über die einzelnen Werkzeugkomponenten mit ihren Funktionen ist im Anhang 13 zusammengestellt.

5.1 DGM-Tools, Import- und Verwaltungswerkzeuge

Die neuste Version des DGM-Werkzeugs ASCII TO GRID, das zum Erzeugen von GRIDs, TIFFs und ERDAS IMAGINE Formaten dient, kann eine beliebige Anzahl von 1m und 5m ASCII-DGM-Dateien des LVA in die vorher genannten Formate umwandeln. Dabei präsentiert es sich in einer einfachen, intuitiven graphischen Oberfläche (s. Abb. 20).

Bevor die erste Version programmiert wurde, wurden durch manuelles Umwandeln mit herkömmlichen ArcGIS Werkzeugen die nötigen Schritte für eine automatische Umwandlung ermittelt. Dabei galt es auch, Inkompatibilitäten der ASCII-Dateien mit den Standardwerkzeugen festzustellen und diese durch eigene Werkzeuge zu beseitigen. Ein Problem beispielsweise war die Koordinatenreihenfolge in den

ASCII-Dateien – wurde diese nicht geändert, waren alle GRIDs nach der Umwandlung um 90° gedreht. Die ArcObjects Funktion für die Rotation eines Rasters brachte allerdings kein befriedigendes Ergebnis.

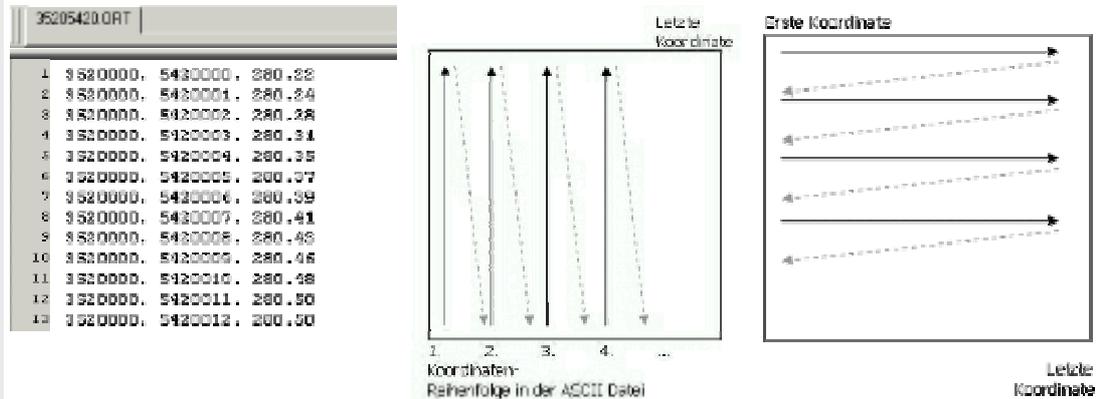


Abb. 19: Koordinatenreihenfolge Ist/Soll

Alle GRIDs wiesen keine exakte Drehung um 90° auf, obendrein verschob sich der Ursprung um etwa einen halben Meter. Bemühungen, diesem Phänomen auf den Grund zu gehen, hatten keinen Erfolg (ESRI Homepage, ESRI Foren, Handbücher, Support, etc.).

Gelöst wurde das Problem letztendlich, indem die einzelnen Zeilen (eine Zeile = ein XYZ-Koordinatentripel) der vorliegenden ASCII-Datei (s. Abb. 19 l. o.) zuerst in ein Array gelesen wurden und anschließend in der benötigten Reihenfolge (s. Abb. 19 r. o.) aus dem Array ausgelesen und in eine zweite temporäre ASCII-Datei, inklusive weiterer Informationen für die Weiterverarbeitung, geschrieben wurden.

Ein weiteres größeres Problem, das eine Überarbeitung der ersten Version zur Folge hatte, war die unterschiedliche Anzahl von Koordinaten in den ASCII-Dateien. Bis nach der Fertigstellung der ersten Version lieferte das LVA Kacheln innerhalb der Landesgrenzen. Die dritte Datenlieferung enthielt zum ersten Mal Kacheln, die über die Landesgrenzen hinausgingen bzw. auf ihr lagen. Diese „Randkacheln“ waren nicht mit den erwarteten 1.002.001 XYZ-Koordinaten, wie die Kacheln im Landesinnern gefüllt, sondern enthielten weniger bis keine Koordinaten. Da der Algorithmus der ersten Version des Import-Werkzeuges aber eine vollständig gefüllte ASCII-Datei erwartet, musste der Algorithmus überarbeitet werden. Bei dieser Gelegenheit wurde die Methode der GRID Erstellung geändert:

In der zweiten Version wird zuerst ein NODATA-GRID mit der entsprechenden Georeferenz und einer Ausdehnung von 1.000 x 1.000m erstellt. Anschließend werden, wie in der ersten Version, alle Zeilen aus der betreffenden ASCII-Datei in ein Array gelesen, allerdings wurden die Z-Werte inklusive der XY-Koordinaten im Gegensatz zur vorhergehenden Version verarbeitet.

Im nächsten Schritt wird jeder Wert aus den Array gelesen und entsprechend seiner XY-Koordinate in das NODATA-GRID geschrieben. Sind nicht für jede Rasterzelle im GRID Werte vorhanden, bleiben an diesen Stellen NODATA-Werte zurück. Die Werte die über das 1.000 x 1.000m große GRID hinausgehen, werden weggelassen, denn wie an anderer Stelle beschrieben, umfasst die tatsächliche Ausdehnung der ASCII-Dateien 1.001 x 1.001m. Der 1m breite Streifen am nördlichen und östlichen Ende jeder Kachel überlappt sich mit der nächsten Kachel, somit geht keine Information verloren.

Das ASCII TO GRID Werkzeug ist so konzipiert, dass es auch Änderungen in der Formatierung durch das LVA kompensieren kann, sofern jedes Koordinatentripel in einer eigenen Zeile steht und sich der Dateiname, wie eingangs gezeigt, aus der Gauß-Krüger-Koordinate der linken unteren Ecke zusammensetzt. Da diese Art der Namensgebung ein bewährtes System ist und sich u.a. bei den Orthobildern bereits bewährt hat, ist eine Änderung von dieser Seite nicht zu erwarten.

Änderungen ergaben sich allerdings zwischen dem Testdatensatz und den aktuellen Lieferungen in der Spaltenanordnung der Koordinaten. Während die XY-Koordinaten des Testdatensatzes in Spalte i zu finden waren, begannen die Koordinaten der Lieferungen in Spalte n. Das Gleiche gilt für die Z-Werte. Um diesen Variationen flexibel zu begegnen, zeigt die Bedienoberfläche INPUT KONFIGURATION (Abb. 20 r.) individuell einstellbare Werte für die Koordinatenpositionen und Höhenwerte in den zu verarbeitenden ASCII-Dateien. (s. Abb. 20 r.)

Damit häufig vorkommende Einstellungen, wie diejenige für 1m und 5m Dateien, nicht bei jedem Start neu eingestellt werden müssen, wurden vorgefertigte Profile für diese Dateien erstellt.

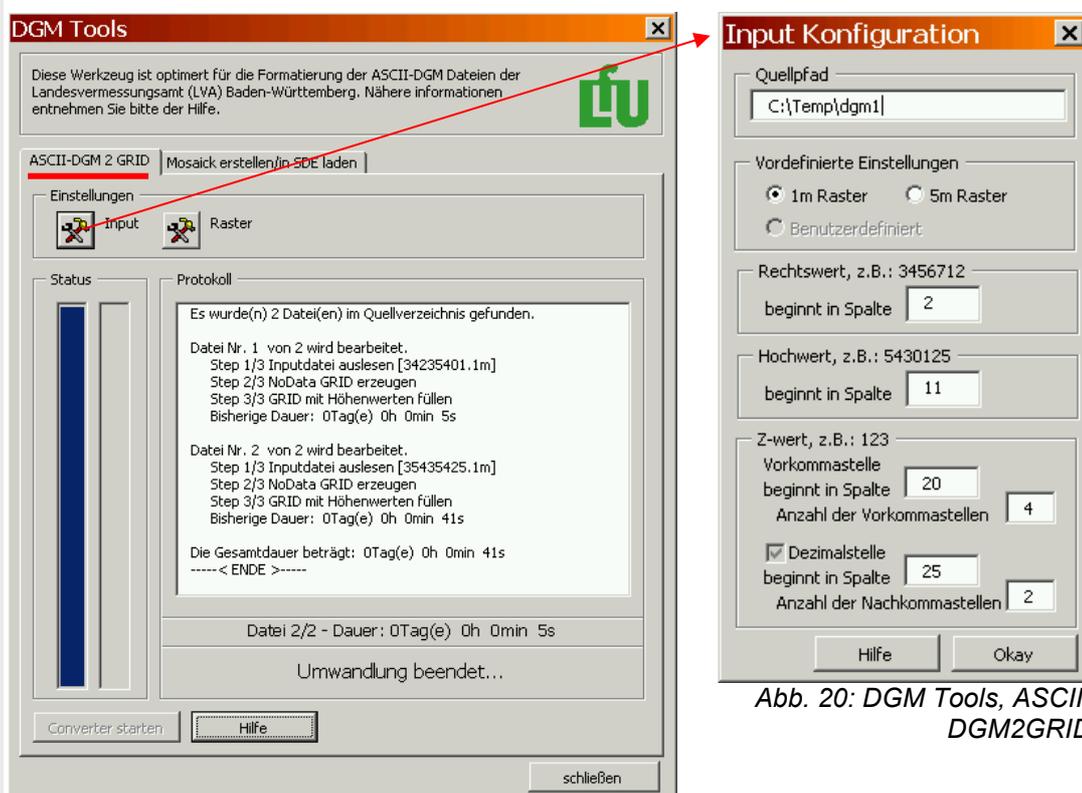


Abb. 20: DGM Tools, ASCII-DGM2GRID

Weiter bietet das Werkzeug die Möglichkeit, verschiedene Ausgabeformate zu erzeugen. Während in der Regel GRIDs erstellt werden, ist es auch möglich, TIFF- oder das ERDAS IMAGINE-Format zu erzeugen.

Es besteht auch die Möglichkeit, zwischen der Erstellung von Integer- und Floating Point-GRID zu wählen. In der Regel sind die Z-Werte Dezimalwerte mit zwei Nachkommastellen. Standardeinstellung ist somit Floating Point-GRID.

Weitere typische Elemente wie Hilfedialog, Statusanzeige, detailliertes Protokoll usw. helfen dem Benutzer und geben ihm die Möglichkeit, jederzeit über den Fortschritt des laufenden Prozesses informiert zu werden und notfalls eingreifen zu können.

Zur Leistung des Werkzeugs ist zu sagen, dass im Laufe der Entwicklung einige Optimierungen durchgeführt wurden, sodass die Erstellungsgeschwindigkeit, von knapp 50 Sekunden für ein 1m GRID inklusive Pyramid Layer auf etwa 30 Sekunden gesenkt werden konnte. Dies wurde z.B. dadurch erreicht, dass einige Tests für die Dimensionierung des Arrays durchgeführt wurden, in den die einzelnen Zeilen der ASCII-Datei geschrieben werden. Es standen verschiedenen Möglichkeiten zur Verfügung, die Koordinaten aus der ASCII-Datei auszulesen.

Mit der Funktion READALL wird die ganze Datei gelesen. Sie kann anschließend mit der Funktion SPLIT in kleinere Stücke zerlegt werden. Mit Hilfe von REDIM wird ein Array bei jedem Durchlauf auf die gewünschte Größe erweitert oder gekürzt. Als schnellste Version ergab sich die Funktion REDIM, die zu Beginn des Algorithmus einmal durchgeführt wurde. Dabei wurde der Array auf einen Wert gesetzt, der groß genug ist, um die maximale Anzahl von Koordinaten (1.002.001) in einer ASCII-Datei aufzunehmen.

5.2 Mosaik-Erstellung und SDE-Loader

Hinter dem zweiten Karteireiter des DGM-Tools (s. Abb. 21) verbirgt sich das Werkzeug zum Erstellen eines Mosaiks und zum optionalen Laden in die Datenbank über die SDE. Betrachtet man den Prozessablauf (Anhang 13, Workflow) ist dieser Schritt der GRID-Erstellung nachgeordnet.

Unter dem Erstellen eines Mosaiks versteht man das Zusammenführen einzelner GRIDs zu einem großen GRID. Beim Ergebnis der Mosaikbildung unterscheidet man zwei Arten. Bei der einen entsteht wieder ein *dateibasiertes* GRID, bei der anderen Version entsteht ein *SDE-Raster* in der Datenbank. Für beide Arten existieren bereits Werkzeuge in ArcGIS, die sich aber in einigen Merkmalen von den selbstprogrammierten Werkzeugen unterscheiden.

Für das Erstellen eines dateibasierten Mosaiks kann man das Spatial Analyst Werkzeug RASTERCALCULARTOR (s. Abb. 22, ArcMap → Spatial Analyst) benutzen. Mit Hilfe des abgebildeten Befehls lassen sich maximal 23 GRIDs in einem Schritt zusammenführen.

Die begrenzte Anzahl ist der Grund dafür, dass ein eigenes Werkzeug programmiert wurde, mit dem ebenfalls keine unbegrenzte Anzahl von Rastern zu Mosaiken zusammengefasst werden kann, aber dennoch deutlich mehr als mit dem RASTERCALCULATOR. In Tests wurden in einem Prozess bis zu 45 1m GRIDs zu

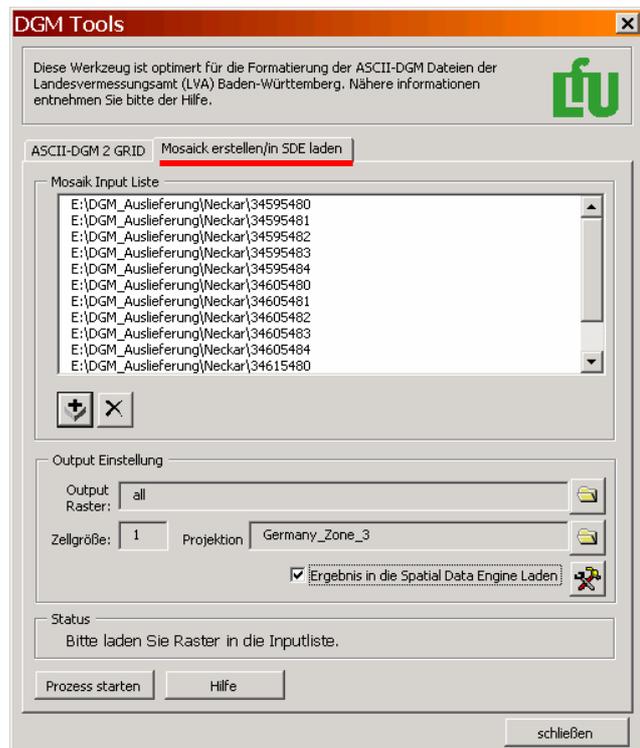


Abb. 21: DGM Tools, Mosaic/SDE-Loader

einem Mosaik zusammengefasst. Die dafür nötigen Parameter sind neben den obligatorischen Input-GRIDs die Angabe des Speicherorts und der Name des Output-GRIDs.

Anders gestaltet sich die Situation bei Mosaiken in der Datenbank. Beim Laden von mehreren GRIDs in die Datenbank über die SDE werden diese automatisch zu *einem* Raster zusammengeführt.

Dabei muss das erste GRID eine Projektion besitzen. Ist dies nicht der Fall, muss eine Projektion mittels eines entsprechenden Dialogs zugewiesen werden.

GRIDs können mit dem Werkzeug RASTER TO GEO-DATABASE in die Datenbank geladen werden. Das Werkzeug steht im ArcCatalog und in der ArcToolbox zur Verfügung. Gegenüber dem selbstprogrammierten Werkzeug hat dies den Vorteil, dass man nicht an die 2GB Grenze der 32bit Rechnerarchitektur für eine Datei gebunden ist (Kapitel 6). Somit lassen sich bis zu getesteten 150GB in die Datenbank mit der SDE laden und zu einem Mosaik verbinden.

Einen Nachteil bringt das ArcGIS Werkzeug allerdings mit sich. Beim Laden von mehr als einem Raster werden an das Ergebnis raster Zeilen an den nördlichen Rand angehängt. Für die eigentliche Bearbeitung ist dies nicht weiter störend. Beim ersten Auftreten des Fehlers wurden die zusätzlichen Zeilen mit der Verzerrung durch die Projektion erklärt, die beim Laden in die SDE notwendig ist, allerdings konnte diese Erklärung mit Tests widerlegt werden. Weitere Theorien und auch der Wechsel von ArcGIS 8.2 zu ArcGIS 8.3 brachten keine neuen Erkenntnisse.

Lädt man zwei oder mehr GRIDs in ein bestehendes SDE-Raster, sollte man sich sicher sein, dass alle Werte der zu ladenden GRIDs korrekt sind. Denn im Falle einer Überlappung werden die *zuletzt geladenen* Werte für eine Rasterzelle gespeichert, d.h. bestehende und korrekte Werte des Rasters in der Datenbank könnten mit falschen Rasterwerten *überschrieben* werden.

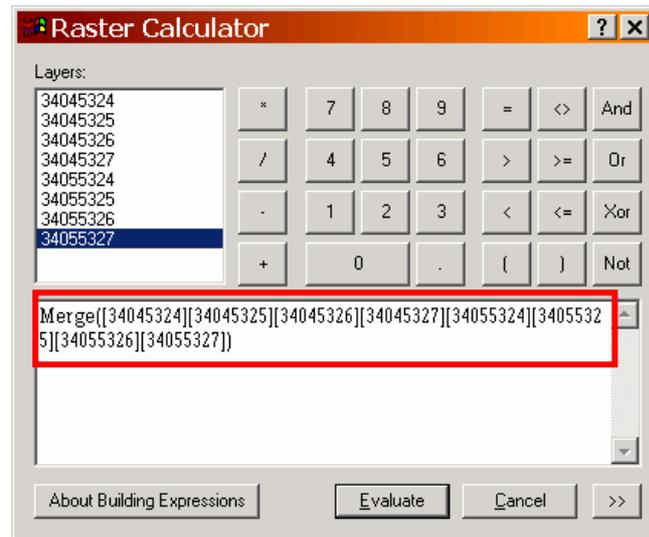


Abb. 22: Spatial Analyst, Raster Calculator mit Mosaikbefehl und Rastern

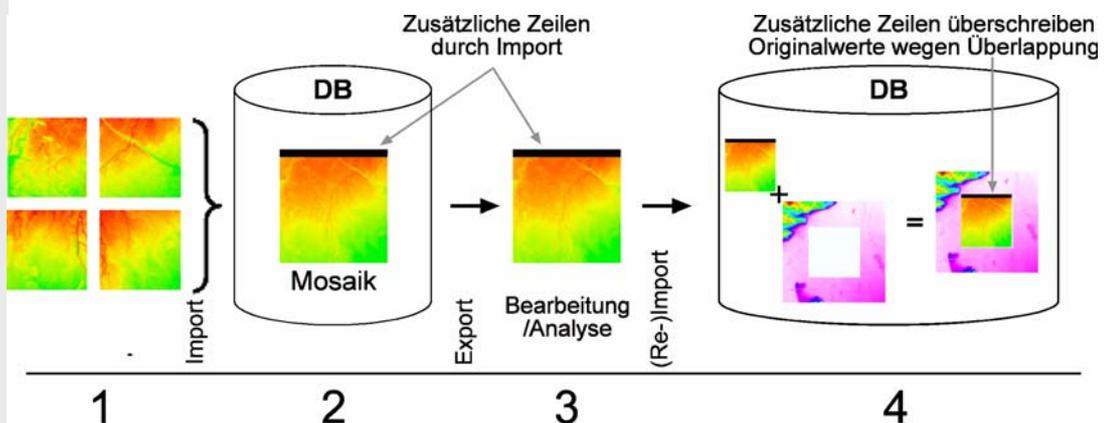


Abb. 23: Überschreiben von korrekten Werten bei Mosaiken

Um das Problem dieser zusätzlichen Zeilen zu umgehen, wurde das in Abbildung 21 gezeigte Werkzeug (Mosaik erstellen/in SDE laden) für das Erstellen von Mosaiken und das Laden in die Datenbank geschrieben. Denn die zusätzlichen Zeilen am oberen Rand eines Rasters treten nicht auf, wenn nur *ein* Raster in die Datenbank geladen wird. Aus diesem Grund fügt das programmierte Werkzeug die einzelnen GRIDs zusammen, *bevor* sie durch das Programm in die Datenbank geladen werden. Auf diese Weise hat auch das Raster in der Datenbank die erwartete Anzahl von Zeilen und Spalten. Da aber das entstandene Mosaik vor dem Laden in die Datenbank aus Dateien bestand, ist man wieder an die 2GB Grenze gebunden, die oben erwähnt wird und deren Zustandekommen in Kapitel 6 näher beleuchtet wird.

Beide Werkzeuge haben trotz ihrer ähnlichen Aufgabe durch die genannten unterschiedlichen Merkmale ihre Berechtigung. Liegt das Mosaik, das in die SDE geladen wird unter 2GB, ist es sicherer, das programmierte Werkzeug zu benutzen, da es keine ungewollten Werte erzeugt. Für das Laden von Gebieten über 2GB, z.B. für ein landesweites Raster, kommt nur das ArcGIS Werkzeug, RASTER TO GEODATABASE in Frage.

Die Frage, welches Werkzeug wann zu benutzen ist, erübrigt sich, wenn man sich den oben genannten Fehler beim Import bewusst macht. Als letzte Hilfestellung zum Verständnis dieses Fehlers wird empfohlen, eine Anzahl von aneinandergrenzenden GRIDs mit dem ArcGIS Werkzeug RASTER TO GEODATABASE in die Datenbank zu laden und die Anzahl der Zeilen des resultierenden SDE-Rasters mit den errechneten Zeilen zu vergleichen. Letztere ergeben sich durch Aufaddieren der Zeilen der einzelnen Input GRIDs. [Graphische Übersicht s. Anhang 13: Programmierte Werkzeuge]

5.3 Rasterkatalog

Ein Rasterkatalog wird benutzt, um mehrere aneinandergrenzende Raster vereint anzuzeigen, ohne sie zu einem Mosaik und somit zu einer großen Datei zusammenzuführen zu müssen. Im ArcCatalog werden Rasterkataloge als einfache Tabellen angezeigt. Die Datensätze in einer Rasterkatalog-Tabelle definieren, welche Raster in den Rasterkatalog eingebunden werden. Beim Anzeigen des Katalogs wird immer vom ersten zum letzten Tabelleneintrag geladen. Die Raster eines Katalogs können unterschiedliche Formate aufweisen und sich überlappen. Die überlappenden Regionen haben den Vorteil, dass im Gegensatz zum Mosaik keine Informationen verloren gehen, da jedes Raster individuell vorliegt. Als Tabelle dient jedes unterstützte Tabellenformat inklusive Textdateien. Die Tabelle kann auch in einer Datenbank vorliegen. Die einzige Bedingung ist das Vorhandensein, der fünf Spalten IMAGE, XMIN, YMIN, XMAX and YMAX. Die Spalte IMAGE enthält den Namen und Pfad des Rasters, die restlichen Spalten enthalten die Ausdehnung (s. Abb. 25).

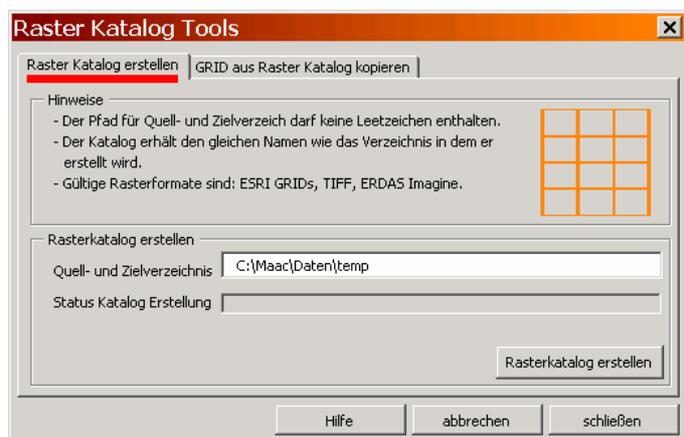
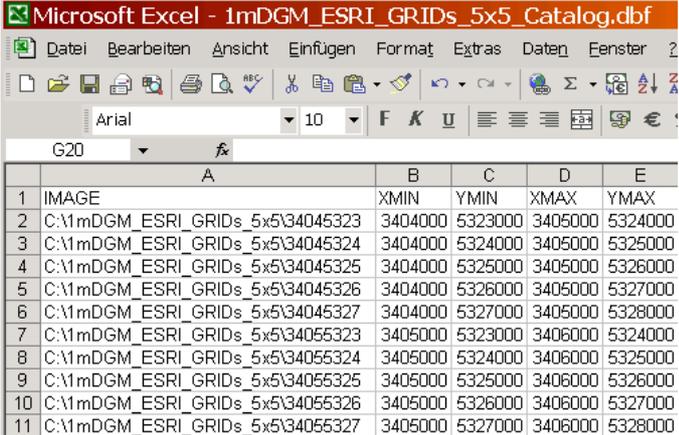


Abb. 24: Rasterkatalog Werkzeug

Weitere Spalten können vorhanden sein, werden allerdings nicht für den Rasterkatalog berücksichtigt. Um eine schnelle Übersicht über eine große Anzahl von Rastern zu bekommen, ist ein Rasterkatalog äußerst hilfreich. Auch wenn es darum geht, Informationen in überlappenden Gebieten zu behalten, da diese – wie bereits erwähnt – in einem Mosaik verloren gehen würden. Um die Anzeigegeschwindigkeit möglichst hoch zu halten, unterstützt ein Rasterkatalog mit Hilfe von einigen Einstellungen, ab wie vielen Rastern auf dem Bildschirm ein Gitter mit Platzhaltern gezeichnet werden soll oder ob die eigentlichen Raster angezeigt werden (s. Abb. 26).

Bei jeder Lieferung wurde für die ESRI GRIDs nach dem Erstellen mit den DGM-Tools ein Rasterkatalog erstellt, um mit ArcMap eine Übersichtskarte für das Intranet im PDF-Format zu erstellen. Ein Beispiel einer solchen Karte kann im Anhang 2 eingesehen werden.

Ein Rasterkatalog kann manuell erstellt werden, wenn die oben genannten Spalten vorhanden sind. Da dies bei einer großen Anzahl von Rastern recht zeitintensiv ist und es in den meisten Fällen Probleme beim Anzeigen *selbstediteter* Rasterkataloge gibt, wurde für diesen Vorgang das Werkzeug RASTER KATALOG TOOLS → RASTERKATLAOG ERSTELLEN (s. Abb. 24) geschrieben.



	A	B	C	D	E
1	IMAGE	XMIN	YMIN	XMAX	YMAX
2	C:\1mDGM_ESRI_GRIDs_5x5\34045323	3404000	5323000	3405000	5324000
3	C:\1mDGM_ESRI_GRIDs_5x5\34045324	3404000	5324000	3405000	5325000
4	C:\1mDGM_ESRI_GRIDs_5x5\34045325	3404000	5325000	3405000	5326000
5	C:\1mDGM_ESRI_GRIDs_5x5\34045326	3404000	5326000	3405000	5327000
6	C:\1mDGM_ESRI_GRIDs_5x5\34045327	3404000	5327000	3405000	5328000
7	C:\1mDGM_ESRI_GRIDs_5x5\34055323	3405000	5323000	3406000	5324000
8	C:\1mDGM_ESRI_GRIDs_5x5\34055324	3405000	5324000	3406000	5325000
9	C:\1mDGM_ESRI_GRIDs_5x5\34055325	3405000	5325000	3406000	5326000
10	C:\1mDGM_ESRI_GRIDs_5x5\34055326	3405000	5326000	3406000	5327000
11	C:\1mDGM_ESRI_GRIDs_5x5\34055327	3405000	5327000	3406000	5328000

Abb. 25: Bsp. Für einen Rasterkatalog in Excel

34045327	34055327	34065327	34075327
34045326	34055326	34065326	34075326
34045325	34055325	34065325	34075325
34045324	34055324	34065324	34075324

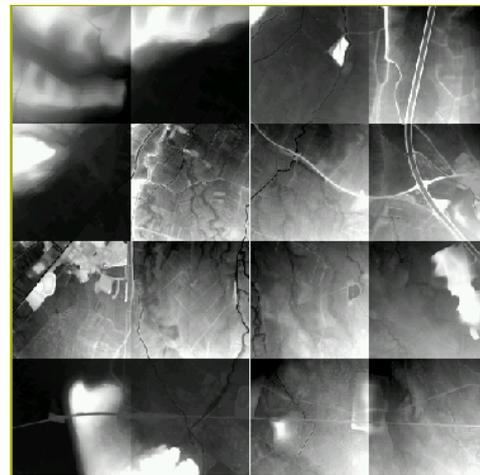


Abb. 26: Rasterkatalog: l. Platzhalter, r. Raster

Das Werkzeug RASTERKATALOG ERSTELLEN, erstellt einen Rasterkatalog aus allen in einem Quellverzeichnis gespeicherten Rastern. Für die Erstellung werden absolute Pfade verwendet. Hinter dem zweiten Karteireiter verbirgt sich das Werkzeug GRID AUS RASTERKATALOG KOPIEREN (s. Abb. 27).

Es wurde notwendig dieses Werkzeug zu programmieren, da sich das manuelle Zusammenstellen der GRID-Kacheln aus einem Verzeichnis, die für die Weiterverarbeitung von Dritten angefordert wurden, als äußerst mühsam erwies. Mit Hilfe der angegebenen Koordinaten und des ArcCatalog mussten bei der

manuellen Zusammenstellung die GRID-Kacheln einzeln an ihrem Dateinamen identifiziert und kopiert werden. Bereits für eine kleine Anzahl von Dateien ist dies ein erheblicher zeitlicher Aufwand. Mit Hilfe des Programms GRID AUS RASTERKATALOG KOPIEREN, konnte dieser Vorgang weitgehend automatisiert und so beschleunigt werden.

Voraussetzung ist, dass der Rasterkatalog in der Datenbank vorliegt und die entsprechende Laufwerksverbindung zu den GRID-Kacheln besteht. Sind diese Voraussetzungen gegeben, kann man sich durch die Angabe zweier Koordinaten, die eine Boundingbox des betreffenden Gebiets beschreiben und dem Namen des Rasterkatalogs, die betreffenden GRID-Kacheln, die innerhalb oder auf der Grenze des Gebiets liegen, in ein Zielverzeichnis kopieren lassen. Aus diesen kopierten Kacheln könnte man sich anschließend ein Mosaik erstellen lassen.

Das Werkzeug enthält wie alle Werkzeuge eine Hilfe, die eine kurze Anleitung und mögliche Fehlerquellen bei der Anwendung enthält.

Ein Statusbalken gibt Auskunft über den Fortschritt des Kopierprozesses. Das Werkzeug ist außerdem so programmiert, dass es ohne Probleme für andere Projekte verwendet werden kann. So könnten auch Orthobilder, die in einem Verzeichnis hinterlegt und in einem Rasterkatalog zusammengefasst sind, mit diesem Werkzeug bearbeitet werden.

Im Arbeitsprozess (Anhang 13, Workflow) hat das Werkzeug seinen Platz als möglicher Output für eine Weiterverarbeitung der DGM-Daten.



Abb. 27: Kacheln aus Rasterkatalog kopieren

5.4 Raster Clipper

Mit Hilfe des RASTER CLIPPERS können definierbare Bereiche aus einem Datenbankraster kopiert und exportiert werden.

Da in kaum einer Anwendung Interesse an der Bearbeitung der gesamten Landesdaten besteht, ist es nötig, Rasterdaten aus der Datenbank exportieren zu können. Für diesen Zweck wurde der Raster Clipper programmiert. In den meisten Fällen werden Teilgebiete, wie zum Beispiel ein Kreis, eine Gemeinde oder ein Gewässereinzugsgebiet bearbeitet. Durch die Eingabe zweier Eckkoordinatenpaare für eine Boundingbox, einem Namen für das Exportraster sowie einem Zielverzeichnis wird der betreffende Ausschnitt als ESRI GRID, TIFF, ERDAS

IMAGINE oder JPEG exportiert. Dieser Ausschnitt kann optional im RASTER CLIPPER weiterverarbeitet werden. Als Möglichkeiten der Weiterverarbeitungen stehen folgende Operationen auf den zu exportierenden Ausschnitt zur Verfügung:

- Schummerung
- Hangneigung
- Höhenlinien
- Exposition

Bei allen Operationen kann man individuelle Parameter angeben. So ist es möglich, bei der Schummerung den horizontalen und vertikalen Beleuchtungswinkel zu variieren und optional Schatten berechnen zulassen. Bei Höhenlinien kann man die Äquidistanz und optional eine Höhenbezugsbasis angeben.

Weiter besteht die Möglichkeit, sich direkt nach dem Erstellen die Ergebnisse in ArcMap anzeigen zu lassen. Alle Rasteroperationen können in den gleichen Dateiformaten erstellt werden wie der zu exportierende Ausschnitt (s. Abb. 28).

Die eben genannten Operationen sind als Beispiele für Raster- (Schummerung, Hangneigung und Exposition) und Vektor- (Höhenlinien) Outputs implementiert worden, da sie die wohl am häufigsten verwendeten Operationen darstellen. Man kann den Raster Clipper mit Hilfe von ArcObject um alle Operationen erweitern, die im Spatial- und 3D-Analyst vorhanden sind, sofern man die Lizenzen besitzt. [Graphische Übersicht s. Anhang 13]

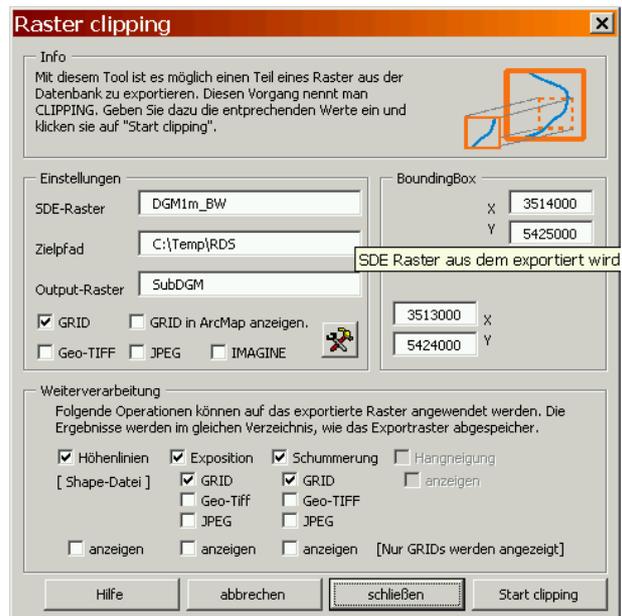


Abb. 28: Werkzeug, Raster Clipper

5.5 Dummy Builder

Im Laufe der Diplomarbeit wurden durch das LVA etwa 8% der DGM-Daten geliefert. Im GRID-Format belegt diese Menge ungefähr 8GB auf der Festplatte. Diese Menge stellt für die Verwaltung in der Datenbank kein größeres Problem dar. Aus diesem Grund wurde der DUMMY BUILDER programmiert. Mit ihm kann man georeferenzierte Pseudo-GRIDs erzeugen. So wurde eine Datenmenge angelegt, die in etwa der zu erwartenden 140GB entspricht.

Auf diese Weise wurde das Verhalten und Arbeiten der Datenbank mit dieser großen Datenmenge für das ganze Land getestet. Genauer ist Gegenstand von Kapitel 6.

Bei der Erstellung der Pseudo-GRIDs lassen sich einige Parameter angeben.

- Anzahl der GRIDs, in X- und Y-Richtung.
- Anzahl der Spalten und Zeilen, wobei man durch den Algorithmus auf maximal 5.000 x 5.000 Zeilen und Spalten beschränkt ist. Der Algorithmus lief auch bei 8.000 x 8.000 für kleine GRID Mengen. Bei dem Versuch, GRIDs mit 10.000 x 10.000 Zeilen und Spalten zu erstellen, stürzte das Programm ab. Möglichst große Gitter lassen sich besser in ArcGIS handhaben. ArcGIS kommt mit wenigen aber großen Dateien besser zurecht als mit sehr vielen kleinen.
- Georeferenziert werden die GRIDs durch die Angabe der linken unteren Ecke des linken untern GRID. Alle weiteren GRIDs werden relativ mit Hilfe der Zeilen- und Spaltenangaben entsprechend georeferenziert.
- Damit beim Laden der GRIDs in die Datenbank über SDE die Komprimierungsalgorithmen in der Datenbank möglichst authentisch arbeiten, werden die Dummy GRIDs mit Zufallswerten aus einem definierbaren Intervall, hier von 75 bis 1500, gefüllt (s. Abb. 37).

Alle genannten Parameter für den DUMMY BUILDER, sind im Quellcode des entsprechenden VB-Moduls einstellbar. Da es sich um ein Werkzeug handelt, das eine relativ einmalige Anwendung ist, wurde keine aufwendige graphische Oberfläche erstellt. [Graphische Übersicht s. Anhang 13]

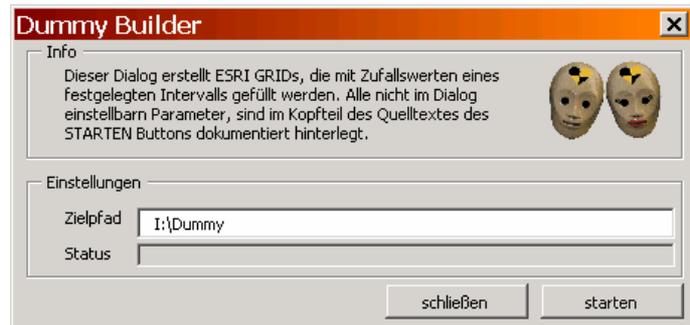


Abb. 29: Werkzeug, SDE Dummy Builder

5.6 Raster 2 ASCII

Das jüngste Werkzeug (RASTER 2 ASCII) wurde aufgrund von Anfragen nach XYZ-Daten programmiert. Es ist ein Programm, das aus einer beliebigen Anzahl ESRI GRIDs in einem Verzeichnis ASCII-Dateien mit oder ohne Headerinformationen erstellt. Im Grunde entspricht der dabei entstehende Output den ASCII-Dateien, die vom LVA als Input dienen. Der Grund, warum statt der Inputdaten aus den ESRI GRIDs der entsprechende Output erzeugt wird, liegt an der recht umständlichen Suche der entsprechenden Dateien, die auf viele CDs und DVDs verteilt sind, sofern sie nicht auf den Server kopiert wurden. Anstelle der manuellen Suche kann man das angefragte Gebiet mit dem RASTER CLIPPER exportieren und anschließend das exportierte GRID mit dem RASTER 2 ASCII Werkzeug in eine ASCII-Datei umwandeln.

XYZ-Daten werden in diesem Fall als Input für CAD-Programme benutzt oder allgemein als Schnittstelle zu anderen Programmen und Systemen, die mit den gängigen Rasterformaten inkompatibel sind.



Abb. 30: Raster2ASCII Werkzeug

6 Möglichkeiten einer blattschnittfreien Speicherung (landesweiter Zugriff)

6.1 Grenze durch 32Bit Architektur

Einige Anwendungen, wie zum Beispiel Hochwassersimulationen für den Hochwassergefahrenschutz, benötigen für ihre Berechnungen große zusammenhängende Bereiche in Form von Gewässereinzugsgebieten. Für diese Anwendungen werden einzelne Raster zu größeren zusammengefasst (*mosaicing*) oder aus dem landesweiten SDE-Raster in der Datenbank ausgeschnitten und exportiert (*clipping*).

Die Zusammenfassung von GRIDs zu einem großen GRID kann nicht beliebig erfolgen. Der Kern des Problems liegt in der 32Bit Architektur und dem damit verbundenen adressierbaren Speicherraum eines Rechners. Dieser adressierbare Speicherraum beläuft sich bei 32Bit Systemen ungefähr auf 4GB (Genaueres weiter unten), wovon laut ESRI, in ArcGIS 2,147GB für die Rasterspeicherung adressiert werden können. Bei den vorliegenden 1m ESRI GRIDs (floating point) à 3,85MB ohne Pyramid Layer bedeutet dies rechnerisch eine maximale Anzahl von 557 GRIDs, die zu einem Mosaik zusammengeführt werden können.

Ein GRID-Datensatz besteht aus mehreren Dateien (s. Abb. 31), Insgesamt kann dieser Datensatz eine Größe von 2GB überschreiten, die *einzelnen* Dateien in ihm allerdings nicht. Der Grund für diese Speicherbarriere lässt sich folgendermaßen erklären:

Im Allgemeinen kann man mit n Bits 2^n Integerzahlen (s. Glossar) im Binärsystem darstellen. Also kann eine 4Bit Binärzahl 2^4 oder 16 Integerzahlen darstellen, eine 8Bit Binärzahl ergibt 2^8 oder 256 Integerzahlen, usw.

Würde man nun von einem 32Bit-auf ein 64Bit-System umsteigen, würde der Wertebereich der Integerzahlen von $2^{32} = 4,3e9$ (32Bit) auf $2^{64} = 1,8e19$ (64Bit) ansteigen, dies entspricht einer Zahl im einstelligen Exabyte Bereich (entspr. 1Mrd Gigabytes oder 1.024 Petabytes). Aus diesem Grund kann ein 64Bit System einen wesentlich größeren Integerbereich zur Verfügung stellen.

Betrachtet man folglich Speicheradressen vereinfacht als „spezielle Integerwerte“, kann dementsprechend ein Prozessor, der mehr Integerwerte kennt, einen größeren Speicherraum adressieren. So errechnen sich die oben genannten 4GB eines 32 Bit Systems aus $2^{32} = 4,3\text{Mrd}$.

Diese 4,3Mrd. Bytes betrachten wir gerundet als 4GB. Nach dieser Rechnung könnte somit ein 64Bit System theoretisch bis zu 18 Exabytes adressieren. Der

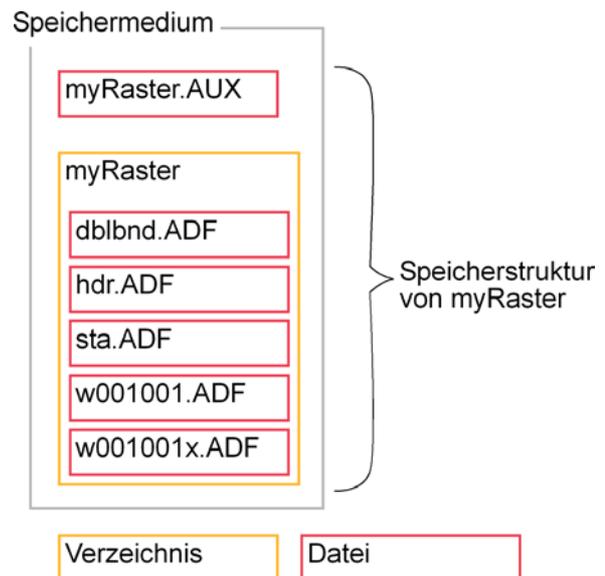


Abb. 31: Dateizusammensetzung eines ESRI GRID

tatsächlich adressierbare Raum ist in der Praxis wesentlich kleiner, aber im Vergleich zu einem 32Bit System eine überaus ausreichende Erweiterung, mit der die oben genannte 2GB Barriere gelöst werden könnte.

Für eine 64Bit Lösung müsste allerdings ArcGIS nicht nur 64Bit-Plattformen unterstützen, sondern es müsste die 64Bit Technik auch selbst nutzen. Die derzeitige Version ist für 32Bit-Systeme ausgelegt, eine Änderung für kommende Versionen ist zu diesem Zeitpunkt (2003) nicht bekannt.

Durch die 2GB Grenze ist ein blattschnittfreier und vor allem landesweiter Zugriff auf 1m-DGM-Daten als ESRI GRID nicht möglich, es müssen demzufolge andere Lösungen gefunden werden.

Welche Möglichkeiten für die Realisierung eines landesweiten, blattschnittfreien Zugriffs auf die DGM-Daten vorhanden sind und wo man auf Grenzen stößt, zeigen die nächsten Abschnitte.

6.2 Lösungen für großräumige Bearbeitungen

Die nachstehende Rechnung zeigt die Mosaik-Möglichkeiten mit den vorliegenden DGM-Daten:

1m ESRI GRID (floating, ohne Pyramid Layer)

Maximale Anzahl allgemein:
$$\frac{\text{Maximaladressierbarer Speicher in MB}}{\text{Größe eines 1m bzw 5 m DGM GRID}}$$

$$1\text{m-DGM: } 1\text{GRID} = 3,85\text{MB} \rightarrow \frac{2147\text{MB}}{3,85\text{MB}} = 557\text{GRIDs}$$

Die Rechnung zeigt, dass im Idealfall aufgrund der 2GB Grenze maximal 557 1m-DGM-GRIDs zu einem Mosaik zusammengefasst werden können. Häufig liegen Kacheln, die zu einem Mosaik zusammengefasst werden sollen, nicht direkt nebeneinander, dies soll an dieser Stelle als NORMALFALL bezeichnet werden. Im Gegensatz zum IDEALFALL (s. Abb. 32) entstehen dabei so genannte NODATA-BEREICHE.

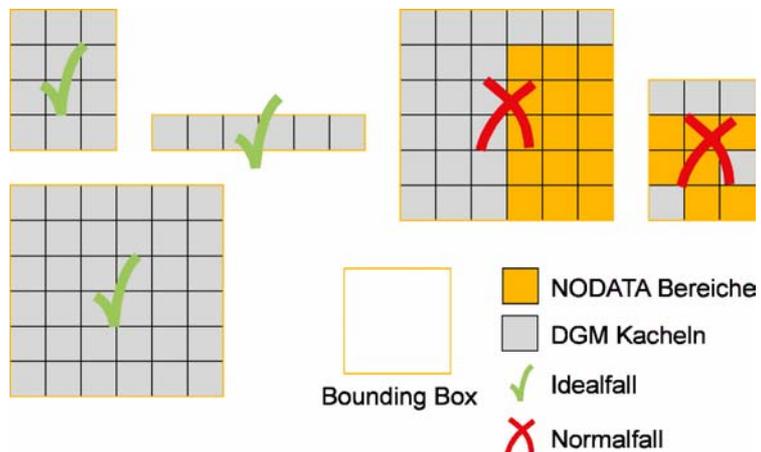


Abb. 32: NODATA Anordnungen

Wie NODATA-Werte entstehen und wie sie helfen, die 2GB Grenze bedingt zu erweitern, stellt Kapitel 6.2.1 vor.

Im Falle der Hochwassersimulation haben Versuche gezeigt (Anhang 5), dass das größte Gewässereinzugsgebiet (GEG) innerhalb der Grenzen des Landes Baden-Württemberg die 2GB Grenze nicht überschreitet.

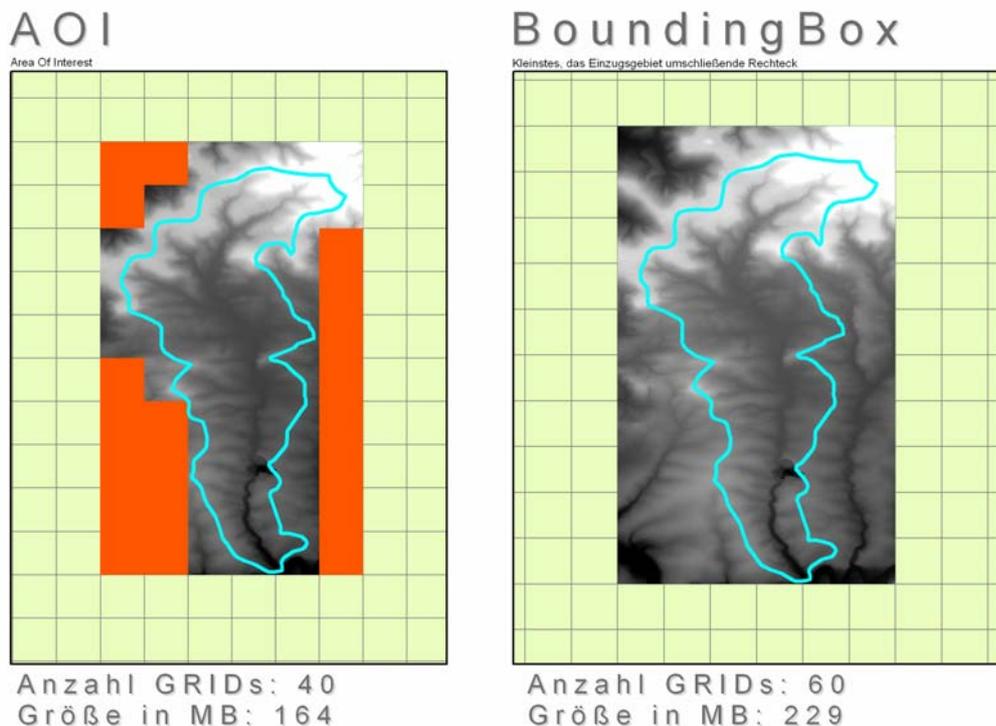
Allerdings werden sicherlich mit der Zeit andere Aufgabenstellungen diese Grenze überschreiten. Daher muss für diesen Fall eine Lösung angeboten werden, auch im Hinblick auf die Aufgabenstellung dieser Arbeit, die einen „blattschnittfreien Zugriff auf landesweite Geländemodelle“ fordert. Mehrere Möglichkeiten wurden in diesem Zusammenhang erarbeitet.

6.2.1 NODATA-Bereiche

Ein Mosaik entsteht, durch Zusammenfassung mehrerer GRIDs. Um ein Mosaik wird – wie bei jedem graphischen Objekt – eine so genannte BOUNDINGBOX gezeichnet, also das minimal umschließende Rechteck. Liegen die GRIDs, die zusammengeführt werden sollen, nicht direkt nebeneinander (s. Abb. 32: Normalfall) entstehen NODATA-Bereiche (Gelb in Abb. 32).

Für verschiedene Anwendungen, oder wie in der Aufgabenstellung gefordert, um einen „landesweiten blattschnittfreien DGM-Zugriff“ zu erreichen, wachsen die Bereiche, die durch DGM-Daten in Form von Mosaiken abgedeckt werden müssen, unter Umständen schnell über 2GB und sind somit nicht mehr mit einem 32Bit System bearbeitbar.

Für diese Fälle kann man je nach Anwendung mit Hilfe der NODATA-Bereiche die Gebiete dennoch als gesamtes Mosaik darstellen, in dem man nur eine AREA OF INTEREST (AOI) aus dem betreffenden Gebiet auswählt. D.h. dass man nur die wirklich benötigten Bereiche im Mosaik mit Werten füllt und die restlichen Bereiche als NODATA-Werte darstellt.



© 2003 Marc Panczak - Marc.Panczak@ifu.fh-wm.de



Abb. 33: GEG Wüstenbach, Auszug eines ArcMap Layouts

Die Erstellung eines solchen Mosaiks, in dem nur die AOI dargestellt wird, ist denkbar einfach. Anstatt alle DGM-Kacheln aus einem Gebiet zu wählen, werden nur die benötigten Kacheln ausgewählt und zu einem Mosaik zusammengefügt. Für kleine Gebiete kann dies manuell erfolgen. Für größere Gebiete wurde das GIS-

DATA EXPORT Werkzeug (LfU-Entwicklung) für ArcView3.2 verwendet. Mit ihm ist es möglich, über ein Polygon Daten eines bestimmten Bereiches auszuwählen.

Dass NODATA-Bereiche keinen zusätzlichen Speicherplatz belegen, zeigt der folgende Versuch:

Um eine physikalische Speichergröße der NODATA-Bereiche ermitteln zu können, wurden AOI-Mosaik und Boundingbox-Mosaik bei gleicher Kachelanzahl miteinander verglichen.

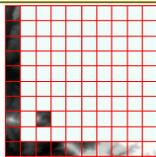
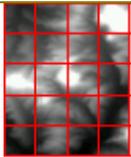
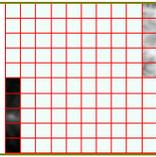
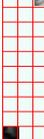
Anzahl der beteiligten GRID	NODATA in km ²	Differenz, AOI zu Boundingbox	AOI (leere Felder sind NODATA-Bereiche)	Boundingbox
20	80	0 MB		
10	90	0 MB		
2	98	0 MB		
2	28	0 MB		

Tabelle 1: Größenvergleich von AOI- und Boundingbox-Mosaiken

Mit dem Versuch wird gezeigt, dass man mit Hilfe von NODATA-Bereichen und einer AOI ein Gebiet abdecken kann, dass größer als 2GB wäre, wenn man alle enthaltenen Kacheln verwenden würde. Durch die Auswahl der tatsächlich benötigten Bereiche ist es möglich, auch größere Gebiete darzustellen.

6.2.2 Maßstabsreduzierung

Eine weitere Möglichkeit die 2GB-Grenze hinauszuzögern, wäre eine Reduzierung der Auflösung, also ein Zurückgreifen auf das 5m-DGM oder die Wahl einer noch größeren Rasterweite.

$$\text{Maximale Anzahl 5m-DGM: } 1\text{GRID} = 0,24\text{MB} \rightarrow \frac{2147\text{MB}}{0,24\text{MB}} = 8.945\text{GRIDs}$$

Wirft man einen Blick auf die vorherige Rechnung, sieht man, dass auch mit der 5m Rasterweite keine landesweite Berechnung durchgeführt werden kann. Erwartet werden etwa 36.000 1km²-Kacheln, aber nach der obigen Rechnung könnte man nur 8.945 zu einem Mosaik zusammenfassen.

Also stellt man die Frage umgekehrt: Wie groß muss die Auflösung sein, um die 2GB Grenze nicht zu überschreiten?

Diese Berechnung birgt eine Schwierigkeit in sich. Stellt man die oben verwendete Rechnung einfach um, muss man vom Speichervolumen *eines* ESRI GRIDs auf seine Rasterweite rückschließen.

Mit der Voraussetzung, dass die Ausdehnung des GRIDs 1.000 x 1.000m ist, wird im Folgenden versucht die Rasterweite näherungsweise zu bestimmen.

Maximale Größe eines GRIDs, bei erwarteten 36.000 GRIDs für Baden-Württemberg, bei einem maximalen Datenvolumen aller GRIDs unter 2.147MB:

$$1\text{GRID} = \frac{2147\text{MB}}{36.000\text{GRIDs}} = 0,059\text{MB}$$

51

Die Rechnung zeigt, dass ein Raster rechnerisch maximal 59KB groß sein darf, sofern man ein Mosaik aus allen Kacheln bilden möchte, um einen landesweiten Zugriff auf die DGM-Daten zu realisieren.

In Versuchen wuchs das Speichervolumen eines ESRI GRIDs mit der Anzahl von Rasterzellen linear, aus diesem Grund ist anzunehmen, dass man mit einer bekannten Anzahl von Rasterzellen und dem dazugehörigen Speichervolumen in einem Dreisatz die maximale Anzahl von Rasterzellen in einem 59KB großen GRID näherungsweise errechnen kann. Nachdem dieser Wert errechnet ist, kann man durch Ziehen der Wurzel die Anzahl von Rasterzellen einer Kante feststellen.

Da alle Kacheln quadratisch sind und eine Fläche von 1km^2 haben, ist eine Kante 1000m lang:

$$\begin{array}{lll} 3,87\text{MB}^* \rightarrow & 3870\text{KB} & 1.000.000\text{Rz (Rasterzellen)} \\ 1\text{KB} & 259\text{Rz} & \\ 59\text{KB} & 15.245\text{Rz} \rightarrow & 123\text{Rz pro Kantelänge} \end{array}$$

$$\frac{1000\text{m}}{123\text{Rz / Kantelänge}} = 8,13\text{m}$$

*Wert für ein 1m ESRI GRID (floating ohne Pyramiden)

Teilt man nun die Kantelänge durch die Anzahl der Rasterzellen, erhält man eine theoretische Auflösung von mindestens 8,13m für ein 59KB großes GRID, mit dem es möglich ist, das gesamte Land in einem Raster zu erfassen und dabei unter 2GB Speichergröße zu bleiben.

In der Praxis bestätigt sich die Theorie nur teilweise. So ergaben Tests, dass ab einer Auflösung von etwa 25m ein entsprechendes Speichervolumen gegeben ist, das die 59KB Größe nicht überschreitet.

Eine Reduzierung des Maßstabs bedeutet gleichzeitig den Verlust eines Großteils der gerade erst durch das 1m-DGM gewonnenen Informationen. Diese Tatsache entspricht nicht dem Ziel, das mit dem Aufbau des 1m-DGM angestrebt wurde, nämlich eine genauere Datenbasis für entsprechende Anwendungen zur Verfügung zu stellen. Von den bisher beschriebenen Möglichkeiten eines landesweiten Zugriffs, ist die Maßreduzierung die am wenigsten geeignete Lösung.

6.2.3 Rasterkatalog

In Kapitel 5.5.3 wurde beschrieben, was ein Rasterkatalog ist, woraus er besteht und wie er erstellt wird. Nun soll untersucht werden, ob ein Rasterkatalog möglicherweise die Aufgaben eines Mosaiks übernehmen kann.

In Anbetracht der 2GB Dateigröße wäre der Rasterkatalog eine Alternative zu einem Mosaik. Anstatt ein zusammenhängendes Gebiet aus einer großen Datei darzustellen, stellen viele einzelne Dateien das Gebiet dar. Aber dies ist leider auch der einzige Vorteil, der sich ergeben würde. Bei der enormen Anzahl von einzelnen Rastern, die im Katalog enthalten wären, hätte ArcGIS Probleme sie in akzeptabler Geschwindigkeit anzuzeigen. In der Regel wird durch die Pyramidenbildung ein Mosaik schneller angezeigt als ein Rasterkatalog.

Ein weiteres Problem stellen Berechnungen aus dem DGM wie z.B. die Erzeugung von Höhenlinien für Baden-Württemberg. Solche SPATIAL- oder 3D ANALYST Operationen sind nur auf einzelne Raster anwendbar und nicht auf einen Rasterkatalog. Somit ist ein Rasterkatalog für einen landesweiten Zugriff unbrauchbar.

6.2.4 TIN

TINs können, wie bereits an früherer Stelle erwähnt, noch nicht in der Datenbank gespeichert werden. Bleibt also die Frage nach einem dateibasierten, landesweiten Zugriff.

An dieser Stelle entsteht die gleiche Problematik, die sich bereits bei den dateibasierten Rastermosaiken gestellt hat. Wie alle Dateien sind auch TIN-Dateien auf eine Größe von maximal 2,147GB beschränkt.

Eine genaue Anzahl von TIN Kacheln, die zusammen ein großes TIN ergeben sollen, ist schwierig anzugeben. Bei Rastern ist die Angabe einfach, da alle das gleiche Datenvolumen auf der Festplatte besitzen. Bei TINs variiert die Dateigröße durch die unterschiedliche Anzahl der Knoten von Kachel zu Kachel, hervorgerufen durch die verwendete Toleranz beim Erstellen und durch die Größe der Ausgangspunktmenge.

Für eine Visualisierung der unterschiedlichen Knotenanzahlen kann man die Dichtebilder zu Hilfe nehmen, die neben den bekannten DGM-Daten vom LVA mitgeliefert werden. Diese Dichtebilder beschreiben in fünf Klassen, wie viele Punkte pro 4m^2 ($2\text{x}2\text{m}$) aufgenommen wurden. Die unregelmäßigen Einfärbungen

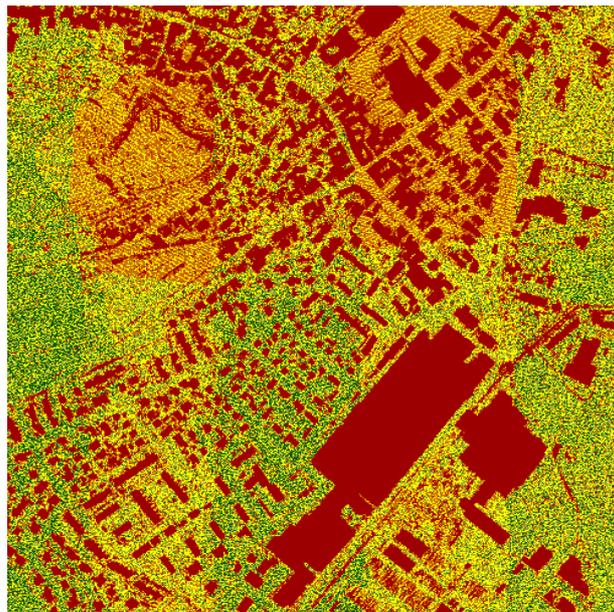


Abb. 34: Dichtebild, Punktdichte/ 4m^2

lassen auf eine unregelmäßige Vermaschung und somit auf eine unregelmäßige Anzahl von Knoten schließen.

Die in vorhergehenden Abschnitten genannten Knotenzahlen in TINs lassen sich bestenfalls relativ zueinander betrachten, da während der Entstehung dieser Arbeit zu wenig Punktwolkendateien vorlagen, um eine repräsentative Aussage über die zu erwartende durchschnittliche Knotenanzahl und somit eine Angabe zur Dateigröße eines TINs machen zu können.

Eines lässt sich bereits zu diesem Zeitpunkt sicher sagen, bei allen bisher durchgeführten Versuchen würde ein landesweites TIN die 2GB Grenze überschreiten.

Eine Lösung wäre eine Glättung durch eine großzügige vertikale Toleranz bei der TIN Erstellung. Diese Lösung kann man vom Effekt mit der Auflösungsreduzierung bei Rastern gleichsetzen – beide Verfahren bedeuten eine Maßstabsreduzierung.

In Kauf nehmen muss man allerdings einen mehr oder weniger großen Verlust von Information, je nachdem wie groß die Toleranz bei der Erstellung gewählt werden muss. Eine derartige Lösung entspricht, wie bereits erwähnt, nicht dem Ziel, dass man mit den hochauflösenden 1m-DGM-Daten verfolgt.

6.2.5 Landesweiter Zugriff auf die Datenbank

Alle bisherigen Methoden führen zu Informationseinbußen, also Generalisierung oder Weglassen von unwichtigen Informationen. Eines haben allerdings alle Methoden gemeinsam: sie sind in ihrem Datenvolumen an die 2GB Grenze gebunden.

In Kapitel 4.2 (Konvertierung in SDE-Raster) wurde bereits darauf aufmerksam gemacht, dass ArcGIS die Speicherung von Rastern mit Hilfe der ArcSDE in einem RDBMS unterstützt.

Die Speicherung von Daten in einer Datenbank unterliegt letztendlich auch den bekannten Grenzen. D.h. der Inhalt von Datenbanken wird auch in Dateien abgespeichert.

In Versuchen soll gezeigt werden, dass es möglich ist, ein landesweites, blattschnittfreies Raster in der Datenbank abzubilden. Dieses Raster kann allerdings wegen der bekannten Grenzen als Datei auf der Festplatte nicht realisiert werden. Grundlage dieses Versuchs ist die Vermutung, dass durch die datenbankinterne Datenverwaltung solch große Datenmengen verarbeitet und gespeichert werden können.

In mehreren Tests soll zuerst eine Datenmenge deutlich über 2GB geladen werden. Diese Menge soll in mehreren Versuchen sukzessive erhöht werden. Als Kontrolle, ob das Raster nicht defekt oder fehlerhaft ist, dient das oben beschriebene Werkzeug RASTER CLIPPER, mit dem ein Teilraster aus dem geladenen, landesweiten SDE-Raster in der Datenbank exportiert werden kann (s. Abb. 35), wodurch ein Zugriff auf ein landesweites DGM-Raster realisiert wird.

Das Laden der einzelnen DGM-Kacheln erfolgt mit dem Werkzeug RASTER TO GEODATABASE im ArcCatalog. Das Konsolenwerkzeug SDELOADER bietet in diesem Fall keine Alternative, da mit ihm zwar TIFF-Dateien geladen werden können, aber kein ESRI GRID Format.

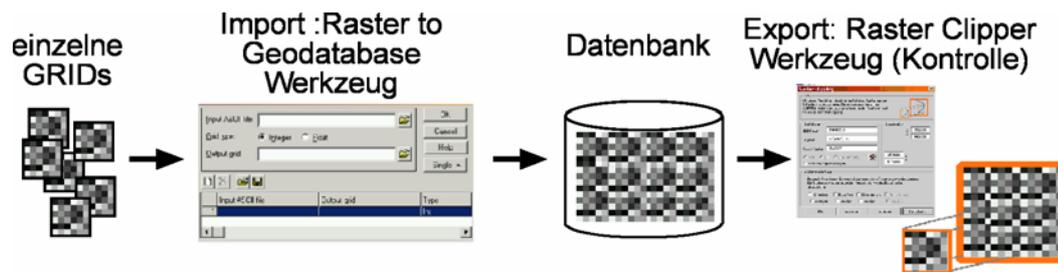


Abb. 35: die einzelnen Schritte des DB Versuchs

Neben dem Versuch, auf ein DGM-Raster über 2GB zugreifen zu können, interessiert ebenfalls, welches Datenvolumen das Raster in der Datenbank einnimmt, um Rückschlüsse auf die zu erwartende Datenmenge für ganz Baden-Württemberg und somit eine Aussage über die Effektivität dieser Art der Datenhaltung machen zu können.

Zu diesem Zweck wurde aus dem in Kapitel 4.2 bekannten Tabellenschema für SDE-Raster ein SQL-Skript formuliert, das für ein beliebiges Raster in der Datenbank die EXTENT Größe aus der Tabelle USER_EXTENTS (Oracle spezifische Tabelle) in Megabyte ausliest (s. Anhang 9). Trotz genauer Beachtung des Tabellenschemas, wich das Ergebnis stets vom dem des TABLESPACE (s. Glossar) ab, der für diese Tests angelegt wurde. Das SQL-Skript wurde daraufhin nicht weiter verwendet. Stattdessen wurde das Speichervolumen mit Hilfe des TABLESPACE, der anfänglich nur zur Kontrolle dienen sollte, ermittelt.

Mit Hilfe dieses TABLESPACE ist es möglich, exakt das Speichervolumen zu bestimmen, das von den geladenen Daten, in der Datenbank benötigt wird. Vor dem Laden der Daten könnte man ihn als eine leere Festplatte bezeichnen. Deshalb entspricht der belegte Speicher nach dem Laden exakt dem Datenvolumen der geladenen Raster.

Insgesamt sind viele Versuche durchgeführt worden, bei denen die beschriebenen Daten in die Datenbank geladen wurden. Viele Versuche schlugen fehl. Grund dafür waren hauptsächlich Fehler in Ladeprofilen und Ladezusammenstellungen, sowie Fehler durch Unterbrechungen im Netzwerk durch Dritte. Die Uploadzeit für die in den ersten Versuchen verwendeten 8,25GB betrug etwa 6-8 Stunden, je nach Auslastung des Netzwerks. Die Uploadversuche wurden daher meist nachts durchgeführt.

Spätere Versuche mit knapp 150GB dauerten inklusive Statistiken und exklusive Pyramid Layern etwa vier Tage. Insgesamt waren diese Versuche am aufwendigsten, da Unterbrechungen, wie die genannten Netzwerkprobleme, in jedem Fall zu einem Abbruch des Prozesses führten, welcher erst am darauf folgenden Morgen bemerkt wurde und wiederholt werden musste. An dieser Stelle werden nur die markanten Versuche dargestellt.

Im **ersten Versuch** wurden alle zu diesem Zeitpunkt vorhandenen 1km² DGM-Kacheln in die Datenbank geladen. Dabei handelt es sich um 8,25GB auf der Festplatte (2.185 GRIDs). Diese belegen nach dem Laden in die Datenbank gut 45GB ohne die Bildung von Pyramid Layern. Die Bildung von Pyramid Layern ist in diesem Zusammenhang sekundär, da Pyramid Layer dazu dienen, eine möglichst rasche Darstellung auf dem Bildschirm zu unterstützen. Das DGM-Raster in der Datenbank hingegen dient als Grundlage für Anwendungen, wie Höhenlinien oder Sichtanalysen oder zum Export von Teilgebieten für eine Bereitstellung der Daten an Dritte. Eine Betrachtung ist deshalb nicht erforderlich und Pyramid Layer können vorerst vernachlässigt werden.

Außerdem würden sie das Speichervolumen im TABLESPACE verfälschen und dadurch eine Größenbestimmung des SDE-Rasters unmöglich machen.

Der Zugriff auf das geladene Raster durch Ausschneiden (clipping) eines Teilrasters mit dem programmierten Werkzeug RASTER CLIPPER, bereitet keine Probleme, solange das zu exportierende Gebiet unter 2GB bleibt.

In weiteren Versuchen wurde getestet, wie sich die Datenbank verhält, wenn eine Datenmenge von der erwarteten Größe des 1m-DGM Baden-Württembergs in der Datenbank abgelegt wird.

Da zur Zeit der Entstehung dieser Diplomarbeit nur ein Bruchteil der gesamten DGM-Daten vorlagen (8%), mussten Daten improvisiert werden. Für diesen Zweck wurde das Werkzeug DUMMY BUILDER programmiert.

Mit Hilfe dieses Werkzeugs lassen sich, wie in Kapitel 5.5. bereits beschrieben, *Dummy ESRI GRIDS* erstellen. Diese lassen sich, neben anderen definierbaren Parametern, mit Zufallswerten für Z-Werte in einem definierbaren Intervall füllen. Wie man sich diese Daten vorzustellen hat, zeigt Abbildung 37:

Rechts sieht man einen Ausschnitt aus einer regulären DGM-Kachel, links den Ausschnitt einer Dummy Kachel.

Im Gegensatz zur regulären Kacheln sind die Werte in der Dummy Kachel nicht stetig, was später bei der Datenkomprimierung in der Datenbank zu berücksichtigen ist.

Mit Hilfe dieses Werkzeugs

wurde eine Datenmenge erzeugt, die in etwa der zu erwartenden Menge für ganz Baden-Württemberg, als 1m Raster entspricht (etwa 140GB)

Diese Daten wurden in einem **weiteren Test** – inklusive der vorhandenen DGM-Daten – in die Datenbank geladen ca. 150GB. Der zeitliche Aufwand für das Laden der Daten betrug ungefähr drei bis vier Tage.

Im ersten Schritt wurden die Dummy-Daten geladen. Der belegte Speicher in der Datenbank dafür beträgt 58,3GB, was ein Datenreduzieren von 47,5% bedeutet. Die schlechtere Komprimierung im Gegensatz zu den regulären DGM-Daten im ersten Versuch ist – wie bereits erwähnt – auf die sprunghaft wechselnden Z-Werte in den Dummy-Kacheln zurückzuführen, bei denen die LZ77 Kompression (s. Anhang

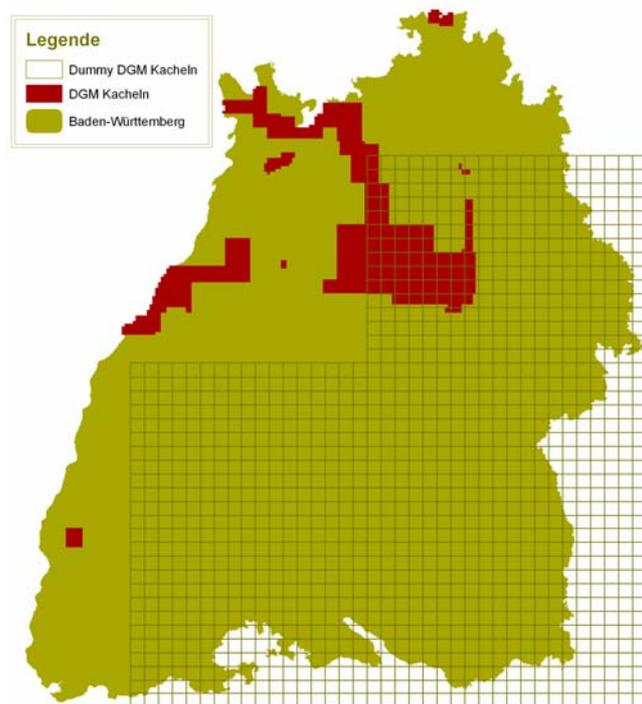


Abb. 36: DGM & Dummy DGM-Kacheln

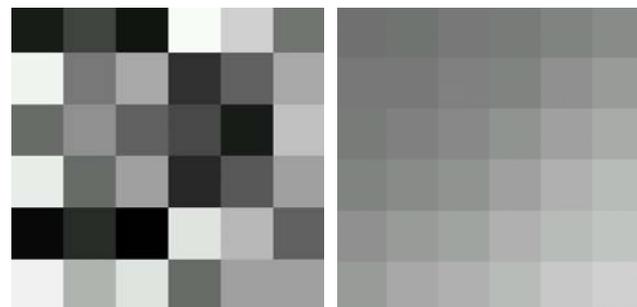


Abb. 37: I. Dummy Raster, r. reguläres DGM-Raster

10) weniger effektiv arbeitet. Ein repräsentativer Wert für die Datenreduktion stellen also die 32% dar, die im ersten Versuch erreicht wurden.

Alle Werte der Tests – Datenvolumen und Datenreduktion – sind in einer Tabelle am Ende des Kapitels zusammengefasst.

Auch bei dem letzten Test wurde auf das landesweite DGM-Raster mit Hilfe des Werkzeugs RASTER CLIPPER zugegriffen und ein Teilraster exportiert, im gleichen Schritt wurden mit dem Werkzeug Folgeprodukte wie Schummerung, Hangneigung, Exposition und Höhenlinien mit allen zur Verfügung stehenden, benutzerdefinierten Parametern in verschiedenen Rasterformaten erfolgreich erzeugt.

Mit dem eben beschriebenen Versuch wird gezeigt, dass es letztendlich doch möglich ist, auf ein landesweites DGM blattschnittfrei zuzugreifen. Das Laden in die Datenbank hat dabei den positiven Effekt, die DGM-Daten verlustfrei auf ca. ein Drittel zu komprimieren. Eine weiterführende Beschreibung der Optimierung für eine bestmögliche Speicherung in der Datenbank, ist im Kapitel 7 *Optimierungen* zu finden.

ersuch 1		
GRID	Festplatte (GB)	Datenbank (GB)
2.185	8,25	40,1
36.000*	~137*	~743*
Versuch 2		
GRID	Festplatte (GB)	Datenbank (GB)
2.185	8,25	22,5
36.000*	~137*	~371,3*
Versuch 3		
GRID	Festplatte (GB)	Datenbank (GB)
2.185	8,25	2,6
36.000*	~137*	~42,6*
Versuch 4		
GRID	Festplatte (GB)	Datenbank (GB)
ca. 37.000 (Originale und Dummies)	128,5	60,6

* errechneter Wert

Tabelle 2: Speicherplatzentwicklung in der Datenbank

6.3 Zusammenfassen der Möglichkeiten

In Kapitel 4.5 wurden die Vor- und Nachteile von GRIDs bzw. TINs zusammengefasst. Abschnitt 6.2.5 brachte eine endgültige Entscheidung für GRID oder TIN.

Alle dateibasierten Lösungen sind an die 32Bit Architektur gebunden, so auch das TIN-Format. Des Weiteren erfordern die meisten Lösungen Informationseinbußen. Somit ist die Speicherung in der Datenbank die einzige Lösung für einen blattschnittfreien und landesweiten Zugriff ohne Informationsverlust. Da zur Zeit nur Rasterformate in der Datenbank abgelegt werden können, fällt die Entscheidung auf ein landesweites SDE-Raster.

7 Optimierungen

7.1 Optimierung von Rasterdaten in der Datenbank

7.1.1 Ladeprofile allgemein

Beim Laden von einem oder mehreren Rastern in die Datenbank kann man in ArcGIS optional ein Ladeprofil aus der Tabelle DBTUNE angeben. Wird kein Profil angegeben, werden Standardwerte aus der gleichen Tabelle verwendet. Diese Standardprofile sehen für Rasterdaten einen EXTENT (s. Glossar) von 50MB für eine Rastertabelle vor (vgl. ArcSDE CONFIGURATION AND TUNING GUIDE FOR ORACLE® 1999, S. 58). Multipliziert man die Extentgröße mit der Anzahl der angelegten Tabellen für ein Raster (vier Tabellen, s. Kapitel 4.2), erklärt sich die Datenmenge von immerhin 200MB, das ein einziges 1km² Raster nach dem Datenbankimport ohne die Angabe eines Ladeprofils belegt.

Da ein 1km² großes GRID auf der *Festplatte* knapp 4MB belegt, ist der Speicherplatz, der durch das Standardprofil reserviert wird, viel zu groß dimensioniert. Egal wie voll ein Speicherblock ist – der von ihm vorgegebene Platz (hier 50MB) wird auf jeden Fall reserviert.

Aus diesem Grund ist es sinnvoll, eigene Ladeprofile anzulegen. Besonders ratsam ist es im Fall der DGM-Daten, da es sich hierbei um Daten handelt, die stetig den gleichen Umfang besitzen.

Die DBTUNE Speicherparameter ermöglichen eine Kontrolle darüber, wie ArcSDE-Clients Objekte in der (Oracle) Datenbank anlegen. Man kann festlegen, wie Speicherplatz für Tabellen und Indices reserviert wird und in welchem TABLESPACE (s. Glossar) eine Tabelle oder Index angelegt wird. Diese und weitere Oracle spezifischen Speicherparameter lassen sich beeinflussen.

Die DBTUNE Speicherparameter werden in der DBTUNE Metadatentabelle in Oracle abgelegt und sind online verfügbar.

```
##XXL
# Dies ist ein Kommentar
F_STORAGE      "PCTFREE 10 PCTUSED 90 INITRANS 4 TABLESPACE SDE_FEATURE
                STORAGE (INITIAL 800M NEXT 800M MINEXTENTS 1 MAXEXTENTS 100
                PCTINCREASE 0
                FREELISTS 4)"
F_INDEX_FID    "PCTFREE 10 INITRANS 4 TABLESPACE SDE_INDEX STORAGE (INITIAL
                800M NEXT 800M MINEXTENTS 1 MAXEXTENTS 100 PCTINCREASE 0
                FREELISTS 4)
                NOLOGGING"
F_INDEX_AREA   [...]
END
```

Abb. 38: Ein Ladeprofil aus der Tabelle DBTUNE

Die Tabelle DBTUNE wird während der Installation von ArcSDE angelegt und mit Standardprofilen gefüllt. Mit dem Konsolenwerkzeug SDEDBTUNE lässt sich die Tabelle exportieren, wonach man sie mit einem einfachen Texteditor um eigene Profile erweitern kann. Anschließend kann man mit dem gleichen Werkzeug die Tabelle wieder nach Oracle importieren.

Die einzelnen Ladeprofile bestehen aus einer Gruppe von Parameterangaben, die unter einem CONFIGURATION KEYWORD zusammengefasst sind. ArcSDE Client Anwendungen und einige Administrator-Werkzeuge greifen mittels diesem

CONFIGURATION KEYWORD auf ein bestimmtes Ladeprofil zurück, wenn ein Objekt angelegt wird, z.B. eine Tabelle.

Wenn ein CONFIGURATION KEYWORD gefunden wurde, werden die Parameter in der zugehörigen Parameter Gruppe in die betreffende SQL-Anweisung, wie CREATE TABLE oder CREATE INDEX eingebettet.

Wird ein vom Benutzer angegebenes CONFIGURATION KEYWORD nicht gefunden, wird auf das entsprechende Standardprofil in der Tabelle DBTUNE zurückgegriffen. Sind diese ebenfalls nicht verfügbar, werden die SQL-Anweisungen zum Anlegen der betreffenden Objekte direkt an Oracle weitergegeben und mit den Oracle spezifischen Parametern angelegt.

Die Struktur einer DBTUNE Datei setzt sich folgendermaßen zusammen: Speicherparameter treten als Kombination von Parametername und Zeichenkette auf. Die einzelnen Speicherparameter werden wiederum durch ein Schlüsselwort zusammengefasst (CONFIGURATION KEYWORD), das zusammen mit einem END die Speicherparameter umschließt. In der Abbildung 38 ist das Schlüsselwort „XXL“, die Parameternamen stehen links darunter. Die Zeichenketten mit den Parametern, umfassen mehrere Zeilen und werden in Anführungszeichen geschrieben.

Lädt man beispielsweise GRIDs mit dem Werkzeug RASTER TO GEODATABASE (ArcCatalog) in die Datenbank, gibt es die Möglichkeit ein Schlüsselwort für ein Ladeprofil anzugeben (s. Abb. 39). Gleiches gilt für das konsolenbasierte Administrator-Werkzeug SDEIMPORT.

Ausführliche Informationen über die Möglichkeiten Parameter zu definieren und anzuwenden findet man im ArcSDE CONFIGURATION AND TUNING GUIDE FOR ORACLE® 1999. (www.esri.com).

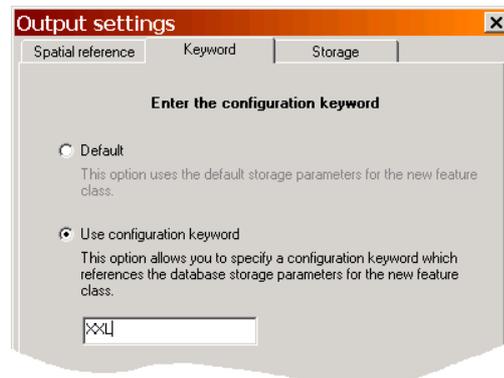


Abb. 39: Ladeprofil in Raster to GeoDB

7.1.2 Ladeprofile für DGM-Daten

Für die DGM-Daten wurde ebenfalls ein eigenes Ladeprofil angelegt und in mehreren Versuchen optimiert. Als Testumfeld wurden die Uploads verwendet, die in Kapitel 6.2.5 beschrieben sind.

Am Ende des gleichen Kapitels zeigt Tabelle 2 die Reduzierung der Datenmenge von Versuch zu Versuch, bis letztendlich keine weitere Minimierung der Daten in der Datenbank durch Änderungen der Speicherparameter erfolgte.

Wie wird diese Minimierung erreicht?

Zur Prüfung des Speicherplatzes hat es sich, am besten bewährt, einen eigene TABLESPACE für die Daten anzulegen, da dadurch der belegte Speicherplatz am genauesten überprüft werden kann.

```
##DGM02
GEOMETRY_STORAGE      "SDEBINARY"
ATTRIBUTE_BINARY      "LONGRAW"
COMMENT "Maac for DGM 1m BW"
RAS_STORAGE           "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL 10M NEXT
10M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4) "
RAS_INDEX_ID          "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL 10M NEXT
10M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4)
NOLOGGING"
BND_STORAGE           "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL 10M NEXT
```

```

BND_INDEX_ID      10M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4)"
                  "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL 10M NEXT
                  10M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4)
                  NOLOGGING"
BND_INDEX_COMPOSITE "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL
10M
                  NEXT 10M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4)
                  NOLOGGING"
AUX_STORAGE       "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL 10M NEXT
                  10M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4)"
AUX_INDEX_COMPOSITE "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL
10M
                  NEXT 10M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4)
                  NOLOGGING"
BLK_STORAGE       "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL 2490M
NEXT
                  2490M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4)"
BLK_INDEX_COMPOSITE "PCTFREE 0 INITRANS 4 TABLESPACE SDE_DGM STORAGE (INITIAL
10M
                  NEXT 10M MINEXTENTS 1 MAXEXTENTS 200 PCTINCREASE 0 FREELISTS 4)
                  NOLOGGING"

END

```

Abb. 40: DGM-Ladeprofil (DGM08)

Das gezeigte Ladeprofil ist das optimierte Ladeprofil, das für die DGM-Daten erarbeitet wurde. Im Folgenden werden die einzelnen Parameter und deren Einstellung für das DGM-Ladeprofil näher erläutert.

- PCTFREE stellt den prozentualen Anteil in jedem Datenblock dar, der als freier Platz reserviert ist. Dieser Platz wird genutzt, wenn Daten, die bereits in diesem Block gespeichert sind, wachsen. Da die Daten im Falle der DGM-Daten, einmal abgelegt, nicht weiter wachsen, wurde der Wert auf 0 gesetzt.
- INITRANS stellt einen Wert dar, der für Transaktionen benutzt wird. Jede Änderung der Daten wird im Datenblock-Header bis zum nächsten COMMIT oder ROLLBACK gespeichert. Wenn man davon ausgehen muss, dass zwei Benutzer öfters gleichzeitig auf ein und den selben Datenblock zugreifen oder eine Tabelle häufig mit parallelen DML-Befehlen (Data Manipulation Language) bearbeitet wird, sollte der INITRANS-Wert entsprechend hoch gesetzt werden. Da dies bei den DGM-Daten in der Regel nicht der Fall ist, wird der Wert entsprechend klein gewählt.
- TABLESPACE ist eine Art virtuelles Laufwerk für die Datenbank. Für die DGM-Daten wurde eine eigene, dieser logischen Einheiten mit dem Namen SDE_DGM eingerichtet. Bei der Angabe dieses Namens im Ladeprofil, werden die Daten nur in diesen TABLESPACE geschrieben. In erster Linie dient er dazu, das Datenvolumen exakt bestimmen zu können. (vgl. LONEY 2001, S.571ff)
- INITIAL stellt die Größe in Megabyte für den ersten EXTENT dar, NEXT für alle weiteren. Da die Daten in ihrem Umfang und Aufbau gleich bleiben, bekommen die beiden Größen den gleichen Wert. Der EXTENT steht für die Speicherblöcke, in die die Daten geschrieben werden. Durch den Datenbankadministrator wurde eine maximale Extentgröße von 2,5GB festgelegt. Um eine möglichst ideale Füllung der EXTENTs mit Daten zu erreichen, wurde die Größe von 2.490MB für die BLK-Tabelle gewählt, da sie die eigentlichen Daten enthält und mit Abstand die größte Tabelle ist. 10MB Speicher wurden freigehalten, um Platz für Blockheader und andere Datenbanksystemkomponenten innerhalb des gleichen EXTENTs zu gewährleisten. (vgl. LONEY 2001, S.506)
- MINEXTENT legt fest, wie viel EXTENTs mindestens angelegt werden sollen. Dieser Wert wurde nicht verändert und auf dem Standardwert 1 belassen.
- MAXEXTENT legt fest, wie viele EXTENTs maximal angelegt werden dürfen. Dieser Wert wurde ebenfalls nicht verändert, da zum einen die erwartete

Datenmenge voraussichtlich diese Grenze nicht erreicht und zum anderen eine größere Anzahl von EXTENTS aus Leistungsgründen nicht ratsam ist.

- PCTINCREASE ist ein „prozentualer Faktor, um den jeder folgende EXTENT geometrisch wachsen wird“ (LONEY 2001, S.128f). Da, wie bereits erwähnt, die DGM-Daten sich nicht ändern, ist in dieser Hinsicht keine Flexibilität erforderlich und der Wert wurde entsprechend auf 0 gesetzt. Dabei ist zu beachten, dass ein PCTINCREASE-Wert gleich 0 dazu führt, dass z.B. beim Löschen von Segmenten die dazugehörigen EXTENTS zwar freigegeben und als frei markiert werden, aber nicht mit benachbarten freien EXTENTS zusammengeführt werden. Dieser Defragmentierungsprozess wird bei einem PCTINCREASE-Wert ungleich 0 automatisch in Oracle ausgeführt. Bei einem Wert gleich 0 kann eine Defragmentierung manuell erzwungen werden.
- FREELIST ist, wie der Name schon sagt, eine Liste. Diese enthält freie Speicherblöcke, in die Daten geschrieben werden können. Wenn ein Datenblock genug Datensegmente enthält und der PCTFREE-Wert erreicht wurde, wird der Block von der Liste entfernt. Veränderungen dieses Faktors zeigten keine Auswirkungen auf die Speicherplatzbelegung. Der Wert 4 entspricht der Standardeinstellung.

Durch das Ladeprofil und die bessere Komprimierung (LZ77) in der Datenbank wird das DGM-Raster in der Datenbank wesentlich kleiner als auf der Festplatte und ermöglicht einen landesweiten blattschnittfreien Zugriff, wie im vorherigen Kapitel beschrieben.

7.2 Integer- vs. Floating Point-GRID

Die Z-Werte, der gelieferten DGM-Daten, sind alle *Dezimalwerte* mit zwei Nachkommastellen. ESRI bietet bei der GRID Erstellung zwei GRID-Typen an, FLOATING POINT-GRID für dezimale Z-Werte und INTEGER-GRID für ganzzahlige Z-Werte.

Beide Arten von GRIDs wurden mit den gelieferten Daten testweise erzeugt. Für die 1m-DGM-Kacheln ergaben sich pro Kachel/Datei folgende GRID-Größen:

GRID	Intege r	Floatin g
Dateivolumen in MB	1,89	3,83
belegter Festplattenplatz in MB	1,92	3,87

(jeweils ohne Pyramid Layers)

Tabelle 3: Vergleich Inter- und Floating-Point-GRID

In diesem Abschnitt wird geprüft, ob es möglich ist, die dezimalen DGM-Daten verlustfrei als Integer-GRID abzulegen und somit ca. 50% des Speichervolumens zu sparen.

Die grundlegende Idee ist eine mathematische Erweiterung der Z-Werte durch eine Multiplikation mit 100, so dass die zwei Nachkommastellen der Z-Werte verschwinden (s. Abb. 41).

Die Programmierung macht in dieser Hinsicht nur wenig Arbeit, da sich die notwendigen Schritte nur gering von der bestehenden Konvertierung zu Floating Point-GRIDs unterscheiden.

Auch das Ergebnis fällt positiv aus, da sich das Datenvolumen, wie bereits erwähnt, halbiert. Diese Halbierung ist auf die unterschiedlichen Datentypen (integer, floating-point) zurückzuführen, mit denen die Daten in den GRIDs abgelegt werden.

In einem GRID stellt die Koordinateninformation den geringsten Teil des Speichervolumens dar. Sie wird in beiden GRID-Arten neben einigen anderen so genannter HEADERINFORMATIONEN als Fließkommazahl mit doppelter Präzision gespeichert. Die doppelte Präzision besteht aus zwei aufeinanderfolgenden 32Bit Variablen. Dies ergibt folgenden Wertebereich (vgl. www.esri.com):

- $2.2250738585072014e-308$ to $1.7976931348623158e+308$

Die durch diesen Wertebereich gegebene Genauigkeit reicht aus, um jeden beliebigen Punkt auf der Erde mit einer Genauigkeit von unter einem Millimeter darzustellen.

Die Speicherung der GRID-Werte (der eigentlichen Z-Werte also) erfolgt mit 32Bit Integer- bzw. 32Bit Floating-Point Variablen, jeweils mit Vorzeichen (signed integer und signed floating point). Diese ergeben folgende Wertebereiche:

- Floating-Point (dezimal): $1.175494351e-38$ to $3.402823466e+38$
- Integer: -2147483648 to $+2147483647$ ($-2*31$ to $(2*31 - 1)$)

Eine Speicherung mit Variablentypen ohne Vorzeichen wäre bei den zu erwartenden Höhenwerten im DGM für Baden-Württemberg denkbar, da alle Höhenwerte im positiven Intervall zwischen 70m und 1500m liegen werden. Allerdings wird nach der *ArcGIS Desktop Hilfe* – „Technical specification for raster data formats“ immer ein Vorzeichen verwendet und keine GRID-Art mit *unsigned* Variablentypen angeboten.

Die Speicherung der DGM-Daten als Integer-GRID ist somit nach einer Erweiterung der Z-Werte mit dem Faktor 100 möglich.

Betrachten wir nun die Auswirkungen für die Weiterverarbeitung.

Das Berechnen von Folgeprodukten und Bereitstellen von DGM-Ausschnitten usw. sind wiederkehrende Prozesse der Weiterverarbeitung, während das Erstellen und Speichern im Allgemeinen einmalige Prozesse darstellen. Somit stellt die Geschwindigkeit bei dem Weiterverarbeitungsprozess einen weiteren ausschlaggebenden Faktor

Original DGM Value	Extended DGM Value (x100)
120.87	12087
120.86	12086
120.86	12086
120.86	12086
120.85	12085
120.85	12085
120.85	12085
120.85	12085
120.85	12085
120.85	12085
120.85	12085
120.85	12085
120.85	12085
120.84	12084

Abb. 41: l. DGM-Datei mit Dezimalwerten, r. Erweiterung der Z-Werte mit 100



Abb. 42: 10km² Testgebiet

dar, der für oder gegen Integer-GRIDs spricht. Der folgende Versuch soll eine Aussage über die Verarbeitungsgeschwindigkeit und die Handhabung der beiden GRID-Arten machen. Zu diesem Zweck wurden zwei 10km² und 100km² große GRIDs (Integer und Floating-Point) erstellt. Auf diese beiden GRIDs wurden verschiedene Operationen angewandt:

- Erstellung von Höhenlinien mit relativ dichter Äquidistanz,
- Erstellung einer Schummerung,
- Berechnung der Exposition,
- Erstellung eines TINs.

Was die Geschwindigkeit bei der eigentlichen Erstellung der Folgeprodukte angeht, konnten keine wesentlichen Unterschiede festgestellt werden. Die Geschwindigkeit des Verarbeitungsprozess insgesamt bleibt allerdings durch die Erweiterung der Z-Werte nicht ohne Folgen. Welche Abweichungen sich ergeben und wie man versuchen kann, diese Änderung zu kompensieren, soll an der Erstellung eines TINs (als Beispiel für einen Weiterverarbeitungsprozess) gezeigt werden.

Im folgenden Beispiel wurde aus dem 10km² Testgebiet drei TINs mit unterschiedlichen Toleranzen aus Integer- und Floating-Point-GRIDs erzeugt. Die Ergebnisse mit ihren Erstellungseigenschaften sind jeweils in der folgenden Abbildung erläutert.

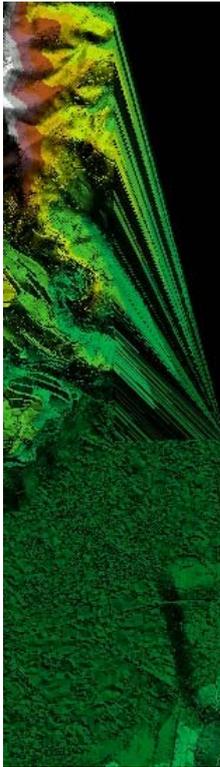
		
Fall A 2D-Ansicht	Fall B 2D-Ansicht	Fall C 2D-Ansicht
TIN aus Floating-GRID erstellt. Toleranz: 0,25m Z-Wert Differenz: 319m (Min-Max)	TIN aus Integer-GRID erstellt. Toleranz: 0,25m Z-Wert Diff.: 31.900m	TIN aus Integer-GRID erstellt. Toleranz: 25m Z-Wert Diff.: 31.900m

Tabelle 4: TIN-Erstellung mit verschiedenen Toleranzen

In Fall A wird aus einem Floating Point-GRID ein TIN mit einer Toleranz von 0,25m erstellt. Diese Situation entspricht dem „Normalfall“, d.h. ohne die Erweiterung der Z-Werte mit dem Faktor 100.

Das TIN aus einem Integer-GRID im Fall B wurde mit der gleichen Toleranz wie Fall A erstellt (0,25m). Hier werden die Auswirkungen der Erweiterung deutlich. Durch die Erweiterung mit dem Faktor 100 liegen alle Z-Werte weiter als die Toleranz von 0,25m voneinander entfernt, das bedeutet, dass jede Koordinate zum Knoten im TIN wird. Dies erklärt die Unruhe im Modell von Fall B.

In Fall C liegt wieder ein TIN vor, das aus einem Integer-GRID erstellt wurde. Hier wurde allerdings bei der Erstellung die Erweiterung um den Faktor 100 berücksichtigt, indem die Toleranz 0,25m ebenfalls um den Faktor 100 erweitert wurde (nun 25,0m). Diese Tatsache erklärt die Ähnlichkeit zu Fall A. Die Z-Werte, die nach der Erweiterung weiter voneinander entfernt liegen, fallen bei der TIN Erstellung (Fall C) wieder in die Toleranz, die um den gleichen Faktor erweitert wurde. Wie man allerdings erkennen kann, sind Fall A und C nicht gleich. Auffällig ist die Ebene im unteren Bildteil, die in Fall C Unruhen aufweist.

Dieser Unterschied liegt daran, dass bei der Erweiterung nur die Z-Werte berücksichtigt wurden, nicht aber die X- und Y-Koordinatenwerte. Es handelt sich folglich um eine Erweiterung in nur eine Dimension. Diese Tatsache führt dazu, dass das Ergebnis verfälscht wird und ein anderes Erscheinungsbild als Fall A zeigt.

Würde man alle Dimensionen, also auch die X- und Y-Koordinatenwerte erweitern, erhielt man ein GRID, das in allen Dimensionen um den Faktor 100 gewachsen ist.

Das würde bedeuten, dass man, wie eben bei der TIN Erstellung gezeigt, die Erstellungsparameter für Höhenlinien, Sichtanalysen usw. ebenfalls erweitern muss. Spätestens bei Messungen, Beschriftungen (z.B. Höhenlinien) und Entfernungsangaben müsste man bei manuellen Arbeiten im Kopf und bei automatischen Verarbeitungen durch Programmierung diese Wertedifferenz jedes Mal umrechnen. Dieser Umstand macht eine Erstellung der 1m-DGM-Daten als Integer-GRID außerordentlich aufwendig. Man muss abwägen, ob der größere Speicherbedarf durch Floating-Point-GRIDs diesen Mehraufwand kompensiert. Tatsächlich handelt es sich bei der erwarteten 1m-DGM-Datenmenge für ganz Baden-Württemberg, nach den letzten Hochrechnungen um ca. 140GB als ESRI GRIDs auf der Festplatte und etwa 50GB als Raster in der Datenbank (je ohne Pyramid Layer). Diese Datenmenge stellt zur heutigen Zeit kein großes Problem dar, schon gar nicht im Bereich der Geodaten. Dazu kommt, dass es meistens „kleinere“ Arbeiten sind, bei denen man in kurzer Zeit in einem kleinen DGM-Gebiet etwas berechnen oder Daten bereitstellen möchte. Besonders in solchen Fällen wäre der Weiterverarbeitungsprozess zu aufwändig. Aus diesem Grund ist eine Speicherung der DGM-Daten als Floating-Point-GRIDs zu empfehlen.

8 DGM-Web-Service

64

Ein normaler Web-Service stellt Daten und Funktionalitäten über das Internet zur Verfügung. Ein Geo- oder GIS-Web-Service, ist ein spezieller Fall eines Web-Services. Er stellt in erster Linie Geodaten und GIS-Funktionalitäten zur Verfügung, wobei zu den Geodaten noch Attribut- und/oder Metadaten hinzukommen können und zu den GIS-Funktionalitäten allgemeinere Funktionalitäten, wie eine Zahlungsabwicklung über das Internet, z.B. beim Kauf von Geodaten. Letzteres würde man eine E-Commerce Anwendung nennen. Ein weiteres Beispiel könnte ein Routenplaner sein, wie man ihn häufig im Internet findet.

Der plattformunabhängige Zugriff auf die DGM-Daten mittels eines Web-Services bietet sich für das ITZ (Informationstechnische Zentrum) der LfU Baden-Württemberg an, da der Verwaltungsaufwand bei Anfragen für DGM-Daten minimiert wird. Ohne selbst über ein GIS zu verfügen, können sich Kunden, mit einfachsten Mitteln (Webbrowser) über den aktuellen Stand der bisher prozessierten DGM-Daten erkundigen und haben zusätzlich die Möglichkeit, abgeleitete Ergebnisse, wie eine Schummerung, als Layer in die Browseransicht hineinzuladen. Für die Realisierung eines DGM-Web-Services wurden zwei Möglichkeiten umgesetzt:

Im ersten Fall handelt es sich um einen Web-Service mit ArcIMS. Mit ihm kann auf die DGM-Daten selbst, sowie auf abgeleitete Ergebnisse zugegriffen werden.

Die zweite Realisierung umfasst die Bereitstellung einer serverseitigen COM-DLL (Component Object Model, Dynamic Link Library; s. Glossar), mit der durch Parameterübergabe vom Client, auf der Serverseite eine Funktion aufgerufen und vom Service bearbeitet und zurückgegeben wird.

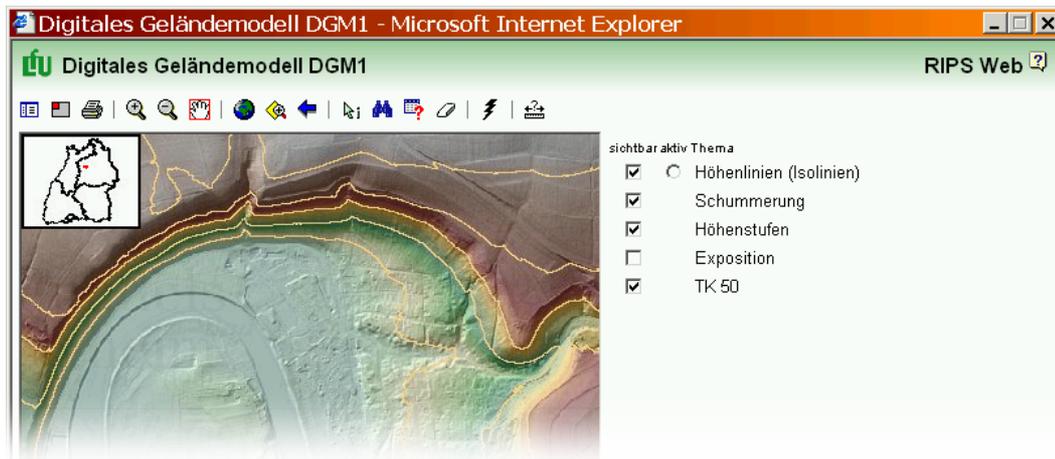


Abb. 43: DGM-Web-Service. Mit ArcIMS

8.1 DGM-Web-Service mit ArcIMS

8.1.1 Was ist ArcIMS?

ArcIMS wurde bereits eingangs dieser Arbeit im Kapitel „Werkzeuge“ vorgestellt. Zusammengefasst ist ArcIMS eine GIS Lösung, die es erlaubt, Vektor-, Rasterdaten und Werkzeuge, wie man sie von GIS-Software kennt, im Internet zu veröffentlichen. Dabei kann die Funktionalität vom einfachen Kartenserver, bis hin zum Online-GIS reichen. Die Architektur wird an dieser Stelle nicht behandelt.

Sie ist mehr als ausreichend in den ArcIMS Programmdokumentationen beschrieben und ist für das Verständnis der Initialisierung des DGM-Web-Services nicht von großer Bedeutung. Falls es nötig ist, werden eventuelle Unklarheiten im nachstehenden Text erläutert.

Viel wichtiger für die Erstellung des Service sind die Werkzeuge, mit denen der Benutzer den Service zusammenstellt.

- **ArcIMS Author** dient dazu, das benötigte Geodatenmaterial zusammenzustellen.
- Mit dem **ArcIMS Administrator** werden die eigentlichen Web-Dienste initialisiert und verwaltet.
- Der **ArcIMS Designer** erlaubt es, die Darstellung der Webseite bis zu einem gewissen Grad zu gestalten.

8.1.2 Voraussetzung für den DGM-Web-Service

Alle in den Web-Service eingebundenen Daten werden mit der SDE in die Datenbank geladen. Dadurch ergibt sich der Vorteil, dass nach einer Aktualisierung der Datenbestände kein neuer Dienst angelegt werden muss, sondern stets auf die aktuellen Daten zugegriffen werden kann.

Auch die abgeleiteten Ergebnisse, die im Prototyp durch Schummerung, Exposition und Höhenlinien für ein 100km² großes Gebiet dargestellt werden (s. Abb. 43), liegen in Zukunft, für das ganze Land berechnet, als SDE-Raster vor und können als Layer im Browserfenster zugeschaltet werden.

Das Anzeigen der Daten im Browser durch ArcIMS erfolgt jeweils für den entsprechenden Ausschnitt mit der benötigten Auflösung, diese Technik wird als "Maps on demand" Konzept bezeichnet. Dieses erlaubt dem Benutzer zu spezifischen Karten mit beliebigem Maßstab zu navigieren und diese Karten mit allen Inhalten und allen dazugehörigen ArcMap-Layouts zu drucken. Dieser Prozess findet dynamisch statt.

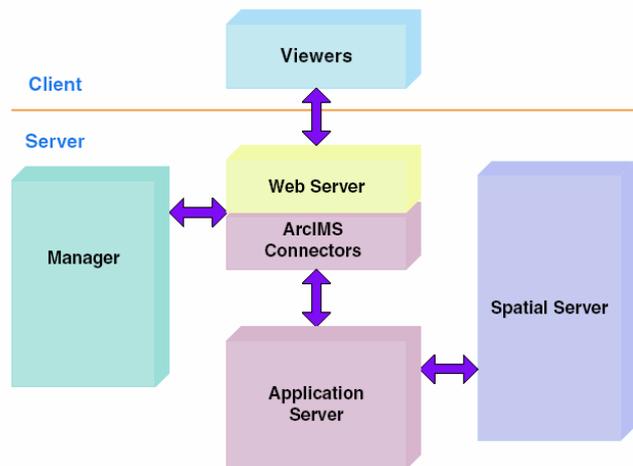


Abb. 44: Architektur ArcIMS

8.1.3 Zusammenstellen des Geodatenmaterials

Im Kapitel 8.1.1 wurden drei Werkzeuge genannt, mit denen man den Web-Service bis zu seiner Veröffentlichung generiert und verwaltet. Damit ist besonders der ArcIMS Author gemeint, in dem das Geodatenmaterial für den Web-Service zusammengestellt wird.

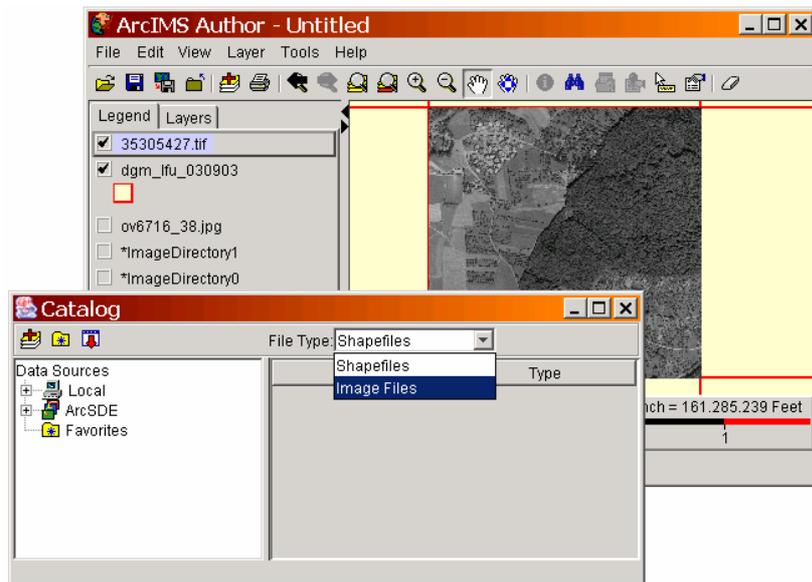


Abb. 45: ArcIMS Author mit dem Catalog Datenauswahlfenster

Die Datensicht

ähnelt dabei der bereits aus ArcMap bekannten Darstellung (s. Abb. 47). ArcIMS Author kann im Grunde alle von ArcGIS unterstützten Datenformate verarbeiten. Das fertige Layout wird in einem MAP CONFIGURATION FILE (*.AXL) gespeichert. Diese Datei besteht aus XML-Code (s. Glossar) mit erweiterten Tags (s. Glossar) für das Kartenlayout.

Bei einigen Datenformaten, wie z.B. SDE-Rastern und SDE-Shapes, die ausschließlich für den DGM-Web-Service verwendet wurden, ergab sich das Problem, dass der Dialog **Catalog** (s. Abb. 45), mit dem die Daten für die Anzeige in ArcIMS Author ausgewählt werden, die eben genannten Formate nicht darstellen konnte.

Einschränkungen gibt es auch bei den dateibasierten Rasterformaten, die im Gegensatz zu ArcMap Worldfiles (s. Glossar) besitzen müssen, um überhaupt zur Auswahl für ein Anzeigen in ArcIMS Author zur Verfügung zu stehen.

Da die resultierende AXL-Datei im XML-Format geschrieben ist, kann man sie mit einem beliebigen ASCII-Editor öffnen und bearbeiten. Jemand, der HTML Kenntnisse besitzt, wird sich schnell zurechtfinden.

Auch für den Webdienst wurde aus den eben genannten Problemen mit diversen Dateiformaten, die AXL-Datei manuell editiert. Der XML-Tag SDEWORKSPACE erlaubt das Einbinden von, in Datenbanken gespeicherten Daten. Auch bei anderen Objekten wurde versucht, sie einzubinden. Imagekataloge, von denen ArcIMS Author vier verschiedene Arten kennt und die ein Äquivalent zum Rasterkatalog in ArcMap sind, bereiteten trotz strenger Befolgung der ArcIMS Dokumentation, die zu jeder Möglichkeit ein Beispiel enthält, Schwierigkeiten. Nach vielen Fehlversuchen, gelang es, sowohl Imagekataloge als auch die für den DGM-Service wichtigen SDE-Raster in einen Dienst einzubinden.

Die fehlende Möglichkeit, die AXL-Datei auf Syntaxfehler, geschweige denn auf logische Fehler zu prüfen, macht die manuelle Bearbeitung der Kartenkonfiguration sehr zeitaufwendig. Fehler machen sich erst bemerkbar, wenn man versucht mit ArcIMS Administrator den Web-Service zu initialisieren. Die dabei entstehenden Fehlermeldungen geben allerdings relativ wenig Aufschluss über die Fehlerursache.

Der Benutzer bekommt in den meisten Fällen lediglich die Meldung „Konnte Service nicht anlegen“ oder „Konnte Service nicht anlegen, bitte Pfade überprüfen“.

```
<?xml version="1.0" encoding="UTF-8"?>
<ARCXML version="1.1">
  <CONFIG>
    <ENVIRONMENT>
      <LOCALE country="DE" language="de" variant="" />
      <UIFONT color="0,0,0" name="dialog" size="12" style="regular" />
      <SCREEN dpi="96" />
    </ENVIRONMENT>
    <MAP>
      <PROPERTIES>
        <ENVELOPE minx="3451814,5" miny="5495846,75"
          maxx="3476007,25" maxy="5518249,25"
          name="Initial_Extent" />
        <MAPUNITS units="decimal_degrees" />
      </PROPERTIES>
      <WORKSPACES>
        <SDEWORKSPACE name="sde" server="itzgs2" instance="port:5151" database=""
          user="test" encrypted="true" password="test" />
      </WORKSPACES>
      <LAYER type="image" name="DGM1m" visible="true" id="0">
        <DATASET name="Test.TEST.DGM1m_10x10.IMAGE" type="image"
          workspace="sde" />
      </LAYER>
    </MAP>
  </CONFIG>
</ARCXML>
```

67

Abb. 46: Map Configuration File. AXL-Datei

Die Überschrift zu diesem Kapitel könnte auch lauten „Klicken oder Programmieren“. Denn in der Regel ist die Eigeninitiative, sprich das manuelle Editieren, die bessere Lösung, hat man doch dabei direkten Einfluss auf das resultierende Ergebnis. Allerdings bietet ArcIMS eine Erweiterung, mit der es möglich ist ArcMAP-Projektdateien (*.MXD), anstelle einer AXL-Datei als Web-Service anzulegen. Diese haben den Vorteil, dass man den gesamten Komfort nutzen kann, den man vom ArcMap Darstellungen in der Datenansicht gewohnt ist. Ohne Probleme lassen sich alle unterstützten Datenformate inklusive in der Datenbank abgelegter Daten einbinden – ohne Editierarbeit. Es sei erwähnt, dass es sich bei MXD-Dateien im Gegensatz zu AXL-Dateien um Binärdateien handelt, und diese nicht mit einem Editor betrachtet oder gar manipuliert werden können. Dennoch hat die Praxis bei dieser Anwendung gezeigt, dass man in einem Bruchteil der Zeit zum gleichen Ergebnis kommt. Deshalb fiel die Wahl auf die Anwendung von MXD-Dateien.

Die ArcMap Server-Erweiterung für MXD-Dateien, ist eine Erweiterung zum ArcIMS Spatial Server. Sie erlaubt dem Anwender, eine Datei vom Format *.MXD oder *.APM (ArcMap Publisher) in ArcIMS zu benutzen. Der Service ist ein *ArcMap Image Service*.

Die besonderen Eigenschaften des ArcMap Servers für ArcIMS 4 sind:

- Möglichst gleiche Funktionalität beim Kartenzugriff, bei der Navigation und der Abfrage, wie mit dem *ArcIMS Image Service*
- Unterstützung des "Maps on demand" Konzepts
- Universeller Zugriff zu allen ArcGIS Karten
- Umfangreiche Layoutmöglichkeiten zur Erstellung qualitativ hochwertiger Karten für das ArcIMS Publikum

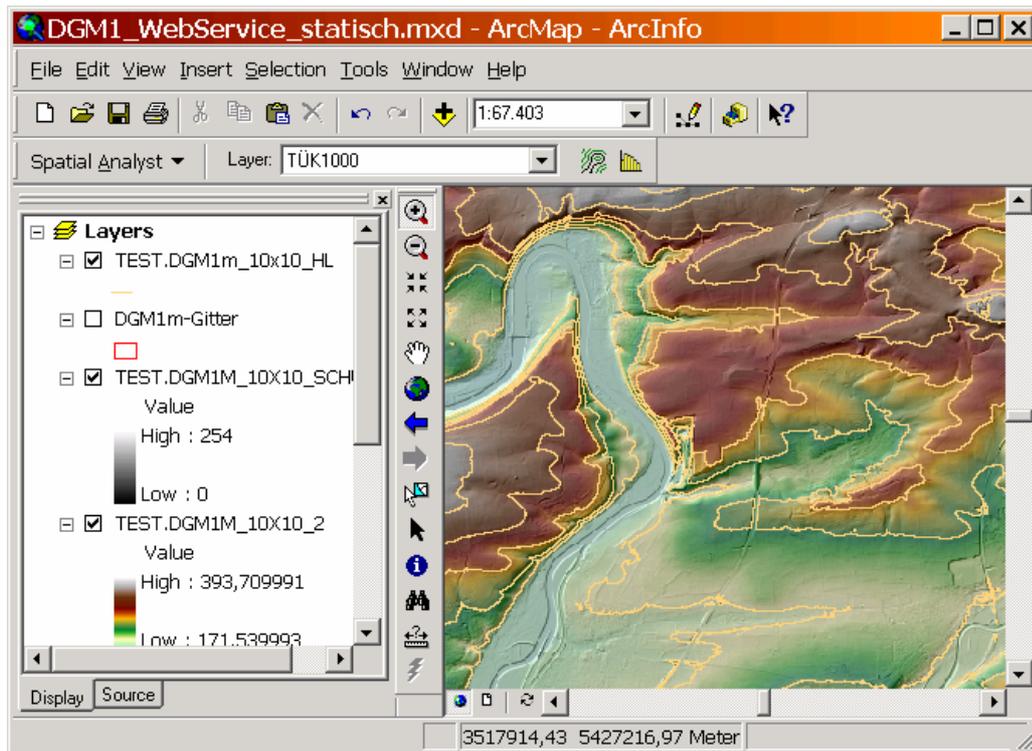


Abb. 47: ArcMap-Projekt; MXD-Datei

8.1.4 Funktionalität und Design



Abb. 48: In den Web-Service eingebundene Werkzeuge

Nachdem der Web-Service im ArcIMS Administrator initialisiert wurde, gelangt man im nächsten Schritt zum ArcIMS Designer.

Der ArcIMS Designer ist für das Erscheinungsbild und die enthaltenen Funktionalitäten des Web-Services verantwortlich. Mit den Assistenten des ArcIMS Designer können, ohne jede Programmierkenntnisse, Webklienten z.B. als ArcIMS Standard HTML Viewer erzeugt werden. Das Aussehen und die Funktionalität einer ArcIMS Website sind damit zunächst festgelegt, können aber nachträglich an spezielle Erfordernisse angepasst werden.

Anpassungen außerhalb des ArcIMS Designers erfolgen für das Design mit HTML beziehungsweise XML, die Anpassung von Funktionalitäten können mit JavaScript und Java realisiert werden.

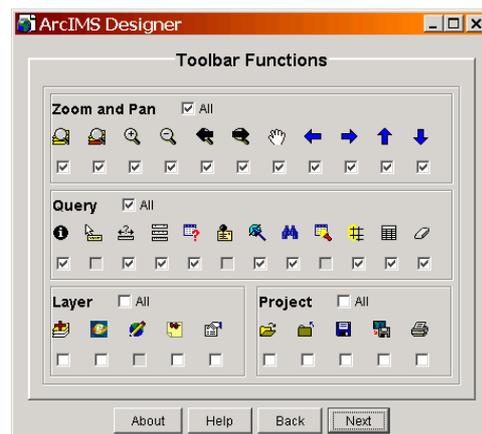


Abb. 49: Funktionen ohne Programmierkenntnisse

8.2 Web-Service durch COM-Komponente

Während der Web-Service mit ArcIMS Zugriff auf die DGM-Daten selbst sowie auf abgeleitete, statische Ergebnisse erlaubt, geht es bei dieser Realisierung darum, Funktionalitäten zur Verfügung zu stellen, die Folgeprodukte dynamisch generieren. Zwei Funktionalitäten wurden in Form einer COM-DLL für einen XML-Web-Service realisiert:

- Mit der Übergabe einer Boundingbox durch den Client, wird für den betreffenden Bereich ein Folgeprodukt, z.B. eine Schummerung erstellt.
- Durch die Übergabe einer Koordinate wird der zugehörige Höhenwert zurückgegeben.

8.2.1 Architektur

Die Architektur des Services ist relativ simpel. Ein dotNET (auch .NET oder .dotNET) XML-Web-Service greift über eine COM-Komponente (DLL) auf ArcGIS Funktionalitäten zu. Dabei ermöglicht die DLL den Zugriff auf die ArcObjects-Bibliothek und somit auf die darin enthaltenen Funktionalitäten.

Was bedeutet das im Einzelnen?

Ein Teilbereich der Microsoft dotNET-Technologie sind XML-Web-Services. Mit ihnen ist es möglich, verteilte Systeme über das Internet zusammenarbeiten zu lassen, um lokal nicht lösbare Aufgaben zu erfüllen.

DotNET selbst, wird von BEER, 2002 S.2, beschrieben, als „ein aufeinander abgestimmtes Ensemble aus Betriebssystemkomponenten, Bibliotheken, Werkzeugen, Web-Services und Servern, das schwer in einem Satz zu definieren ist. Alles zusammen hat zum Ziel, die Programmierung von Windows und Web-Anwendungen bequemer und einfacher zu machen.“

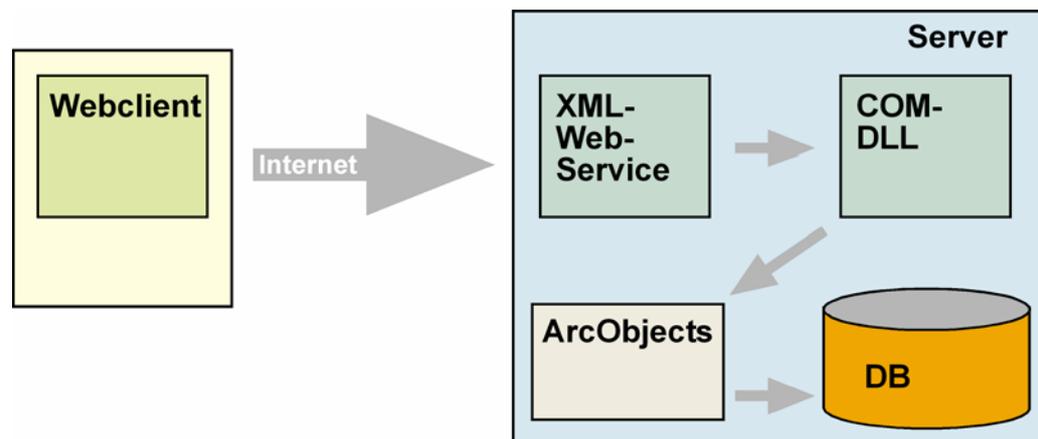


Abb. 50: Architektur des XML-Webservice

Der XML-Web-Service greift auf die DLL zu, indem er eine Funktion in der DLL aufruft und an diese die benötigten Parameter übergibt. Die Funktion in der DLL berechnet die Schummerung, oder gibt einen Höhenwert zurück, in dem sie sich weiterer ArcObjects Funktionen bedient.

8.2.2 Realisierung

Die Aufgabe ist die Bereitstellung einer DLL, die die entsprechende Funktionalität für den Webservice zu Verfügung stellt. Zum einen ist das die Erzeugung einer Schummerung für ein definierbares Gebiet, zum anderen die Rückgabe eines Höhenwerts zu einer definierbaren Koordinate innerhalb von Baden-Württemberg.

Die DLL selbst ist in Visual Basic (VB) programmiert. Enthalten ist eine Funktion, die bereits im RASTER CLIPPER zur Erstellung der Schummerung benutzt wurde. Der Code wurde aus der VBA-Umgebung (Visual Basic for Application) kopiert und an die VB Umgebung mit einigen zusätzlichen Änderungen angepasst.

Die Funktion für den Höhenservice, mit der man zu einer definierbaren Koordinate, den entsprechenden Höhenwert zurückbekommt, wurde extra in VB programmiert. Beide Funktionen wurden in einer DLL kompiliert.

Der Zugriff auf die benötigten Informationen, sprich die DGM-Rasterdaten, erfolgt in beiden Fällen durch die Herstellung einer Verbindung zu dem landesweiten DGM-Raster in der Datenbank. Von ihm werden die Höhenwerte abgefragt oder ein Teilraster exportiert, von dem die angeforderte Schummerung erstellt wird.

8.2.3 Probleme

Die umgesetzte DLL enthält zwei Funktionen.

Die erste Funktion beinhaltet die Generierung einer Schummerung. Durch die Übergabe eines definierten Gebiets an die Funktion, erfolgt die Erstellung der Schummerung. Das Ergebnis der Schummerungsfunktion ist ein TIFF-Bild. Dieses TIFF das mit ArcGIS Funktionen aus der Datenbank exportiert wurde, kann zur Zeit nur in ArcGIS richtig dargestellt werden. Eine Darstellung in einem externen Bildbetrachtungsprogramm hat eine falsche Farbdarstellung zur Folge (s. Abb. 51). Die Ursachen dafür konnten noch nicht beseitigt werden. Aus diesem Grund ist es momentan auch nicht möglich, das Bild für eine Browserdarstellung weiterzuverarbeiten.

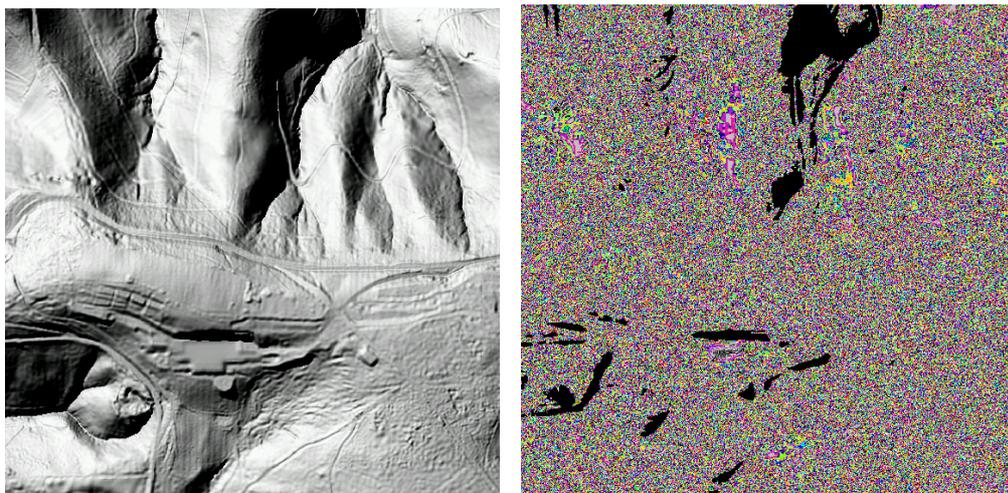


Abb. 51: exportiertes ArcGIS TIFF: l. in ArcGIS, r. in ArcView3.2 dargestellt.

Was den Höhenservice betrifft, konnte dieser bisher nur mit einem Zugriff auf ein dateibasiertes Raster realisiert werden. Beim Zugriff auf ein Datenbankraster

wurde das Testprogramm stets mit einem Fehler beendet. Es konnte daher in der Kürze der Zeit keine Lösung gefunden werden.

Letztendlich konnte mit dem dateibasierten Höhenservice jedoch exemplarisch der Zugriff auf Funktionen mit den 1m-DGM-Daten gezeigt werden. Mit Sicherheit nur eine Frage der Zeit, bis man auch das TIFF-Bild der Schummerung weiterverarbeiten kann.

Die Arbeit hat gezeigt, dass der Wunsch der Landeanstalt für Umweltschutz Baden-Württemberg nach einer flexiblen und blattschnittfreien DGM-Datenverwaltung erfüllt werden kann. Auch der plattformunabhängige Web-Service lässt sich realisieren.

Die dabei notwendigen Werkzeuge, mit denen sich die DGM-Daten von der Lieferung bis zur Bereitstellung und Auslieferung importieren, verwalten, und exportieren lassen, konnten prototypisch realisiert werden. Sie besitzen eine ergonomische Bedienoberfläche und decken alle gängigen Rasteroperationen innerhalb von ArcGIS ab:

- Datenimport. Konvertierung von 1m und 5m ASCII-DGM-Dateien nach ESRI GRID.
- Mosaiken. Zusammenfassen von GRIDs.
- Datenbankimport. Upload von dateibasierten GRIDs in die Datenbank (SDE-Raster).
- Rasterkatalog erstellen.
- Kopieren von Rastern aus einem Rasterkatalog anhand einer Boundingbox.
- Konvertierung von GRIDs nach ASCII-XYZ-Koordinaten.
- Kopieren und Exportieren von Teilrastern aus einem SDE-Raster in verschiedene Rasterformate anhand einer Boundingbox.
- Erstellen von Folgeprodukten in verschiedenen Rasterformaten (GRID, TIFF, ERDAS Imagine und JPEG):
 - Höhenlinien
 - Exposition
 - Schummerung
 - Hangneigung
- Für Testzwecke die Erzeugung von Pseudo-GRIDs mit diversen Parametern.

Auch in Richtung TIN-Datenhaltung konnten, soweit dies möglich war, Erkenntnisse und Vergleiche zur Rasterdatenhaltung gewonnen werden. Wie in der Arbeit beschrieben wurde, können TIN-Datensätze bisher nur als Datei abgespeichert werden. Diese Informationen werden dann mit der neuen Version von ArcGIS (Version 9) nutzbar, die man mit Spannung erwarten kann.

In dieser Version wird es möglich sein, TINs mit Hilfe der ArcSDE in der Datenbank abzulegen. Diese TINs werden über die Technik „TIN on the fly“ verfügen, einer automatischen maßstabsabhängigen Retriangulation, vergleichbar mit den Pyramid Layern bei ESRI GRIDs.

Praktisch zu prüfen wäre dann, ob eine DGM-Datenhaltung als TIN gegenüber einer Rasterdatenhaltung von Vorteil wäre.

Etwas Handlungsbedarf besteht noch in der Ausarbeitung einer Normierung für die Weitergabe der DGM-Daten. Diese Ausarbeitung war nicht Bestandteil der Diplomarbeit, allerdings kamen immer wieder Anfragen für Datenauslieferungen, die es nebenher zu bedienen galt.

Dabei wurde festgestellt, dass eine Normierung der Anfrageart für DGM-Daten sinnvoll wäre, denn bisher reichten die Anfragearten für DGM-Daten von

Ortsbeschreibungen und Koordinaten über konkrete Kachelnamen bis hin zu Polygonshapes.

Ein Vorschlag in dieser Richtung wäre es, den DGM-Webservice mit einzubeziehen und über diesen dem Kunden Werkzeuge für eine Gebietsauswahl zur Verfügung zu stellen. Somit könnte man die Anfragearten einschränken, bzw. besser steuern.

Insgesamt handelt es sich um eine interessante und vor allem aktuelle Thematik. Die Auseinandersetzung mit den Problem großer Datenmengen, wie sie die DGM-Daten darstellen, und die Visual Basic Programmierung waren neue Themengebiete, ebenso wie die Einarbeitung in die Techniken und Anwendungen von Web-Services. Diese Themengebiete, gekoppelt mit der neusten Softwareumgebung, wie sie in der LfU vorhanden ist, stellen eine moderne, innovative Herausforderung dar.

Mit Abschluss dieser Arbeit stehen der LfU mit wenigen Einschränkungen gute Werkzeuge für eine flexible, landesweite und blattschnittfreie DGM-Datenverwaltung, sowie ein Web-Service zu Verfügung.

Glossar & Abkürzungen

API	(Application Program Interface) API, Schnittstelle / Interface zur Anwendungsprogrammierung. Ein Teil im Betriebssystem, das als einheitliches Benutzerinterface für Anwendungen dient. (www.babylon.de)
ArcGIS	S. GIS.
ArcIMS	Image Map Server der Firma ESRI. (www.esri.com)
ArcObjects	ArcObjects ist die Basis von ArcGIS. Es ist eine Sammlung von Softwarekomponenten mit GIS Funktionalität und programmierbaren Schnittstellen. Die ArcObjects Technologie beruht auf dem COM Protokoll.
ArcSDE	S. SDE.
Business Tabelle	Eine Business Tabelle wird jede Tabelle in Oracle genannt, die durch einen ArcSDE Client, das <i>SDETable</i> Konsolenprogramm oder durch die Funktion <i>ArcSDE S API SE_table_create</i> angelegt wurde. (www.esri.de)
COM	Component Object Model. Microsofts Standard zur Erstellung von Software Komponenten. Im Gegensatz zu anderen herstellerunabhängigen Standards wie CORBA ist COM ausschließlich auf Win3.x/95/98/NT (für neuere Betriebssysteme gilt die Weiterentwicklung COM+) ausgerichtet. Ziel des Standards ist es, selbstständige Programmteile zu erzeugen. (www.babylon.de)
DGM	Digitales Geländemodell ist dem DHM untergeordnet und bildet ein Gelände ohne Bebauung und Vegetation ab. Strukturelemente (Bruchkanten, Mulden usw.) sind enthalten.
DHM	Digitales Höhenmodell. Der höhenmäßige Verlauf einer beliebigen Oberfläche wird mathematisch abgebildet.
DLL	Dynamic Link Library. Eine Datei die Routinen der Betriebssysteme Windows und OS/2 enthält.
DOM	Digitales Oberflächenmodell ist dem DHM untergeordnet und bildet ein Gelände mit Bebauung, Vegetation und Strukturelementen (Bruchkanten, Mulden usw.) ab.
Extent	Ein Extent ist eine Stückchen Platz in einer Datendatei, der in Bytes, Kilobytes, oder Megabytes gemessen wird. (vgl. ABBEY 2000, S. 531)
GIS	Geographisches Informationssystem.
GRID	Regelmäßiges Gitter. In dieser Arbeit meist als ESRI GRID bezeichnet, stellt es ein Raster-Bildformat dar.
Integer (-zahl)	Ein Integerwert ist im Computergenre ein Datentyp, der ganze Zahlen abspeichern kann. Aus mathematischer Sicht alle Zahlen, die Elemente der Natürlichen Zahlen sind.
OGC	Open Gis Konsortium. Schlägt Standards im GIS Bereich vor.
OLE	(Object Linking and Embedding) frühe Umsetzung des COM- Standards zur Erstellung von unabhängigen Programm-Komponenten. Die Weiterentwicklung von OLE wird von Microsoft ActiveX genannt. (www.babylon.de)

PDF	Portable Dokument Format. Dokumentenformat der Firma Adobe.
RDBMS	Relationales Datenbank Management System.
SDE	Spatial Data Engine. Ein Produkt aus der ESRI Palette, das auf ein RDBMS aufgesetzt wird und als Schnittstelle zwischen ESRI Produkten und einem RDBMS dient. (www.esri.com)
Spatial Analyst	Erweiterung für ArcGIS im Bereich der Rasterfunktionalität. Stellt Funktionen zur räumlichen Modellierung und Auswertung zur Verfügung.
Tablespace	Container für Datenbanksegmente (Datenbankinterne Organisationseinheit). Man kann Tablespaces mit physischen Dateien gleichsetzen. (www.oracle.de)
Tag	Der Inhalt von HTML-Dateien steht in HTML-Elementen. HTML-Elemente werden durch sogenannte Tags markiert. Fast alle HTML-Elemente werden durch ein einleitendes und ein abschließendes Tag markiert. Der Inhalt dazwischen ist der "Gültigkeitsbereich" des entsprechenden Elements. Tags werden in spitzen Klammern notiert. (www.selfhtml.de)
TIN	Triangulated Irregular Network. Speicherform für dreidimensionale Oberflächen, bei der eine nicht attributierte 3D-Punkmenge über ein Dreiecksnetz vermascht ist. BOLLMANN 2002
UDK	Umweltdatenkatalog der Umweltbehörde.
Worldfile	Ein Worldfile ist eine Zusatzdatei zu einem Rasterbild, das Daten für die Georeferenzierung enthält. (www.uni-tuebingen.de)
XML	eXtensible Markup Language. XML ist eine "vereinfachte" Version der Standard Generalized Markup Language (SGML). W3C-Standard. XML soll es den Web-Site-Programmierern erleichtern, SGML-Anwendungen zu schreiben und dabei eigene Dokumententypen festzulegen. XML ist eine Methode, um strukturierte Daten - z.B. Kalkulationstabellen, Konfigurationsparameter usw. - in einer Textdatei darzustellen. (www.glossar.de)
XY-Data	In ArcView 3.x EventTheme oder Ereignisthema genannt. Eine Tabelle wird als Punktthema in ein ArcGIS Projekt geladen, wobei zwei Spalten Lagekoordinaten enthalten müssen und eine beliebig wählbare dritte Spalte ein Thema (z.B. Höheninformationen) zur Visualisierung.

Tabellen- und Abbildungsverzeichnis

Tab.1: Größenvergleich von AOI- und Boundingbox-Mosaiken	26
Tab.2: Speicherplatzentwicklung in der Datenbank.....	26
Tab.3: Vergleich Inter und Floating Point GRID	26
Tab.4: TIN Erstellung mit verschiedenen Toleranzen	26
Abb. 1: Mehrfachreflexion.....	16
Abb. 2: Schema des Laserscannings. SCHLEYER 2001	16
Abb. 3: Kacheleinteilung mit Gauß-Krüger. Quelle: LVA.....	18
Abb. 4: ArcGIS Produktfamilie. Quelle: www.esri.com	20
Abb. 5: Spatial Analyst Rasteroperation, Pufferung. Quelle: www.esri.com	20
Abb. 6: ArcScene, 3D Darstellung. Quelle: www.esri.com	21
Abb. 7: ArcSDE Position in der Produktfamilie. Quelle: www.esri.com	21
Abb. 8: 1m ASCII DGM-Datei.....	23
Abb. 9: GRID mit Pixelstruktur.....	24
Abb. 10: TIN, 2D-Drahtmodell	24
Abb. 11: Pyramid Layer	26
Abb. 12: ArcToolbox, ASCII to GRID	26
Abb. 13: Rastertabellen und ihre Beziehungen	26
Abb. 14: vertikale Toleranz bei der TIN Erstellung	26
Abb. 15: TIN Erstellung ohne Toleranz.....	26
Abb. 16: TIN Erstellung mit Toleranz.....	26
Abb. 17: TIN Erstellung aus Punktwolken oder Raster, mit und ohne Toleranz	26
Abb. 18: Werkzengleiste DGM-GRID-RASTER in ArcMap.....	26
Abb. 19: Koordinatenreihenfolge Ist/Soll	26
Abb. 20: DGM Tools, ASCII-DGM2GRID	26
Abb. 21: DGM Tools, Mosaic/SDE-Loader	26
Abb. 22: Spatial Analyst, Raster Calculator	26
Abb. 23: Überschreiben von korrekten Werten bei Mosaiken	26
Abb. 24: Rasterkatalog Werkzeug	26
Abb. 25: Bsp. Für einen Rasterkatalog in Excel	26
Abb. 26: Rasterkatalog: l. Platzhalter, r. Raster	26
Abb. 27: Kacheln aus Rasterkatalog kopieren	26
Abb. 28: Werkzeug, Raster Clipper.....	26
Abb. 29: Werkzeug, SDE Dummy Builder.....	26
Abb. 30: Raster2ASCII Werkzeug	26
Abb. 31: Dateizusammensetzung eines ESRI GRIDs	26
Abb. 32: NODATA Anordnungen	26
Abb. 33: GEG Wüstenbach, Auszug eines ArcMap Layouts	26
Abb. 34: Dichtebild, Punktdichte/4m ²	26
Abb. 35: die einzelnen Schritte des DB Versuchs.....	26
Abb. 36: DGM & Dummy DGM-Kacheln	26
Abb. 37: l. Dummy Raster, r. reguläres DGM-Raster	26
Abb. 38: Ein Ladeprofil aus der Tabelle DBTUNE	26
Abb. 39: Ladeprofil in Raster to Geodatabase	26
Abb. 40: DGM-Ladeprofil (DGM08)	26
Abb. 41: l. DGM-Datei mit Dezimalwerten, r. Erweiterung der Z-Werte mit 100	26

Abb. 42: 10km ² Testgebiet.....	26
Abb. 43: DGM-Web-Service. Mit ArcIMS	26
Abb. 44: Architektur ArcIMS	26
Abb. 45: ArcIMS Author mit dem Catalog Datenauswahlfenster	26
Abb. 46: Map Configuration File. AXL-Datei.....	26
Abb. 47: ArcMap-Projekt; MXD-Datei.....	26
Abb. 48: In den Web-Service eingebundene Werkzeuge	26
Abb. 49: Funktionen ohne Programmierkenntnisse	26
Abb. 50: Architektur des XML-Webservice.....	26
Abb. 51: exportiertes ArcGIS TIFF: l. in ArcGIS, r. in ArcView3.2 dargestellt.	26

Quellen- und Literaturverzeichnis

Printmedien

- ABBHEY, Michael; Corey, Michael J. u.a. 2000. Oracle für Einsteiger. München, Wien: Carl Hanser.
- ArcSDE CONFIGURATION AND TUNING GUIDE FOR ORACLE®. 1999. Autor unbekannt. Redlands: ESRI.
- BARTELME, Norbert. 1994. Geoinformation, Modelle, Strukturen, Funktionen. Berlin, Heidelberg: Springer.
- BARTELME, Norbert. 2000. Geoinformation, Modelle, Strukturen, Funktionen. Berlin, Heidelberg: Springer.
- BEER, Wolfgang; Birngruber, Dietrich u.a. 2002. Die .NET-Technologie. Grundlagen und Anwendungsprogrammierung. Heidelberg: Dpunkt.
- FINDEISEN, Dirk. 1990. Datenstruktur und Abfragesprache für raumbezogene Informationen. Bonn: Kirschbaum.
- JÄGER, R. und Kälber, S. 2000. Konzepte und Softwareentwicklungen für aktuelle Aufgabenstellungen für GPS und Landesvermessung. Beitrag in DVW Mitteilungen 10/2000, Landesverein Baden-Württemberg.
- LONEY, Kevin und Theriault, Marlene. 2001. Oracle8i DBA-Handbuch. München, Wien: Carl Hanser.
- SCHLEYER, Andreas. 2001. Das Laserscan-DGM von Baden Württemberg. Erschienen in: Photogrammetric Week '01. 2001. Stuttgart: Wichmann.
- SIGLE, Manfred. 2002. Moderne Verfahren zur Erstellung und Anwendung digitaler Geländemodelle. Vortragsunterlagen, veröffentlicht im Rahmen des DGM-Anwenderseminar des Landesvermessungsamts Baden-Württemberg, Karlsruhe 24.Sptember 2002.
- WEIDENBACH, Markus. 1999. Geographische Informationssysteme und Neue Digitale Medien in der Landschaftsplanung. Berlin: Logos.

Internet

- <http://support.esri.com>: Dokumentationen, Forum uvam., zu Produkten des GIS-Software Herstellers ESRI.
- www.esri.com: GIS-Software Herstellers ESRI.
- www.babylon.com: GIS Begriffe (Wörterbuch/Lexikon). 2001. Autor: Eric, J. UniGIS Salzburg.
- www.lv-bw.de: Landesvermessungsamt Baden-Württemberg. Vorträge und Veröffentlichungen zum Thema: Digitale Geländemodelle und ihre praktische Anwendung. 24.09.2002.
- www.microsoft.de: Seite der Deutschen Niederlassung des Microsoft Software Konzerns.
- www.TopScan.de: Verantwortliche Firma für die Erfassung der Laserscandaten für das digitale Geländemodell von Baden-Württemberg.

Verzeichnis Anhänge

Anhang 1: Hinweise zur Handhabung der DGM-Daten	26
Anhang 2: DGM-Übersichtskarte	26
Anhang 3: TIN Erstellung	26
Anhang 4: Größen- und Formelsammlung für DGM-Operationen	26
Anhang 5: Anwendungsfall Gewässereinzugsgebiet	26
Anhang 6: Anleitung für die Erstellung des DGM-Übersichtsgitters (Shape)	26
Anhang 7: Anleitung, ASCII to GRID mit der ArcToolbox	26
Anhang 8: Anleitung, ASCII to XY [to Shape] to GRID in ArcMap	26
Anhang 9: SQL Abfrage des Speichervolumens eines Rasters in Oracle	26
Anhang 10: Der LZ77 Algorithmus	26
Anhang 11: Raster To Geodatabase, Anleitung, Erfahrungen und Handhabung	26
Anhang 12: Vergleichstest der Datenvolumina	26
Anhang 13: Programmierte Werkzeuge, graphische Übersicht	26

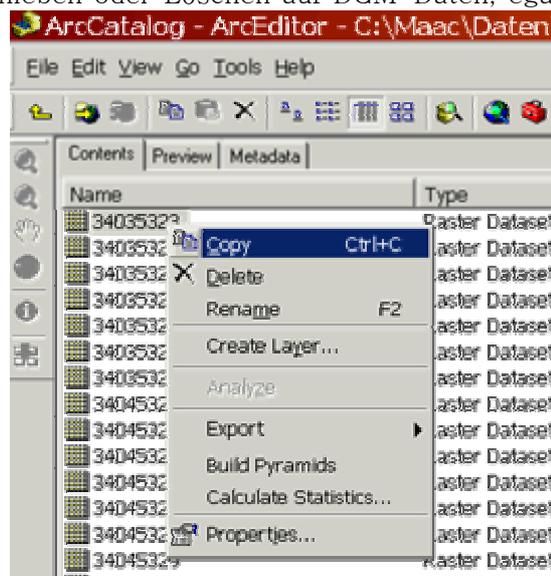
Anhang 1: Hinweise zur Handhabung der DGM-Daten

Beim Umgang mit den Daten wurden einige hilfreiche Erkenntnisse über die Handhabung der Daten in ihren verschiedenen Formaten (ASCII, GRID, ZIP-komprimiert usw.) gewonnen. Diese Hinweise sind teilweise im Text oben enthalten und sollen an dieser Stelle gesammelt zusammengefasst werden.

Kopieren, Verschieben und Löschen

Die Dateioperation „verschieben“ ist in der Version ArcGIS 8.3 nicht als Menüoption realisiert. Im Kontextmenü des ArcCatalog gibt es die Möglichkeit, die entsprechenden Daten erst zum Zielverzeichnis zu kopieren und anschließend die gleichen Daten vom Quellverzeichnis zu löschen. Der Nachteil dieser Krücke sind die zwei Schritte, die der Anwender durchführen muss, gleichzeitig bedeuten mehr Schritte auch eine größere Anzahl von Fehlerquellen. Ärgerlicher sind noch mehr die Wartezeiten, die beim Hantieren mit Daten in großen Verzeichnissen entstehen. Verschieben steht nur als „Drag and Drop“ Funktion ArcCatalog zur Verfügung.

Wenn Operationen wie Kopieren, Verschieben oder Löschen auf DGM-Daten, egal in welchem Format, angewandt werden, ist es unbedingt erforderlich, dass dies mit ArcCatalog geschieht. Besonders dann, wenn einzelne GRIDs verschoben werden, denn alle GRIDs eines Verzeichnisses teilen sich ein Info-Verzeichnis, das alle Informationen über Attributtabelle für ein Coverage oder Raster im selben Verzeichnis beinhaltet. Werden einzelne Dateien über das Betriebssystem (z.B. Windows Explorer) kopiert, verschoben oder gelöscht, werden diese Aktionen nicht im Info-Verzeichnis abgeglichen und führen zu einer Fragmentierung und letztendlich zu Fehlermeldungen.



Dies gilt nicht, wenn man ein *ganzes Verzeichnis* verschiebt. Möchte man ein ganzes Verzeichnis mit Rasterdaten an einen anderen Ort verschieben, kann man dies auch über das Betriebssystem erledigen (z.B. Windows Explorer), da gerade bei Verzeichnissen mit großen Datenmengen der Vorteil der Statusanzeige und der Geschwindigkeit beim Durchführen genutzt werden kann.

Wenn ESRI-GRIDs von einer Sicherungs-CD oder DVD auf eine Festplatte kopiert werden sollen, bedeutet dies, dass nur ganze CDs oder DVDs auf eine Festplatte kopiert werden können.

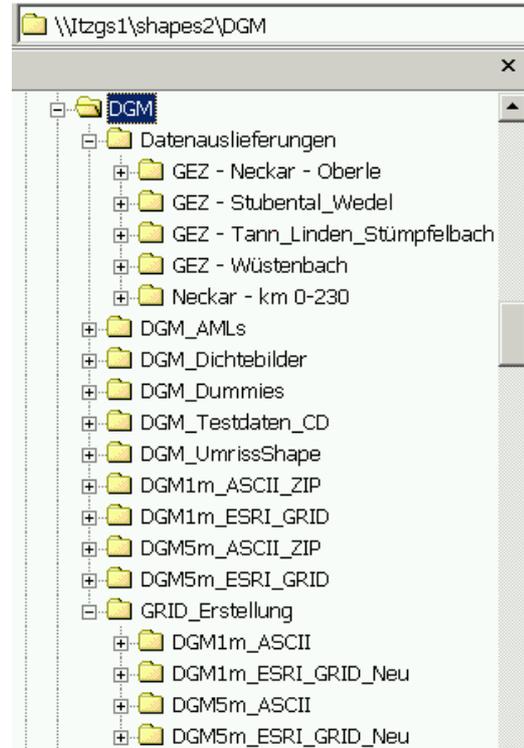
Möchte man ESRI GRIDs von mehreren CDs oder DVDs in ein gemeinsames Verzeichnis auf einer Festplatte kopieren, muss man nach den oben gewonnenen Erkenntnissen vorgehen und zuerst jede CD oder DVD in ein eigenes Verzeichnis auf der Festplatte kopieren. Anschließend wird mit dem ArcCatalog in das gemeinsame Endverzeichnis verschoben.

Datenlieferung, Datenprozessierung und Datenhaltung

Workflow

Die folgenden Punkte beschreiben die Schritte, die durchgeführt werden, wenn eine neue DGM-Datenlieferung durch das LVA bei der LfU eingeht:

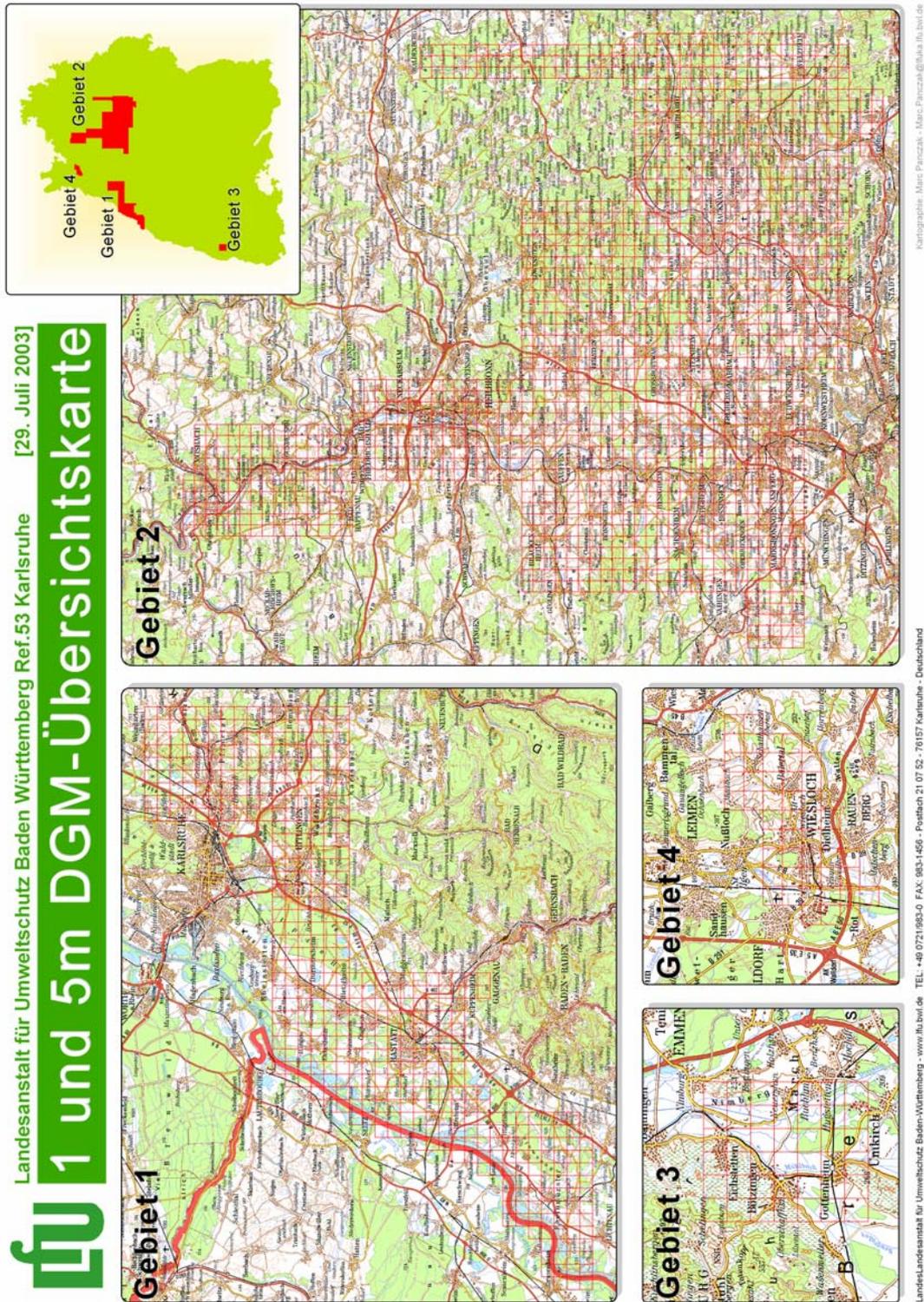
- Die Dichtebilder und gezippten 1m und 5m-DGM-Daten werden in die entsprechenden Verzeichnisse auf dem Datenserver (itzgs2) abgelegt. Dabei bekommt jede Lieferung ihr eigenes Verzeichnis. Die Datenhaltung auf dem Server hat den Vorteil, dass zum einen die Daten bei Gebrauch nicht auf den vielen verschiedenen CDs und DVDs der einzelnen Lieferungen gesucht werden müssen, zum anderen gab es bei einigen DVDs die Schwierigkeit, dass trotz pfleglicher Behandlung der Datenträger nicht mehr lesbar war. In solch einem Fall ist es schwer nachzuvollziehen, welche Daten auf der DVD gespeichert sind, um Ersatz beim LVA anzufordern.
- Die 1m-DGM-Daten werden in das Verzeichnis „GRID-Erstellung“ entpackt (s. Abb.), wo sie nach dem Entpacken als *.1m-Datei (bzw. 5m) im ASCII-Format vorliegen.
- Von diesem Verzeichnis können die Daten mit dem ASCII TO GRID Werkzeug in ESRI GRIDS umgewandelt werden. Als Zielverzeichnis dient das entsprechende Verzeichnis auf dem Server.
- Nachdem die ESRI GRIDS erstellt wurden, werden sie mit dem Werkzeug RASTER TO GEODATABASE (ArcCatalog) in die Datenbank geladen. Vor dem Laden sollte die Ladeliste im RASTER TO GEODATABASE Werkzeug als Batchdatei abgespeichert werden (Anleitung im Anhang: RASTER TO GEODATABASE, Anleitung, Erfahrungen und Handhabung).
- Ist das SDE-Raster in der Datenbank um die neuen DGM-Daten erweitert, können die dateibasierten ESRI GRIDS zu den GRIDS aus den älteren Lieferungen mit Hilfe des ArcCatalogs verschoben werden (s.o.: kopieren, verschieben, löschen).
- Die entpackten ZIP-Dateien für die GRID-Erstellung können gelöscht werden.



Hinweis

- Das Verzeichnis, in dem die dateibasierten ESRI GRIDS abgelegt werden, umfasst derzeit (Oktober 2003) knapp 3.000 GRIDS. Durch die große Anzahl dauert das Öffnen des Verzeichnisses im ArcCatalog recht lange, deshalb ist es ratsam, mehrere Unterverzeichnisse (z.B. pro Lieferung, wie bei den ZIP-Dateien) anzulegen.

Anhang 2: DGM-Übersichtskarte



[Diese Karte ist für DIN A3 konzipiert und hier verkleinert dargestellt.]

Anhang 3: TIN Erstellung

Zur Zeit der Diplomarbeit war leider keine Lizenz für den ArcGIS 8.x 3D-Analyst vorhanden, folglich können hier nur die Erfahrungen mit ArcView 3.x 3D-Analyst wiedergegeben werden.

Um TIN-Datensätze in ArcView erstellen zu können, muss die Erweiterung für den 3D-Analyst eingeschaltet sein (ArcView → Menü Datei → Erweiterungen → 3D-Analyst).

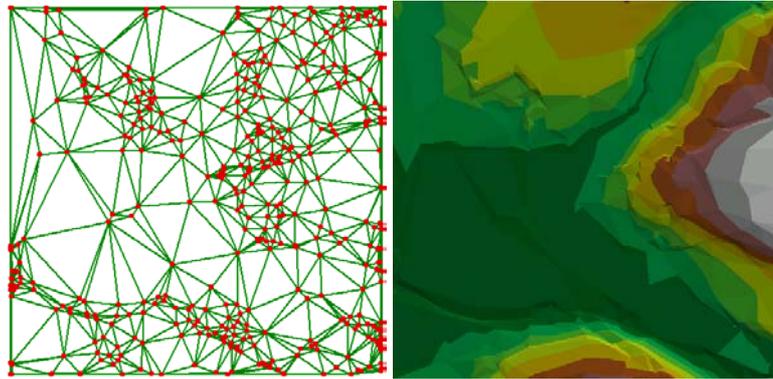
Man kann aus ASCII-XYZ-Koordinaten über ein XY-Ereignisthema ein GRID erstellen, mit dem man anschließend ein TIN erstellt. Sofern ein GRID vorliegt, kann man auch direkt aus dem GRID ein TIN erstellen.

XYZ-Koordinaten → Ereignisthema → GRID → TIN

- Zuerst importiert man die XYZ-Koordinaten als Tabelle nach ArcView. Um die Tabelle importieren zu können, müssen die Daten als DBase-Tabelle oder Textdatei (*.TXT) vorliegen. Die einzelnen Spalten müssen mit Kommas getrennt sein, ebenso die erste Zeile, die als Überschrift fungiert.
- Im nächsten Schritt kann das Ereignisthema angelegt werden.
 - View aktivieren → Menü View → Ereignisthema Hinzufügen.
 - Im darauf folgenden Dialog, wählt man unter „Tabelle“, die zuvor importierte Koordinatentabelle. Im Feld X-Wert wählt man die Spalte mit den Rechtswerten und entsprechend für die Y-Werte, die Spalte mit den Hochwerten.
 - Nach Bestätigung mit OK wird das Ereignisthema dem aktiven View hinzugefügt.
- Als Zwischenprodukt erstellt man nun das GRID.
 - Im View aktiviert man das Ereignisthema, dafür muss es nicht angeschaltet sein (Häkchen), sondern nur gegenüber den restlichen Themen hervorgehoben (Schattierung).
 - Menü Thema → In GRID umwandeln.
 - Im ersten Dialog gibt man den Namen und das Zielverzeichnis des Outputs an.
 - Im zweiten Dialog, werden die Parameter nach Wunsch gesetzt. Hinweis: Eine Änderung der Zellgröße hat erst nach dem Wechseln des Eingabefelds oder drücken der Eingabetaste eine Wirkung auf die Eingabefelder „Number of Rows“ und „Number of Columns“.
 - Im dritten und letzten Dialog für die GRID-Erstellung wird man nach der Spalte gefragt, die die eigentlichen Zellwerte enthält. Die entsprechende Spalte ist auszuwählen (in der Regel „Z“).
 - Mit einem Klick auf OK kann man den Dialog verlassen.
- Nachdem das GRID erzeugt ist, kann man mit der Erstellung des TINs fortfahren. Die ersten Schritte sind der GRID-Erstellung ähnlich.
 - Im View aktiviert man das GRID, dafür muss es nicht angeschaltet sein (Häkchen), sondern nur gegenüber den restlichen Themen hervorgehoben (Schattierung).
 - Menü Thema → GRID in TIN umwandeln.

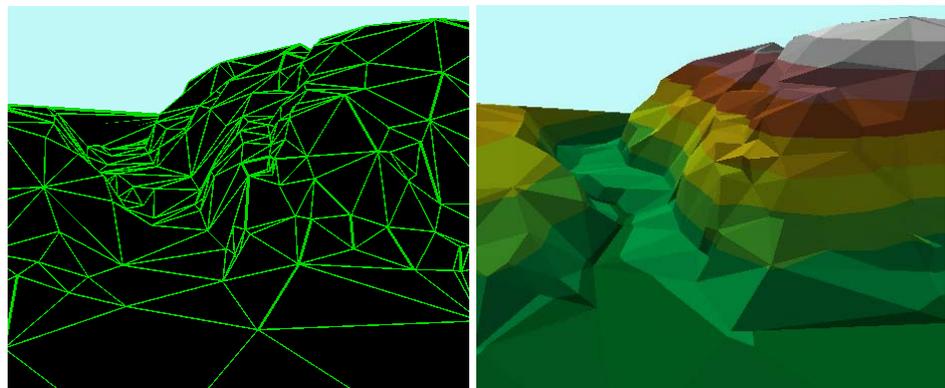
- Der einzige Parameter für die Erstellung wird im folgenden Dialog abgefragt. Es handelt sich um eine vertikale Toleranz, von der es abhängig ist, ob ein Punkt vermascht wird oder nicht. Nur wenn der Punkt außerhalb der Toleranz liegt wird er vermascht.
- Mit einem Klick auf OK startet die TIN Erstellung.

Je nachdem wie die Darstellungsoptionen ausgewählt wurden, sieht das Ergebnis entsprechend den folgenden Beispielen aus:



*(l. 2D Drahtgitterdarstellung mit hervorgehobenen Knoten;
r. 2D Hörschichtendarstellung)*

Die Darstellungen oben sind beides zweidimensionale Darstellungen, da sie im gewöhnlichen Viewer in ArcView angezeigt werden. Erst das Laden in den 3D-Analyst Viewer stellt die Daten in drei Dimensionen dar:



*(l. 3D Drahtgitterdarstellung mit hervorgehobenen Knoten;
r. 3D Hörschichtendarstellung)*

Anhang 4: Größen- und Formelsammlung für DGM-Daten

Konstanten und Größen:

1m ASCII DGM-Datei komprimiert:	4,18MB* (3,3 – 4,6MB*)
1m ASCII DGM-Datei unkomprimiert:	28,058MB*
5m ASCII DGM-Datei komprimiert:	0,181MB* (0,148 – 0,205MB*)
5m ASCII DGM-Datei unkomprimiert:	1,066MB*
1m ESRI GRID (floating, ohne Pyramid Layer)	3,85MB*
5m ESRI GRID (floating, ohne Pyramid Layer)	0,24MB*
Benötigter Speicherplatz pro ASCII-Datei bei der 1m GRID-Erstellung	31,91MB*
Benötigter Speicherplatz pro ASCII-Datei bei der 5m GRID-Erstellung	1,31MB*
Größe einer NODATA Kachel (1km ²)	0MB*

*Die Größenangabe bezieht sich auf den belegten Speicher auf der Festplatte.

Berechnungen

Berechnung des benötigten Speicherbedarfs bei der GRID-Erzeugung von n-GRIDs (ASCII-Datei + ESRI GRID ohne Pyramid Layer):

1m GRID:

Anzahl der ASCII DGMs \times 31,91 = benötigter Speicherbedarf

5m GRID:

Anzahl der ASCII DGMs \times 1,31 = benötigter Speicherbedarf

Berechnung des benötigten Speicherbedarfs für ESRI GRIDs für die Datenhaltung:

1m GRID:

Anzahl der ESRI GRIDs \times 3,85 = benötigter Speicherbedarf

5m GRID:

Anzahl der ESRI GRIDs \times 0,24 = benötigter Speicherbedarf

Berechnung der theoretisch maximalen Anzahl von ESRI GRIDs, die zu einem ESRI GRID-Mosaik zusammengeführt werden können:

$$\frac{\text{Maximaladressierbarer Speicher in MB}}{\text{Größe eines 1 bzw 5 m DGM GRID}}$$

$$\text{1m-DGM: } 1\text{GRID} = 3,85\text{MB} \rightarrow \frac{2147\text{MB}}{3,85\text{MB}} = 557\text{GRIDs}$$

$$\text{5mDGM: } 1\text{GRID} = 0,24\text{MB} \rightarrow \frac{2147\text{MB}}{0,24\text{MB}} = 8.945\text{GRIDs}$$

Anhang 5: Anwendungsfall Gewässereinzugsgebiet

Bei Hochwassersimulationen werden durch die hohe Genauigkeit der vorliegenden neuen DGM-Daten in Zukunft wesentlich genauere Prognosen möglich sein.

Um Simulationen rechnen zu können, werden dazu GRIDs mit der Ausdehnung des zu untersuchenden Gewässereinzugsgebiets (GEG) benötigt. Aufgrund der 32Bit Rechnerarchitektur sind GRIDs auf 2,175GB Speichervolumen beschränkt (näheres wurde ausführlich in der Diplomarbeit beschrieben). Es gilt zu untersuchen, ob im Rahmen der Hochwassersimulation mit Problemen bei der GRID-Größe zu rechnen ist oder nicht.

Hierzu wurde aus dem Geodatenbestand der Landesanstalt für Umweltschutz (LfU) das größte GEG innerhalb der Landesgrenzen Baden-Württembergs ermittelt. Als größtes GEG stellte sich Stubental-Wedel mit einer Gesamtfläche von 215,523km² heraus.



Übersicht über das GEG Stubental-Wedel (Teil eines ArcMap Layouts)

Die beiden folgenden Abbildungen zeigen dasselbe Gebiet.

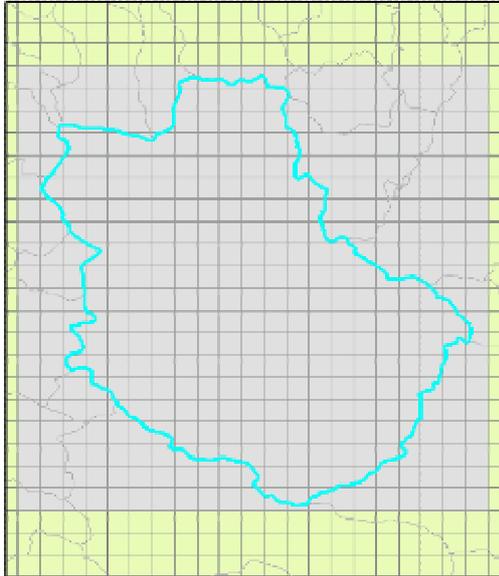
Links, als Boundingbox-Mosaik (das kleinste, das Objekt umschließende Rechteck) werden alle 1km² Kacheln in das Mosaik aufgenommen, die mit einer Boundingbox das GEG abdecken. Somit sind in der Regel auch Kacheln enthalten, die bereits außerhalb des zu untersuchenden Gebiets liegen. In der Tat sind dies in diesem Beispiel knapp 35% der Kacheln, die ohne Überschneidung außerhalb des Gebiets liegen.

Um zu demonstrieren, wie man sich im Falle einer Überschreitung der Bearbeitungsgrenze von ca. 2GB oder rechnerischen 557 Kacheln behelfen kann, wurde in der Abbildung unten rechts ein AOI-Mosaik (Area Of Interest) erstellt. In ihm sind, im Gegensatz zum Boundingbox-Mosaik, nur die Kacheln enthalten, die komplett innerhalb des GEG liegen oder von diesem angeschnitten werden, also nur

der Bereich, dem das Interesse gilt – eine „Area Of Interest“. Die so entstehenden NODATA-Bereiche belegen keinen zusätzlichen Speicher, wie andere Tests ergeben haben. Somit können auch, allerdings begrenzt, größere Gebiete zusammenhängend dargestellt werden.

BoundingBox

Metriken, zur Flächengröße und Mosaikgröße

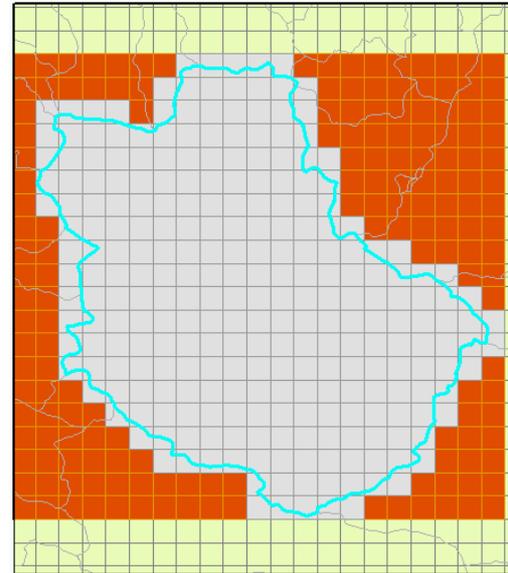


Anzahl GRIDS: 400
Größe in GB: 1,540 (berechnet)

BoundingBox-Mosaik Stubental-Wedel

AOI

Area Of Interest



Ds: 261
B: 1,005 (berechnet)

AOI-Mosaik Stubental-Wedel

Anhang 6: Anleitung für die Erstellung des DGM-Übersichtsgitters (Shape)

Der nachstehende Text ist eine Anleitung für das Erstellen einer DGM-Shapedatei, die die aktuell vorhandenen DGM-Kacheln in der LfU präsentiert.

Diese Shapedatei wird zum einen als Download ins Intranet gestellt, zum anderen wird die Shapedatei durch die SDE in die Datenbank geladen, wo sie die dort vorhandene DGM-Shapedatei ersetzt. Auf diese Shapedatei in der Datenbank greift der DGM-Web-Service zu. Im Web-Service wird das DGM-Gitter als optionaler Layer dargestellt.

Um die Shapedatei zu erzeugen, wird die Übersichtsdatei (Shape) der Digitalen Orthophotos (DOP, Name: ShapeUIS_ADDs.Blattschnitt_DOP) genutzt, da die Kacheln der DOP zu den DGM-Kacheln kongruent sind und die DOP Kacheln bereits landesweit vorliegen.

Anleitung zur Erstellung des DGM-Übersichtsgitters:

Im ersten Schritt müssen alle DGM-Kachelnamen in einer DBase-Tabelle vorliegen, wie in der Abbildung gezeigt. Dieser Schritt ist nötig, damit die Liste der DGM-Kacheln mit der Liste der DOP verschnitten werden kann. Dafür gibt es verschiedene Möglichkeiten, eine ist im Folgenden gezeigt:

- In einem Konsolenfenster wird in das Verzeichnis gewechselt, in dem alle DGM-Kacheln als ZIP-Datei, TXT-Datei oder GRID usw. vorliegen. Wichtig ist nur, dass die Dateien den Namen der Kacheln enthalten (z.B. 34505460.zip).
- Als nächstes wird mit einem DOS-Befehl der Inhalt der Konsole in eine Datei umgeleitet:
 - `Dir -> C:\DGM_KachelNamen.txt`
 - Der Befehl `dir`, listet alle Dateien in einem Verzeichnis auf. Diese Liste wird durch den Pfeil in die dahinter stehende Datei, mit dem angegebenen Pfad umgeleitet.

	A
1	DGMs
2	34035323
3	34035324
4	34035325
5	34035326
6	34035327
7	34035328
8	34035329
9	34045323
10	34045324
11	34045325
12	34045326
13	34045327

Diese Textdatei kann in Excel geöffnet und so editiert werden, dass beim Abspeichern als DBase-Tabelle (*.DBF) das Layout, dem auf der Abbildung oben entspricht. Zu beachten ist die Zeilenüberschrift in der ersten Zeile, ohne die die DBase-Tabelle von ArcGIS nicht als bearbeitbare Tabelle erkannt wird.

Nachdem die DBase Tabelle erstellt wurde, wird in ArcMap die eigentliche Shapedatei erstellt:

- ArcMap öffnen.
- Thema laden:
 - Add → Database Connections → Connection to SDE.sde → ShapeUIS_ADDs.Blattschnitt_DOP
- Rechtsklick auf das Thema:
 - Joins and Relates → Joins → Dialog Join Data
(selbsterklärend) Die DBase-Tabelle, mit den DGM-Namen wird

mit der Attributtabelle des DOP-Shapes verbunden (Join). Der Dialog wird mit einem Klick auf OK geschlossen.

- Anschließend Rechtsklick auf das DOP-Thema in ArcMap:
 - Export → All Features, lokaler Speicherort wählen.Aus irgendeinem Grund muss zuerst die gesamte Attributtabelle als Shapedatei exportiert werden, ansonsten funktioniert die anschließende Selektion nicht. Wünschenswert wäre es, gleich die Selektion der entsprechenden Datensätze vornehmen zu können.
- Thema vom lokalen Speicherort erneut laden.
 - Add → C:\...\DOP_Blattschnitt_lokal.shp
- Rechtsklick auf das Thema:
 - Open Attribut Table → Options → Select By Attributes... →
- SQL-Anweisung: "OBJECT_ID" = "DGMS" → Apply → Close
- Rechtsklick auf das Thema:
 - Data → Export Data → Dialog Export Data: Export "Selected features", Output Speicherort eingeben.

Die Bereitstellung der Datei erfolgt, wie bereits eingangs erwähnt, als Download auf der entsprechenden Intranetseite. Für eine aktuelle Darstellung im DGM-Web-Service muss die gleiche Datei die veraltete Version in der Datenbank ersetzen.

Anhang 7: Anleitung, ASCII to GRID mit der ArcToolbox

Die manuelle Erzeugung von GRIDs mit der *ArcToolbox*, war einer der ersten Schritte, die zu Beginn der Diplomarbeit gemacht wurden. Verschiedene Möglichkeiten für die GRID-Erstellung mit ArcGIS sollten herausgefunden werden, um das daraus resultierende Wissen eventuell später in den Werkzeugen für eine automatische Verarbeitung einfließen zu lassen.

Vorbereiten der Daten

Eine direkte Erzeugung eines ESRI-GRIDs aus ASCII-Dateien ermöglicht die ArcToolbox. Dafür werden lediglich die Koordinaten der linken unteren Ecke oder die Mitte der linken unteren Zelle der DGM-Kachel benötigt. Mit Hilfe der richtigen Zellgröße wird anschließend das GRID erstellt.

Die ASCII-Datei muss dafür folgende Kopfzeilen enthalten:

Header Informationen

```
<NCOLS xxx>
<NROWS xxx>
<XLLCENTER xxx | XLLCORNER xxx>
<YLLCENTER xxx | YLLCORNER xxx>
<CELLSIZE xxx>
{NODATA_VALUE xxx}
Zeile 1
Zeile 2
...
Zeile n
```

Bedeutung:

NCOLS	Ist die Anzahl der Spalten des GRIDs.
NROWS	Ist die Anzahl der Zeilen des GRIDs.
XLLCENTER XLLCORNER	Ist entweder die X-Koordinate der linken unteren Ecke oder der Mitte der linken unteren Zelle des GRIDs.
YLLCENTER YLLCORNER	Ist entweder die Y-Koordinate der linken unteren Ecke oder der Mitte der linken unteren Zelle des GRIDs.
CELLSIZE	Steht für die Kantenlänge der Zellen.
NODATA VALUE	Für Zellen, für die kein Wert vorliegt, kann ein Wert definiert werden. Diese Angabe ist optional und ist standardmäßig auf den kleinsten Wert des verwendeten Datentyps gesetzt.
XXX	Stellen die Zahlen der erwarteten Werte dar. Die eigentlichen Zellwerte sind durch ein Leerzeichen oder einen Zeilenumbruch voneinander getrennt.

Die Schritte im Einzelnen

- Die ASCII-Datei in eine TXT-Datei umbenennen und in einem Editor öffnen (hier Ultraedit).
- Kopfzeilen eintragen

```
NCOLS 201
NROWS 201
XLLCORNER 3455000
YLLCORNER 5430000
CELLSIZE 5
```

- Die XY Koordinaten entfernen (Suche nach: „_34*__“ ersetze durch: „“)*. Die XY-Koordinaten werden nicht gebraucht, da mit den Informationen aus den Kopfzeilen das GRID aufgebaut werden kann.
- Die Dezimalstellen bleiben unverändert
(* alle Unterstriche “_” sind als Leerzeichen zu verstehen)

Das Ergebnis sieht folgendermaßen aus:

```
NCOLS 201
NROWS 201
XLLCORNER 3455000
YLLCORNER 5430000
CELLSIZE 5
116.4
116.7
117.0
117.2
117.5
...
```

ASCII to GRID

- ArcToolbox öffnen → Conversion Tools → ASCII to GRID.
- Unter INPUT ASCII FILE die Quelldatei auswählen.
- Der GRID TYPE ist „float“.
- Output GRID: Ausgabenname und -ort des zu erzeugenden GRIDS angeben. Der Pfadname darf keine Leerzeichen enthalten und das GRID darf nicht mit einem Sonderzeichen oder einer Zahl beginnen, außerdem nicht mehr als 13 Zeichen im Namen enthalten.
- Mit OKAY bestätigen.



ArcToolbox mit geöffnetem Konverter WERKZEUG

Wichtig

1)

Folgender Fehler ergibt sich bei der Konvertierung mit der ArcToolbox, wenn man die falsche Zellgröße wählt. (s. Abb.) Da alle Werte relativ zu einem Wert (linke untere Ecke oder Mitte der linken unteren Zelle) anhand der Zellgröße ausgerichtet werden, ergibt sich nicht wie bei der Erzeugung über einen XY-Daten-Layer (s.u.) eine andere Auflösung, sondern eine andere GRID-Größe.



*Zellgrößenfehler bei GRID Erzeugung mit der ArcToolbox.
Bild: Kaiser Platz Karlsruhe*

2)

Die Bildkoordinaten müssen nach Zeilen sortiert

(Hochwert sortiert) vorliegen. Die Daten, die von der LVA kommen, sind nach Spalten sortiert (Rechtswert sortiert). Die Folge davon ist, dass das Bild der erzeugten GRIDs um 90° gedreht ist.

Das Werkzeug Ultraedit bietet eine Sortierfunktion, mit der es möglich ist, die Daten umzusortieren, allerdings nicht bis zum gewünschten Ergebnis. Die Daten müssten, wie Versuche gezeigt haben:

- Nach Hochwerten sortiert werden – aufsteigend (1. Sortierschlüssel)
 - Nach Rechtswerten sortiert werden – absteigend (2. Sortierschlüssel)
- Diese Art von Sortierfunktion unterstützt zum Beispiel Excel, dort lassen sich aber nicht mehr als 65.536 Zeilen importieren, bei mehr Zeilen wird abgebrochen. Folglich muss ein geeignetes Werkzeug gefunden oder, wie man später sehen wird, selbst ein passendes Werkzeug programmiert werden.

Anhang 8: Anleitung, ASCII to XY [to Shape] to GRID in ArcMap

Die manuelle Erzeugung von ESRI GRIDS mit ArcMap bzw. ArcCatalog, war eine der ersten Aufgaben, die zu Beginn der Diplomarbeit durchgeführt wurde. Der Zweck bestand darin, die Möglichkeiten kennen zu lernen, ESRI GRIDS zu erzeugen und die beste Möglichkeit für den späteren automatisierten Weg herauszufinden.

Im Nachfolgenden wird aus den bekannten ASCII-Daten über XY-Datenlayer (s. Glossar) zuerst eine Shapedatei erstellt und aus dieser ein ESRI GRID erzeugt.

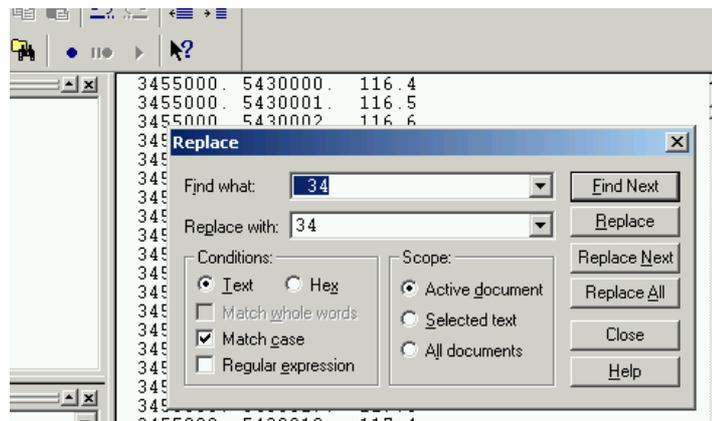
Vorbereiten der Daten

- Um die Daten im ArcCatalog darstellen zu können, sollten die ASCII-Dateien die Extension *.TXT haben. Man kann auch die Endung als neuen Dateityp in ArcCatalog integrieren.
- Die Dateien dürfen nur *einen* Punkt im Dateinamen haben, da sie sonst nicht sichtbar sind.
- ArcCatalog und ArcMap erkennen nur durch Tabulator getrennte Textdateien und nicht, wie ArcView, kommagetrennte Koordinatendateien.
- Vor- und Nachkommastellen bei Dezimalzahlen müssen durch ein Komma getrennt sein.
- Hat man eine ASCII-Datei mit XYZ-Koordinaten für die Anzeige in ArcGIS formatiert, kann es passieren, dass die Datei in ArcCatalog erkannt und angezeigt wird, aber die XY-Pulldownmenüs zum Hinzufügen von XY-Daten in ArcMap leer bleiben. In diesem Fall ArcMap einfach schließen und wieder öffnen. Die Spalten der Texttabelle sollten nun auch in ArcMap angezeigt werden.

Die Schritte im Einzelnen

Die folgenden Schritte beziehen sich auf die vom LVA gelieferten 1m und 5m-DGM-Daten für Baden-Württemberg:

- Dateiendung der ASCII-Datei in TXT umbenennen und im Editor öffnen (hier TextPad). Evtl. muss der Schreibschutz entfernt werden.
- In die erste Zeile die späteren Überschriften einfügen (x_y_z)*.
- Überflüssige Leerzeichen am Zeilenanfang entfernen (ersetze „_34“ durch „34“)*.
- Überflüssige Dezimaltrennzeichen entfernen (ersetze „._“ durch „.“)*.
- Doppelte Leerzeichen entfernen (ersetze „__“ durch „_“)*.



Replace WERKZEUG in TextPad

- Um die Dezimalzahlen später zu erhalten, die Punkte durch Kommas ersetzen (ersetze „.“ durch „,“)*.
- Da nur durch Tabulator getrennte Dateien erkannt werden, müssen Leerzeichen durch Tabulatoren ersetzt werden (ersetze „ “ durch „/t“. wobei ein Häkchen bei „Regular Expressions“ gesetzt werden muss. TextPad spezifisch)*.
*(alle Unterstriche „_“ sind als Leerzeichen zu verstehen)

ASCII to XY

- ArcMap öffnen.
- Menüleiste → Tools → Add XY-Data.
- Die entsprechende Datei auswählen.
- Die Spalten mit den X- beziehungsweise Y-Koordinaten bestimmen.
- Mit OKAY bestätigen.

101

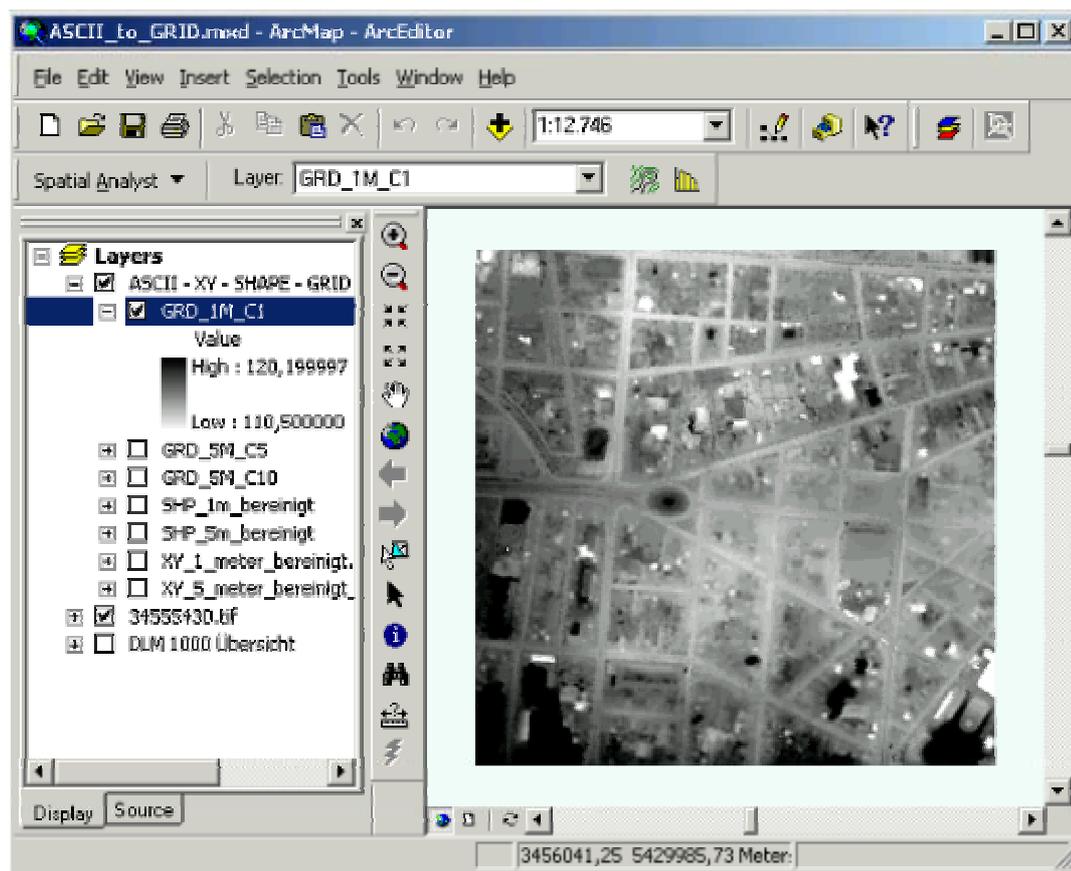
XY-Layer to Shape

Die Erzeugung der Shapedatei ist für die Erstellung des GRIDs nicht unbedingt erforderlich. Aus einem XY-Layer kann man ebenso direkt ein GRID erstellen.

- Mit der rechten Maustaste das Kontextmenü des XY-Data-Layers aufrufen.
- Data → Export Data.
- Name der zu erzeugenden Shapedatei wählen.
- Mit OKAY bestätigen.

XY-Layer/Shape to GRID

- In der Tooleiste „Spatial Analyst“ auf den gleichnamigen Button klicken:
Spatial Analyst → Converter → Features to Raster.
- Den Layer wählen, aus dem das GRID erstellt werden soll (XY-Layer oder Shapedatei).
- FIELD: Die Spalte wählen, mit deren Werten das GRID dargestellt werden soll. In diesem Fall sind das die Höhenwerte der Z-Spalte.
- CELLSIZE: Die Zellengröße ist mindestens 1 bei einem 1m Raster, sofern die Einheit des Views auf Meter eingestellt ist. (ArcMap → Menüleiste → View → Data Frame Properties → General → Display → Meters)
- Den Ausgabename und -ort des GRIDs festlegen und mit OKAY bestätigen.



Das erstellte GRID in ArcMap angezeigt

Zur Darstellung der GRIDs in ArcMap eignet sich folgende Einstellung:
GRID Kontextmenü (rechter Mausklick) → Layer Properties → Symbology → Show:
Stretched → Schwarz und Weiß.

Anhang 9: SQL Abfrage des Speichervolumens eines Rasters in Oracle

```

REM #####
REM ## Marc Panczak
REM ## Diplomarbeit 2003
REM ##
REM ## letzte Aenderung 23.06.2003
REM #####

REM ## diverse Einstellungen
REM #####
set heading off
REM set feedback off
set termout on
set pagesize 1000
set linesize 65
set echo on
spool C:\Maac\TestDIR\Tabellengrößen_mit_SQL\TabGr.log
cl scr

REM ## Tabstatistic von ORA fuer User
REM #####
select * from cat;

REM ## Tabstatistic von ORA fuer User
REM #####
desc user_extents;
REM select * from user_extents;

REM ## Tabellen der Raster beschreiben
REM #####
desc sde.RASTER_COLUMNS;
REM select * from sde.RASTER_COLUMNS where rastercolumn_id = 'DGM4X4';

REM ## Speicherallocation für Raster XXX abfragen
REM #####
select 'DGM4X4', sum(bytes)/1048576 as "SPEICHER" from user_extents
  where SEGMENT_NAME = 'SDE_BLK_528' or
         SEGMENT_NAME = 'SDE_AUX_528' or
         SEGMENT_NAME = 'SDE_BND_528' or
         SEGMENT_NAME = 'SDE_BLK_528' or
         SEGMENT_NAME = 'SDE_RAS_528' or
         SEGMENT_NAME = 'DGM4X4';

REM ## den Speicherplatz einzeln ausgeben - XXX
REM #####
select bytes, Segment_name from user_extents
  where SEGMENT_NAME = 'SDE_BLK_528' or
         SEGMENT_NAME = 'SDE_AUX_528' or
         SEGMENT_NAME = 'SDE_BND_528' or
         SEGMENT_NAME = 'SDE_BLK_528' or
         SEGMENT_NAME = 'SDE_RAS_528' or
         SEGMENT_NAME = 'DGM4X4';

REM ## Speicherallocation für Raster DGM1_CD1 abfragen
REM #####
select 'DGM4X4' as Tabelle, sum(bytes)/1048576 as "SPEICHER" from user_extents
  where SEGMENT_NAME = 'SDE_BLK_567' or
         SEGMENT_NAME = 'SDE_AUX_528' or
         SEGMENT_NAME = 'SDE_BND_528' or
         SEGMENT_NAME = 'SDE_BLK_528' or
         SEGMENT_NAME = 'SDE_RAS_528' or
         SEGMENT_NAME = 'DGM4X4';

REM ## diverse Einstellungen
REM #####
spool off

```


Anhang 10: Der LZ77 Algorithmus

Es gibt viele Varianten der LZ-Algorithmen. Trotz der großen Vielfalt haben alle LZ-Algorithmen einige gemeinsame Eigenschaften:

- Verlustfreie Datenkompression.
- Die LZ-Algorithmen benötigen zur Dekompression keine zusätzlichen Informationen über die Daten. Vielmehr wird diese Information während des Dekodiervorganges gewonnen. Das ist der Hauptunterschied der LZ-Algorithmen gegenüber anderen verlustfreien Verfahren (z.B. Huffman);
- Die LZ-Algorithmen eignen sich auf Grund dieser Eigenschaften sehr gut, um transparent für den Benutzer implementiert zu werden (z.B. das Modem-Übertragungsprotokoll v42.bis).

Der LZ77-Algorithmus wurde 1977 (daher der Name) publiziert, und war der erste von Abraham LEMPEL und Jacob ZIV entwickelte LZ-Algorithmus.

Begriffe:

Eingabedatenstrom: Die Folge von Zeichen, die komprimiert werden sollen.

Kodierungsposition: Die Position des Zeichens im *Eingabedatenstrom*, das gerade kodiert werden soll.

Datenfenster: Ein Intervall der Größe n , der den zuletzt verarbeiteten n Zeichen des *Eingabedatenstromes* entspricht.

Zeiger: Ein Zeiger auf die Position einer Zeichenkette im *Datenfenster*; im Gegensatz zum normalen Gebrauch eines Zeigers verstehen wir hier unter *Zeiger* das Tupel (Position, Länge).

Vorausschaupuffer: Die Zeichensequenz von der Kodierungsposition bis zum Ende des Eingabedatenstroms.

Wörterbuch: Eine Tabelle aus Zeichenketten. Zu jeder Zeichenkette im *Wörterbuch* gehört genau ein *Kode* als eindeutiger Index in Bezug auf das Wörterbuch.

Präfix: Eine beliebig lange Folge von Zeichen, die genau einem Zeichen vorangestellt ist.

Der Algorithmus in Einzelschritten:

Setze die Kodierungsposition auf den Anfang des Eingabedatenstroms.

(*) Finde die längste Übereinstimmung im Datenfenster mit der an der Kodierungsposition beginnenden Zeichenkette im Vorausschaupuffer.

Gib den Zeiger auf die Übereinstimmung im Datenfenster und das erste nicht passende Zeichen im Vorausschaupuffer aus.

Ist das Ende des Eingabedatenstroms noch nicht erreicht, dann bewege die Kodierungsposition (und somit auch das Datenfenster) um $l+1$ Zeichen weiter (l bezeichnet die Länge der zuletzt gefundenen Zeichenkette) und fahre bei * fort.

Beispiel

Kleinbuchstaben: Inhalt des Datenfensters

Großbuchstaben: Inhalt des Vorausschauuffers

Ausgabe: (Position, Länge) Zeichen

Position	0	1	2	3	4	5	6	7	8
Eingabedatenstrom	A	A	B	C	A	A	A	B	C

Kodierungsvorgang

Kodierungsposition	Datenfenster: Vorausschauuffer	Position: Übereinstimmung	Ausgabe
0	:AABCAAABC	-	(0,0)A
1	a:ABCAAABC	0:A	(0,1)B
3	Aab:CAAABC	-	(0,0)C
4	aabc:AAAABC	0:AA	(0,2)A
7	aabcaaa:BC	3:BC	(2,1)C(*)

(*) Hier wird anstatt der gefundenen Zeichenkette "BC" nur der *Kode* für "B" und das Zeichen "C" ausgegeben, da der Dekodierer immer einen *Kode* und ein Zeichen in Folge erwartet. Wäre hier nicht das Ende des *Eingabedatenstromes* erreicht, würde der *Kode* für "BC" ausgegeben werden.

Dekompression

Der Dekompressionsvorgang ist trivial: Das Datenfenster wird auf dieselbe Weise gehandhabt wie bei der Kodierung. Der Dekodierer gibt zuerst die durch (Position, Länge) angegebene Zeichenkette aus, anschließend das dem Zeiger nachgestellte Zeichen.

Eingabedatenstrom | (0,0)A | (0,1)B | (0,0)C | (0,2)A | (2,1)C

Dekodierungsvorgang

Eingabe	Datenfenster:Ausgabe
(0,0)A	-:A
(0,1)B	a:AB
(0,0)C	aab:C
(0,2)A	aabc:AAA
(2,1)C	aabcaaa:BC

Quelle (Internet): Universität Karlsruhe, Fakultät Informatik, www.ira.uka.de.

Anhang 11: Raster To Geodatabase, Anleitung, Erfahrungen und Handhabung

Der folgende Text enthält neben Erfahrungen und Informationen zur Handhabung eine Anleitung für das ArcCatalog Werkzeug RASTER TO GEODATABASE in Verbindung mit den Laserscan DGM-Daten für Baden-Württemberg.

Das Werkzeug dient zum Laden von Rasterdaten in die Datenbank. Die Anleitung ist relativ ausführlich, wer eine Kurzanleitung sucht, wird unter dem Stichwort RASTER TO GEODATABASE in der ArcGIS DesktoPHilfe fündig. Allerdings wird dabei nicht auf die individuellen Eigenarten und Probleme der vorliegenden DGM-Daten eingegangen.

Anleitung

Vorbereitung

Ein Problem stellt die große Anzahl von Dateien (DGM-Daten) in den Verzeichnissen dar. Diese große Anzahl verursacht unter Umständen sehr lange Wartezeiten beim Öffnen von Verzeichnissen und Hinzufügen von GRIDs in dem entsprechenden Werkzeug für den Ladevorgang.

Um bei den einzelnen Ladevorgängen nicht wiederholt diese Wartezeiten in Kauf nehmen zu müssen, wird empfohlen, beim ersten Ladevorgang in ArcCatalogs RASTER TO GEODATABASE die Liste der zu ladenden GRIDs als Batchdatei abzuspeichern (RASTER TO GEODATABASE → Batch → Save File).

Die GRIDs, die es zu laden gilt, können in mehreren solcher Batchdateien abgespeichert sein.

Diese Batchdateien können bei zukünftigen Ladevorgängen anstelle der Auswahl der einzelnen GRIDs geladen werden. Diese Batchdateien können bei Änderungen mit einem einfachen Editor modifiziert werden. Ursachen für Änderungen können beispielsweise eine Pfadänderung zum Speicherort der GRIDs im Netzwerk sein oder ein anderer Name für die Datenbankverbindung.

Laden

Die folgenden Schritte erklären die Vorgehensweise für den Ladevorgang der DGM-Daten mit Hinweisen zu möglicherweise auftretenden Problemen:

- ArcCatalog öffnen
- Datenbankverbindung zur gewünschten Datenbank herstellen
- Wenn Daten aus Batchdateien geladen werden sollen (s. Vorbereitung), ist es notwendig, zuerst ein einzelnes Raster zu laden. Dieses Raster erhält
 - die gewünschten Projektion,
 - das richtige Ladeprofil (DGM02)
 - und den SDE-Rasternamen, der für die in der Batchdatei abgespeicherten GRIDs vorgesehen ist. Standardmäßig ist hier F1 eingestellt. Dazu sind folgende Schritte nötig:
 - Rechtsklick auf die entsprechende Datenbankverbindung → IMPORT → RASTER TO GEODATABASE

- INPUT RASTER → Datei auswählen. Hat man eine Datei dem Dialog hinzugefügt, kann man unter CHANGE SETTINGS die genannten Parameter einstellen. HINWEIS: Idealerweise handelte es sich bei dieser „Musterdatei“ um ein GRID, das ein Teil der eigentlich zu ladenden GRIDs ist. Dadurch enthält das resultierende SDE-Raster kein ungewolltes, zusätzliches Raster.
- Nachdem das Musterraster geladen wurde, können nun die eigentlichen Daten geladen werden. Dazu wird wieder der RASTER TO GEODATABASE Dialog geöffnet (Klick auf: Datenbankverbindung → IMPORT → RASTER TO GEODATABASE), anschließend klickt man auf BATCH → OPEN FILE und wählt die entsprechenden Batchdateien (*.BAT) aus. Beim ersten Ladeversuch, wenn noch keine Batchdatei vorliegt, werden über INPUT RASTER → DATEI(EN) ÖFFNEN die entsprechenden DGM-Daten/GRIDs direkt hinzugefügt.
- Nachdem alle Dateien hinzugefügt wurden, muss in der Regel die erste Zeile der Ladeliste unter BATCH gelöscht werden, da sie keine Datei enthält, aber standardmäßig vorhanden ist. Hierzu die erste Zeile markieren (Linksklick auf die „1“ in der ersten Zeile) → Klick auf DELETE ROW.
- Um die #-Zeichen aus der Spalte OUTPUT RASTER CATALOG zu entfernen, klickt man in die gleichnamige Spaltenüberschrift, so dass die gesamte Spalte markiert wird. Anschließend kann im Textfeld „SELECT AN EXISTING RASTER CATALOG OR CREATE A NEW ONE“ das #-Zeichen entfernt werden.
- Versucht man den Namen des OUTPUTRASTERS zu ändern, treten Fehler gleich zu Beginn des Ladens auf und der Vorgang wird abgebrochen. D.h. keine Änderung des OUTPUT Namens vornehmen. Stattdessen muss hier der Name des zuvor geladenen Musterrasters stehen. Dadurch werden die zu ladenden Raster an das bestehende angehängt.
- Zum Ändern der Optionen für STATISTIK und PYRAMID LAYER muss die gesamte Spalte INPUT RASTER markiert werden. Dazu klickt man in die gleichnamige Spaltenüberschrift. Nach etwas Geduld klickt man auf CHANGE SETTINGS.
 - Die Optionen BUILD STATISTICS und BUILD PYRAMIDS unter STORAGE werden ausgeschaltet, da diese Optionen nur für das Ergebnisraster interessant sind und nicht für Zwischenstufen.
 - Unter SPATIAL REFERENCE kann es sein, dass hier UNKNOW steht, dies ändert sich aber ohne eigenes Zutun nach dem Verlassen des Dialogs. Nach dem Schließen, wird im RASTER TO GEODATABASE Dialog die Projektion des Musterrasters aus der Datenbank übernommen und angezeigt.
 - Unter KEYWORD wird ebenfalls das Ladeprofil des Musterrasters automatisch übernommen, allerdings schadet es auch nicht, es nochmals einzutragen. (→ DGM02)
- Der CHANGE SETTINGS-Dialog kann durch einen Klick auf OKAY geschlossen werden.
- Der Ladevorgang kann mit OKAY im RASTER TO GEODATABASE Dialog gestartet werden.

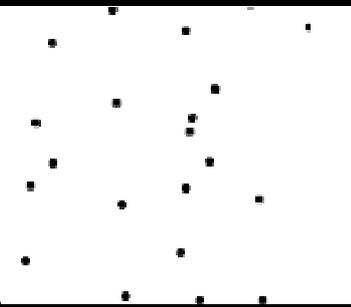
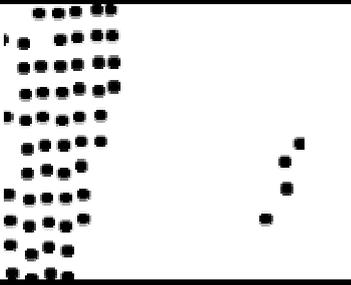
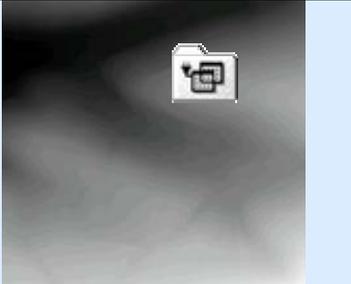
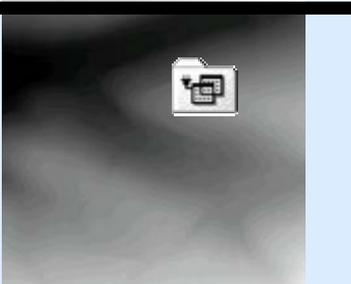
Erfahrungen und Informationen zur Handhabung

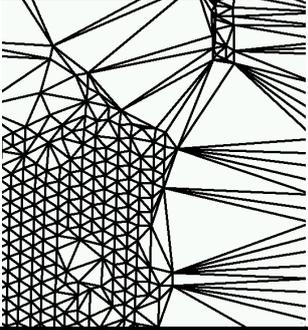
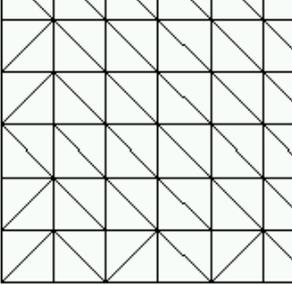
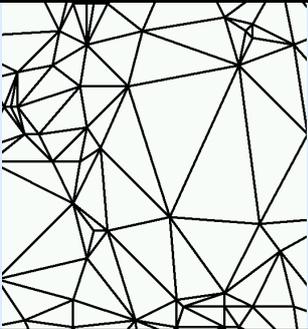
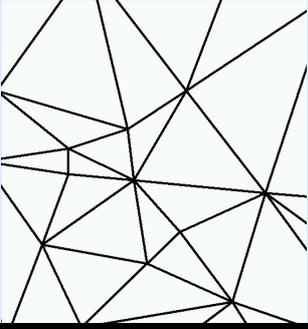
- Das Öffnen und Hinzufügen von großen Rasterdateien in das Ladewerkzeug dauert relativ lange.

- Das Laden einer großen Anzahl von Rastern in das Ladewerkzeug dauert lange. Während dieses Ladevorgangs erfolgt keine Rückmeldung der Anwendung über den Fortschritt des Prozesses. Während des Öffnens und Hinzufügens sollten keine weiteren Prozesse offen sein. Das Öffnen des Taskmanagers oder des Dialogs für die aktive Netzwerkverbindung können Aufschluss darüber geben, ob der Rechner noch arbeitet.
- Ist eine große Menge von Rastern dem Ladewerkzeug hinzugefügt worden, sollte die Ladeliste als Batchdatei (*.BAT) abgespeichert werden. Im Falle eines Fehlversuchs oder bei zukünftigen Ladeprozessen werden die Wartezeiten für das Öffnen von Verzeichnissen und Hinzufügen von Rasterdateien stark verkürzt (s. Anleitung oben).
- Werden Rasterdateien aus einer Batchdatei zur Ladeliste des RASTER TO GEODATABASE-Werkzeugs geladen, sollten diese an ein *bestehendes* Raster in der Datenbank angehängt werden. Dieses bestehende Musterraster sollte über die richtige Projektion und das richtige Ladeprofil verfügen. Denn bei einer Ladezusammenstellung aus einer Batchdatei lässt sich nachträglich die Projektion nicht einstellen oder ändern. Eine Änderung des Namens führte bei allen Versuchen zu einem Absturz, deshalb sollten bei der Abspeicherung der Ladeliste zu einer Batchdatei keine Änderungen des Outputnamens gemacht werden – der Standardname ist „F1“, dieser sollte übernommen werden. Eine Änderung des SDE-Rasternamens kann nach erfolgreicher Beendigung des Ladevorgangs im ArcCatalog vorgenommen werden. Dort kann der Name auch länger als acht Zeichen sein, was für den Ladedialog aus Kompatibilitätsgründen nicht der Fall ist.
- Das Erstellen einer Statistik und von Pyramid Layer muss aus Zeitgründen für alle bis auf das letzte zu ladende Raster ausgeschaltet werden. Es genügt, wenn beim Laden des letzten Rasters die Statistik und Pyramid Layer für das SDE-Raster erstellt werden.
- Der Ladevorgang selbst, dauert pro 1km² Kachel à 3,87MB relativ kurz (wenige Sekunden), da es sich allerdings um mehrere Tausend dieser Kacheln handelt, summiert sich die Ladezeit auf mehrere Tage. Ein Probelauf mit 150GB dauerte ca. 4Tage inklusive Pyramid Layer und Statistik, wobei Ausfälle und Fehlversuche durch Netzwerkprobleme nicht dazugezählt wurden.
- Bei dieser großen Datenmenge ist es meist nicht möglich, die einzelnen Kacheln lokal auf einem Rechner zu speichern und man ist gezwungen die Kacheln über das Netzwerk zu laden, was allerdings bei der geringen Größe der Kacheln kaum Leistungseinbußen bringt.

Anhang 12: Vergleichstest der Datenvolumina

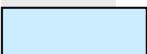
112

Punkt- wolke*	Original, ASCII-Datei	~17MB** je nach Koordinatenanzahl	
Punkt- wolke*	Bereinigt, ASCII-Datei	~14MB** je nach Koordinatenanzahl	
GRID*	Dateibasiert, ESRI- GRID [integer], interpoliert	1,18MB** [gezippt: 0,68MB]	
GRID*	Datenbank gestützt, ESRI-GRID, interpoliert	3,86MB** [gezippt: 0,94MB]	
DB Raster*	Datenbank gestützt, ESRI-GRID [floating], interpoliert	In Arbeit	

TIN ⁺	Aus Punktwolke, ohne Toleranz	14,1MB** 217.405 Knoten 434.763 Dreiecke	
TIN ⁺	Aus ESRI-GRID* (interpoliert), ohne Toleranz	65,2MB** 1.002.001 Knoten 2.000.000 Dreiecke	
TIN ⁺	Aus Punktwolke, mit Toleranz (±0,25m)	1,81MB** 27.622 Knoten, 55.121 Dreiecke	
TIN ⁺	Aus ESRI-GRID* (interpoliert), mit Toleranz (±0,25m)	0,20MB** 2.571 Knoten 5.064 Dreiecke	
TIN ⁺	Datenbank gestützt	Ist erst ab ArcGIS9.0 verfügbar (2004)	



Basisdaten oder
irrelevante Methode



Mögliche Methode

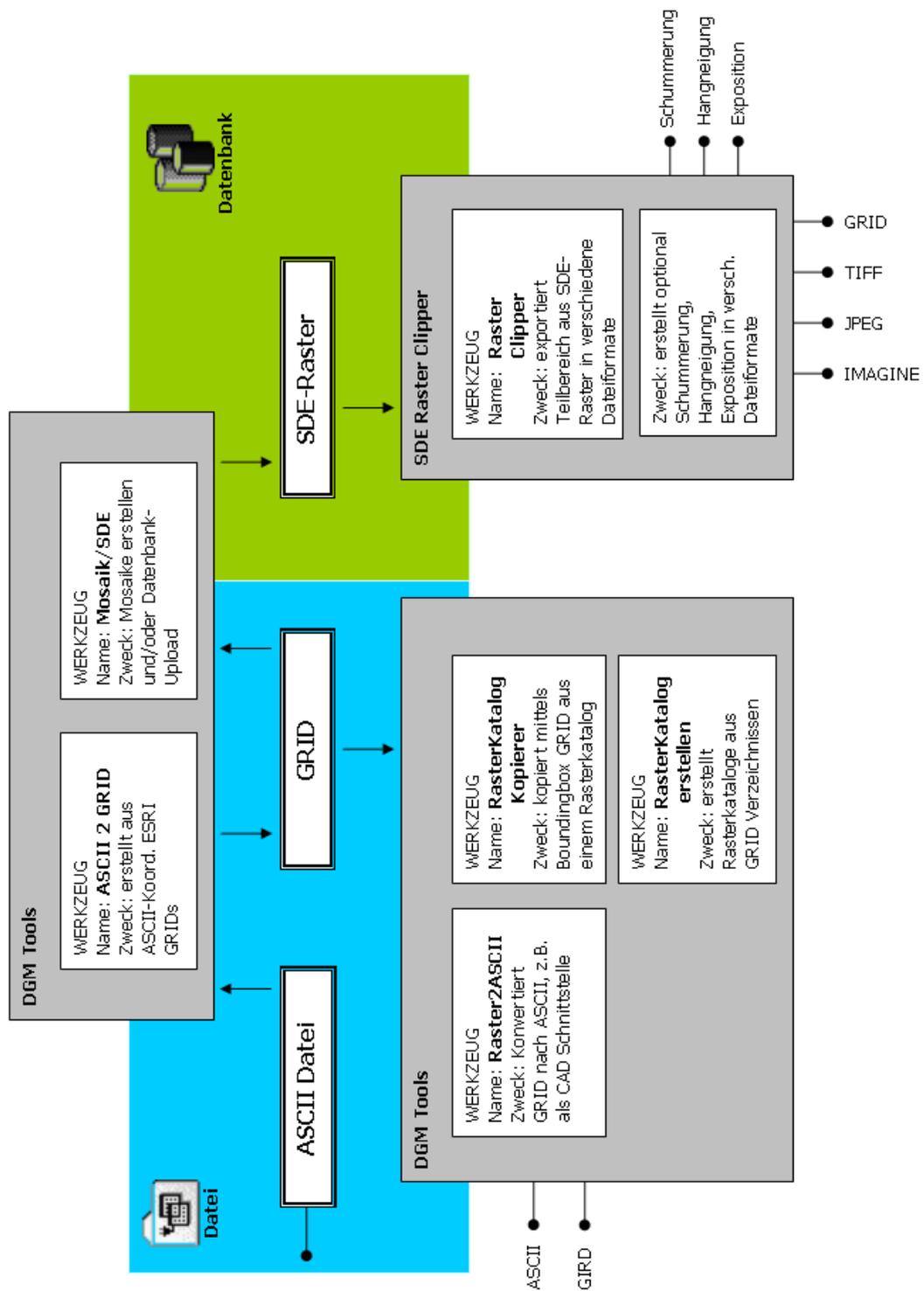
* 1m-DGM (1km² = 1000 Zeilen x 1000 Spalten)

** Größe auf dem Datenträger.

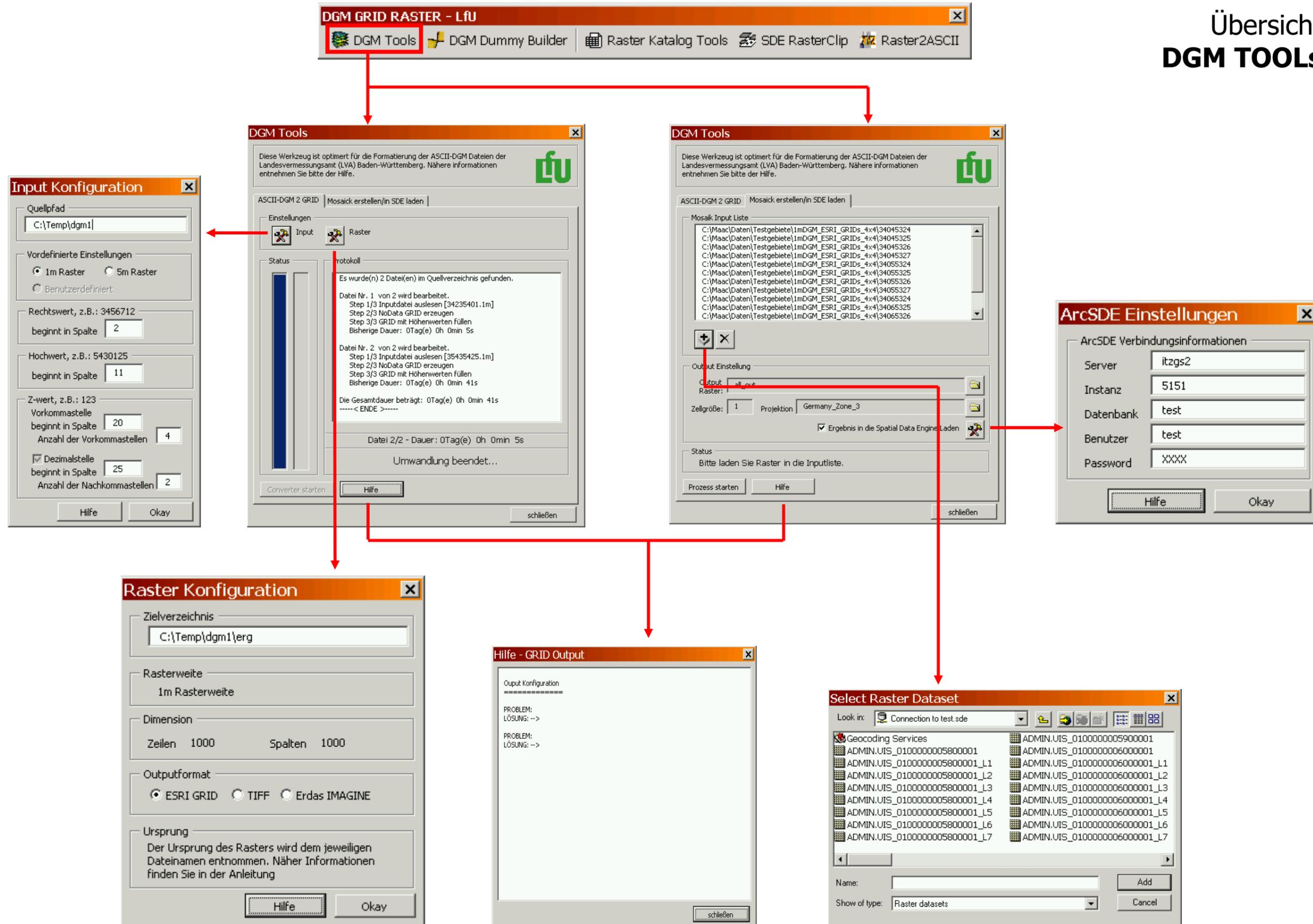
+ Die Größenangaben bei den getesteten TINs sind relativ zueinander zu betrachten, da die Größenangaben (u.a.) abhängig sind von der Anzahl der Koordinaten in der Ausgangsdatei bzw. der entstehenden Knoten.

Anhang 13: programmierte Werkzeuge, graphische Übersicht

Workflow



Übersicht DGM TOOLS



Übersicht Raster Katalog Tools



Übersicht SDE Raster Clipper & DGM Dummy Builder

