



# **Verfahren zur automatisierten Generalisierung flächenhafter Geofachdaten**

## **Masterarbeit**

im Rahmen des Universitätslehrganges  
„Geographical Information Science & Systems“ (UNIGIS MSc)  
am Interfakultären Fachbereich für GeoInformatik (Z\_GIS)  
der Paris Lodron-Universität Salzburg

eingereicht von

**Dipl.-Ing. (FH) Joachim Schuff**

104460, UNIGIS Msc Jahrgang 2016

zur Erlangung des Grades

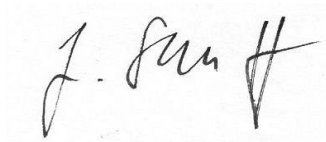
„Master of Science (Geographical Information Science & Systems) - MSc(GIS)“

## Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbständig verfasst habe, dass ich sie zuvor an keiner anderen Hochschule und in keinem anderen Studiengang als Prüfungsleistung eingereicht habe und dass ich keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Stellen der Masterarbeit, die wörtlich oder sinngemäß aus Veröffentlichungen oder aus anderweitigen fremden Äußerungen entnommen wurden, sind als solche kenntlich gemacht.

Freiburg, 04.02.2019

*Ort, Datum*

A handwritten signature in black ink, appearing to read 'J. Sam H', written over a horizontal line.

*Unterschrift*

## Kurzfassung

Aus Effizienzgründen werden heutzutage neue Kenntnisstände zu geowissenschaftlichen Themen in zentral fortgeschriebenen Datensätzen eingearbeitet. Solche Geofachdaten liegen originär meist in hoher Auflösung, also großem Maßstab vor. Für zahlreiche (Web-)Anwendungen und Fragestellungen besteht jedoch der Bedarf an einer vereinfachten, generalisierten Ausgabeform dieser Daten. Nach den weitreichenden Fortschritten bei der modellbasierten Generalisierungsautomatisation topografischer Karten im letzten Jahrzehnt wächst nun die Nachfrage nach ganzheitlichen Lösungsansätzen für die Generalisierung flächenhafter Geofachdaten. Eine besondere Herausforderung ist dabei die Umsetzung mit vorhandener Funktionalität aus verfügbarer GI-Software. In der vorliegenden Arbeit wurde auf Grundlage vorheriger Sichtung existenter Funktionen und Methoden ein praxisgerechtes automatisches Verfahren erarbeitet, mit dem die modellbasierten Originaldaten mittels Standardsoftware für verschiedene Ausgabezwecke und Zielmaßstäbe so generalisiert werden können, dass die Ergebnisse sowohl fachlichen als auch kartografischen Anforderungen genügen und reproduzierbar sind. Die Tauglichkeit und thematische Übertragbarkeit des Verfahrens wurde anhand verschiedener flächenhafter Geofachdatensätze des Landesamtes für Geologie, Rohstoffe und Bergbau Baden-Württemberg demonstriert. Hierzu dienten beispielhafte Datensätze aus den Fachbereichen Geologie, Bodenkunde und Hydrogeologie, die in umfangreichen Testreihen mit unterschiedlicher Parametrisierung des Verfahrens evaluiert wurden. Der Schwerpunkt lag dabei auf den komplexen Eingangsdaten der Geologie, die neben flächenhaften Objekten auch durch topologisch abhängige, linienhafte tektonische Strukturen geprägt sind und im Ablauf gesondert als invariante Bereiche berücksichtigt werden müssen.

Dem Verfahren wurde ein hybrider Lösungsansatz zugrunde gelegt, bestehend aus vektorieller und rasterbasierter Verarbeitung: im ersten Teil des Gesamtprozesses werden die flächenhaften Objekte auf Vektorbasis aggregiert und mit Hilfe berechneter, geometrischer Kenn- und Distanzwerte anhand verschiedener Verfahren eliminiert bzw. aufgelöst. Erhaltenswerte Elemente unterliegen dabei der Typisierung hin zu linien- oder punkthaften Objekten. Diesem Ablauf vorausgesetzt ist eine inhaltlich hierarchische Struktur der Eingangsdaten, die eine semantische, modellbasierte Zusammenfassung ermöglicht und anhand einer datenspezifischen Attributanreicherung die Berücksichtigung weiterer thematischer Aspekte beinhaltet. Im zweiten Teil des Verfahrens werden die Geometrien der aggregierten Objekte im Rasterformat vereinfacht und geglättet. Als Kernfunktionalität wurde hierzu ein zellulärer Automat eingesetzt. Das Ergebnis dieses Teilschrittes wird final wieder in einen Vektordatensatz konvertiert. Die Bündelung aller Teilschritte zu einem automatischen, zentral steuerbaren Gesamtprozess erfolgte durch Implementierung mit der Software FME.

## Abstract

For reasons of efficiency, new levels of knowledge on geoscientific topics are now incorporated in centrally updated data sets. Such geotechnical data are usually available in high resolution, i.e. on a large scale. For numerous (web) applications and questionings, however, there is a need for a simplified, generalized output form of this data. After the far-reaching progress in model-based generalization automation of topographic maps in the last decade, the demand for holistic solutions for the generalization of categorical data is now growing. A particular challenge is the implementation with existing functionality from available geoinformation software. In the present work, a practice-oriented automatic procedure was developed on the basis of a previous review of existing functions and methods. The model-based original data can be generalized by means of standard software for various output purposes and target scales in such a way, that the results satisfy both technical and cartographic requirements and are reproducible. The suitability and thematic transferability of the method was demonstrated on the basis of various categorical geo-data sets of the state bureau for geology, mineral resources and mining of Baden-Württemberg (Landesamt für Geologie, Rohstoffe und Bergbau Baden-Württemberg - LGRB). For this purpose exemplary data sets from the fields of geology, soil science and hydrogeology were used. They were evaluated in extensive test series with different parameterisation of the applied method. One focus was on the complex input data of the geology, which are characterized not only by polygonal objects but also by topologically dependent, linear tectonic structures which have to be considered separately as invariant areas.

The procedure was based on a hybrid approach, consisting of vectorial and raster-based processing: in the first part of the overall process, the areal objects are aggregated on a vector basis and eliminated or resolved with the aid of calculated geometric characteristics and distance values using various procedures. Elements that are worth preserving are subject to collapse into linear or point-like objects. The prerequisite for this process is a hierarchical structure of the input data in terms of content, which enables a semantic, model-based aggregation and includes the consideration of further thematic aspects based on data-specific attribute enrichment. In the second part of the procedure, the geometries of the aggregated objects are simplified and smoothed within raster format. A cellular automaton was used as core functionality. The result of this sub-step is finally converted back into a vector data set. The bundling of all sub-steps to an automatic, centrally controllable overall process took place by implementation with the software FME.

## **Danksagung**

Ein ganz herzlicher Dank geht an Herrn Dr. Andreas Hagemeyer, der die Arbeit im LGRB von fachlicher Seite fundiert mitbetreute und mir jederzeit mit Rat und Tat zur Seite stand. Ebenso bedanke ich mich bei Herrn Dr. Matthias Franz für das gewissenhafte Lektorieren der Arbeit sowie Frau Claudia Rodat für ihre Hinweise und Tipps.

Neben den Kollegen geht mein Dank auch an das UNIGIS-Team der Universität Salzburg sowie im Speziellen Frau Wallentin für die Betreuung während der Masterarbeit.

Herrn Dupke von der Firma Conterra gebührt besondere Erwähnung für die Umsetzung einiger technischer Besonderheiten mit der Software FME.

Zu guter Letzt danke ich Isabella und Lennard für ihre Geduld - die erlittenen Abstriche im Familienleben werde ich wiedergutmachen...

# Inhaltsverzeichnis

<b>1 Einleitung.....</b>	<b>1</b>
1.1 Ausgangssituation.....	1
1.2 Rahmenbedingungen.....	2
1.3 Leitfrage und operative Teilziele.....	3
1.4 Vorgehen.....	4
1.4.1 Anforderungskategorien.....	4
1.4.2 Anwendungsszenarien der Generalisierung im LGRB-Umfeld.....	5
1.5 Aufbau der Masterarbeit.....	7
<b>2 Grundlagen der automatisierten Generalisierung.....</b>	<b>8</b>
2.1 Überblick zu Literatur und dem Stand der Forschung.....	8
2.2 Wesentliche Komponenten und Aspekte der Generalisierung.....	10
2.2.1 Modellgeneralisierung und kartografische Generalisierung.....	10
2.2.2 Semantische und geometrische Generalisierung.....	11
2.2.3 Operatoren und Algorithmen.....	12
2.2.4 Kartografische Bedingungen.....	13
2.2.5 Messungen und Kenngrößen.....	15
2.2.6 Räumliche oder maßstäbliche Bezugsebenen.....	16
2.2.7 Evaluierung.....	16
2.2.8 Kartografische Relationen.....	17
2.2.9 Generalisierungskonflikte.....	20
2.2.10 Leiter- oder Sternansatz.....	20
2.2.11 Datenanreicherung.....	21
2.2.12 Konzeptioneller Rahmen.....	21
2.2.13 Orchestrierung des Gesamtprozesses mittels Workflow.....	22
<b>3 Verfügbare Verfahren und Methoden.....</b>	<b>23</b>
3.1 Situation im Landesamt für Geologie, Rohstoffe und Bergbau.....	23
3.2 ArcGIS.....	24
3.2.1 ModelBuilder als Framework für die Generalisierungsautomatisation.....	24
3.2.2 GeoScaler: Lösungsansatz mittels Arc-Toolbox.....	24
3.3 PostGIS (und QGIS mit Python).....	26
3.3.1 Generische PostGIS-Funktionen mit Generalisierungsfunktionalität.....	26
3.3.2 Schnittstelle zum OpenSource-Client QGIS.....	29

3.3.3	Inward-Outward-Buffering.....	30
3.4	FME.....	31
3.4.1	Zusammenstellung wichtiger FME-Transformer mit Generalisierungsfunktionalität.....	31
3.4.2	Handhabung einzelner Ablaufschritte innerhalb einer Workbench.....	34
3.5	Zusammenfassende Beurteilung hinsichtlich vorhandener Generalisierungsfunktionalität. . .	35
<b>4</b>	<b>Datengrundlagen.....</b>	<b>37</b>
4.1	Allgemeine Betrachtung der GeoLa-Geofachdaten.....	37
4.1.1	Herausstellungsmerkmal Geologie.....	38
4.1.2	Semantische Aggregation im vorliegenden Kontext.....	39
4.2	Daten der bodenkundlichen Karte 1 : 50 000.....	40
4.2.1	Semantische Vorbedingungen.....	40
4.2.2	Besonderheiten.....	41
4.3	Daten der Geologischen Karte 1 : 50 000.....	42
4.3.1	Bezugsmatrix der geologischen Themenebenen.....	43
4.3.2	Besonderheit quartär-geologischer Einheiten.....	43
4.3.3	Semantische Vorbedingungen.....	44
4.4	Datenproblematik bei der Ausgangssituation.....	46
4.4.1	Quartärflächen versus abgedeckte geologische Karte.....	46
4.4.2	Datenumfang und Datenqualität.....	46
4.4.3	Auswahl des Sternansatzes als Ableitungsmethode.....	46
4.5	Analyse der Ausgangsdaten.....	47
4.5.1	Komplexitätsunterschiede der Eingangsgeometrien.....	48
4.5.2	Lokalisierung der unterschiedlichen Generalisierungssituationen/Konfliktfälle.....	49
4.5.3	Generische Problemfelder.....	51
4.5.4	Festlegung von Testbereichen.....	54
<b>5</b>	<b>Methodenimplementierung.....</b>	<b>55</b>
5.1	Generalisierung der bodenkundlichen Daten mittels Python/PostGIS-Geoprozessierung.....	55
5.1.1	Polygonaufteilung mittels Skelettlinien.....	56
5.2	Anwendung der GeoScaler-Toolbox.....	57
5.2.1	Funktionsweise des zellulären Automaten (CA).....	57
5.2.2	Einbettung in der Toolbox.....	58
5.2.3	Generalisierungstests mit dem CA-Modul.....	59
5.2.4	Entkapselung der CA-Prozessierung.....	61
5.3	Lösungsansatz durch FME-Workflow.....	63
5.3.1	Ablaufschema.....	65

5.3.2 Datenanreicherung.....	66
5.3.3 Objektauflösung.....	66
5.3.4 Generalisierungsstrategie für die Formvereinfachung auf Makroebene.....	67
5.3.5 Bestimmung und Handhabung variabler Werte (Parameter) im Ablauf.....	68
5.3.6 Modul für die statistische Ausgabe von Kennwerten.....	68
5.3.7 Vorgehen bei der Anwendungssimulation des Workflows.....	69
<b>6 Ergebnisse.....</b>	<b>70</b>
6.1 Darstellung der Generalisierungsergebnisse.....	70
6.1.1 Beispiel 1 - semantische Aggregation und Formvereinfachung (B1).....	71
6.1.2 Beispiel 2 - Punkt- und Linientypisierung (Kollaps - C1, C2).....	72
6.1.3 Beispiel 3 - Auswirkung der Formvereinfachung auf Makro-Ebene (C4).....	73
6.1.4 Beispiel 4 - Einbezug invarianter Bereiche durch tektonische Lineaturen (D1).....	74
6.1.5 Beispiel 5 - Erhalt geologischer (Graben)Strukturen (D3).....	76
<b>7 Diskussion.....</b>	<b>77</b>
7.1 Thematische Übertragbarkeit.....	77
7.1.1 FME-Lösungsansatz in der Anwendbarkeit auf die bodenkundlichen Ausgangsdaten.....	77
7.2 Weitere Fachthemen.....	79
7.2.1 Anwendbarkeit auf die Generalisierung von Deckschichten.....	79
7.2.2 Ingenieurgeologische und rohstoffgeologische Daten.....	80
7.3 Fazit.....	81
7.4 Ausblick - Erweiterbarkeit und Entwicklungspotenzial.....	82
<b>Literaturverzeichnis.....</b>	<b>83</b>
<b>Tabellenverzeichnis.....</b>	<b>88</b>
<b>Abbildungsverzeichnis.....</b>	<b>89</b>
<b>Anhang.....</b>	<b>90</b>
Anhang A - Auflistung der Werkzeuge mit Generalisierungsfunktionalität in ArcGIS Desktop.....	90
Anhang B - Prä-prozessierung der Ausgangsdaten.....	92
Anhang C - Python-Skript sowie alternative FME-Umsetzung zur Polygon-Partitionierung.....	93
Anhang D - CA-Testreihe.....	95
Anhang E - PythonCaller-Funktion zur semantisch bedingten Objektauflösung.....	96
Anhang F - Statistikausgabe im FME-Workflow.....	97



# 1 Einleitung

## 1.1 Ausgangssituation

Die geologische und bodenkundliche Landesaufnahme sowie die daran anknüpfenden angewandten Themenbereiche der Hydrogeologie, Ingenieurgeologie und Rohstoffgeologie liefern wesentliche Grundlagen für die Kenntnis und die Nutzung der natürlichen Ressourcen eines Landes. Die Erstellung einer landesweiten homogenisierten Datengrundlage für die Basisthemen Geologie und Bodenkunde war eine der vordringlichen Aufgaben des Landesamts für Geologie, Rohstoffe und Bergbau in Baden-Württemberg (LGRB) im Rahmen der integrierten geowissenschaftlichen Landesaufnahme (GeoLa) im vergangenen Jahrzehnt (LGRB 2015). Diese Daten liegen nun flächendeckend seit dem Jahr 2015 themenabhängig für den Maßstabsbereich 1 : 25 000 bis 1 : 50 000 vor, unterliegen der ständigen Überarbeitung und Verfeinerung und sind wesentliche Grundlage für die Ableitung weiterer fachspezifischer Daten der angewandten Bereiche. So können aus den Ausstrichflächen der geologischen Einheiten an der Oberfläche basierend auf deren Eigenschaften z.B. durch Zusammenfassung oder mittels Analyse Karten anderer Fragestellungen generiert werden: Beispiele hierfür sind die ingenieurgeologische Karte der Geogefahren (LGRB 2017) oder die hydrogeologische Karte von Baden-Württemberg (LGRB 2018). Diese Daten werden produktbezogen in Form von Geodaten oder Geodatendiensten der Öffentlichkeit zur Verfügung gestellt. Deren Verwendung und Darstellung in kleineren Maßstäben ist aus Gründen der Erkennbarkeit allerdings ohne vorherige kartografische Generalisierung nicht praktikabel. Die vorliegenden amtlichen Übersichtskarten kleinerer Maßstäbe basieren als individuelle kartografische Erzeugnisse auf analogen Grundlagen bzw. Altdaten und sind nicht mehr konform mit den derzeit vorliegenden Datenbeständen. Die Karten wurden hierfür im Fachbereich normalerweise einmalig einem eigenständigen, zu meist nicht klar regelbasiert definierten herkömmlichen fachlich-kartografischen Generalisierungsprozess unterzogen. Dieses Vorgehen ist aber heutzutage unter dem Aspekt der Aktualität und der damit verbundenen Fortschreibungsnotwendigkeit von Grundlagendaten nicht mehr zufriedenstellend und zweckdienlich. Ein strategisches Ziel wäre nun die automatisierte Vereinfachung/Ableitung/Generalisierung der jeweiligen Flächendaten aus dem grundlegenden Basisdatenmodell für die Erzeugung valider, wiederholbarer Generalisierungsergebnisse. Bei einer vorgegebenen themenspezifisch vorzunehmenden Zusammenfassung von Klassen in gleichartige übergeordnete Kategorien (semantische Generalisierung) liegt der Fokus der Arbeit auf der Entwicklung einer nachvollziehbaren und übertragbaren Prozessablaufkette zur Vereinfachung und Anpassung der neugebildeten Raumeinheiten mittels verfügbarer und einfach zu implementierender Generalisierungsverfahren. Die dabei auftretenden Schwierigkeiten und Teilprobleme werden identifiziert und skizziert und der gewählte Lösungsansatz beschrieben.

## 1.2 Rahmenbedingungen

Im konkreten Fall wird von folgenden Rahmenbedingungen ausgegangen:

- Ausgangsbasis sind die flächenhaft vorliegenden GeoLa-Geofachdaten unterschiedlicher Themenbereiche (Geologie, Bodenkunde, Hydrogeologie, Ingenieurgeologie, Rohstoffgeologie). In einem ersten Untersuchungs- und Umsetzungsschritt sollen die bodenkundlichen und geologischen Einheiten im Fokus der Betrachtung stehen.
- Generalisierungsziel ist die Auswahl und Anzeige geometrisch vereinfachter und inhaltlich zusammengefasster Kategorien unterschiedlicher Gliederungsstufen mit vorab festgelegtem Maßstabsbezug.
- Zentrales Element für die inhaltliche Zusammenfassung (semantische Aggregation) ist ein Themenbaum, welcher dem Nutzer unterteilt nach den Fachgebieten eine hierarchisch gegliederte Auswahl der im Land vorkommenden fachspezifischen Einheiten liefert.
- Die Generalisierungsvorgänge sollen möglichst anhand eines einfach zu steuernden, vollautomatisch ablaufenden Workflows in der zur Verfügung stehenden Arbeitsumgebung und Geodateninfrastruktur bewerkstelligt werden.
- Das Generalisierungsverfahren soll für eine künftig zyklische Ausführung in bestehende Abläufe zur Produktgenerierung integrierbar sein und zu wiederholbaren Ergebnissen führen. Die Performanz der Lösung steht dabei nicht im Vordergrund - die Generalisierung zur Laufzeit in webbasierten Umgebungen ist nicht Gegenstand der Arbeit<sup>1</sup>.
- Inhaltlich wird für die Generalisierungslösung sowohl der Erhalt strukturgebender oder charakteristischer fachlicher Objekte wie auch des topologischen Bezugs der Objekte zueinander angestrebt. Der Grad der möglichen Umsetzung wird aber durch den damit verbundenen notwendigen Aufwand definiert und ist unter dem jeweiligen Aspekt des Zielmaßstabs zu beurteilen.
- Die zu generalisierenden flächenhaften Geofachdaten werden in einer PostGIS-Datenbank bereit gehalten. Zielsetzung ist die Übernahme der Ergebnisse aus der Generalisierungsautomatisation für die einzelnen Hierarchiestufen als jeweils eigenständiger Geodatensatz in die Datenbank-Umgebung. Abruf und Anzeige der gewünschten Themen werden vordergründig auf Basis von WMS/WFS-Diensten bewerkstelligt.

---

<sup>1</sup> Die Ausgabe einer generalisierten Version von Originalgeometrien zur Laufzeit ist aufgrund der Größe des angefragten Gebietes (im Maximalfall die Landesfläche Baden-Württembergs) und der darin enthaltenen Geometrien zumeist nicht performant möglich. Die Aggregation in die nächsthöhere Hierarchieebene ist zwar mit einer Reduktion der Anzahl von Einzelobjekten verbunden, führt aber ohne Formvereinfachung gleichzeitig zu einer Erhöhung der Komplexität der neuentstandenen Geometrien - auf der obersten Hierarchieebene kann dies zu geometrischen Gebilden führen, die trotz räumlicher Indizierung ohne eine Tilestruktur nicht mehr zeitgerecht abrufbar sind.

### 1.3 Leitfrage und operative Teilziele

Die nun im Kontext dieser Arbeit relevante Forschungsfrage kann folgendermaßen formuliert werden: In wie weit kann ein mittels herkömmlicher GI-Software erstellter Prozessablauf den komplexen Vorgang zur ganzheitlich automatisierten Generalisierung von kategorischen Flächendaten abbilden und wie kann die logische Abfolge der Teilschritte in einem solchen Ablauf beschaffen sein?

Als operative Teilziele können genannt werden:

- Identifizierung der im Speziellen vorhandenen Teilprobleme bzw. unterschiedlichen Problemniveaus zugrundeliegender Daten (z.B. Umgang mit invarianten Teilbereichen eines Polygons; unterschiedlicher geometrischer Homogenitätsgrad bei der Klassenneubildung; allgemeine Generalisierungsstrategie etc.).
- Vergleich, Einschätzung und Auswahl von Werkzeugen und Funktionen vorhandener gebräuchlicher GI-Systeme, die bei der Erarbeitung eines geeigneten Generalisierungsverfahrens Verwendung finden können.
- Generierung einer (eventuell in Form eines Frameworks) möglichst fachthemenunabhängigen stringenten Prozessablaufkette zur geometrischen Generalisierung, welche auch die Implementierung der fachlich vorab definierten semantischen Generalisierung (Aggregation) im Gesamtablauf beinhaltet.
- Beurteilung der Ergebnisse unter Berücksichtigung unterschiedlicher Anwendungsszenarien und Zielmaßstäbe sowie Einschätzung hinsichtlich Übertragbarkeit und Weiterverwertbarkeit des Prozessablaufs.

## 1.4 Vorgehen

### 1.4.1 Anforderungskategorien

Beim Betrachten der verschiedenen Eingangsdaten lassen sich unterschiedliche Anforderungsstufen an die Problemlösung identifizieren. Damit verbunden sind jeweils unterschiedliche Zielsetzungen oder Anwendungsfälle, in deren Mittelpunkt entweder die Daten (modellorientiert) oder deren Visualisierung (grafikorientiert) stehen:

#### A) Themenunabhängige Flächengeneralisierung ohne weitergehende geometrische Transformation

Zielsetzung ist eine einfache Zusammenfassung und Reduzierung vorhandener Einheiten durch Neuklassifizierung im Rahmen einer semantischen Aggregation. Es finden keine über die Auflösung von Polygongrenzen hinausgehenden geometrischen Operationen statt. Das Generalisierungsergebnis führt im besten Falle zu weniger und größeren Flächeneinheiten, die allerdings bei der maßstabsbezogenen verkleinerten Anzeige eine Interpretation nur schwerlich gewährleisten dürfte. Diese Lösung vermeidet zwar Informationsverlust, kann aber sehr komplexe Geometrien zur Folge haben, die im Abfrage-/Darstellungsfall zusätzlich Performanzprobleme verursachen können. Diese Kategorie stellt einen notwendigen Teilschritt für weitere geometrische Generalisierungsoperationen dar.

#### B) Themenunabhängige geometrische Flächengeneralisierung ohne Einbezug von Bedingungen

Aufbauend auf dem initialen Teilschritt der Kategorie A werden im Nachgang auch konkrete geometrische Generalisierungsoperationen durchgeführt. Dabei kann es sich neben Maßnahmen zur Linienvereinfachung oder Stützpunktreduktion auch um das Eliminieren von Flächen handeln. Dabei sollen aber keine lagebezogenen Informationen verloren gehen, vielmehr werden die Daten nach ihrer Klassifizierung separat in einzelnen Gruppen generalisiert und ohne topologischen Bezug zueinander in einem Datensatz bereit gestellt. Für die Anzeige in einer Flächenkarte mit allen Klasseneinheiten in der Form von gegenseitigen Überlagerungen nach dem Prinzip „von alt nach jung“ ist diese Anforderungsstufe bedingt geeignet.

#### C) Themenunabhängige Flächengeneralisierung unter Einbezug allgemeiner Bedingungen

In der nächsten Stufe findet dann eine tatsächliche Vereinfachung der Information statt, konkret werden Flächen auf Basis von Mindestgrößen unterdrückt bzw. weitere charakteristische Generalisierungsvorgänge wie geometrische Zusammenfassung, Auswahl oder Vereinfachung angewandt, um somit einer für die Anzeige in einer Kartendarstellung geringerer Auflösung angepassten Lösung gerecht zu werden. Topologische Bezüge innerhalb der Flächendaten bleiben gewahrt. Das Generalisierungsergebnis ist für kartografische Ausgaben geeignet, hält aber nur bedingt einer detaillierten fachlichen Betrachtung und Überprüfung stand.

#### D) Themenunabhängige Flächengeneralisierung unter Einbezug fachlicher Bedingungen

Es fließen weitere zusätzliche Faktoren fachlicher Art in die Generalisierung mit ein: Wahrung topologischer Abhängigkeiten zu Informationsebenen innerhalb des Datenmodells (z.B. Tektonik, Talverläufe) sowie struktureller Gegebenheiten bzw. Wertung und Gewichtung fachlicher Information. Das Ergebnis wird den fachlichen Anforderungen an eine Übersichtskarte vollumfänglich gerecht.

#### E) Themenübergreifende Flächengeneralisierung unter Berücksichtigung aller Bedingungen

In der komplexesten Anforderungsstufe wird über die bisher berücksichtigten allgemeinen und fachlichen Bedingungen hinaus zusätzlich die Wahrung der topologischen Abhängigkeiten zwischen verschiedenen Themenebenen innerhalb eines Gesamtmodells angestrebt.

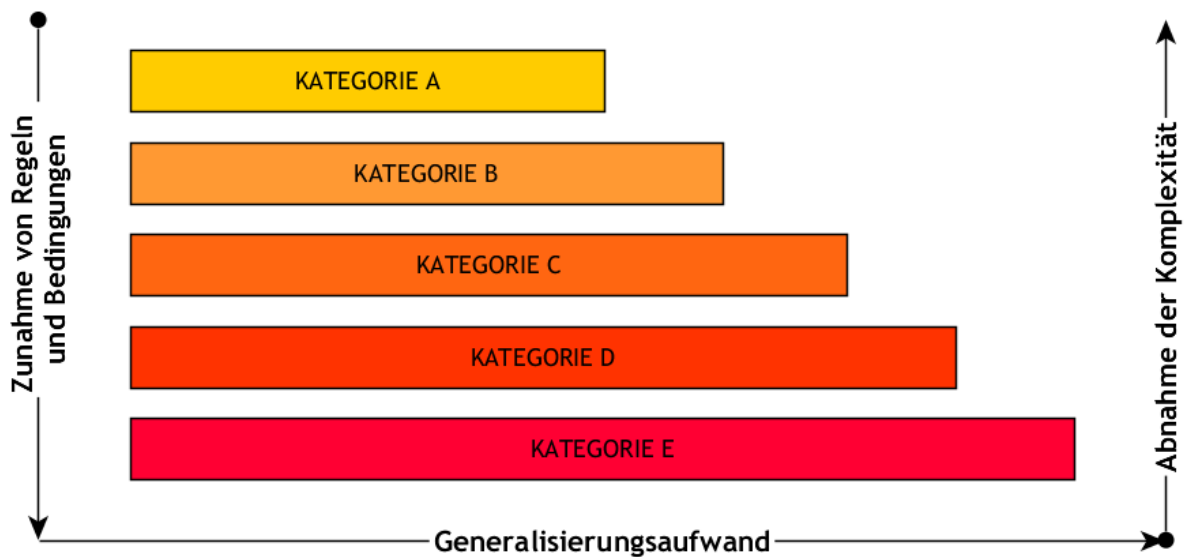


Abb. 1: Anforderungsstufen der Generalisierungsautomatisierung von Geofachdaten im LGRB

#### 1.4.2 Anwendungsszenarien der Generalisierung im LGRB-Umfeld

Die Ableitung der Daten aus dem Basisdatenmodell höchster Auflösungsstufe dient in der Regel einem konkreten Anwendungsziel. Dabei muss es sich nicht zwangsläufig um die Auswahl und Vereinfachung von Datenobjekten zum Zweck ihrer Visualisierung in zumeist geringer aufgelösten Karten handeln - „die grafische Darstellung steht heute nicht mehr ausschließlich im Zentrum der Generalisierung. Im Zuge des vermehrten Einsatzes von Geoinformationssystemen zur Verarbeitung von Geodaten verliert die grafische Darstellung an Bedeutung und wird zum Kommunikationshilfsmittel für den Menschen bei der Betrachtung von Geodaten herabgestuft. Denn Datenverarbeitung und Analyse kann mit Hilfe der EDV zum Teil auch ohne grafische Darstellung möglich gemacht werden“ (Bobzien 2005).

Dennoch steht der Anwendungsfall der kartografisch motivierten Generalisierung auch bei der Automation nach wie vor im Fokus. Im Umfeld der flächenhaften Geodaten des LGRB lassen sich den im vorigen Abschnitt aufgeführten Anforderungskategorien komplementär entsprechende Anwendungsszenarien aus der Praxis zuordnen:

- A) Die Zusammenfassung und Neukategorisierung von Einheiten ohne entsprechende Grenzauflösung wird in der Regel im Zusammenhang mit (bodenkundlichen) Auswertekarten lediglich auf der attributiven Ebene durchgeführt, ein Maßstabswechsel ist dabei nicht vorgesehen. Normalerweise werden hierfür entsprechende Look-Up- oder Schlüsselstabellen verbunden oder dauerhaft angehängt. Bei der Visualisierung werden die Polygone als farbdifferenzierte Klassen entweder ohne oder mit Umrandungslinie angezeigt, je nachdem ob die ursprünglich zugrundeliegende Raumeinheit von Interesse für die Aussage ist.
- B) Die Motivation zur Durchführung einer geometrischen Vereinfachung auf Ebene der Kategorien unter Entkoppelung der topologischen Konsistenz im gesamten Polygonmosaik liegt in der Reduktion der Komplexität von Geometrien. Die Daten der unterschiedlichen Fachbereiche dienen jeweils als

Grundlage eines WMS- (oder WFS-)Dienstes, der in einem WebGIS-Kartenmodul (hier: LGRBWissen<sup>2</sup>) die gefilterte Abfrage und Anzeige von Verbreitungsflächen fachspezifischer Einheiten ermöglicht. Diese Anforderung sollte als Mindestvoraussetzung für die in dieser Arbeit avisierte automatisierte Generalisierungslösung erfüllt werden.

- C) Im konkreten Fall ist hier die Erzeugung vereinfachter Übersichtskarten für die verschiedenen Fachbereiche angedacht. Die soweit existenten Produkte wurden durch fachliche manuelle Generalisierung einmalig für die kartografische Umsetzung und analoge Veröffentlichung erzeugt und wurden zumeist erst nachträglich digitalisiert. Die Ablösung dieser veralteten und mit den heute vorliegenden Daten nicht mehr kongruenten Karten ist eines der vordringlichsten Ausgangsszenarien für die Entwicklung der automatisierten Generalisierung.
- D) Aus dem Basismodell erfolgt eine als Modellgeneralisierung durchgehend automatisierte Ableitung zur Bereitstellung eines fachlich verifizierten Geodatenproduktes für den Vertrieb. Die Daten sind als (vereinfachte) Ausgangsdaten in bestimmten Zielmaßstäben für die weitere analytische und visuelle Verwendung durch Dritte geeignet. Ein konkreter Beispielfall wäre die Lieferung entsprechender Daten für die Bundesanstalt für Geowissenschaften und Rohstoffe, die als übergeordnete Fachbehörde auf nationaler Ebene auch für die Erstellung länderübergreifender fachspezifischer (geologischer/bodenkundlicher) Übersichtskarten zuständig ist. Diese Anforderungsstufe kann als Maximalziel für die vorliegende Arbeit definiert werden.
- E) Die Erfüllung dieser Anforderungsstufe wäre in der LGRB-Praxis ein fachlich-topologisches Konstrukt, welches die bestehenden Abhängigkeiten aller GeoLa-Fachdaten zueinander konsistent durch alle Zielmaßstäbe hindurch gewährleistet. Dies würde in etwa den schon in den nationalen Kartografiebehörden vorhandenen gesamtheitlichen agenten-basierten Generalisierungslösungen topografischer Datenbestände entsprechen. Aufgrund der zum Teil unterschiedlichen fachspezifischen semantischen Ausgangssituation (vor allem zwischen Bodenkunde und Geologie/Hydrogeologie) ist aber die praktische Umsetzung - wenn überhaupt - nur schwierig möglich.

---

<sup>2</sup> LGRBWissen ist ein als datenbankgestütztes Informationssystem konzipiertes Internetportal, in welchem das LGRB den Abruf seiner gesammelten Sachinformationen zu den GeoLa-Fachthemen und Datenprodukten ermöglichen möchte. Eine Kernkomponente ist dabei die integrierte räumliche Suche sowie die Übersichtsdarstellung der jeweils zugehörigen Geodaten in einer Kartenkomponente. Die Freischaltung einer ersten Version ist für Sommer 2019 geplant.

## 1.5 Aufbau der Masterarbeit

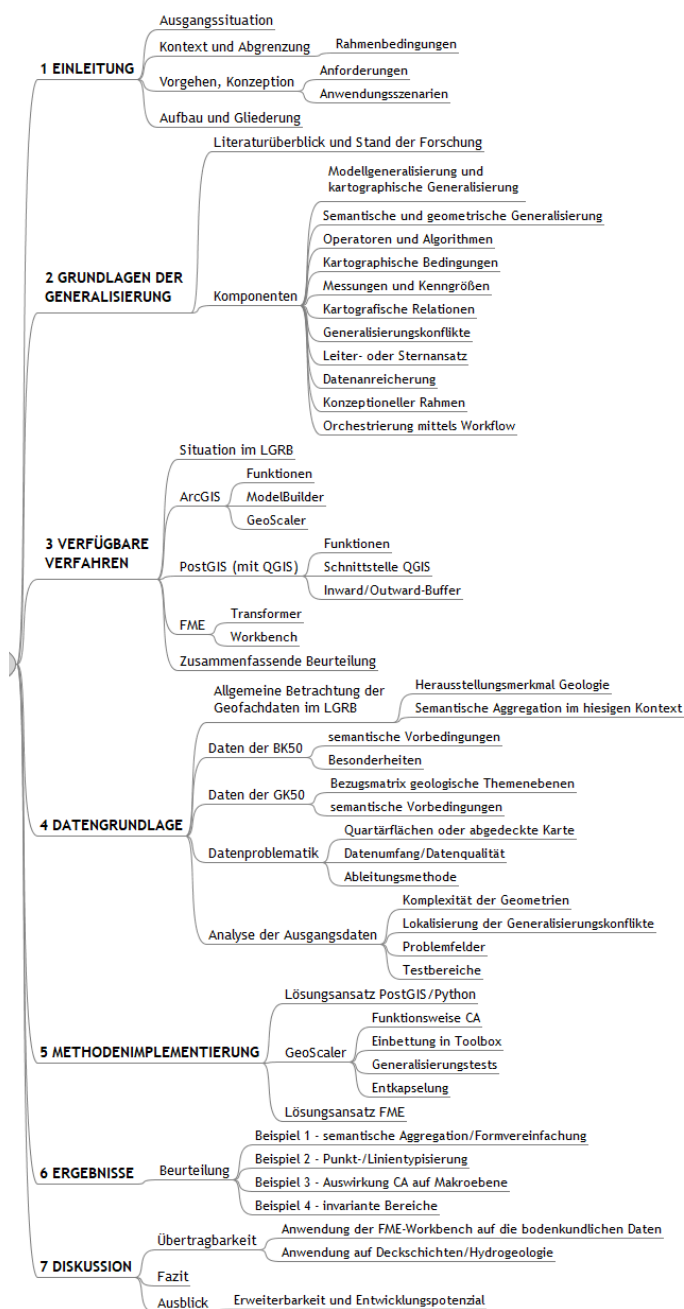


Abb. 2: Überblick Gliederung der Masterarbeit

wältigung dieser Konflikte anhand einer bedingungs-basierten Ablaufsteuerung und dessen konkrete praktische Umsetzung steht im Mittelpunkt des Kapitel 5 zur Methodenimplementierung.

Im Kapitel 6 werden die Ergebnisse zu den vorherig skizzierten Problemstellungen und Konfliktszenarien in Bildreihen gegenübergestellt und erläutert. In der abschließenden Diskussion des Kapitel 7 folgen Aussagen zur themenbereichsübergreifenden Übertragbarkeit des Lösungsansatzes an konkreten Beispielen sowie Fazit und Ausblick.

Diesem einführenden Kapitel 1 folgen insgesamt sechs weitere Hauptgliederungspunkte, die sich mit den allgemeinen theoretischen Aspekten, ihrem praktischen Bezug, den verfügbaren Methoden und schließlich deren konkreter Anwendbarkeit und Umsetzung befassen.

In Kapitel 2 werden die wesentlichen Aspekte der kartografischen Generalisierungsautomatisierung im Allgemeinen beleuchtet. Dabei werden die für den Praxisteil der Arbeit relevanten Komponenten etwas detaillierter beschrieben bzw. auf deren spätere Anwendung hingewiesen. Die den Automatisierungsablauf steuernden Funktionen und Algorithmen werden an Hand ihrer Verfügbarkeit in marktüblicher bzw. im Durchführungsumfeld der Masterarbeit vorhandener Software (ArcGIS, PostGIS, FME) vorgestellt und ihre Eignung im Kontext der Arbeit zusammenfassend gegenübergestellt (Kapitel 3).

In Kapitel 4 werden zuerst die dem Vorhaben zu Grunde liegenden Daten einer genaueren Untersuchung unterzogen und beschrieben. Dies beinhaltet auch zu einem umfangreicheren Teil die Beschäftigung mit den semantischen Vorgaben, die für die Durchführung des eigentlichen Generalisierungsablaufs das wesentliche Rahmenkonstrukt liefert. Es folgt eine Analyse der maßgeblichen Generalisierungssituationen und -konflikte, die sich im Zusammenhang mit der vorliegenden spezifischen semantischen und geometrischen Datensituation ergeben. Die Auseinandersetzung mit dem methodischen Konstrukt zur Be-

## 2 Grundlagen der automatisierten Generalisierung

Die Kartografie als eines der Hauptfächer der Geowissenschaften (Wikipedia 2019) hat in den letzten 25 Jahren einen zum Teil dramatischen Wandel erlebt - von einer durch konventionelle Techniken der Kartenerstellung und Reproduktionsgrafik geprägten, sehr eigenständigen Wissenschaft hin zu einem durch sich ständig verändernde und weiterentwickelnde digitale Technologien enorm dynamischen Spannungsfeld. In dieser neuen Welt der „Web-Services“, „Map-Mashups“ und „Volunteered Geographic Information“ sucht die Kartografie nach einem substanziellen zeitgemäßen Selbstverständnis. Eine Karte soll in erster Linie effizient raumbezogene Informationen kommunizieren, was auch die Veränderung von Geometrie und Semantik der zugrundeliegenden Daten beinhaltet (Gartner & Schmidt 2010). Der Vorgang der Generalisierung war und ist nach wie vor eine der Hauptaufgaben der klassischen Kartografie. Die Übertragung kartografischen Wissens im Zusammenhang mit der Generalisierungsautomatisation stellt dabei eine nicht einfach zu lösende Herausforderung dar. Hier manifestiert sich auch in besonderer Form eine grundsätzliche Anforderung, mit der sich die Kartografie im digitalen Zeitalter schwer tut: die durch den handelnden (Red)Akteur individuell und subjektiv gelagerten Gedanken- und Arbeitsprozesse in ein fundiertes formalisiertes Regelkonstrukt zu gießen, um die Automatisierung eines Prozesses mittels Soft- und Hardware nachbilden zu können.

### 2.1 Überblick zu Literatur und dem Stand der Forschung

Eine wesentliche Motivation zur Generalisierungsautomatisation in der digitalen Kartografie liegt darin, entsprechende kartografische Erzeugnisse schneller, kostengünstiger und flexibler erstellen zu können, als dies durch herkömmliche Techniken möglich ist. Im Bereich der nationalen topografischen Kartenwerke gab es deshalb schon frühzeitig die Bestrebungen nach einer automatisierten Ableitung der Daten aus einem Basismodell in ein Modell geringerer Auflösung. Darüber hinaus wäre eine automatische Generalisierung der Datenbestände einer der Schlüssel, um geografische Informationen einmalig zu sammeln und zu pflegen und sie nach Bedarf und Zweck und so oft als nötig zu verwenden, um Daten/Karten mit dem erforderlichen Inhalt und Detaillierungsgrad dynamisch abzuleiten (Stoter et al. 2016). Eine Vollautomatisierung des Generalisierungsprozesses in seiner Gesamtheit ist allerdings nur sehr schwierig umzusetzen. Dem entsprechend wurden lange Zeit nur Teilabläufe stringent als vollautomatisierte Schritte beschrieben. Es gibt eine Reihe von Studien, die ihre Aufmerksamkeit vor allem der Entwicklung und Untersuchung von Teilprozessen der Generalisierung steuernder Operatoren und Algorithmen widmeten, wobei den Methoden zur Liniensimplifizierung die wohl größte Aufmerksamkeit entgegengebracht wurde - beispielhaft genannt seien hier an vorderster Stelle Arbeiten von Saalfeld (1999) zum Douglas-Peucker-Algorithmus, wie auch Corcoran, Mooney & Winstanley (2011), Haurert & Wolff (2010), Tong et al. (2015) sowie Ai et al. (2017).

Vor allem in Europa sind in den nationalen kartografischen Instituten - kurz NMA (National Mapping Agencies) - die Arbeiten zur Implementierung ganzheitlicher Generalisierungsautomatisation in ihren jeweiligen Geodateninfrastrukturen relativ weit fortgeschritten (Stoter et al. 2016). Voraus ging dieser Entwicklung ein nahezu zwei jahrzehntelanges Zusammenspiel von behördlichen und wissenschaftlichen Gruppen, die in enger Abstimmung mit Softwareanbietern sukzessive Anforderungen und Bedürfnisse für einen vollautomatisierten Produktionsablauf definierten und verfeinerten (Duchêne et al. 2014, 331). Wurden noch zu Beginn des 21. Jahrhunderts in einigen NMA nur Teilschritte automatisiert, so konnten schon ein Jahrzehnt später die ersten durchgehend automatisiert generalisierten topografischen Kartenwerke vorgestellt werden, exemplarisch genannt seien Stoter et al. (2014a), Lafay et al. (2015) sowie Duchêne et al. (2014). Dieser mittlerweile fast



unüberschaubaren Anzahl von Beiträgen im Bereich topografischer Karten und Datenbestände steht ein eher dünn gesätes Feld an Vorhaben gegenüber, welche sich konkret mit der Generalisierungsautomatisierung von Verbreitungsdaten qualitativer thematischer Karten befassen. Als vordergründig behandelte Themenbereiche der Generalisierungsproblematik im Kontext von kategorischen Flächendaten wären dabei Landnutzung, Bodenkarte und Geologische Karte zu nennen - alle drei beschäftigen sich mit dem gleichen Ausschnittsvolumen der sich an der Erdoberfläche abbildenden jeweiligen Kategorien. Im Zusammenhang mit der Landnutzung stehen allerdings Fernerkundungsdaten und dementsprechend rasterbasierte Methoden zur Analyse und Verarbeitung der Grundlagen im Mittelpunkt (Thiemann, Sester & Bobrich 2010), während bei Bodenkarte (Fuchs 2002) und geologischer Karte (Sayidov & Weibel 2016) vor allem semantische Aspekte wie auch spezifische themenrelationale Zusammenhänge im Vordergrund stehen.

Einige der wegweisenden Beiträge im Umfeld der Generalisierungsautomatisierung thematischer bzw. kategorischer Polygonkarten sowie anderweitig relevante frühere Arbeiten liefert die Zeitleiste in Abb. 3.

Mit als Erste widmeten sich gegen Ende des 20. Jahrhunderts Bader & Weibel (1997) der Beschreibung von Konfliktsituationen bei der Polygongeneralisierung, während Peter (2001) eine vollständige Auflistung von wichtigen Mess- und Kenngrößen erarbeitete. Eine Grundlagenarbeit zu kartografischen Relationen und Strukturen in kategorischen Karten lieferten Steiniger & Weibel (2005b), im gleichen Jahr stellten die beiden Autoren auch ein konzeptionelles Framework für die Generalisierung von thematischen Karten vor (Steiniger & Weibel 2005a). Vorher hatte Galanda (2003) mit seiner Dissertation über den Einsatz von agentenbasierten Generalisierungssystemen einen maßgeblichen Einfluss auf die nachfolgende Entwicklung in den NMA. Bobzien (2005) erweiterte in einer der wenigen deutschsprachigen Arbeiten das theoretische Fundament zur Kenntnis über das Wesen der Generalisierung und Harrie & Weibel (2007) lieferten eine erste Bestandsaufnahme hinsichtlich der Modellierung des Gesamtprozesses bei der Generalisierungsautomatisierung.

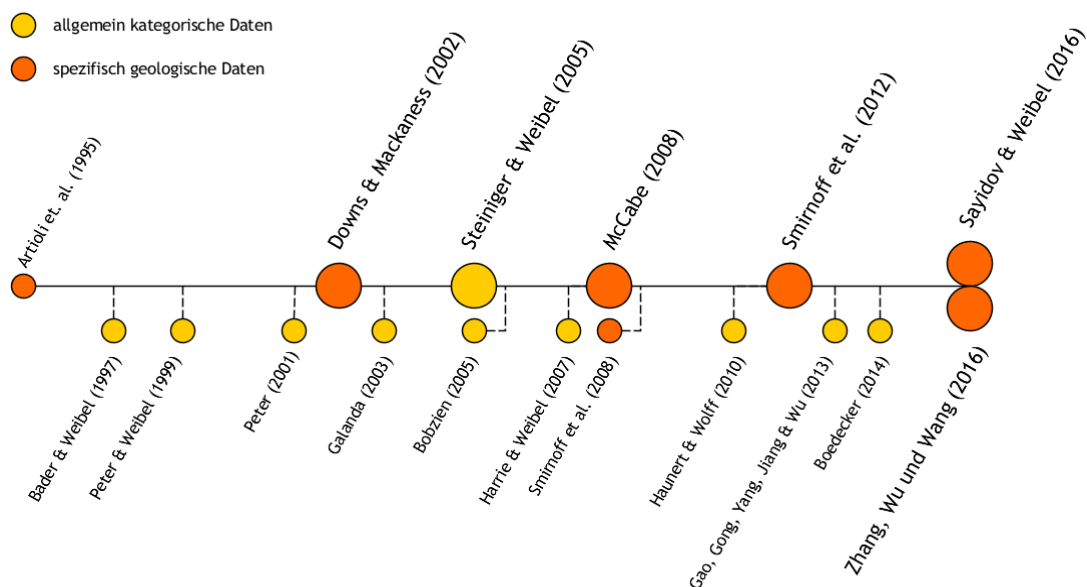


Abb. 3: Zeitleiste mit wissenschaftlichen Beiträgen zur Generalisierungsautomatisierung geologischer Karten

Nur wenige Autoren haben sich im speziellen und explizit mit den theoretischen und praktischen Problemen bei der Generalisierungsautomatisierung von geologischen Karten beschäftigt. Eine erste praxisnahe Studie hierzu erstellten Artioli et al. (1995). Etwas später stellten Downs & Mackaness (2002) einen kontextuell sen-

sitiven, regelbasierten teilautomatisierten Prozessablauf zur Generalisierung eines geologischen Ausgangsdatensatzes im Maßstab 1 : 50 000 nach 1 : 250 000 vor.

McCabe (2008) entwickelte und beschrieb einen Generalisierungs-Workflow, der basierend auf der Anwendung einer Reihe von ArcGIS-Tools eine geologische Karte im Gebiet Santa Rosa in drei aufeinanderfolgenden Maßstabsstufen (1 : 50000, 1 : 100000 und 1 : 250000) erstellte. Eine wesentliche Folgerung aus seiner Arbeit besteht in der Annahme, dass gerade bei geologischen Karten zur Verbesserung des Generalisierungsergebnisses zusätzlicher Aufwand in eine vorab durchzuführende Datenanreicherung angebracht ist. Ebenfalls 2008 wurde erstmals von Smirnoff, Paradis & Boivin (2008) die rasterbasierte Methode der zellulären Automaten im Umfeld der Generalisierung thematischer Karten angewandt und beschrieben, was zur Entwicklung einer umfangreichen ArcGIS-Toolbox führte (Smirnoff et al. 2012). Auch Sayidov & Weibel (2016) befürworteten einen bedingungsbasierten Ansatz, der durch die Integration von vektor- und rasterbasierten Generalisierungsmethoden implementiert werden könnte. Zhang, Wu & Wang (2016) wiederum propagieren ein agentenbasiertes Rahmenkonzept unter Einbezug der semantisch-hierarchischen Belange von geologischen Karten. Bei einer möglichst praxisnah anzustrebenden Umsetzungslösung zur automatisierten Generalisierung themenspezifischer Flächendaten stellt sich letztlich auch die Frage hinsichtlich der aktuell in GI-Software bzw. räumlichen RDBMS implementierten Generalisierungsfunktionen. Zurückgegriffen werden kann hier auf einschlägige Software-Dokumentationen von FME (SAFE Software 2018), ArcGIS (ESRI 2018c) sowie PostGIS (PostGIS Project Steering Committee 2018). Die Untersuchungen von Stoter et al. (2009) und Podolskaya et al. (2009) liefern methodische Ansätze zur Evaluierung von Generalisierungsergebnissen.

## 2.2 Wesentliche Komponenten und Aspekte der Generalisierung

### 2.2.1 Modellgeneralisierung und kartografische Generalisierung

Im übergeordneten Sinn versteht man unter dem Begriff Generalisierung die Verallgemeinerung, Vereinfachung oder Abstraktion eines Sachverhaltes. Im Umfeld der Geowissenschaften wird dieses Prinzip auf die modellhafte Übertragung von Objekten oder räumlichen Phänomenen der realen Welt in ihre Repräsentation in Datenmodellen bzw. deren Visualisierung an Hand von Karten angewendet. Dieser Vorgang findet im Wesentlichen mittels Auswahl, Zusammenfassung, Klassifizierung oder Betonung von Information unter Auslassung des Unwichtigen statt. Ausgangspunkt für die Übertragung der Wirklichkeit in ein geografisches Modell ist die sogenannte Objektgeneralisierung oder Erfassungsgeneralisierung, die gestützt auf Methoden der Primärdatenerfassung zu einer je nach Zweck mehr oder weniger stark reduzierten Informationsübernahme in ein Basismodell bzw. eine Grundkarte führt. Ausgeführt wird dies in der Regel im jeweiligen Fachbereich, in welchem das beabsichtigte Modell angesiedelt ist (topografisches Basismodell, geologisches Basismodell, etc. - dementsprechend auch als topografisches oder thematisches Generalisieren bezeichnet) (Chen et al. 2014). Die beiden Begriffe Basismodell und Grundkarte stehen dabei synonym für die beiden im Bereich der Kartografie miteinander/ nebeneinander existierenden Herstellungs- und Darstellungsebenen analoger und digitaler Ausprägung. Die so entstandenen Grundlagen bedürfen für die abgeleitete Verwendung bzw. Darstellung in kleineren Maßstäben einer weiteren Vereinfachung oder Auflösungsreduktion - im Falle des Basismodells als Modellgeneralisierung, bezogen auf die Darstellung wiederum als kartografische Generalisierung bezeichnet. Erstere unterliegt eher statistischen oder mathematischen Überlegungen, letztere stellt den grafischen Aspekt der Lesbarkeit eindeutig in den Vordergrund. Generalisierung bezieht sich also sowohl (in einem ersten Schritt) auf die Objekte und ihre Beziehungen, als auch (im darauf folgenden Schritt) auf deren kartografische Darstellung. Die Modellgeneralisierung als Bearbeitung von Objektmodellen ist mit der Erfassungsgene-

ralisierung vergleichbar, aber mit dem Unterschied, dass dieser Vorgang sich nicht auf das Objekt selbst bezieht, sondern auf das Objektmodell, aus dem ein neues Objektmodell geringerer semantischer und geometrischer Auflösung abgeleitet werden soll (Hake, Grünreich & Meng 2002, 168).

Die kartografische Generalisierung führt zu digitalen kartografischen Modellen oder Folgekarten und ist weitestgehend eine Aufgabe des Kartographen. Sie beruht beim digitalen kartografischen Modell auf den semantischen Bewertungen und graphikbedingten Einschränkungen, denen das zugrundeliegende Objektmodell zu unterziehen ist. Beim Standardfall der klassischen Generalisierung entsteht die Folgekarte unmittelbar aus einer anderen Karte größeren Maßstabs (Folgekarten-Prinzip) (Hake, Grünreich & Meng 2002, 168f).

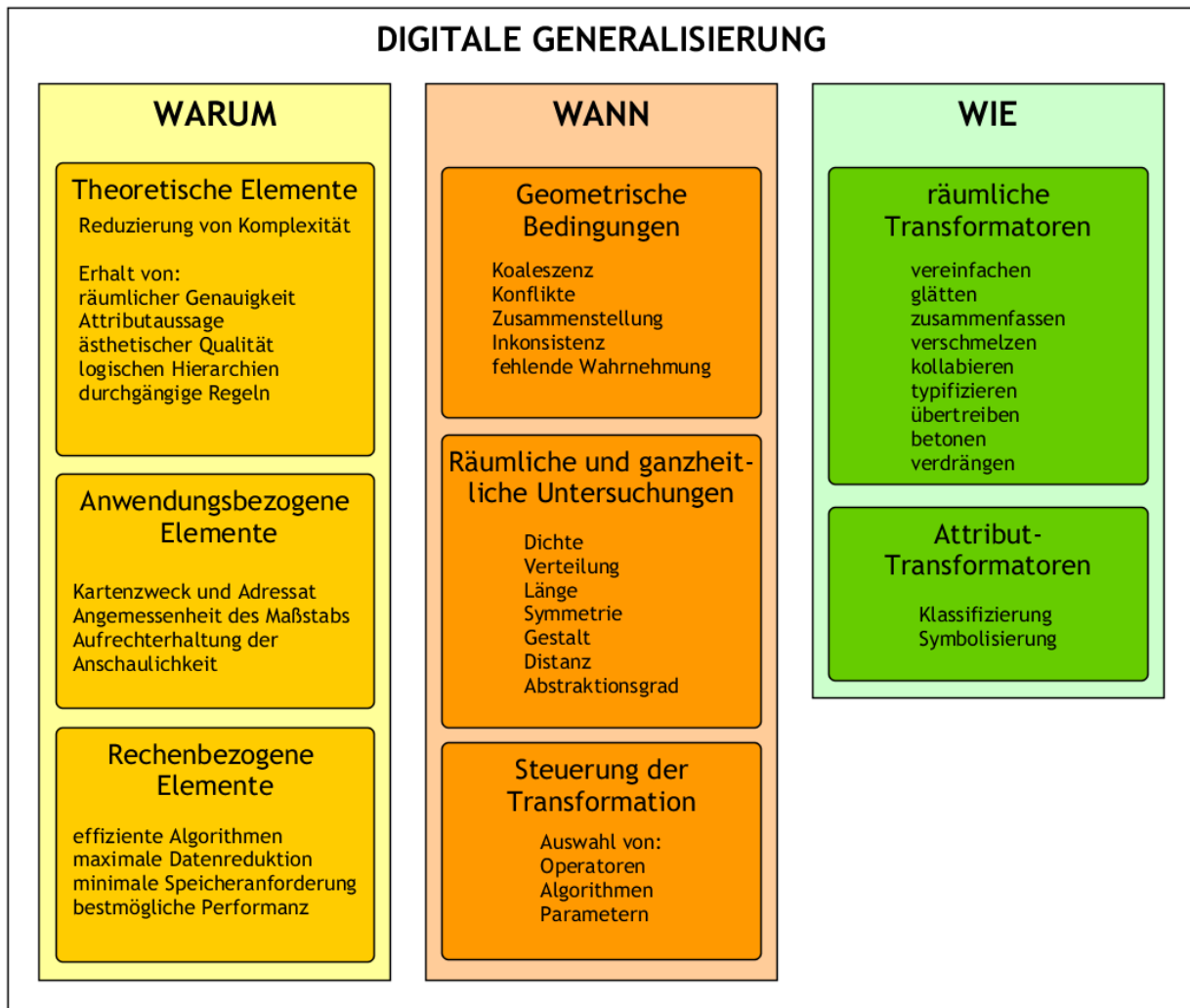


Abb. 4: Konzeptioneller Rahmen der digitalen Generalisierung nach McMaster & Shea (1992)

### 2.2.2 Semantische und geometrische Generalisierung

Strukturell lässt sich die kartografische Generalisierung in die beiden Teilvorgänge der semantischen Generalisierung und der sich normalerweise daran anschließenden geometrischen Generalisierung aufteilen. Die semantische (sachbezogene) Generalisierung ist die kartografiespezifische Bezeichnung für das methodische Verfahren der begrifflichen und sprachlichen Verallgemeinerung. Das begriffliche Generalisieren erfordert eine Klassifikation bzw. eine Begriffshierarchie des Inhalts der zu bearbeitenden Grundlagendaten, die auf ihre Verwendung, z. B. in der Wissenschaft, der Verwaltung, der Wirtschaft oder in anderen Bereichen ausge-

richtet ist. Häufig existiert bereits eine entsprechende Klassifizierung, z. B. als Zeichenschlüssel topografischer Karten, in Gestalt einer geologischen Formationstabelle oder der Bodenklassifikation (Bollmann & Koch 2001).

Die geometrische (raumbezogene) Generalisierung umfasst alle elementaren Generalisierungsvorgänge, von denen das Vereinfachen (Simplifizierung, Glättung), das Vergrößern und die Verdrängung zur geometrischen Generalisierung im engeren Sinne gezählt werden. Das Auswählen (Objektauswahl), z. B. von Flüssen ab einer festgelegten Mindestlänge, das Klassifizieren bzw. Typisieren (Geometrieübergang) und das Bewerten (z. B. Betonen) werden als Generalisierungsvorgänge von begrifflich-geometrischem Charakter betrachtet. Der Generalisierungsvorgang der Zusammenfassung kann rein geometrischer Natur sein, z. B. die Vereinigung einzelner, dicht beieinander liegender Waldflächen zu einer Fläche betreffen, oder aber aus einer begrifflichen Zusammenfassung resultieren (z. B. wenn Einzelhaussignaturen zu einem Bebauungsblock integriert werden). Hierbei bedingen sich geometrische und begriffliche Zusammenfassung wechselseitig (Bollmann & Koch 2001).

Abb. 4 liefert nach McMaster & Shea (1992) eine grafisch-strukturierte Zusammenschau mit den bestimmenden Aspekten bezüglich des Warum (Gründe und Zielsetzung), dem Wann (Anwendungsfälle) sowie dem Wie (Verfahren) bei der digitalen Generalisierung.

### 2.2.3 Operatoren und Algorithmen

Die Benennung und Definition von Operatoren entstand in der frühen Generalisierungsforschung, indem abstrakte Beschreibungen einzelner Aktionen des Kartographen während der manuellen Generalisierung extrahiert wurden. Sie erlauben die Dekonstruktion des Gesamtprozesses in individuelle, handhabbare Teilprozesse (Harrie & Weibel 2007). Ein Generalisierungsoperator ist also eine abstrakte Beschreibung einer atomaren Generalisierungsfunktionalität, der als Schlüssel eingesetzt werden kann, um unterschiedliche Generalisierungsalgorithmen zu vergleichen und zu klassifizieren. Ein Operator wird dabei durch verschiedene Algorithmen implementiert. Generalisierungsalgorithmen sind formalisierte mathematische Konstrukte zur Lösung eines Generalisierungsproblems und dienen der konkreten Umsetzung und Anwendung eines Generalisierungsvorgangs durch Objekt- oder Attributtransformation (Foerster, Stoter & Köbben 2007, 3). Als prominentes Beispiel für die häufige Verwendung in der GIS-Praxis sei hier stellvertretend der Douglas-Peucker-Algorithmus (Saalfeld 1999) zur Linienvereinfachung angeführt.

McMaster & Shea (1992) unterscheiden für die digitale Kartografie 10 räumliche Transformationen: vereinfachen, glätten, zusammenfassen, verschmelzen, kombinieren, kollabieren, verfeinern, betonen, übertreiben und verdrängen. Zusätzlich werden die beiden attributiven Transformationen Klassifizieren und Symbolisieren genannt, welche Änderungen auf semantischer Ebene bewirken. Über die darüber hinaus bekannten Typisierungen und Benennungen von Operatoren besteht kein allgemeiner Konsens in der Kartografie.

Die später durch Foerster, Stoter & Köbben (2007) vorgenommene formalisierte Klassifizierung der Generalisierungsoperatoren strukturiert dann auch hinsichtlich Operatoren für die Modellgeneralisierung bei globaler Anwendung an einem Datensatz sowie Operatoren der kartografischen Generalisierung bei individueller objektbezogener Anwendung:

Modellgeneralisierung	kartografische Generalisierung
(Klassen)Selektion (Class Selection)	Betonung (Enhancement)
Klassifizierung (Reclassification)	Verdrängung (Displacement)
Kollaps (Collapse)	Eliminierung (Elimination)
Zusammenfassung (Aggregation)	Typifizierung <sup>3</sup> (Typification)
Vereinfachung (Simplification)	
Verschmelzung (Amalgamation)	

Tab. 1: Zuordnung der Operatoren nach Foerster, Stoter & Köbben (2007)

Abhängigkeiten zwischen den einzelnen Generalisierungsschritten verlangen Analysen der aktuellen Situation sowie die Entwicklung/Anwendung einer auf den jeweiligen Fall abgestimmten Strategie mit der Auswahl der korrekten Operatoren und ihrer Anwendung in der richtigen Reihenfolge. Einzelne in GIS verfügbare Operatoren bzw. in ihrer Form als einschlägige Funktionen werden im Kapitel 3.5 unter dem besonderen Kontext der Generalisierung von kategorischen Flächendaten aufgeführt und untersucht.

#### 2.2.4 Kartografische Bedingungen

Die Anwendung dieser elementaren Vorgänge führt zu zwei charakteristischen Arbeitsweisen: dem intuitiv geprägten Vorgehen des Kartographen oder dem rechnergestützten Generalisieren basierend auf Regeln und Vorschriften. Dieses ursprünglich regelbasierte Generalisieren entwickelte sich zu einem bedingungs-basierenden Ansatz, in welchem nun der Fokus nicht mehr darauf gelegt wird, wie man die Generalisierung durchführt, sondern darauf, was die Generalisierung erreichen soll (Ruas & Plazanet 1996). Der Unterschied besteht also darin, dass eine Einschränkung oder Bedingung nicht an eine bestimmte Aktion gebunden ist. Alternativ kann nach (Weibel & Dutton 1999) eine Einschränkung als Entwurfsspezifikation betrachtet werden, an die sich ein Objekt, eine Gruppe von Objekten oder eine Karte halten sollte. Es kann unterschieden werden zwischen strikten (hard) und weichen (soft) Bedingungen, wobei bei letzteren auch leichte Verstöße gegen die Erfüllung des Formalismus toleriert werden. Bedingungen werden durch Mess- und Kennwerte beschrieben (s. 2.2.5), der Grad der Verletzung einer Bedingung kann danach bewertet werden, inwieweit der Ist- von einem vorgegebenen Sollwert abweicht (Steiniger & Weibel 2007).

Da sich die für ein Objekt definierten Einschränkungen oft widersprechen, ist eine Generalisierungslösung immer ein Kompromiss zwischen verschiedenen Einschränkungen und ihrem Erfüllungsgrad. Zum Beispiel ist es unmöglich, eine Formverzerrung eines Polygons zu vermeiden, wenn der Polygonumriss wegen zu vieler Details im Zielmaßstab unleserlich ist und daher eine Vereinfachung erfordert (Galanda 2003). Im Zusammenhang mit der Generalisierung weisen kartografische Bedingungen eine der drei folgenden Funktionen auf (Sayidov & Weibel 2016):

- **Konflikterkennung:** Identifizierung von Bereichen, die generalisiert werden müssen
- **Konfliktlösung:** Auswahl geeigneter Operatoren und Algorithmen für die Beseitigung erkannter Konflikte

3 Die Typifizierung bezeichnet das Ersetzen einer Menge gleichartiger Objekte durch eine Teilmenge (Sester 2000) und ist nicht zu verwechseln mit der Typisierung. In dieser Arbeit wird der Terminus „typisieren“ synonym für den Generalisierungsoperator des „Kollaps“ verwendet. Dieser bezieht sich explizit auf den Geometrietypübergang Polygon/Linie oder Polygon/Punkt.

- Evaluierung: Kontrolle der Anwendungsauswirkung von Algorithmen auf den Erfüllungsgrad einer Konfliktlösung

Kartografische Bedingungen bzw. Einschränkungen lassen sich nach Petzold, Burghardt & Bobzien (2006) den vier folgenden Gruppen zuordnen:

- Grafische Einschränkungen: Limitierungen in der Wahrnehmung von Kartenobjekten werden unter Berücksichtigung des Kartenzwecks, Maßstabs und Ausgabemediums in Mindestgrößen übersetzt.
- Topologische Einschränkungen: sie stellen den Erhalt oder zumindest die konsistente Modifikation von topologischen Beziehungen sicher.
- Strukturelle Einschränkungen: beinhalten Kriterien zur Wahrung der räumlichen und semantischen Struktur von Daten. Erstere in Bezug auf typische Formen und Muster von Objekten und Objektgruppen, Letztere hinsichtlich ihrer logischen Beziehungen bei der Aggregation.
- Prozedurale Einschränkungen: beziehen sich auf den Generalisierungsprozess

### 2.2.5 Messungen und Kenngrößen

Messungen sind ein Schlüsselement im Zusammenhang mit der Formalisierung von Einschränkungen und Bedingungen bei Generalisierungsprozessen. Sie werden zur Konflikterkennung benötigt, um Transformationsoperationen auszuwählen und zu steuern und letztlich auch die Ergebnisse hinsichtlich Qualität und Plausibilität zu bewerten. Dabei unterscheidet Peter (2001, 4) zwischen mandatorischen und optionalen Messungen, wobei letztere nur bei bestimmten Datentypen und in bestimmten Situationen Anwendung finden sollten, da sie die Performanz verringern bzw. unangemessene oder irreführende Kenngrößen sich nachteilig auf den Generalisierungsablauf auswirken können. Tabelle 2 liefert eine Kategorisierung von Messungen sowie ausgesuchte Beispiele nach Peter (2001).

Tab. 2: Kategorisierung von Mess- und Kennwerten (nach Peter 2001)

Typ	Beschreibung	Anwendung
Größenmaße	finden Anwendung als eigenständige Indikatoren wie auch als Bestandteil komplexerer Messungen bei kategorischen Daten; kommen auf allen räumlichen Ebenen (Mikro, Meso, Makro s. 2.2.6) zum Einsatz	Flächengröße, Perimeter, relative Flächengröße, Anzahl/Flächengröße, Anzahl Polygone unter Mindestgröße einer Kategorie, Länge der gemeinsamen Grenze mit Nachbar-Polygon, gewichtete relative Länge, Differenz zur minimalen Flächengröße, Flächenkoeffizient
Distanzmaße	werden verwendet, um Konflikte innerhalb und zwischen Polygonen zu identifizieren und zu lösen	Minimalentfernung, Vertexabstand
Formmaße	dienen auf Mikroebene hauptsächlich der Identifizierung charakteristischer Formen einzelner Objekte; auf Makroebene zur Berechnung von Formindikatoren ganzer Kategorien (für die Evaluierung des Generalisierungsergebnisses)	Zero Core Area, Shape Index, Gruppenformindex, Anzahl entfernungsbezogener Konflikte
Topologismaße	dokumentieren die topologischen Beziehungen von Objekten auf Meso-Ebene; definieren Validität von Objekten; kein Messwert im eigentlichen Sinn	Nachbarschaft, Inselepolygon, ungewollte Nachbarschaft, Selbstüberschneidung
Dichte- und Verteilungsmaße Muster- und Ausrichtungsmaße	beziehen sich auf räumliche Strukturen höherer Ordnung und beschreiben Eigenschaften der Objekte auf Makroebene	Relative Fläche einer Kategorie, relative Flächengröße einer Kategorie anhand ihrer konvexen Hüllen; Cluster (Verdichtungsraum) Muster, Ausrichtung von Polygonen
Semantikmaß	kennzeichnet ein Objekt auf Grundlage seiner semantischen Bedeutung	Flag

Laut Stoter et al. (2014b) wird die derzeit noch umfassendste Klassifizierung von Messungen zur Qualitätsbewertung von Mackaness & Ruas (2007) vorgeschlagen und stellt zwei Aspekte in den Vordergrund:

1. eine Messung kann sich auf intern (Messung von Objekten gleicher Größe innerhalb eines Datensatzes) oder extern (Messung von Objekten zwischen zwei Maßstäben) beziehen;
2. eine Messung kann Mikro (Messung an einzelnen Objekten oder Teilen eines Objekts), Meso (Messung an Objektgruppen) oder Makro (Messung an allen Objekten) sein.

Von den in Tab. 2 aufgeführten Kenngrößen fanden mit Ausnahme der in den Kategorien Dichte-/Verteilungsmaße und Muster-/Ausrichtungsmaße genannten eine Mehrzahl konkrete Anwendung im eigenen entwickelten Verfahren (s. Kap. 4.5.3.1 c).

### 2.2.6 Räumliche oder maßstäbliche Bezugsebenen

Die vorhergehend hinsichtlich der Messungen genannten Bezugsebenen Mikro, Meso und Makro sind im Umfeld der automatisierten Generalisierung weit verbreitet und bezeichnen die drei unterschiedlichen geometrischen Abstraktionsebenen (Galanda 2003, 14):

- *Mikro* (Einzelobjektebene): Die Mikroebene wird an einzelnen geografischen Objekten wie einem Fluss oder einer einzelnen geologischen Flächeneinheit (Polygon) usw. fest gemacht. Es geht hierbei um die unabhängige Generalisierung von Objekten ohne Einbezug der Nachbarschaft.
- *Meso* (Gruppenobjektebene): Die Meso-Ebene ist einer Gruppe von Objekten gewidmet, z.B. adjunkte oder disjunkte Polygone unterschiedlicher/gleicher Kategorie und deren kontextueller Generalisierung. Ein Objekt auf der Mesoebene kann aus Objekten bestehen, die entweder zur Mikro- oder zur Mesoebene gehören.
- *Makro* (Kartenebene): Die Makroebene bezieht sich auf eine Population von Objekten, z.B. alle Polygone eines Datensatzes (gesamtes Polygonmosaik). Das Hauptaugenmerk besteht darin, die Generalisierung der Verteilung von Kategorien bzw. den kartografischen Gesamteindruck der Inhalte in der Karte zu steuern und zu kontrollieren.

Die Begrifflichkeiten mikro, meso, makro finden im weiteren Verlauf der Arbeit an verschiedenen Stellen Anwendung.

### 2.2.7 Evaluierung

Evaluierung ist der Prozess der Untersuchung und Überprüfung, ob die Eigenschaften resultierender Daten/Karten für eine bestimmte Aufgabe ausreichend sind. Der Zeitpunkt im Ablauf kann vor, während und nach der Generalisierung durchgeführt werden und ist durch die jeweils unterschiedlichen Zwecke Feinabstimmung, Kontrolle und (End)Beurteilung motiviert (Stoter et al. 2014b).

Es besteht allgemeiner Konsens darüber, dass die Beurteilung des Ergebnisses ein unverzichtbarer Bestandteil im ganzheitlichen Generalisierungsprozess ist. Die Bewertung hilft, Parameter zu optimieren und die besten Algorithmen oder die beste Reihenfolge von Algorithmen bei der Generalisierung auszuwählen (Stoter et al. 2014b). Die von Bard (2004) für die Evaluierung von Generalisierungsergebnissen vorgenommene Aufteilung unterscheidet die drei Kategorien



- **Bewertung der Bearbeitung**: mit Fokus auf die Erkennung von kartografischen Fehlern und Inkonsistenzen. Sie wird in der Endphase des Generalisierungsprozesses durchgeführt und dient bei Bedarf zur Unterstützung einer anschließenden Bearbeitung bzw. im gesamtheitlichen Automatisierungsprozess als Kontrollpunkt, um den Ablauf mit geänderten Parametern im erneuten Durchlauf empirisch zu optimieren.
- **Bewertung des Generalisierungsgrades**: versucht einen Wert zu finden, der die Gesamtqualität der generalisierten Daten widerspiegelt und verschiedene Generalisierungslösungen vergleicht, um die optimale Lösung zu ermitteln oder schlechte Generalisierungslösungen bei spezifischen Generalisierungsaufgaben zu identifizieren.
- **Beschreibende Bewertung**: liefert allgemeine Informationen über die an den Daten vorgenommenen Änderungen. Diese Informationen können verwendet werden, um die Qualitätsbeschreibung (z.B. Metadaten) des Endprodukts zu verbessern (z.B. was wurde entfernt oder hervorgehoben? Wie stark haben sich die Daten verändert?)

Grundsätzlich bestehen bei der Bewertung von Generalisierungsergebnissen zwei Ansätze: visuell-basiert sowie quantitativ. Die visuelle Bewertung ist subjektiv, qualitativ und zeitaufwendig, während die quantitative Bewertung tendenziell nur bestimmte Merkmale untersucht und daher auch nur für die Beurteilung spezifischer Anforderungen geeignet ist. Insofern ist es schwierig, hierfür durchgeführte Messungen oder Kennwerte anzupassen bzw. in einen ganzheitlichen automatisierten Generalisierungsprozess zu integrieren (Stoter et al. 2014b). Aufgrund der fachspezifischen Besonderheiten der GeoLa-Daten ist im vorliegenden Fall eine visuell-qualitative Bewertung zielführend, die Generalisierungsergebnisse werden den Experten aus den Fachbereichen vorgelegt - sie achten insbesondere darauf, dass der inhaltliche Kontext und die Aussagekraft erhalten bleibt.

### 2.2.8 Kartografische Relationen

Der geografische Kontext eines Objektes bestimmt im Wesentlichen den Ablauf seiner Transformation innerhalb der Generalisierung. „Damit ein Generalisierungssystem in Anlehnung an die Arbeitsweise eines Kartographen ebenfalls qualitativ hochwertige Generalisierungsergebnisse liefern kann, muss es den geografischen Kontext analysieren und die Wahl und Anwendung von Generalisierungsoperationen und Algorithmen dem Kontext entsprechend steuern.“ (Boedecker 2011, 18). Der Koordination solcher Relationen oder Beziehungen kommt also eine zentrale Bedeutung zu, da ansonsten entstehende kartografische Konflikte die Generalisierung erschweren bzw. das Generalisierungsergebnis verschlechtern. Durch die Ermittlung von Relationen können Daten charakterisiert, Konflikte ermittelt, Algorithmen ausgewählt und das Generalisierungsergebnis evaluiert werden.

Es wird unterschieden zwischen horizontalen, vertikalen sowie „Update“-Relationen. Horizontale Relationen bestehen innerhalb einer Maßstabsebene bzw. des „Level of Detail“ und repräsentieren gemeinsame strukturelle Eigenschaften. Vertikale Beziehungen erscheinen zwischen homologen Objekten und Objektgruppen über verschiedene Kartenmaßstäbe hinweg (Neun & Steiniger 2005, 1). Update-Relationen beschreiben zeitliche Veränderungen an den Objekten.

Die Zusammenhänge zwischen den drei Arten von Beziehungen erläutern (Bobzien et al. 2008, 3-4) bezogen auf die Generalisierung in einer „Multi-Representation-Database“<sup>4</sup> wie folgt: horizontale Beziehungen unter-

4 Eine Multi-Representation-Database (MRDB) ist eine räumliche Datenbank, welche die selben Objekte in unterschiedlichen Auflösungen und Genauigkeiten speichert und gleichzeitig miteinander verlinkt (Hampe, Sester & Harrie 2004). Dies ist bei den GeoLa-Geo-

stützen den automatisierten Generalisierungsprozess. Das Ergebnis dieses Prozesses wird in verschiedenen Auflösungen gespeichert, die durch eine vertikale Beziehung verbunden sind. Diese Beziehung wird für die automatische inkrementelle Aktualisierung benötigt, die durch Updates ausgelöst wird; sie wird durch eine Update-Relation zwischen verschiedenen zeitlichen Zuständen modelliert. Vertikale und Update-Relationen können in der Praxis auch durch entsprechende Attributierung in den Daten verankert werden (s. 2.2.11). Die drei zuvor (s. Abb. 5) als Kernbestandteile der Generalisierung definierten kartografischen Bedingungen, Messungen und Relationen bedingen sich dabei gegenseitig. Relationen werden dem Generalisierungsprozess als Einschränkung oder Bedingung auferlegt, anhand von Messungen werden diese Einschränkungen beschrieben, um die Entwicklung der Relationen und damit auch den Erfüllungsgrad der Bedingungen im weiteren Ablauf der Generalisierung zu überwachen (Steiniger & Weibel 2007).

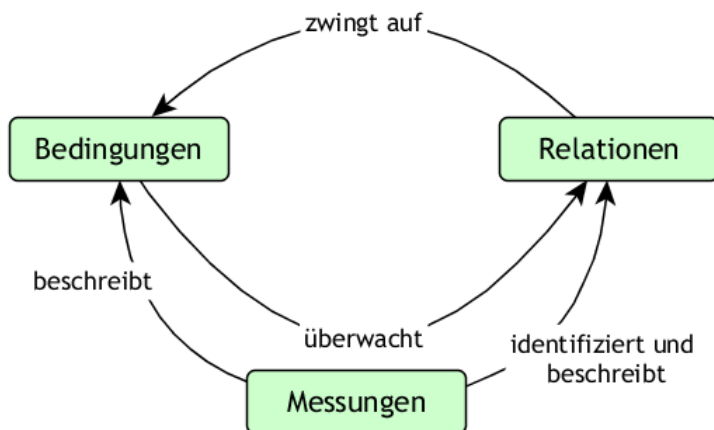


Abb. 5: Interaktionsmodell zwischen Bedingungen, Messungen und Relationen nach Steiniger & Weibel (2007)

Nach Steiniger & Weibel (2005b, 5) kann zwischen geometrischen, topologischen, statistischen, semantischen und strukturellen Relationen unterschieden werden - wobei die drei erstgenannten generischer, also allgemeiner Natur, die beiden letztgenannten spezifischer Art sind.

Kartografische Relation	Geometrische Domäne						Eigenschaft der Relation				
	PP	PL	PA	AA	AL	LL	h	v	l	r	g
<b>Geometrische Relation</b>											
<i>Minimaldistanz, disjunkt</i>	x	x	x	x	x	x	x		x		
<i>Minimaldistanz, überlagernd</i>			x	x	x		x		x		
<i>Minimalgröße</i>				x	x	x	x		x		
<i>Parallelität</i>				x	x	x	x		x		
<i>Schnittwinkel</i>				x	x	x	x		x		
<i>Relative Position, disjunkt</i>	x	x	x	x			x		x		
<i>Relative Position, überlagernd</i>				x	x		x		x		
<i>Formveränderung</i>				x	x	x		x			x
<b>Topologische Relation</b>											
<i>disjunkt</i>	x	x	x	x	x	x	x		x		
<i>berührt</i>		x	x	x	x	x	x		x		
<i>kreuzt</i>					x	x	x		x		
<i>innerhalb</i>		x	x	x	x	x	x		x		
<i>enthält</i>		x	x	x	x	x	x		x		
<i>überlappt</i>				x		x	x		x		
<i>gleich</i>		x		x		x	x		x		
<i>schneidet</i>	x	x	x	x	x	x	x		x		
<i>links-/rechtsseitig</i>		x			x	x	x		x		
<b>Statistische Relation</b>											
<i>Größen- und Mengenverhältnis</i>	x			x		x		x			x
<b>Semantische Relation</b>											
<i>Kantenpriorität</i>				x	x		x				x
<i>Barriere</i>				x	x		x				x
<i>Existenzabhängigkeit</i>	x	x	x	x	x	x	x				x
<i>Positionserhaltung</i>	x	x	x	x	x	x	x				x

P=Punkt, L=Linie, A=Polygon, h=horizontal, v=vertikal, l=lokal, r=regional, g=global

Tab. 3: Bezugsmatrix kartografischer Relationen nach Boedecker (2014)

In der Bezugsmatrix kartografischer Relationen aus Tab. 3 sind diejenigen Relationen farblich hervorgehoben, welche im Zusammenhang mit der Generalisierung kategorischer Flächendaten bzw. genauer der geologischen Daten eine konkrete Rolle spielen (s. Kap. 4.3.1). Die umrandeten geometrischen Relationen der Messwerte Minimalgröße und Minimaldistanz stehen dabei im Fokus der Betrachtung.

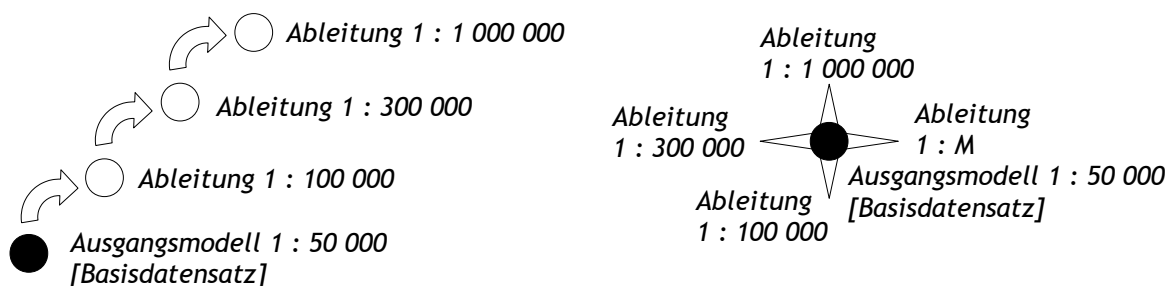
### 2.2.9 Generalisierungskonflikte

Generalisierungskonflikte stellen sich als Verletzung von Bedingungen und Einschränkungen der räumlichen und semantischen Eigenschaften von Objekten sowohl im Selbstbezug wie auch im Gruppenkontext dar. Die Erkennung dieser Konflikte ist ein integraler Bestandteil der Generalisierung und liefert wichtige Informationen, welche die spätere Konfliktlösung durch Generalisierungsoperatoren steuern. Laut Bader & Weibel (1997, 2) ist die Entwicklung geeigneter und robuster Modelle zur Quantifizierung und Identifizierung von Größen- und Nachbarschaftskonflikten in polygonalen Karten keine triviale Aufgabe. Als Beispiel genannt wird das scheinbar einfache Problem der Identifizierung zu kleiner Polygone: ein Polygon kann z.B. eine ausreichende Flächengröße haben, aber langgestreckt und schmal sein und zu einer weniger wichtigen Kategorie gehören und dadurch dennoch eliminiert werden. Oder umgekehrt: ein längliches Objekt stellt sich im geologischen Kontext als Singularität dar, wird aber aufgrund der Mindergröße aufgelöst. Größen- und Nachbarschaftsbeziehungen können also auf unterschiedliche Weise modelliert und gemessen bzw. interpretiert werden. In Kap. 4.5.2 werden solche und andere Fälle der in den vorliegenden GeoLa-Daten vorhandenen Situationen und Konflikte gelistet und beschrieben.

### 2.2.10 Leiter- oder Sternansatz

Die Vorgehensweise bei der automatisierten Ableitung von Daten aus einem Ausgangs- oder Basismodell kann auf unterschiedlichen Strategien beruhen. Es werden hierbei die beiden Methoden Leiteransatz und Sternansatz unterschieden. Beim Leiteransatz werden die Daten jeder Ableitungsstufe jeweils auf Basis der vorhergehenden Datensatzes generiert - quasi in Art einer inkrementellen Generalisierung, während beim Sternansatz alle Ableitungen unterschiedlicher Auflösung bzw. Maßstabs direkt aus den ursprünglichen Ausgangsdaten erzeugt werden. Nach Ross (2009) ist es möglich, dass ein Ansatz für manche Aufgaben besser geeignet ist als der andere. So kann beispielsweise der eine Ansatz schneller und insgesamt effizienter sein, während der andere bei gleichem Generalisierungsszenario Ergebnisse liefert, die optisch ansprechender oder genauer sind. Es ist auch möglich, dass ein Ansatz dem anderen einfach überlegen ist und in den meisten oder allen Fällen verwendet werden sollte. Eine dritte Möglichkeit ist, dass es zwischen den beiden Ansätzen nur einen geringen oder gar keinen erkennbaren Unterschied gibt.

Abb. 6: Schema des Leiter- und des Sternansatz



In der Studie von Ross (2009) werden in Bezug auf Liniensimplifizierungs-Operatoren (Douglas-Peucker und Bend-Simplify) Aussagen hinsichtlich kartografischer Ausgabe, Laufzeit (Effizienz), Datenkomplexität und Datenmenge getroffen, die aber nicht auf andere Operatoren gleichermaßen übertragbar sind. In jedem Fall ist die Bewertung der Ergebnisse abhängig von den Ausgangsdaten, dem beabsichtigten Generalisierungsziel und

den dabei verwendeten Algorithmen. Einer der Schlüsse, die man im Allgemeinen aber ziehen kann, betrifft die Performanz, die beim Leiteransatz mit jeder Stufe zunimmt - so wie umgekehrt die Datenmenge abnimmt. Beim Sternansatz ist immer die gleich hohe Ausgangsdatenmenge zu prozessieren, was sich entsprechend auf die Laufzeit der Ableitungen auswirkt. Ein zu berücksichtigender Nachteil beim Leiteransatz besteht in der Gefahr, dass vorhandene Fehler weitervererbt werden.

Aus dem von Buttenfield, Stanislawski & Brewer (2011) durchgeführten Experiment zum Stern- und Leiteransatz kann nicht geschlossen werden, dass die beiden Strategien bei jedem Abbildungsmaßstab austauschbar wären. Bei größeren Maßstäben, bei denen sich kleinere Details stärker zu unterscheiden beginnen, könnte der Sternansatz für Referenz- oder topografische Karten bevorzugt werden, während der Leiteransatz für die Erstellung von thematischen oder Webkartendatenbanken besser geeignet ist. Im ersten Fall geht man davon aus, dass zusätzliche Daten überlagert werden, und im zweiten Fall wäre das Ziel ein kleineres Datenvolumen für einen schnelleren Abruf, wie er für die Darstellung im Web gefordert ist.

Es kann auch eine gemischte Strategie in Erwägung gezogen werden, d.h. Sternversionen auf zwei oder drei Schlüsselskalen zu erzeugen und kartografische Datenprodukte unter Verwendung des Leiteransatzes aus diesen Schlüsselskalen abzuleiten. Kap. 4.4.3 widmet sich der Auswahl der geeigneten Methode im Kontext der vorliegenden Ausgangssituation.

### 2.2.11 Datenanreicherung

Unter Datenanreicherung versteht man die Bereitstellung zusätzlicher Informationen sowie deren Nutzung für die Anwendungssteuerung von Operatoren in einem automatisierten Generalisierungsprozess. Dadurch sollen die komplexen Entscheidungsprozesse innerhalb des Prozessablaufs mit strukturellem Wissen über den räumlichen und semantischen Kontext der Objekte ergänzt werden (Neun 2007). Die Ausgangsdaten in einem Basismodell sollten vorab durch fachlich bezogene Zusatzinformationen angereichert werden, um damit nur schwierig abzubildende bzw. zu programmierende Entscheidungskonstrukte im Generalisierungsframework zu ersetzen. Dabei kann es sich um die zusätzliche Attributierung der Ausgangsdaten mit Mess- und Kennwerten oder sonstigen klassifizierenden Eigenschaften handeln. Dies bietet den Vorteil, dass solche datenspezifischen Eigenschaften wiederholt im Ablauf Verwendung finden können.

Im Fall der geologischen Grundlagendaten ist die Attributierung der Ausgangsdaten nicht ausreichend, um einige der komplexeren Operationen durchführen zu können. Vor allem in Bezug auf die Relation zwischen geologischen Flächeneinheiten und den tektonischen Lineationen ist eine nachträgliche Klassifizierung an den Störungen für die Übernahme bzw. Gewichtung in die nachfolgenden Abbildungsebenen/Hierarchiestufen zu integrieren<sup>5</sup>.

### 2.2.12 Konzeptioneller Rahmen

Bisher gibt es bei kategorischen Daten nur vereinzelte integrierte Ansätze um die einzelnen Komponenten zu kombinieren und als ganzheitliches Framework zu strukturieren (Downs & Mackaness 2002; Zhang, Wu & Wang 2016).

Basierend auf der Literatur können solche Verfahrensmodelle weitestgehend in prozessorientierte und objektorientierte Rahmenkonzepte unterteilt werden. Erstere versuchen die Strukturierung des gesamten Generalisierungsprozesses basierend auf der Charakterisierung von Struktur und Semantik der Objekte. Der Vorteil dabei ist in der klar präzisierten Struktur zu sehen, die stringente und einfacher zu realisierende Verfahrensabläufe ermöglicht. Die zweite Methode adressiert die verschiedenen Ebenen der Modell- bzw. Kartenob-

<sup>5</sup> Im vorliegenden Fall konnte die Datenanreicherung für die tektonischen Linien nur für ausgewählte Testgebiete (s. 4.5.4) durchgeführt werden. Eine landesweite Umsetzung ist aktuell in Bearbeitung.

jekte (Mikro, Meso, Makro). Auf jeder dieser Ebenen sind entsprechende Kenntnisse und Regeln erforderlich. Dabei ist es sehr schwierig, das fachliche und kartografische Wissen in eine digitale Formalisierung zu überführen. Der Mehrwert bzw. Vorteil kommt vor allem bei der Generalisierung mit fachspezifischer Zielsetzung basierend auf einer klar umrissenen hierarchischen Ordnung zum tragen (Zhang, Wu & Wang 2016, 6375). In sogenannten Expertensystemen wird das Verhalten menschlicher Experten, wann und wo eine bestimmte Aktion bei der Konfliktlösung angewandt wird, software-technisch nachgeahmt. Der Erfolg eines Expertensystems steht in direktem Zusammenhang mit dem darin enthaltenen Wissen. Je besser das Problem definiert ist (d.h. je mehr "Lehrbuchwissen" existiert), desto besser sind die Regeln, die abgeleitet werden können, und desto besser sind die Lösungen, die erreicht werden können (Galanda 2003). Expertensysteme sind allerdings nicht flexibel und dynamisch genug, um Konfliktlösungen eines höheren Anforderungslevel (namentlich Kategorie E aus Abb. 1) zu liefern. Erwähnenswert wäre hier der AGENT-Ansatz (Barrault et al. 2001), bei dem die Orchestrierung mit Hilfe von Meso-Agenten auf der Basis kartografischer Vorgaben erfolgt. Ein solches Verfahrenskonzept lässt sich allerdings nicht innerhalb der hier zur Verfügung stehenden Soft- und Hardwareumgebung umsetzen.

### 2.2.13 Orchestrierung des Gesamtprozesses mittels Workflow

Abschließend zum Erläuterungsteil über die wesentlichen Prinzipien, Terminologien und Bestandteile im Zusammenhang mit der kartografischen Generalisierung sei in diesem Abschnitt noch auf die Möglichkeiten zur automatisierten Auswahl, Verkettung und Parametrisierung dieser in den vorangegangenen Abschnitten aufgeführten Generalisierungskomponenten eingegangen. Der Beitrag von Petzold, Burghardt & Bobzien (2006, 3) widmet sich als einer der Ersten exklusiv der Frage der Orchestrierung der Einzelkomponenten in einem Generalisierungsworkflow. Als Grund für den Mangel an entsprechenden wissenschaftlichen Publikationen zu diesem Thema wird hier genannt, dass Experimente nur mit einem umfangreichen Satz von Generalisierungsfunktionalitäten durchgeführt werden können, die oft nicht auf einer Plattform verfügbar sind. Die Verwirklichung eines solchen methodisch-technischen Konstrukts in der Art eines Frameworks stellt in diesem Zusammenhang eine besondere Herausforderung dar.

Workflows erlauben die Definition des Ablaufs und die Orchestrierung von Prozessen anhand eines grafischen Editors mit der Visualisierung in Form eines Diagramms (Prozessmodellierung). Das ist vergleichbar mit einer Programmiersprache - wobei eine solche grafische Oberfläche es auch Anwendern ohne Programmierkenntnisse ermöglicht, Workflows zur Steuerung der Generalisierungsprozesse zu erstellen. Im Gegensatz zu Batch-Workflows werden z.B. bedingte Sprungbefehle, Schleifen und gleichzeitige Prozesse unterstützt. In diesem Sinne schließen Workflows die Lücke zwischen manueller und vollautomatischer Orchestrierung. Ein Workflow-Management-System besteht aus einer Workflow-Engine und einem Workflow-Editor. Die Workflow-Engine führt den Workflow aus, der entweder interaktiv mit dem Workflow-Editor verbunden ist oder automatisch definiert wird (Petzold, Burghardt & Bobzien 2006, 3). Das Kap. 5.3 beschäftigt sich mit der konkreten praktischen Anwendungsumsetzung mittels der Software FME als Framework für einen Generalisierungsworkflow, der mehrere Teilprozesse in modularem Aufbau beinhaltet.

### 3 Verfügbare Verfahren und Methoden

#### 3.1 Situation im Landesamt für Geologie, Rohstoffe und Bergbau

Das Landesamt für Geologie, Rohstoffe und Bergbau (LGRB) verfügt über eine im Lauf der Jahre gewachsene Geodateninfrastruktur (GDI), die vereinfacht im Wesentlichen auf drei Säulen fußt: ArcSDE/PostGIS für die Geodatenhaltung, ArcGIS Desktop für das Geodatenmanagement sowie der UMN Mapserver für internetbasierte Geodatendienste. Intern werden die Geofachdaten mittels ArcSDE (ArcGIS for Server) basierend auf einer PostGIS-DB vorgehalten, am GIS-Arbeitsplatz stehen ArcGIS Desktop-Clients in der Version 10.4 sowohl in der Ausprägung Professional wie auch Standard zur Verfügung. An einzelnen Arbeitsplätzen ist zusätzlich auch das kostenfreie OpenSource-GIS QGIS im Einsatz. Da auch Geobasisdaten und sonstige Fachdaten ebenfalls in PostgreSQL/PostGIS vorgehalten werden und viele interne Anwendungen und technische Abläufe anhand SQL-basierter räumlicher Abfragen gesteuert werden - wobei sich hier der Einsatz zumeist vordergründig auf die Abfrage-Implementierung und -bereitstellung von Geodaten in Web-Anwendungen und Kartendiensten konzentriert - kommt diesem DBMS im LGRB eine zentrale Bedeutung zu. In den letzten Jahren hat die Software FME in dieser Infrastruktur im Arbeitsbereich des GIS-Administrators eine immer größer werdende Bedeutung als zentrale Drehscheibe bei der Geodatenverarbeitung erlangt. Sei es bei der Konvertierung, Geoprocessing oder Qualitätssicherung von Daten, der Konfektionierung von Geodaten-Produkten oder allgemein der Professionalisierung von Automatisationsabläufen, in all diesen Fällen erweist sich FME zumeist als äußerst performante und stringent zu bedienende Softwarelösung.

Das LGRB steht gewissermaßen stellvertretend für eine mittelgroße Behörde mit einer weitestgehend eigenständigen GIS-Infrastruktur. Die Bestrebungen zur Zentralisierung von IT-Dienstleistungen im öffentlichen Sektor schreiten auch im geowissenschaftlichen Bereich mit seiner Vielzahl fachspezifischer Anwendungen und Fachsoftware weiter voran - künftige Einschränkungen hinsichtlich der Wahl der Methoden und Verfahren sind hier nicht auszuschließen. Umso wichtiger ist beim genannten Vorhaben eine möglichst stringente, einfache technische Umsetzung basierend auf allgemein verfügbarer Software. Die Ausstattung mit den vor genannten GI-Systemen sollte einen durchaus ausreichenden technischen Rahmen für die Umsetzung einer Generalisierungslösung bieten.

	ArcGIS Desktop	QGIS	PostGIS	FME
Erscheinungsjahr	1999	2002	2001	1996
akt. Version im LGRB (10/18)	10.4	3.2	2.4	2018

Tab. 4: Im LGRB genutzte Software mit Generalisierungsfunktionalität

## 3.2 ArcGIS

ArcGIS von ESRI ist sowohl im Herkunftsland USA wie auch in Europa im Bereich Desktop-GIS unbestrittener Marktführer und ersetzt mittlerweile auch bei einigen staatlichen topografischen Behörden bisher genutzte kartografische Spezialsysteme (Hardy 2014). Durch die Zusammenarbeit mit Akteuren solcher NMA (z.B. SwissTopo) resultieren erweiterte Funktionalitäten, die über die bloße Bereitstellung einzelner Algorithmen hinaus die Bewältigung komplexerer Abläufe im Herstellungsprozess zu generalisierender topografischer Karten ermöglichen. Die Funktionen mit Bezug zur Generalisierung finden sich in den beiden Arc-Toolboxen „Cartography“ und „DataManagement“ im jeweiligen Toolset „Generalization“. Zusätzlich existiert in der „Editing“-Toolbox die Einzelfunktion Generalisieren, die analog zur „Linie vereinfachen“-Funktionsoption „Point remove“ den Douglas-Peucker-Algorithmus zur Vereinfachung von Linienobjekten innerhalb einer Editingsitzung bereit stellt. Darüber hinaus<sup>6</sup> gibt es eine Reihe weiterer Generalisierungswerkzeuge in der Toolbox „Production Mapping“, die als ArcGIS-Erweiterung aber einer gesonderten Lizenzierung unterliegt. Eine komplette tabellarische Zusammenstellung der Werkzeuge mit Generalisierungsfunktionalität ist dem Anhang A zu entnehmen.

### 3.2.1 ModelBuilder als Framework für die Generalisierungsautomatisation

Der ModelBuilder ist eine in ArcGIS Desktop (bzw. ArcGIS Pro) implementierte Anwendung, mit welcher Ablaufmodelle zur Geoprocessing erstellt, bearbeitet und verwaltet werden können. In solchen Workflows können Sequenzen von in der Software verfügbaren Geoverarbeitungswerkzeugen aneinander gereiht werden, wobei die Ausgabe des einen Werkzeugs als Eingabe für ein anderes Werkzeug dient. Die Bearbeitung und Erstellung dieser Workflows wird mittels Drag & Drop in einer grafisch angelegten Benutzeroberfläche gehandhabt. Für die statischen Prozessabläufe werden Variablen (Parameter) und Aufgaben durch die GUI des ModelBuilders visualisiert. Die Erstellung eines dynamischen Workflows allerdings ist erst nach Verlassen der GUI und anhand des Exports eines Modells als zu überarbeitendes Python-Skript möglich.

### 3.2.2 GeoScaler: Lösungsansatz mittels Arc-Toolbox

Die GeoScaler-Software entstand als Arc-Toolbox im Rahmen einer Projektarbeit des Natural Resources Canada innerhalb der Targeted Geoscience Initiative (TGI) der Kanadischen Regierung. Ursprünglich im Jahr 2009 als ArcGIS-Plugin für die Version 9.3 entwickelt, wurde das Programm in zweiter Generation als Arc-Toolbox basierend auf Python verbessert und neu programmiert und in 2012 der Allgemeinheit im Internet zur Verfügung gestellt (Huot-Vézina et al. 2012). Es ermöglicht die Generalisierung geologischer Karten bzw. stellt ein Framework für die Generalisierung von Polygonen und Symbolen zur Verfügung und berücksichtigt dabei wesentliche geologiespezifische Besonderheiten. Die Software besteht u.a. aus zwei zentralen Modulen, die für die Generalisierung von geologischen Verbreitungs- bzw. Grundgesteinskarten entwickelt wurden. Die Toolsets innerhalb der Toolbox beinhalten Python-basierte Skripte, die in vorgegebener Reihenfolge unter Verwendung interaktiv festgelegter Parameter ausgeführt werden müssen, um das endgültige Ziel zu erreichen: eine lesbare und semantisch korrekte geologische Karte für den beabsichtigten Zielmaßstab. Die angewendeten ArcGIS-Verfahren und die entsprechenden Eingabeparameter wurden während intensiver Forschungsarbeiten und Tests am kartografischen und photogrammetrischen Labor des Quebec Geoscience Centres (LCNP) sorgfältig ausgewählt.

<sup>6</sup> Im weiteren Sinne beinhalten auch die ArcGIS-Toolsets „Analyzing Patterns“ und „Mapping Clusters“ innerhalb der SpatialStatistics-Toolbox potenzielle Funktionen für die Bewältigung von Generalisierungsaufgaben im Hinblick auf strukturelle Einschränkungen im Gesamtkartenkontext. Von besonderem Interesse ist auch die Möglichkeit der Partitionierung der Ausgangsdaten, um große Datenmengen in kleinere, verwaltbare Sätze von Features zu unterteilen.



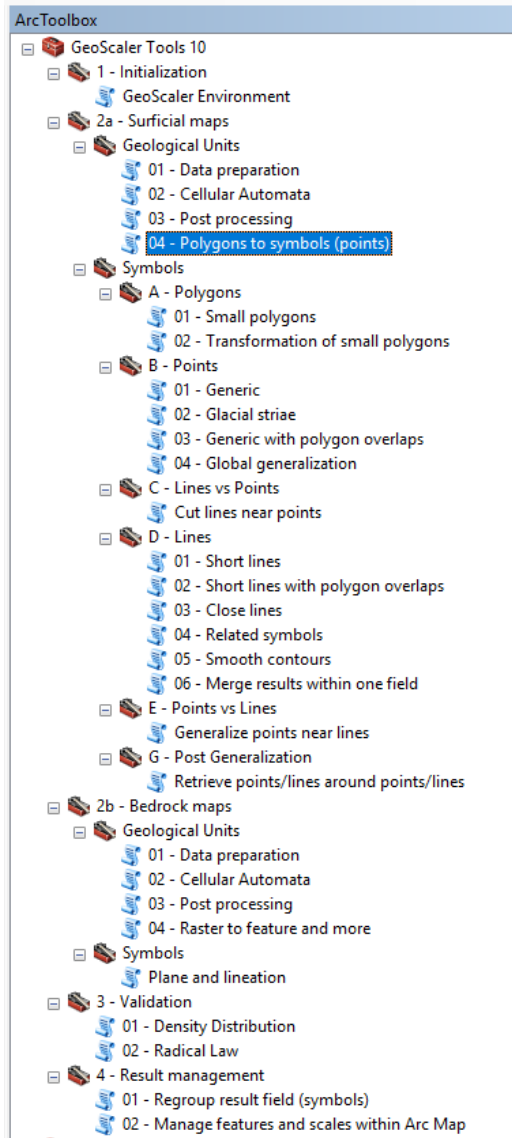


Abb. 7: Aufbau der GeoScaler-Toolbox in ArcGIS Desktop

Die Verfahrensweise für die Flächengeneralisierung ist mit der Anwendung des zellulären Automaten (s. Kap. 5.2.1) bei beiden Kartentypen identisch, im Gegensatz zur Verbreitungskarte können bei der Grundgesteinskarte aber auch die tektonischen Gegebenheiten in Linienform als topologische Relation mit einbezogen werden, dem entsprechend müssen in einem Generalisierungsablauf in solchen Kontaktbereichen Flächen als invariant behandelt werden. Der Ablauf für Grundgesteinskarten ermöglicht darüber hinaus auch die selektive Berücksichtigung von Flächen, die durch ihre Form oder ihre geologische Aussage als erhaltenswert gekennzeichnet sind. Die Toolbox beinhaltet neben der Polygonvereinfachung durch den zellulären Automaten hinaus eine umfangreiche Sammlung weiterer Arbeitsschritte, mit der auch kartografische Linien- und Punktelemente aus der Ausgangskarte in vereinfachter Form in den designierten Zielmaßstab überführt werden können. Die Toolbox in ihrer Gesamtheit deckt somit weitestgehend die unterschiedlichen kartografischen Aspekte der Generalisierung geologischer Karten ab und ermöglicht mittels interaktiver Steuerung bei Verwendung gleicher Eingabeparameter reproduzierbare Generalisierungsergebnisse. In der Praxis ist die Durchführung der einzelnen Arbeitsschritte in sequentieller teilautomatisierter Weise eine relativ komplexe und zeitaufwändige Aufgabe, die aber immerhin ein klar vorgegebenes und strukturiertes Handlungsgerüst für die Generalisierungsabläufe bietet.

### 3.3 PostGIS (und QGIS mit Python)

PostGIS ist die räumliche Erweiterung des mächtigen OpenSource-Datenbankmanagementsystems PostgreSQL und bietet eine umfangreiche Anzahl von Funktionen zur Analyse und Verwaltung von Geodaten, die zu einem Großteil die Aufgabenbereiche herkömmlicher Desktop-GIS abdecken. Die Bereitstellung entsprechender Schnittstellen bzw. die enge Kopplung und Implementierung in andere OpenSource-Programme (z.B. QGIS) ermöglicht eine weite Verbreitung dieses objektrelationalen DBMS und macht die Nutzung von PostGIS als Geodatendrehscheibe auf unkomplizierte Weise möglich.

Die direkte Ausführung von Analysen in der Geodatenbank ermöglicht die Automatisierung und Vereinfachung von Arbeitsabläufen, da der Im- und Export in andere Anwendungen dadurch überflüssig wird. Die SQL-basierte Methode zur Erstellung von Analysen ist in besonderem Maße geeignet, im Vorfeld definierte Anweisungen performant auf große Datenbestände anzuwenden. In der Praxis sind PostGIS-basierte Thin-Client- und Webanwendungen sehr gebräuchlich, hier muss im Zusammenspiel mit PostGIS nur die Nutzerschnittstelle dargestellt werden, da Speicherung und Analyse in der Datenbank stattfinden kann.

Im Umfeld der Generalisierungsautomatisation gibt es einige Beispiele aus dem praxisnahen Anwenderumfeld (Peter 2012; Dick & Retterath 2012), in denen automatisierte sequentielle Teilabläufe mittels PostgreSQL/PostGIS beschrieben werden. Das ist vor allem im Zusammenhang mit der Vorprozessierung und der Laufzeitgenerierung von Daten zur visuellen Bereitstellung in webbasierten Karten mittels Kartenservern ein wiederkehrendes Thema.

#### 3.3.1 Generische PostGIS-Funktionen mit Generalisierungsfunktionalität

Die von PostGIS bereit gestellten Datenbankfunktionen lassen sich nach Kapitel 8 des offiziellen Manuals (PostGIS Project Steering Committee 2018) unterschiedlichen Kategorien von Werkzeugen für den Umgang mit Geometrieobjekten zuordnen: so gibt es Funktionen für Ein- und Ausgabe, Konstruktion, Manipulation, Prozessierung sowie Validierung von Geometrien.

Nachfolgende Auflistung liefert eine Auswahl von PostGIS-Funktionen, die in Zusammenhang mit der Generalisierung von übergeordnetem Interesse sind bzw. mit deren Hilfe die aus 2.2.3 bekannten Generalisierungsoperatoren direkt oder indirekt im System bereit gestellt werden können.

Tab. 5: PostGIS-Funktionen mit Bezug zur Generalisierung (PostGIS Project Steering Committee 2018)

Operator	Funktion	Kurzbeschreibung
Vereinfachung	ST_Simplify	Gibt eine vereinfachte Version der Ausgangsgeometrie zurück. Verwendet den Douglas-Peucker Algorithmus.
	ST_SimplifyPreserveTopology	Wie ST_Simplify, berücksichtigt aber die topologische Konsistenz einer Geometrie und vermeidet die Erzeugung invalider Geometrien.
	ST_SimplifyVW	Gibt eine vereinfachte Version der Ausgangsgeometrie zurück. Verwendet den Visvalingam-Whyatt Algorithmus.
	ST_RemoveRepeatedPoints	Gibt eine Geometrie zurück, deren Stützpunkte anhand einer Schwellwertdistanz reduziert wurden.
Vergrößerung	ST_Buffer	Gibt eine Geometrie zurück, welche alle Punkte innerhalb einer gegebenen Entfernung von der Eingabegeometrie beinhaltet. Kann sowohl als positiver- wie auch als negativer Puffer angewendet werden.
Zusammenfassung	ST_Union ST_LineMerge	Gibt eine Geometrie zurück, welche der mengentheoretischen Vereinigung der Geometrien entspricht.
Typisierung	ST_Centroid	Gibt den geometrischen Schwerpunkt der Geometrie zurück.
	ST_ApproximateMedialAxes*	Errechnet die genäherte Mediale Achse einer Flächengeometrie
	ST_StraightSkeleton*	Berechnet aus einer flächenhaften Geometrie ein "Gerippe" aus Geraden.
Vergrößerung, Verschiebung	ST_TransScale	Umwandlung einer Geometrie entsprechend den gegebenen Skalierungsfaktoren und Versatzwerten.
Verschmelzung	ST_ConvexHull	Gibt die minimale konvexe Geometrie zurück, die alle Geometrien innerhalb einer Menge umschließt.
	ST_ConcaveHull	Gibt die konkave Hülle einer Geometrie zurück, welche alle Geometrien einer Menge einschließt.
	ST_Envelope	Gibt das umgebende Rechteck einer Geometrie zurück
	ST_ClusterWithin (ST_ClusterIntersecting)	Gibt den Array von GeometryCollections zurück, wobei jede GeometryCollection eine Sammlung von Geometrien darstellt, die durch nicht mehr als den angegebenen Abstand getrennt sind.
	ST_MinimumBoundingCircle	Gibt das kleinstmögliche Kreispolygon zurück, welches eine Geometrie zur Gänze beinhaltet.
* Spezialfunktionen aus SFCGAL-Bibliothek		

Darüber hinaus gibt es eine ganze Reihe von Funktionen in unterschiedlichen Aufgabenbereichen, die für die Erstellung eines designierten Generalisierungsablaufs von grundsätzlicher Bedeutung wären - sei es zur Ermittlung topologischer Bezüge oder von Kenngrößen, zur Konstruktion von Geometrien, für die performante Handhabung der Daten im Allgemeinen oder zur Überprüfung der Validität (Tab. 6).

Tab. 6: Für die Generalisierungsprozessierung relevante Funktionen (PostGIS Project Steering Committee 2018)

Aufgabenfeld	Funktion	Kurzbeschreibung
Handhabung	ST_Accum, ST_Collect	Gruppieren eine Menge an Geometrien zu einer Einzelgeometrie in Form einer Geometriesammlung bzw. eines Geometriearrays
	ST_Dump, ST_DumpRings	Gibt einen Satz an Einzelgeometrien bzw. Innere Ringe zurück, aus denen eine Eingangsgeometrie zusammengesetzt ist bzw. welche eine Eingangsgeometrie beinhaltet.
	ST_InteriorRingN, ST_ExteriorRing	Gibt den Nten innenliegenden Linienzug des Ringes einer Polygoneometrie zurück. Gibt einen Linienzug zurück, welcher den äußeren Ring einer Polygoneometrie darstellt.
	ST_SnapToGrid	Fängt alle Punkte der Eingabegeometrie auf einem regelmäßigen Gitter
Validierung	ST_IsSimple, ST_IsValid, ST_IsValidReason, ST_IsValidDetail, ST_MakeValid	Prüfen eine Eingangsgeometrie auf Validität bzw. geben Hinweise auf Zustand oder Gründe für Invalidität.  Versucht aus einer invaliden Geometrie eine valide Repräsentation zu erzeugen
	Einzelobjekt: ST_Area, ST_Length, ST_Perimeter	Geben den entsprechenden Kennwert (Flächengröße, Länge, Umfang) eines Geometrieobjekts (Polygon bzw. Linie) zurück
Messung	Objektgruppe: ST_Distance, ST_HausdorffDistance, ST_Azimuth	Ermitteln Abstandswerte bzw. Winkel zwischen Geometrieobjekten
	ST_Intersection,  ST_Difference, ST_SymDifference, ST_Split	Geben Geometrien zurück, welche der mengentheoretischen Verschneidung von Geometrien entsprechen. Gibt die Teile zweier Geometrien zurück, die sich nicht überlagern Gibt eine Sammelgeometrie zurück, die beim Auftrennen einer Geometrie entsteht
Topologie	ST_Within, ST_DWithin, ST_Relate, ST_Contains, ST_Touches, ST_Intersects, ST_Equals, ST_Disjoint	Testen von Geometrien auf ihren räumliche Bezug. Rückgabewert True (1) oder False (0)
	ST_MaxDistance, ST_LongestLine, ST_ClosestPoint	Geben charakteristische Geometrien oder Kennwerte mit speziellem Bezug zur Eingangsgeometrie zurück
Konstruktion	ST_BuildArea, ST_MakePolygon	Erzeugt eine Flächengeometrie aus den einzelnen Linien bzw. dem Linienkonstrukt einer gegebenen Geometrie

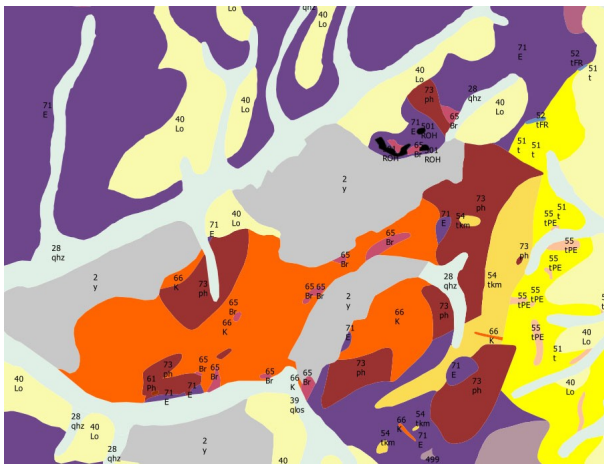

### 3.3.2 Schnittstelle zum OpenSource-Client QGIS

Die seit QGIS-Version 1.8 vorinstallierte Erweiterung DB-Manager ermöglicht neben der klassischen Verwaltung und Administration von Datenbankinhalten auch direkte räumliche Tabellenabfragen sowie die unmittelbare Anzeige der Abfrageergebnisse im QGIS-Kartenfenster. Dieser SQL-basierte Verarbeitungsansatz bietet besonders im Hinblick auf die testweise Implementierung von Geoprocessingsschritten durchaus Vorteile: so können die Auswirkungen von geoanalytischen Operationen am geometrischen Objekt quasi direkt überprüft und anschließend mit geänderten Parametereinstellungen wiederholt durchgeführt werden, bis die gewünschten Resultate erzielt sind.

Durch die Verschachtelung und Aneinanderreihung von räumlichen Funktionen können zum Teil komplexe Verarbeitungsschritte innerhalb einer einzigen SQL-Abfrage ausgeführt werden - dies ist ansonsten im GIS-Client nur durch entsprechendes (Python-)Skripting oder die Nutzung eines Moduls zur Erstellung von Workflows zur Geoprocessingung wie dem ArcGIS ModelBuilder oder dem QGIS Graphic Modeler möglich.

In Abb. 8 ist beispielhaft eine solche verschachtelte räumliche Abfrage aufgeführt: aus einem Ausschnitt der geologischen Basisdaten wird eine Zusammenfassung von gleichartig klassifizierten Polygonen auf geometrisch/semantischer Ebene mittels verbundener Zuordnungstabelle bewerkstelligt. Gleichzeitig kommen zusätzliche räumliche Operationen zur Vereinfachung der Zielgeometrien zur Anwendung.

Abb. 8: In QGIS visualisierte verschachtelte räumliche SQL-Abfrage

	
Geometrien im Ausgangszustand	Geometrie nach Abfrage (in grün)
Abfrage	Erläuterung
<pre>Select ( ST_Dump(   ST_RemoveRepeatedPoints(     ST_Buffer(       ST_Buffer(         ST_Union(a.geom),         50), -50),15)) ).geom as geom from geola.geo_gfl a left join geola.geo_all_tb b on a.gle = b.gle where ST_Intersects(a.geom,   ST_SetSRID(     ST_GeomFromText('POLYGON((3400750 ...))', 31467)) and b.hchy_7 = 348</pre>	<p><i>Entfernen von Stützpunkten</i>  <i>Outward-Buffer mit 50 m</i>  <i>Inward-Buffer mit 50 m</i>  <i>Vereinigung aller Eingangsgeometrien</i></p> <p><i>Eingangstabelle</i>  <i>Verbindung mit Schlüsselstabelle</i>  <i>Verschneidungsoperator v-Umriss</i>  <i>setzt Referenzsystem des Vorgabe-Polygons</i>  <i>vorgegebenes Polygon</i></p> <p><i>Einschränkung auf Objekte der Klasse 348</i></p>

### 3.3.3 Inward-Outward-Buffering

In der oben verwendeten Abfrage (Abb. 8) findet der simple, gleichwohl sehr effektive Ansatz des Inward-Outward-Buffering (I/O-Buffer) seine Anwendung. Dazu wird ein Polygon zunächst mittels der PostGIS-Funktion *ST\_Buffer(geom, n)* vergrößert, anschließend zieht man denselben Wert durch *ST\_Buffer(geom, -n)* wieder vom Objekt ab. Man erzielt damit eine Vereinfachung der Form, da man mit dem ersten Puffer Einstülpungen oder Löcher im Polygon einfach auffüllen kann. Durch den zweiten Puffer wird sichergestellt, dass die Lage und Ausdehnung der äußeren Form wieder auf den Ausgangswert zurückgesetzt wird. Abhängig von der Größe der Pufferung können so auch größere Lücken in Ringpolygonen gefüllt werden, wobei man damit aber auch den äußeren Ring zum Teil sehr stark glättet, was unter Umständen nicht gewünscht ist (Dick & Retterath 2012). Der I/O-Buffer ist in der Verwendung mit einem minimalen Pufferdistanzwert eine weit verbreitete Validierungsmethode unsauberer Flächengeometrien. Bei der Auflösung von Objektgrenzen kann es bei komplexen Geometrien durch eine fehlerhafte Snapping-Toleranzwert-Einstellung zum Entstehen von Artefakten in Form von inneren Ringen kommen - solche Fälle können durch die Anwendung des I/O-Buffers auf einfache Art und Weise bereinigt werden. Im Generalisierungsumfeld dient der I/O-Buffer auch als Methode zur Konflikterkennung bei Polygonkarten (Gao et al. 2012).

So können durch die Anwendung in umgekehrter Abfolge (Outward-Inward) schmale Korridore bei komplexen Polygonen erkannt und aufgelöst werden. Die Pufferung nach außen ergibt ein neues größeres Polygon, wodurch Einstülpungen geglättet und Korridore verbreitert werden. Die Pufferung nach innen mit gleichem Distanzwert stellt die annähernde Originalgröße wieder her. Mit zunehmender Pufferbreite werden so mehr Details des ursprünglichen Polygons entfernt. Zusätzlich zur Konfliktlösung verbessert sich auch das Gesamtbild des Polygons, was zum gewünschten Vereinfachungseffekt bei der maßstabsbezogenen Generalisierung führt (Gao et al. 2013). Auf die Makro-Ebene bezogen, also in der Anwendung im nachbarschaftlichen Kontext von Polygonen arbeitet der I/O-Buffer topologisch inkorrekt.

### 3.4 FME

Die Software FME (Feature Manipulation Engine) der Firma Safe Software Inc. ist laut Eigenwerbung das weltweit meistgenutzte sogenannte Spatial ETL-Werkzeug (extract, transform, load) zur Integration, Bearbeitung und Qualitätssicherung beliebiger räumlicher Informationen. Es ist eine äußerst leistungsfähige und flexible Softwarelösung, mit der auf sehr einfache Weise komplexe Verarbeitungsprozesse von räumlichen und nicht-räumlichen Daten modelliert und automatisiert ausgeführt werden können. Die Stärken des Programms liegen einerseits in der Vielzahl der verarbeitbaren Formate sowie in der einfachen intuitiven Bedienbarkeit über eine grafisch-interaktive Benutzerschnittstelle, mit deren Hilfe vollständig auf klassischen Programmiercode verzichtet werden kann. Programmierkenntnisse sind zwar einerseits zur Handhabung des Systems nicht zwingend erforderlich, andererseits verfügt FME über Schnittstellen zur Implementierung von externem Python-Programmcode oder programmatischer Erweiterungen in Java oder C++ anhand einer SDK.

Das zentrale Programm-Modul zur Entwicklung eines Prozessablaufs ist die sogenannte FME-Workbench, sie stellt die Plattform für die verfügbaren Formate und Funktionen bereit. Geometrische Objekte werden als Feature Type mittels eines Readers eingelesen (extract), anhand sogenannter Transformer verändert (transform) und über einen zu definierenden Writer wieder in ein Ausgabeformat nach Wahl zurückgeschrieben (load). Die Transformer liefern in einem FME-Workflow dabei die benötigten Werkzeuge zur Manipulation der eingelesenen Objekte (Features), wobei es sich nicht zwangsläufig um Geometrieobjekte handeln muss.

#### 3.4.1 Zusammenstellung wichtiger FME-Transformer mit Generalisierungsfunktionalität

FME verfügt eine Reihe verschiedener Transformer, die explizit für die Generalisierungsautomatisation geeignet sind. Dabei kann unterschieden werden zwischen Werkzeugen, die von vorneherein als Generalisierungsfunktion angelegt sind und Werkzeugen, die im Zusammenhang mit dem Gesamtprozess der Generalisierungsautomatisation Anwendung finden. Erstere werden in nachfolgender Auflistung vorgestellt und kurz in ihrer Wirkungsweise beschrieben (SAFE Software 2018).

##### Generalizer

Dieser Transformer ist vielmehr eine Werkzeugsammlung und beinhaltet gleich mehrere unterschiedliche Algorithmen aus verschiedenen Bereichen der (Linien)Generalisierung wie Simplifizierung, Glättung, Anpassung und Messung. Für die Liniensimplifizierung stehen wahlweise eine Reihe bekannter und weniger bekannter Algorithmen<sup>7</sup> zur Verfügung: u.a. Douglas-Peucker, Deveau und Wang. Bei den Glättungsalgorithmen sind die Methoden von McMaster, die McMaster Weighted Distance sowie NURBfit hinterlegt. Als Mess-Algorithmus wird die Inflection-Point-Methode zur Bestimmung der Sinuosität einer Linie bereit gestellt. Als Anpassungsalgorithmus schließlich findet die orthogonale lineare Regression Anwendung, sie ermöglicht die Minimierung der orthogonalen Distanz einer generalisierten Linie zu ihrem Original.

##### Sherbend Generalizer

Eigenständiger Transformer, der die Generalizer-Werkzeugsammlung zur Liniensimplifizierung um den Sherbend-Algorithmus ergänzt und hier eine Reduktion unnötiger Details (Linienbiegungen) gegebenenfalls auch unter Wahrung des räumlichen Bezugs zu Punktdaten gewährleistet. Der Algorithmus eignet sich dadurch besonders bei natürlichen Linienverläufen.

##### Area Amalgamator

<sup>7</sup> Leider fehlt in der FME-Dokumentation (SAFE Software 2018) zu den Transformern der Bezug auf die wissenschaftlichen Grundlagen bei den angewendeten Algorithmen, so dass deren genaue Funktionsweise nur schwer nachvollziehbar ist.

Der AreaAmalgamator ermöglicht eine Vereinfachung von Polygonen auf Basis einer Zusammenfassung bzw. Verschmelzung nahegelegener geometrischer Details. Diese geometrischen Details können sowohl im gleichen Objekt (Selbstbezug), zwischen verschiedenen Objekten (binär) sowie in beiden Fällen vorkommen und gehandhabt werden. Der interne Ablauf vollzieht sich in folgenden Schritten: Input-Polygone werden verdichtet (Parameter "Maximale Dreiecksbreite") und danach überlagert mit der konvexen Hülle, der Unterschied zwischen der konvexen Hülle und den Polygonen wird trianguliert. Die Dreiecke, die eine bestimmte Größe (Parameter "Maximale Dreieckslänge") nicht überschreiten, werden aufgelöst und bilden Verbinder, die dann mit den ursprünglichen Polygonen in Amalgame aufgelöst werden.

#### Displacer

Der Displacer kommt zum Einsatz bei der Lösung von Nachbarschaftskonflikten zwischen zwei Objekten, wobei als Parameter eine Mindestdistanz zwischen einer unveränderlichen Basisgeometrie und einem Kandidaten Anwendung findet. Anhand eines weiteren Parameters (Stiffness) wird der Grad der zugelassenen Veränderung der Kandidatengeometrie gesteuert.

#### Center Line Replacer, CenterPointReplacer

Sorgt für die Reduktion von Eingangspolygonen auf ihre mediale Achse oder Skelettlinie bzw. den Schwerpunkt. Dies ist im Fall des CenterLineReplacer insbesondere im Zusammenhang mit dem Auffüllen von Löchern oder zu entfernenden länglichen Flächengebilden interessant. Wichtig bei der Typisierung (Kollaps) von Polygonen zu Linien- oder Punktgeometrien.

#### DonutHoleExtractor

Löst den inneren Ring aus einem Eingangspolygon und ermöglicht somit im Zusammenspiel mit weiteren Transformern (s. Testfilter) das Löschen von Flächen, die eine Mindestgröße unterschreiten.

#### AreaOnAreaOverlayer

Dieser Transformer führt eine Flächenüberlagerung durch, so dass alle Eingabebereiche gegeneinander verschnitten und die daraus resultierenden Flächenmerkmale erzeugt und ausgegeben werden können. Bei der unabhängigen Generalisierung einzelner Kategorien ist hierdurch im Nachgang ein „Verebnen“ der sich topologisch inkorrekt überlagernden Geometrien möglich.

#### Chopper

Ermöglicht das Aufteilen von Geometrien in Punkte, Linien oder Bereiche. Kann verwendet werden, um komplexe Objekte grob zu vereinfachen.

FME bietet mit einer Vielzahl verschiedener „Calculator“-Transformatoren die Voraussetzung, um die in einem Generalisierungsablauf notwendigen Messungen auszuführen. Basierend auf diesen Messungen kann der weitere Ablauf bestimmt werden z.B. durch Auswahl der geeigneten Transformationsoperatoren. Es stehen zur Verfügung (Auswahl):

- Area Calculator, Length Calculator

Berechnet die Fläche bzw. Länge einer Objektgeometrie und speichert den Wert in einem Attribut.

- Angularity Calculator, Circularity Calculator

Berechnet die "Winkligkeit" bzw. „Zirkularität“ einer Objektgeometrie. Die Angularität gibt den Grad der Krümmung eines Objekts an. Je höher der Wert, desto gekrümmter die Geometrie. Die Zirkulari-



tät ermöglicht Aussagen über die Länglichkeit eines Objekts (perfekter Kreis = Wert 1 , Linie = Wert 0)

- Density Calculator

bestimmt das Dichtemaß einer Gruppe von Objektgeometrien und wird als numerischer Kandidatenwert pro Flächeneinheit berechnet; je größer der Wert, desto größer die Objektdichte nach Anzahl der Kandidaten/Untersuchungsfläche.

- LeftRightSpatialCalculator

Berechnet die relative Position der Eingabe-Objektkandidaten zu den Basis-Kandidaten.

- Statistics Calculator

Berechnet Statistiken basierend auf einem bestimmten Attribut oder einer Reihe von Attributen von Ausgangsgeometrien.

Das FME-Hub<sup>8</sup> bietet als zusätzliches Repository über die Standard-Transformer hinaus spezielle „custom transformer“, die in der Regel mehrere bestehende Transformer zu einem neuen Transformer-Paket bündeln und für die Nutzung in der eigenen Workbench bereitstellen:

MultiAreaLayerGeneralizer

Funktioniert wie der Generalizer, berücksichtigt aber mehr als einen Layer. Dieser Transformator hält topologisch saubere Beziehungen zwischen den Merkmalen aufrecht. Voraussetzung ist eine saubere Topologie der verschiedenen Ebenen zueinander.

Vertex Limiter (Custom Transformer)

Ermöglicht die iterative Reduktion der Anzahl von Stützpunkten unter einen vorgegebenen Grenzwert.

---

8 <https://hub.safe.com>

### 3.4.2 Handhabung einzelner Ablaufschritte innerhalb einer Workbench

Die mitunter komplexe Funktionalität der Transformer in der Workbench wird dadurch transparenter, dass der Status einzelner Objekte an jeder beliebigen Stelle des Workflows im sogenannten „Data Inspector“ beobachtet und validiert werden kann. Dadurch lassen sich gezielt einzelne Teilschritte in direkter vorher/nachher-Betrachtung überprüfen.

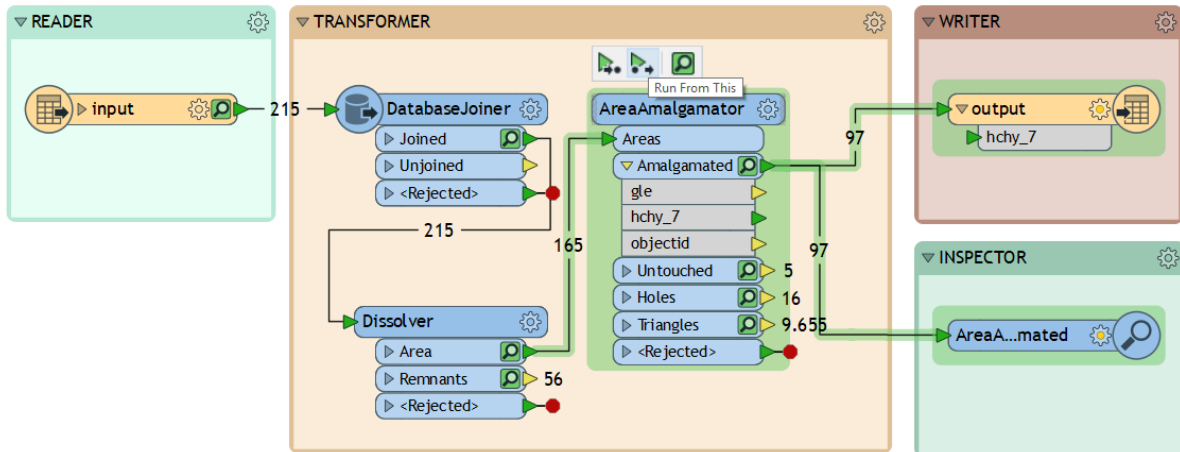


Abb. 9: Beispiel einer FME-Workbench bei der erneuten Ausführung mit aktiviertem FeatureCaching

Die testweise Anpassung oder Änderung von Parametern oder Optionen in einem Transformer mit wiederholter Durchführung ist ein häufiger Arbeitsschritt bei der Modellierung in der Workbench - hier ermöglicht das zur Laufzeit ausgeführte Feature-Caching die Ablage der Daten jedes einzelnen Prozessschrittes im Arbeitsspeicher und die wiederholte Ausführung des Prozessablaufs ab oder bis zu einem dieser Punkte mit geänderten Einstellungen (siehe Abb. 9). Gerade bei Tests mit größeren Datenmengen oder in längeren Prozessketten ist die Zeitersparnis durch diese Option sehr hilfreich.

### 3.5 Zusammenfassende Beurteilung hinsichtlich vorhandener Generalisierungsfunktionalität

ArcGIS bündelt vorhandene Funktionalität in den eigens bereit gestellten Toolsets „Generalization“, die unter den jeweils separaten Toolboxes „DataManagement“ und „Cartography“ sowie darüber hinaus mit erweiterter Lizenzierung im „ProductionMapping“ verortet sind. Erstere enthalten vordergründig allgemeine Methoden, die mit den Operatoren Auflösung bzw. Zusammenfassung und Eliminierung gleichgesetzt werden können. Die beiden Letzteren beinhalten darüber hinaus explizit kartografie-spezifische Funktionalitäten, die vor allem bei der Generalisierung topografischer Kartenwerke hilfreich sind. Implementiert wurden diese Funktionen auch als Resultat aus der Zusammenarbeit mit größeren Akteuren staatlich-behördlicher Kartografie (NMA), die zum Teil ArcGIS in Ablösung bisher genutzter Spezial-Software bei der Generalisierung einsetzen (Hardy 2014; Baig, Rahman & Duncan 2013; Šimková & Talhofer 2017). Ein grundsätzlicher Nachteil proprietärer Systeme liegt dabei in der zum Teil nicht nachvollziehbaren Implementierung von Algorithmen bzw. in der oft fehlenden Dokumentation.

Einblick und Umgang mit dem umfangreichen Repertoire an PostGIS-Funktionalität in ihrer Anwendung bei der Generalisierung erfordern ein etwas höheres Maß an Bereitschaft, sich mit den nativen Funktionen zu beschäftigen und so - laut Ramsey (2010), einem der Mit-Initiatoren und Programmierer der PostgreSQL-Erweiterung, den schwierigen Schritt auf ein Anwendungslevel zu schaffen, in welchem PostGIS hinsichtlich der Nutzung seiner Möglichkeiten den proprietären GI-Systemen in nichts nachsteht. Es mangelt dabei nicht an Praxisbeispielen im Internet mit Beschreibungen von Teilschritten oder Lösungsansätzen zu Teilproblemen der Generalisierung (Santilli 2012; Dick & Retterath 2012; Zimmermann 2015), diese stehen aber zumeist im Zusammenhang mit der performanten On-The-Fly-Generalisierung beim Abruf von Datenbeständen und deren Darstellung innerhalb web-basierter Geo-Anwendungen. Ein auf PostGIS basierender gesamtheitlicher Generalisierungsansatz ist allerdings nicht auszumachen.

FME wiederum weist einen ebenso mächtigen Funktionsumfang auf wie die zuvor genannten Systeme, die Stärken der Software liegen aber in der Modellierung und Automatisierung wiederkehrender Geoprozesse. Die in großer Zahl vorhandenen Transformer zeichnen sich anhand ihrer verwendeten Parameter und Eigenschaften durch eine mehr oder weniger ausgeprägte Komplexität aus. Einige der Transformer beinhalten eine Reihe verschiedener Optionen, die eine völlig unterschiedliche Zielsetzung haben und in ihrer jeweiligen Anwendung zu komplett anderen Resultaten führen. Das Verständnis über die Wirkweise der Transformer anhand der offiziellen Hilfe ist nicht immer einfach zu erlangen. Hier bietet das FME Knowledge Center als Internet-Plattform dem Anwender ein Community-Forum, in welcher recherchiert und die Anwendung oder Kopplung von Transformern anhand praktischer Beispiele erfahren werden kann. Dennoch ist für den situationsbezogenen Einsatz und der entsprechenden diesbezüglichen Anpassung oftmals wirkliches Expertenwissen erforderlich, was ohne einschlägige Schulung nur schwerlich zu erlangen ist.

PostGIS liefert keine diesbezüglich vorkonfektionierten Funktionen, bietet aber ein überaus mächtiges Funktionalitätenrepertoire, aus welchem sich durchaus die relevanten Teilschritte für eine gesamtheitliche Generalisierung zusammenbauen ließen. Allerdings fehlt hierfür ein entsprechendes Softwarevehikel in Art eines Workflow-Systems, in welchem sich ein solcher Prozessablauf in adäquater Weise modellieren ließe. Dies wiederum bietet die Software FME zusammen mit einem ebenfalls äußerst breiten Spektrum an sogenannten „Transformern“, die gleichermaßen die wesentlichen Teilprozesse einer Generalisierungsautomatisation abdecken und darüber hinaus die Möglichkeit bieten, Funktionalität aus anderen GIS zu implementieren. Den Ausspruch von Weibel (2016) „Durch Kreativität, geschickte Kombination und Parametrisierung herkömmlicher GIS-Operationen kann man recht viel erreichen“ gilt es nun in den nächsten beiden Kapiteln zu belegen.

Die in Tab. 7 aufgeführten und mit \* gekennzeichneten ArcGIS-Funktionen sind explizit für die Generalisierung von topografischen Kartenwerken konzipierte Werkzeuge, wobei einige dieser Funktionen aber auch im Umfeld der Generalisierung von Polygonmosaiken angewendet werden können. Grundsätzlich bilden sich alle einschlägigen Generalisierungsoperatoren in den einzelnen GIS anhand unterschiedlicher Funktionalitäten ab.

Tab. 7: GIS-spezifische Generalisierungs-Funktionalität mit Zuordnung zu einschlägigen Operatoren

ArcGIS	PostGIS (ST_)	FME
<b>zusammenfassen, aggregieren</b>		
Dissolve, ThinRoadNetwork*, DelineateBuiltUpAreas*, PruneHydro*, ThinHydroFeatures*	Union, MemUnion, LineMerge,	Dissolver, Aggregator
<b>vereinfachen, glätten</b>		
SimplifyPolygon, SimplifyLine, SmoothPolygon, SmoothLine, SimplifyBuilding*, Generalize, GeneralizeSharedFeatures*	Simplify, SimplifyVW, SimplifyPreserveTopology, RemoveRepeatedPoints	Generalizer, VertexLimiter, SberbendGeneralizer, Chopper, MultiAreaLayerGeneralizer
<b>vergrößern</b>		
Buffer, IncreasePolygonArea*, IncreaseLineLength*, ExtendPolygonSides*	Buffer, TransScale	Bufferer, Scaler
<b>verschieben, verdrängen</b>		
ModifyUnderlyingPolygon*	TransScale	Displacer
<b>verschmelzen, clustern</b>		
AggregatePoints, AggregatePolygons	ConvexHull, ConcaveHull, Envelope ClusterWithin, MinimumBoundingCircle	AreaAmalgamator, AreaOnAreaOverlay, HullAccumulator
<b>typifizieren</b>		
MergeDividedRoads*		
<b>eliminieren</b>		
Eliminate, EliminatePolygonPart, ConvertPolygons*, DeleteDangles*, SplitNarrowPolygons*, DeletePolygonsAndExtendLines*	Difference SymDifference	SliverRemover AreaGapAndOverlayCleaner
<b>kollabieren (typisieren)</b>		
FeatureToPoint, CollapseDualLinesToCenterlines*, CollapseRoadDetail*, ProductionCenterline*	Centroid, ExteriorRing, Points, ApproximateMedialAxes*	CenterPointReplacer, CenterLineReplacer

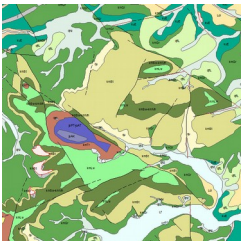
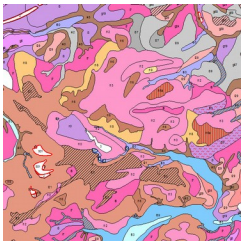
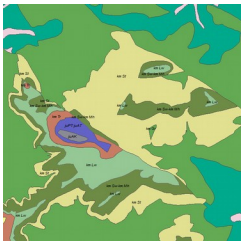


## 4 Datengrundlagen

### 4.1 Allgemeine Betrachtung der GeoLa-Geofachdaten

Die fortschreibungspflichtigen, amtlichen Geodaten des LGRB<sup>9</sup> liegen nach unterschiedlichen Fachbereichen geordnet in ESRI-FeatureDatasets vor. Grundlage des SDE-Applications-Servers ist eine PostgreSQL-DB, in welcher die Geodaten im PostGIS-Format (PG\_Geometry) gespeichert sind. Sie unterliegen der ständigen Überarbeitung und Fortschreibung durch den jeweiligen Fachbereich.

Ein intern im LGRB formuliertes mittelfristiges Ziel ist die Umsetzung eines durchgehend automatisierten Generalisierungsprozesses in der möglichst übertragbaren Anwendung auf die Flächendaten der Bereiche Bodenkunde, Geologie und der daran anhängigen angewandten Bereiche Hydrogeologie, Rohstoffgeologie und Ingenieurgeologie (s. Tab 8). In einem ersten Schritt und - als Ausgangspunkt für weitere Entwicklungen - wurden in der vorliegenden Arbeit nun die beiden Basisthemen aus dem GeoLa-Projekt - Geologische Karte und Bodenkarte - hinsichtlich ihrer semantischen Aufbereitung und geometrisch vereinfachten Abbildung als Verbreitungsflächen untersucht und verarbeitet. Die in diesem Kapitel und bei den darauf folgenden Praxis-Tests gemachten Erkenntnisse sollen als Rahmen dienen für die Konzeption einer prototypischen Lösung für ein automatisiertes Generalisierungsverfahren zur Ableitung von Übersichtskarten / Datensätzen als Bestandteil des LGRB-Produktportfolios und seiner Web-Dienste.

Die Zielsetzung einer übertragbaren modellbasierten Generalisierungslösung muss im Zusammenhang mit den vorliegenden GeoLa-Ausgangsdaten betrachtet werden. Hier zeigen sich z. T. fundamentale Unterschiede der Anforderungen hinsichtlich einer begrifflichen wie auch geometrisch-topologischen Vereinfachung in den verschiedenen geowissenschaftlichen Bereichen und deren jeweiligen Daten.

Geologie	Bodenkunde	Hydrogeologie	Ingenieurgeologie	Rohstoffgeologie
■	■	■	□	□
				
Geologische Karte 1 : 50 000 (GK-BW)	Bodenkarte 1 : 50 000 (BK-BW)	Hydrogeologische Karte 1 : 50 000 (HK-BW)	Gefahrenhinweiskarte 1 : 50 000 (GGHK-BW)	Karte der oberflächennahen mineralischen Rohstoffe 1 : 50 000 (KMR50)
Verbreitungsflächen geologischer Einheiten bzw. geologischer Gruppeneinheiten	Verbreitungsflächen bodenkundlicher Einheiten bzw. Bodenformgruppen	Verbreitungsflächen hydrogeologischer Einheiten und Deckschichten	Verbreitungsflächen unterschiedlicher Geogefahrenflächen	Verbreitungsflächen prognostizierter oberflächennaher mineralischer Rohstoffe
Tektonik	(Quartär-Geologie)	Geologie	Geologie	Geologie

Tab. 8: GeoLa-Fachdatenthemen

<sup>9</sup> Regierungspräsidium Freiburg, Landesamt für Geologie, Rohstoffe und Bergbau (Hrsg.) 2018: Geodaten der Integrierten geowissenschaftlichen Landesaufnahme (GeoLa). [https://lgrb-bw.de/aufgaben\\_lgrb/geoLa/produkte\\_geoLa](https://lgrb-bw.de/aufgaben_lgrb/geoLa/produkte_geoLa)

In der Auflistung der relevanten Themen aus Tab. 8 bildet der Themenbereich Geologie mit den geologischen Verbreitungsflächen die Bezugsgrundlage und Datenbasis für nahezu alle anderen Fachthemen. Ein bestimmendes Merkmal der geologischen Karte ist eine direkte topologische Abhängigkeit der Ausstrichs-/Verbreitungsflächen mit der zugrundeliegenden (linearen) tektonischen Situation. Der Themenbereich Bodenkunde wiederum stellt als zweites eigenständiges Basisthema anhand der Bodenkarte dank der detaillierten Beschreibung der Ausgangsgesteine der Böden - gegliedert nach Stratigraphie, Petrographie und Genese - eine wertvolle Grundlage für die Abgrenzung und Kennzeichnung quartärer Deckschichten auf geologischen Karten dar. Die hydrogeologische Karte wird in direkter geometrischer Ableitung durch Reklassifizierung aus der geologischen Karte erzeugt. Dies trifft in ähnlicher Weise auch für Einzelthemen der ingenieur-geologischen Gefahrenhinweiskarte zu. Die rohstoffgeologischen Vorkommen basieren zwar ebenso auf der Geologie, sind aber aufgrund eigenständiger Abgrenzung sozusagen geometrisch entkoppelt. Die beiden letztgenannten Themen beziehen sich dabei auf inselhafte Verbreitungsflächen, während GK, BK und HK jeweils als flächendeckendes Polygonmosaik vorliegen.

#### 4.1.1 Herausstellungsmerkmal Geologie

In der Form ihrer Darstellung und in der Methodik ihrer Erstellung sind sich die GeoLa-Basisthemen geologische Karte und Bodenkarte ähnlich. „Naturgemäß korreliert die räumliche Verbreitung verschiedener Böden mit der räumlichen Verbreitung der entsprechenden, unterhalb der Böden anstehenden Ausgangsgesteine, dabei können jedoch einerseits ähnliche Böden aus unterschiedlichen Ausgangsgesteinen hervorgehen, andererseits auch unterschiedliche Böden aus ähnlichen Ausgangsgesteinen“ (Wikipedia 2018). Unterschiede bei der Generalisierung geologischer Karten im Vergleich zu anderen Karten sind bezüglich des Inhalts, der Zielsetzung und des Zwecks sowie der Ausprägung auszumachen (Zhang, Wu & Wang 2016). So sollte hinsichtlich des Zwecks nicht nur die Koordination der räumlichen Verteilung von geologischen Elementen beim Maßstabswechsel Gegenstand einer Generalisierung sein, sondern auch die Verdeutlichung und Heraushebung geologischer Besonderheiten als Zielsetzung Eingang finden. Ein möglichst themenübergreifender Generalisierungsansatz sollte sich in der Praxis wohl zuvorderst auf die Verbreitungsflächen an der Oberfläche bei den Themenbereichen Geologie, Bodenkunde und Hydrogeologie konzentrieren - die Gemeinsamkeit hierbei ist die flächendeckende Verbreitung. Inhaltlich ist ein stärkerer Einbezug festgelegter semantischer Gegebenheiten erforderlich, was einen speziellen bzw. erweiterten Ansatz erforderlich macht. Bezüglich der Ausprägung der Generalisierung ist bei geologischen Karten eine selektivere Handhabung von Bedingungen und Einschränkungen notwendig. Um wesentliche prioritäre geologische Merkmale hervorzuheben, muss gegebenenfalls der Grad der Vereinfachung bei den geologischen Elementen ohne Fokus erhöht bzw. anhand Datenanreicherung eine wichtige zusätzliche Grundlage für den weiteren Generalisierungsablauf geschaffen werden.

Bei den anderen GeoLa-Fachthemen zeigen sich mehr oder weniger inhaltlich-geometrische Abhängigkeiten vom geologischen Basisthema, die Komplexität der Generalisierungsaufgabe ist aber sowohl in semantischer wie auch geometrisch-topologischer Hinsicht innerhalb des Geologiethemas am höchsten. Als Ausgangspunkt für eine prototypische Generalisierungslösung sind die geologischen Daten aufgrund der geometrischen Komplexität (s. Kap. 4.5.1) und der hier in der Mehrzahl aufzufindenden speziellen Konfliktszenarien (s. Kap. 4.5.2) am ehesten geeignet. Sie sollten die grundsätzlichen Lösungsstrategien bei Anwendung auf einfachere Daten beinhalten.

Sayidov & Weibel (2016) unterscheiden speziell bei geologischen Ausgangsdaten verschiedene Komplexitätstufen: im einfachsten Fall sind nur wenige geologische Einheiten mit einfachen Geometrien beteiligt, die le-

diglich mittels Liniensimplifizierung vereinfacht werden. Die zweite Anforderungsstufe wäre die Aggregation oder Verschmelzung verteilter zusammengehöriger Objekte zu einer neuen Einheit. Stufe 3 beinhaltet längliche Polygone ähnlicher Anordnung und Ausrichtung<sup>10</sup>, die zusätzlich zu einer Verschmelzung auch einen Strukturhalt erfordern. Im nächsten Level wäre der Einbezug topologischer Abhängigkeiten zu berücksichtigen, wie sie z.B. bei den dendritischen Talflächenverläufen im quartären Bereich und der topografischen Information des Gewässernetzes gegeben sind. Die oberste Anforderungsstufe versammelt alle bzw. mehrere der vorher genannten Anforderungen in einem Generalisierungsszenario. Dies erfordert im Umkehrschluss die Orchestrierung einer Vielzahl womöglich zueinander abhängiger und gegebenenfalls widersprüchlicher Generalisierungsoperationen.

#### 4.1.2 Semantische Aggregation im vorliegenden Kontext

Für die GeoLa-Geofachdaten wurden intern Schlüssellisten aufgesetzt, die es ermöglichen sollen, bestimmte Objekte mit gleichen Eigenschaften zusammenzufassen. Ein für den jeweiligen Fachbereich vorab fest umrissener tabellarischer Themenbaum liefert das zentrale Element für die inhaltliche Gliederungsstruktur. Dieser taxonomische Rahmen bildet schließlich auch das jeweilige semantische Gerüst für die Ableitung der einzelnen Kartenthemenvisualisierungen, wie sie in vereinfachter Form für die unterschiedlichen Generalisierungsszenarien und Maßstabsziele Anwendung finden sollen. Kap. 4.3.3 beschäftigt sich eingehend mit der zugrundeliegenden geologischen Datensituation und ihren semantischen Hierarchien.

Ein zu beachtender Teilaspekt dieser hierarchischen Gliederung ist die mögliche unterschiedliche Wertigkeit einzelner Karteneinheiten, die gegebenenfalls durch die Aggregation und Überführung in die nächste Hierarchiestufe verloren geht - falls in dieser Hierarchiestufe keine Unterscheidbarkeit der ursprünglichen Merkmale eines geometrischen Objekts mitgeführt wird.

Entweder muss bei einer Auflösung mit anderen Neuklassenelementen der Verlust der Aussagekraft/Information über spezifische geologische Eigenheiten in Kauf genommen oder die entsprechende Einheit muss als Ganzes in die nächste Hierarchiestufe übernommen werden - also als eigene neue Klasse singular übernommen werden. Hinterlegt werden sollte ein Gewichtungsfaktor für jede Einheit, um eine Priorisierung der Einheiten innerhalb einer Hierarchiestufe zu ermöglichen und dadurch auch die Generalisierungsabläufe gezielter steuern zu können. Dieser Kennwert (s. 4.5.3.1 a und 5.3.2) entscheidet über Verbleib und entsprechende Transformation bzw. gegebenenfalls, wie und ob ein Objekt mit welchen benachbarten Objekten aufgelöst oder mittels Geometrieübergang als Punkt oder Linie typisiert wird.

<sup>10</sup> Typisch für geologische Karten, da geologische Ausstriche einer bestimmten vertikalen Mächtigkeit im reliefiertem Gebiet zu schmalen, bandartigen Objekten werden.

## 4.2 Daten der bodenkundlichen Karte 1 : 50 000

Die Ausgangsdaten der landesweiten bodenkundlichen Einheiten (rund 200 000 Objekte) beruhen auf einer umfassenden, landesweit an 26 Bodengroßlandschaften orientierten regionalen Kategorisierung mit rund 4000 Kartiereinheiten. Diese entsprechen den unterschiedlichen Bodenformgruppen. Für die Darstellung im Geodatendienst (Maßstabsbereich 1 : 10 000 bis 1 : 125 000) werden sie in ca. 125 übergeordnete Leitbodengesellschaften zusammengefasst.

### 4.2.1 Semantische Vorbedingungen

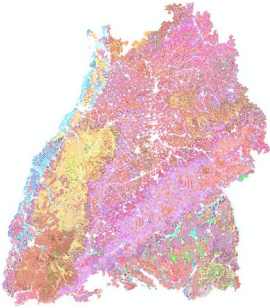
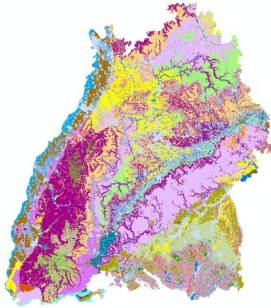

„Kleinmaßstäbige Bodenübersichtskarten werden in der Regel aus groß- oder mittelmaßstäbigen Bodenkarten durch eine Zusammenfassung der dort dargestellten Bodeneinheiten abgeleitet. Dabei umfasst die grafische Bearbeitung neben der Unterdrückung von Kleinstflächen auch eine Generalisierung der Konturen. Die Inhaltsbeschreibungen konzentrieren sich auf dominierende Bodenparameter. Um in diesem Prozess der Aggregation und Generalisierung auch überregional eine vergleichbare Vorgehensweise zu sichern und damit zu gewährleisten, dass so entstandene Bodenkarten auswertbar und vergleichbar bleiben, ist eine einheitliche pedoregionale Gliederung erforderlich. Diese muss sowohl die Ausprägung der Bodenparameter als auch die regionale Verbreitung der Böden berücksichtigen. In der Bodenkundlichen Kartieranleitung (Bundesanstalt für Geowissenschaften und Rohstoffe 2005) werden dazu die folgenden sieben Aggregierungsstufen von der Einzelfläche einer Bodenform bis zur regionalen Dimension beschrieben:

1. Bodenform
2. Bodenformengesellschaft
3. Leitbodengesellschaft
4. Leitbodenassoziation
5. Bodenlandschaft
6. Bodengroßlandschaft
7. Bodenregion

Für die Aggregierungsstufen 1 bis 4, die die Bodenformen und deren Vergesellschaftungen betreffen und für die Stufen 5 bis 7, die sich an landschaftlichen Grenzen orientieren, werden Kriterien zur inhaltlichen und räumlichen Zusammenfassung bzw. zur gegenseitigen Abgrenzung angegeben. Ziel ist es, die Böden unabhängig von ihrer regionalen Verbreitung immer nach vergleichbaren Vergesellschaftungsmerkmalen zusammenzufassen. Bodengroßlandschaften (BGL) folgen in der Systematik der pedoregionalen Gliederung auf die Hierarchieebene der Bodenregionen (BR) und unterteilen diese weiter. Sie selbst setzen sich aus Bodenlandschaften zusammen, die durch eine gemeinsame geologisch-paläogeografische Entwicklung und folglich durch gleiche pedogenetische Prozesse geprägt sind“ (Bundesanstalt für Geowissenschaften und Rohstoffe 2012). Das für LGRBWissen (s. 1.4.2 B) maßgebliche Aggregationsschema basiert auf den unter Punkt 5 und 6 genannten Stufen, die hierfür vorgesehenen Zusammenfassungen gründen jeweils auf den in der zugehörigen Sachdatenbank verankerten Schlüsselzuordnungen von der detailliertesten Ebene der bodenkundlichen Einheiten (BKE) zu den jeweiligen übergeordneten räumlichen Einheiten auf Bezugsebene der Bodenlandschaft-



ten (BL) bzw. Bodengroßlandschaften (BGL). Tab. 9 zeigt die designierte Themenhierarchie in ihrer Auswirkung auf die Anzahl der räumlichen Objekte und distinkten Einheiten sowie den Zielmaßstab(sbereich).

Hierarchiestufe	BKE	BL	BGL
Abbildung	 Bodenformgesellschaften	 Bodenlandschaften	 Bodengroßlandschaften
Anzahl Polygone	ca. 200 000	ca. 55 000	ca. 16 000
Anzahl distinkter Einheiten	4003	181	26
Designierter Maßstabsbereich	1 : 10 000 - 1 : 125 000	1 : 125 000 - 1 : 350 000	1 : 350 000 - 1 : 1 500 000

Tab. 9: Ausgangsdaten und Hierarchiestufen des Themas Bodenkunde

#### 4.2.2 Besonderheiten

Siedlungsbereiche, Gewässerflächen und Rohstoffgewinnungsstellen bilden sich als gesondert klassifizierte Themenobjekte im Polygonmosaik ab. Bei der vorgesehenen Generalisierung zu Bodenlandschaften sollen diese Kategorien nicht mehr ausgewiesen werden. Sie weisen keine Zuordnung zu einer vorhandenen bodenkundlichen Einheit auf und müssen bei der geometrischen Aggregation einer bestehenden Kategorie zugeschlagen bzw. je nach topologischem Bezug auf mehrere Nachbarflächen aufgeteilt werden (s. Kap. 4.5.3.1 b). Ein weiteres Merkmal der Bodenkarte sind einzelne Polygone mit komplexer Geometrie. Diese werden oftmals durch Abschwemmmassen in fein verzweigten Talbereichen in Mittelgebirgen abgebildet, die sich im Gelände über mehrere Kilometer erstrecken können.

Die beiden Sachverhalte bilden auf Grund des häufigen Vorkommens solcher „Nullflächen“ sowie der zum Teil erheblichen Größe der Einzelobjekte eine zentrale Problemstellung bei diesem Fachthema. Im Vergleich mit den geologischen Flächeneinheiten weist die BK ansonsten aber weniger Beschränkungen auf und ist deshalb als Generalisierungsgegenstand etwas einfacher in der Handhabung.

### 4.3 Daten der Geologischen Karte 1 : 50 000

Thematisch und topologisch sind die Ausgangsdaten der geologischen Karte 1 : 50 000 von Baden-Württemberg - im Weiteren verkürzt als GK-BW (LGRB 2018) bezeichnet - strukturiert nach den Einzelerebenen

- a) geologische Grundflächen (Polygone)
- b) geologische Überlagerungsflächen (Polygone)
- c) geologische Linien und Bänke (Linien)
- d) tektonische Flächen (Polygone)
- e) tektonische Linien (Linien)
- f) Erz- und Mineralgänge (Linien)
- g) punkthafte geologische Objekte (Punkte)

Einige der vorhergehend gelisteten untergeordneten Themenobjekte (d, g) liegen noch nicht in abgeschlossener bearbeiteter landesweiter Konsistenz vor, die offizielle kartografische Darstellung der GK-BW verzichtet deshalb zum jetzigen Zeitpunkt noch auf diese Informationsebenen. Nachfolgender Kartenausschnitt verdeutlicht die außerordentliche Komplexität einer geologischen Karte im Zusammenspiel der einzelnen Layer, im nächsten Abschnitt wird auf die Einzelheiten näher eingegangen.

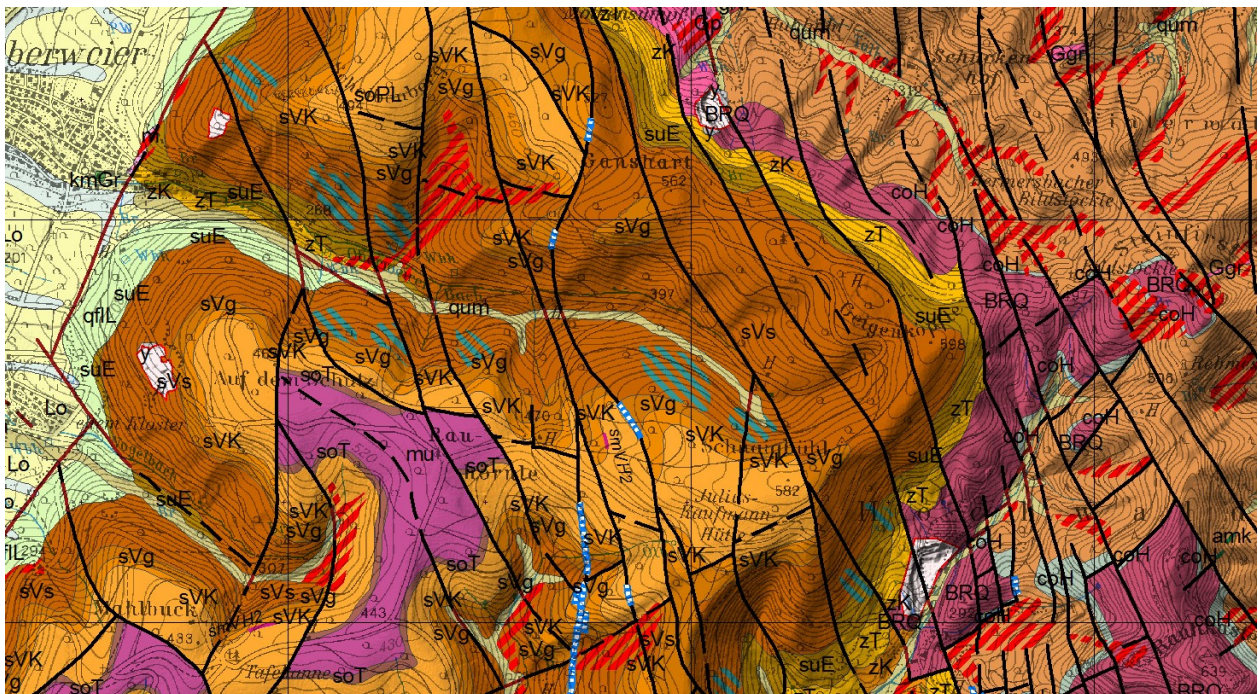


Abb. 10: Ausschnitt aus der GK50 (Arbeitsversion) im Randbereich des mittleren Schwarzwalds

### 4.3.1 Bezugsmatrix der geologischen Themenebenen

Zwischen den oben genannten Themenlayern der GK-BW bestehen konkrete Abhängigkeiten - zum Einen als direkte topologische Relationen, zum anderen aber auch (lediglich) anhand des semantischen Bezugs zwischen den Einzelebenen. Die geologischen Grundflächen weisen hierbei zu allen anderen Ebenen einen konkreten topologischen Bezug auf, weitere direkte Abhängigkeiten sind bei tektonischen Lineationen im Bezug auf tektonische Flächen sowie Erz- und Mineralgänge, indirekt und im minderen Maße auch in ihrem Lagebezug zu den punkthaften Objekten festzustellen. Der wichtigste topologische Bezug der GK-BW indes liegt zwischen den geologischen Grundflächen und Segmenten der tektonischen Linien vor. Letztere bilden in vielen Fällen gemeinsame Geometrien mit dem Umrisslinien der geologischen Grundflächen.

Eine weitere geometrisch-topologische Abhängigkeit besteht zwischen den geologischen Grundflächen und einzelnen geologischen Linien. Diese Linien sind oftmals vom stratigrafischen Rang geologische Bänke, die - wie tektonische Linien - partiell gemeinsame Geometrien mit den Umrisslinien der geologischen Grundflächen aufweisen. Von untergeordneter Bedeutung sind die topologischen Beziehungen von tektonischen Flächen zu tektonischen Linien sowie zu Erz- und Mineralgängen. Letztere treten in Natur gehäuft im Bereich tektonischer Verwerfungen auf, die als Schwächezonen die Ausbildung von Mineralgängen fördern. Tektonische Flächen, oft im Bereich von Grabenstrukturen anzutreffen, werden in der Regel von tektonischen Linien begrenzt.

	a	b	c	d	e	f	g
a	0						
b	x	0					
c	x		0				
d	x			0			
e	x			x	0		
f	(x)				x	0	
g	(x)		(x)		(x)	(x)	0

Tab. 10: Bezugsmatrix geologischer Themenebenen

### 4.3.2 Besonderheit quartär-geologischer Einheiten

Innerhalb des Themenlayers der geologischen Grundflächen sind die quartär-geologischen Einheiten als Besonderheit zu verstehen. Sie bilden als Deckschichten ein quasi autonomes Gebilde, das unabhängig vom Grundgestein als eigenständiger Geometriedatensatz im Generalisierungsablauf behandelt werden kann. Die quartär-geologischen Daten weisen dabei einen themenübergreifenden Topologie-Bezug zu den bodenkundlichen Einheiten auf, innerhalb des im Einführungskapitel angesprochenen Datenintegrationsprojektes GeoLa im LGRB wurden diese Einheiten durchgehend als konsistente Bereiche (d.h. gemeinsame Objektgeometrien in BK-BW und GK-BW) gewahrt.

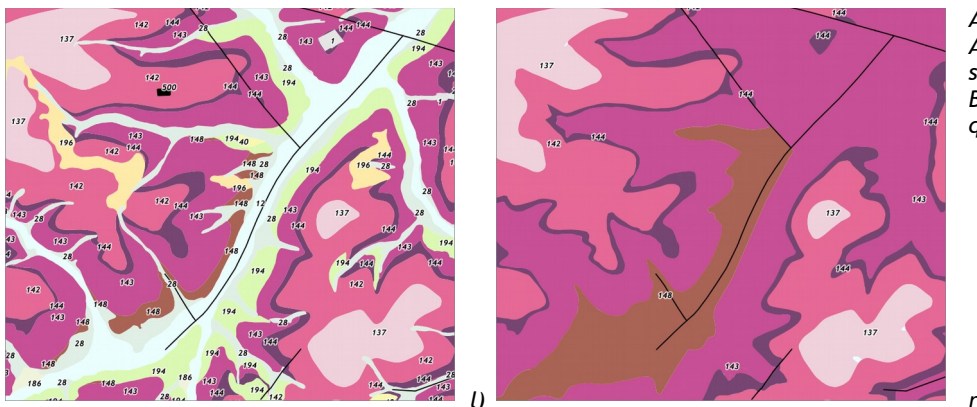

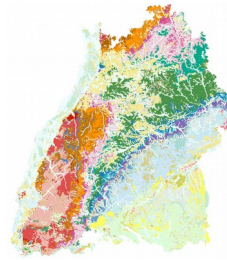
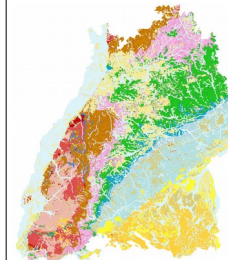
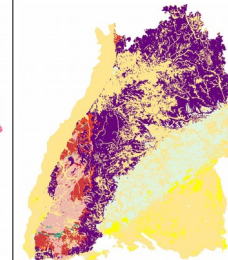


Abb. 11: verkleinerter Ausschnitt einer Datenvisualisierung aus der GK-BW mit (l) und ohne (r) quartäre Deckschichten

Das Augenmerk bei vorliegender Arbeit liegt auf der Generalisierung geologischer Grundflächen in ihrer Form als abgedeckte geologische Karte (GKoD<sup>11</sup> - s. Abb. 11). Da die GKoD zusammen mit den Deckschichten und den Sonderflächen (Rohstoffabbau, Rutschungen) wieder die GK-BW ergibt, kann die Generalisierung auch für die zwei getrennten Datensätze (GKoD-Grundflächen und geologische Deckschichten) durchgeführt und die einzelnen Ergebnisse im Nachgang wieder zusammengefügt/überlagert werden.

### 4.3.3 Semantische Vorbedingungen

Wie bereits in Kap.4.1.2 geschildert, gibt es es eine klare Vorgabe aus dem Fachbereich, wie die auf der Ausgangsebene der GK-BW vorkommenden Einheiten für die nachfolgenden Hierarchiestufen abzubilden und zusammenzufassen sind. Eine Übersicht zur Datenlage beim Übergang von der detailliertesten Stufe in die jeweils nächsthöhere Stufe liefert die nachfolgende Tabelle<sup>12</sup>.

Hierarchiestufe	GLE (GE4)	GE3	GE2	GE1
Abbildung				
Anzahl Polygone	141487	96144	74983	36477
Anzahl distinkter Einheiten	419	79	31	11
Designierter Maßstabsbereich	1 : 10 000 - 1 : 75 000	1 : 75 000 - 1 : 150 000	1 : 150 000 - 1 : 500 000	1 : 500 000 - 1 : 3500 000

Tab. 11: Kennwerte der Daten zur geologischen Karte in ihren einzelnen Hierarchien

Die Geologiedaten gliedern sich in eine 4-stufige Hierarchieabfolge, die ausgehend von der Geologischen Karte 1 : 50 000 aus ursprünglich +- 450 Flächeneinheiten auf der detailliertesten Stufe besteht. Diese Einheiten haben jeweils einen 1:1-Bezug zu einer übergeordneten Klasseneinheit der nächstfolgenden Stufe, die sich wiederum 1:1 zu den weiteren Kategorien höherwertigen Ranges zuordnen lässt. Diese Taxonomie basiert auf einer lithostratigrafischen Zusammenfassung in der Abfolge Formation (GE4) → Untergruppe (GE3) → Gruppe (GE2) → Hauptgruppe (GE1). Dabei gibt es bezüglich der Zuordnung und Zusammenfassung durchaus problematische Festlegungen, die einer stringenten und konfliktfreien Umsetzung bei der automatisierten Generalisierung hinderlich sind.

Die jeweilige Klassenneubildung wird dabei unter Berücksichtigung der für die folgenden Hierarchiestufen vorgesehenen Maßstabsbereiche vorgenommen. Diese Maßstabsbereiche orientieren sich auch an den Gegebenheiten, wie sie im Zusammenhang mit der Visualisierung in Darstellungsdiensten (Web Mapping Services - WMS) und deren Verwendung und Anzeige in WebGIS bzw. Web Mapping-Anwendungen eine Rolle spielen. Die dort vordefinierten Zoomstufen mit ihren je nach zugrundeliegendem initialem Referenzkoordinatensystem

11 Geologische Karte ohne Deckschichten

12 Die Zahlenwerte zur Flächenanzahl wurden auf Grundlage der landesweiten, nicht abgedeckten GK-BW-Daten anhand SQL-basierter Aggregation in der DB erstellt. Die Verwendung einer landesweiten abgedeckten Datenbasis würde von einem Ausgangswert von ca. 45 000 distinkten Polygonen ausgehen.

(normalerweise Spherical Mercator) ungeraden Maßstabszahlen liefern den Rahmen für die Zuweisung von relevanten Maßstabsbereichen.

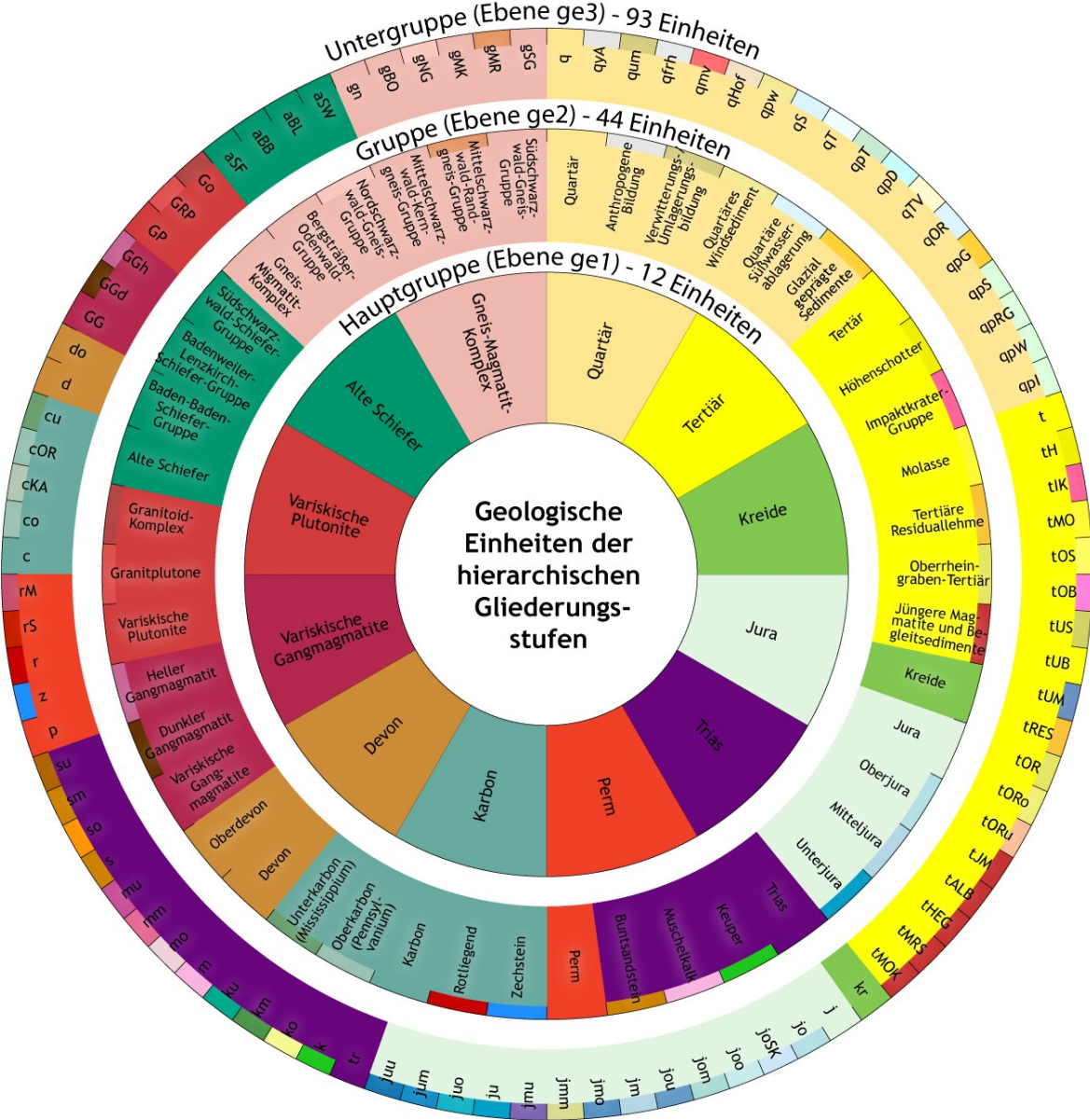


Abb. 12: Die drei Gliederungsebenen mit ihren geologischen Einheiten und deren Farbgebung

Beim Übergang von Hierarchiestufe 2 (GE2) in Stufe 3 (GE1) ergeben sich auf der semantischen Ebene durch die Reduzierung von vorher 31 Einheiten auf 11 Zieleinheiten<sup>13</sup> noch einmal weiträumige Zusammenfassungen von Geometrien, die bei der Auflösung gemeinsamer Grenzen zu extrem großflächigen und außerordentlich komplexen Geometrien führen. Die daraus resultierende Auswirkung auf die Performanz im Generalisierungsprozess ist nur eine von mehreren Fragestellungen, die sich bei der Gesamtbetrachtung ergeben und im nächsten Abschnitt näher erläutert werden.

13 In Abb. 12 sind alle Generallegendeneinheiten der GK-BW aufgeführt, daraus erklärt sich die Abweichung in den Angaben. Einige der Einheiten treten in Baden-Württemberg als Ausstrichsfläche an der Erdoberfläche nicht zu Tage oder sind - wie im Fall der (Haupt)Gruppe Kreide - gänzlich ohne Vorkommen.

## 4.4 Datenproblematik bei der Ausgangssituation

### 4.4.1 Quartärflächen versus abgedeckte geologische Karte

Die Ausgangsdaten der Geologischen Karte 1 : 50 000 stellen sich zum Zeitpunkt dieser Arbeit noch in der Form eines in den Talbereichen stark untergliederten und an den Begebenheiten der Bodenkarte und holozänen Sedimenten orientierten Flächenmosaiks dar. Der abgedeckten geologischen Karte wiederum liegt eine deutlich homogenere, vereinfachte und konsistente Flächenaufteilung zugrunde. Das Fehlen der mitunter feingliedrigen dendritischen holozänen Flächenfortsätze der Deckschichten hat vor allem hinsichtlich der Prozessierung größerer (landesweiter) Datensätze einen deutlich positiven Nebeneffekt. In der künftig landesweit vorliegenden Ausgangsdatenbasis einer solchen Grundgebirgskarte sind darüber hinaus die bisher ausgewiesenen Rohstoffgewinnungsflächen sowie Gewässerflächen aufgelöst, so dass im Generalisierungsprozess grundsätzlich keine nicht-zuordenbaren Objekte existieren<sup>14</sup>. Die Bearbeitung dieser geologischen Ausgangssituation hin zu einer von den geringmächtigen jüngeren quartären Ablagerungen abgedeckten geologischen Karte (GkoD) ist zum Zeitpunkt dieser Arbeit noch nicht abgeschlossen - einige fertig gestellte Teilbereiche Baden-Württembergs standen aber zur Verfügung bzw. wurden zu diesem Zweck explizit vorbereitet (s. Kap. 4.5.4).

### 4.4.2 Datenumfang und Datenqualität

Grundsätzlich muss bei der hier geforderten Generalisierungslösung der Performanz bzw. der Laufzeit als Aspekt keine oberste Priorität zukommen. Die zugrunde liegenden Daten werden in ihrer Verwendung für die Produkterstellung in maximal halbjährlichen Zyklen durch automatisierte FME-Prozesse aus der Editierumgebung der ArcGIS-SDE herausgeschrieben. Das avisierte Verfahren wäre als zusätzlicher Schritt in die bestehenden Abläufe zu integrieren. Die Ausführung kann als Batchjob ohne Probleme über einen längeren Zeitraum (über Nacht) ablaufen. Dem Zusammenhang zwischen Datenumfang und Performanz sollte dennoch im Vorfeld schon Aufmerksamkeit geschenkt werden. Die in kleineren Testbereichen als effizient und erfolgreich durchgeführten Prozessabläufe können sich bei einer Anwendung auf die Gesamtdaten als extrem zeitaufwändige oder im schlimmsten Fall gar undurchführbare Operationen herausstellen. Hier sollte in jedem Fall bestehendes Optimierungspotenzial ausgeschöpft werden.

Die Ausgangsdaten werden gemeinhin als valide vorausgesetzt, dennoch kann es in der Praxis - gerade bei der späteren Übernahme aktualisierter Ausgangsdaten - im Ablauf zu Problemen aufgrund von topologischen oder attributiven Unsauberkeiten kommen<sup>15</sup>. Diese Problematik legt nahe, im designierten Framework möglichst direkt nach dem Einlesen der Ausgangsdaten zusätzlich auch eine Geometrie-Validierung (mit automatischer Reparatur) zu hinterlegen. Bei nicht automatisch zu bereinigenden Invaliditäten sollte eine sofortige Prozessterminierung erfolgen, um einen unvermittelten Abbruch im fortgeschrittenen Stadium des zeitaufwändigen Gesamt Ablauf zu verhindern.

### 4.4.3 Auswahl des Sternansatzes als Ableitungsmethode

Ausgehend von einem Basisdatenmodell oberster Auflösungsstufe - in diesem Fall den Fachdaten im Planungsmaßstab 1 : 50 000 - ergibt sich die Frage nach der Vorgehensweise bei der Ableitung der Daten aus dem Ausgangsmodell in Modelle geringerer Auflösung bzw. kleineren Maßstabs. Wie in 2.2.10 beschrieben, ist

<sup>14</sup> Bei der bodenkundlichen Karte sind es aber genau diese in den Ausgangskategorien Siedlungsfläche, Auf- oder Abtragsfläche ausgewiesenen Kategorien, die bei einer Generalisierung keine sachliche Zuordnung zur Zielkategorie der übergeordneten Hierarchiestufe vorweisen und dementsprechend einen gesonderten Problemfall darstellen (s. 4.5.3.1 b).

<sup>15</sup> Im Anhang B wird in einem kleinen Exkurs im Detail auf diese Problematik eingegangen.

die Ausgangsdatenmenge beim Sternansatz immer gleich, wodurch sich im Gegensatz zum Leiteransatz kein Performanzgewinn beim Durchlaufen einzelner Maßstabsstufen erzielen lässt.

Dies betrifft allerdings in geringerem Maße einen rasterbasierten Generalisierungsworkflow, wie in diesem Fall beim GeoScaler-Ansatz, der in 3.2.2 vorgestellt wurde. Hier ist die eigentliche Algorithmen-Anwendung bzw. Prozessierung in beiden Fällen zellgrößen-basiert gleich.

Ein weiteres Argument für den Sternansatz liegt in der eindeutigen semantischen n:1-Aggregation der Einheiten, die auch zwischen den Stufen GE4 zu GE2 bzw. GE4 zu GE1 eindeutig definiert ist. Das erlaubt es, immer vom Ausgangsdatensatz mit der größten Auflösung auszugehen. Gegen den Leiteransatz spricht dagegen die Fehlerfortpflanzung in hierarchischen Systemen. Das Generalisierungsergebnis der ersten Stufe ist Eingangsdatensatz für die 2. Stufe. Damit arbeitet man mit den veränderten, geglätteten und gegebenenfalls fehlerhaften Daten weiter. Bei der Erzeugung der unterschiedlichen Ebenenlevel der geologischen Daten ist deshalb die Vorgehensweise nach dem Sternansatz zu bevorzugen. Die vorgegebene semantische Ableitung basiert zwar jeweils auf der vorhergehenden Detailstufe gleich eines Leiteransatzes, diese jeweiligen semantisch-klassifizierenden Zusammenfassungen sollten aber auf der Grundlage der Ausgangsdaten beruhen, da ansonsten auch die Gefahr besteht, dass nicht alle Einheiten oder geometrischen Objekte in den Generalisierungsvorgang einbezogen werden. Die Verwendung des Leiteransatzes wäre hier nur dann möglich, wenn in den Ursprungsdaten alle Objekte entsprechende Kennzeichnungen bezüglich ihres Vorkommens bzw. Erscheinungsbilds in den designierten Datenableitungen aufwiesen.

#### 4.5 Analyse der Ausgangsdaten

Die jeweils flächendeckend vorliegenden Verbreitungseinheiten der bodenkundlichen und der geologischen Karte sind über die in den Tabellen 9 und 11 genannten Kennwerte einer genaueren Betrachtung hinsichtlich ihrer inneren Struktur zu unterziehen.

Die in der Datenbank vorliegenden Grundlagenobjekte können direkt anhand einiger weniger SQL-Abfragen auf diverse Summenwerte in ihrer jeweiligen Aggregationsstufe abgefragt werden, ohne dass hierfür schon eine zeitaufwändige Geoprozessierung durchgeführt werden muss. Diese Statistiken liefern schon im Vorfeld Hinweise auf Parameter, die bei der späteren Generalisierungslösung von Beachtung sind. Im Fall der geologischen Daten ermöglicht die vorausblickende Analyse - exemplarisch ausgeführt auf den vorliegenden Gesamtdatensatz - schon Rückschlüsse auf Problemkategorien, die bei der späteren Generalisierungslösung bzw. bei einer erforderlichen Datenanreicherung unter besonderem Augenmerk stehen.

(VTX=Stützpunkte, LGTH=Perimeter [km]), AREA=Fläche [km<sup>2</sup>])

			GE_ID	ANZAHL	VTX	LGTH	AREA	VTX/LGTH
GE1	1	Alte Schiefer (aSF)	1196	129	16587	225,02	19,59	0,0136
	2	Variskische Gangmagmatite (GG)	1139	3799	159598	3009,76	51,63	0,0189
	3	Karbon (c)	1108	457	65500	876,46	74,33	0,0134
	4	Perm (p)	1059	1098	246526	3838,42	264,33	0,0156
	5	Variskische Plutone (GP)	1150	4354	756479	10814,12	1195,27	0,0143
	6	Tertiär (t)	196	5290	1189699	13524,49	1207,33	0,0114
	7	Gneis-Migmatit-Komplex (gn)	1209	5907	1030834	15044,62	1679,08	0,0146
	8	Jura (j)	370	22051	6768395	60618,24	4735,86	0,009
	9	Trias (tr)	649	30088	10514249	114943,47	9629,28	0,0109
	10	Quartär (q)	4	63111	26439077	216263,78	16534,36	0,0082

Tab. 12: statistische Werte auf Aggregationsebene GE1

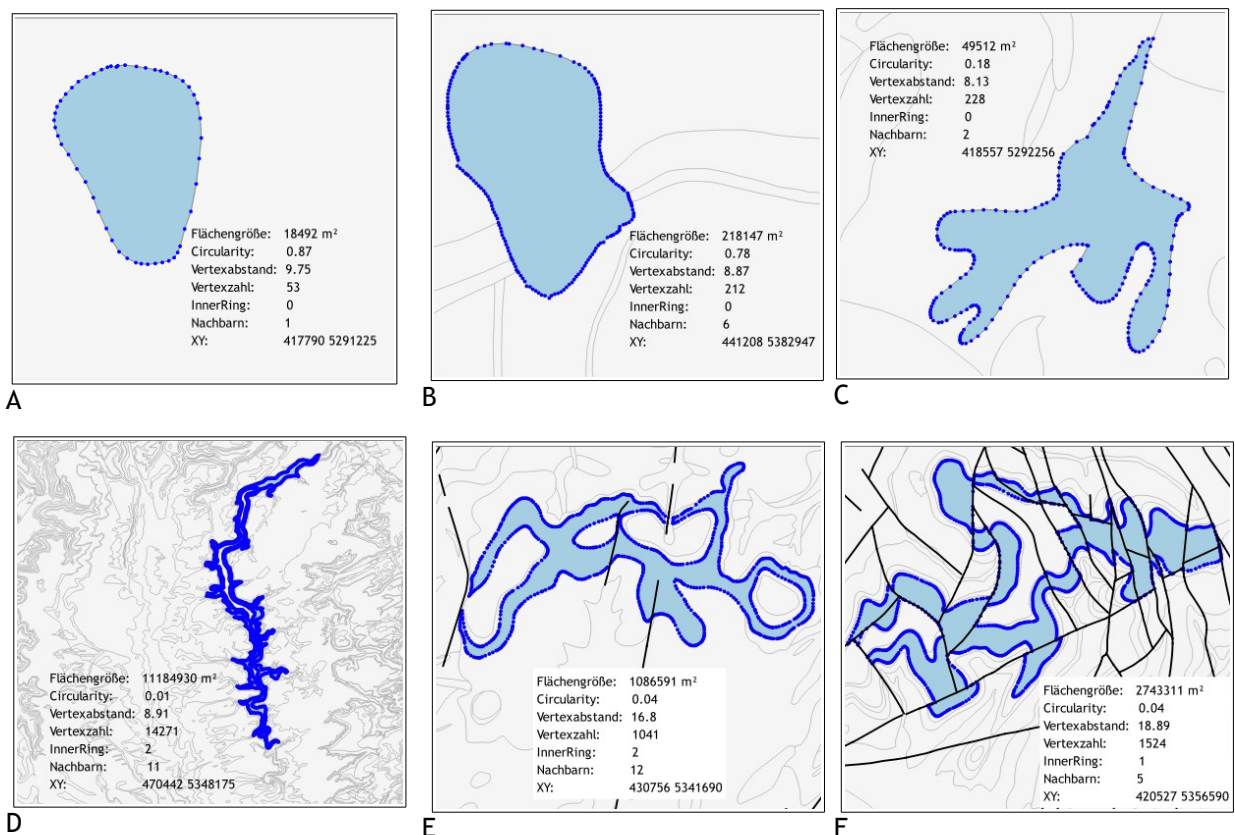
In diesem Fall kann aus Tab. 12 die Kategorie der Variskischen Gangmagmatite als Besonderheit identifiziert werden. Die Kennwerte aus der Tabelle weisen auf eine Vielzahl kleinteiliger Objekte hin, die gegebenenfalls im Zielmaßstab der Hierarchiestufe GE1 weitestgehend der Auflösung unterliegen.

Durch die in Abschnitt 4.4.1 geschilderte Problematik eignen sich die vorliegenden geologischen Daten der GK-BW allenfalls in den Teilbereichen mit durchgeführter Abdeckung für die Untersuchung hinsichtlich einer Generalisierungsautomatisierung.

#### 4.5.1 Komplexitätsunterschiede der Eingangsgeometrien

Bei den vorliegenden Objekten innerhalb des Polygonmosaiks besteht eine relativ große Bandbreite bezüglich der Komplexität ihrer Geometrien - das reicht von

- einfachen, kleinen, relativ runden Flächen, die lediglich einen inneren Ring innerhalb einer anders kategorisierten Fläche bilden,
- feingranularen, einfachen Geometrien mit einer größeren Anzahl von Nachbarn,
- organisch wirkenden Gebilden mit Einbuchtungen und Kurven,
- sich über eine größere Ausdehnung erstreckende, dendritische Gebilde,
- umringartigen Verbreitungsflächen mit einer Vielzahl an inneren Ringen und komplexen Nachbarschaftsbeziehungen,
- bis zu stärker durch die Tektonik zerklüfteten Strukturen, die wie ein ungeordnetes Puzzle wirken.



Tab. 13: Geometrien mit Stützpunktzahl und innerem/äußerm Topologiebezug

Die Auswirkungen dieser unterschiedlichen Komplexität bzw. ihr Einfluss auf die Performanz bei der Prozessierung der Objekte ist offensichtlich: jeder einzelne Stützpunkt geht in die transformativen Rechenprozesse einer Generalisierung mit ein, insofern kann sich eine vorherige Herabsetzung der Objektkomplexität schon anhand einer Stützpunktreduktion deutlich positiv auf die Performanz auswirken.

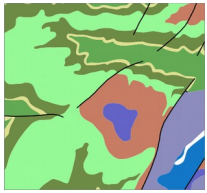
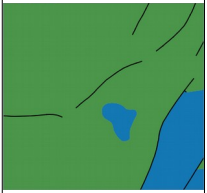

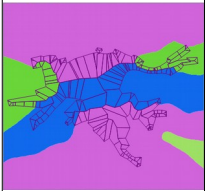
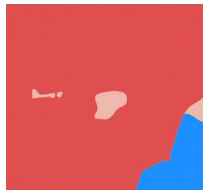
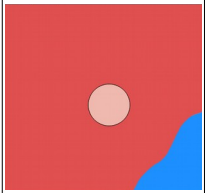
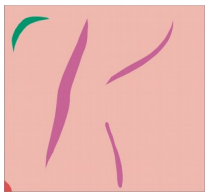

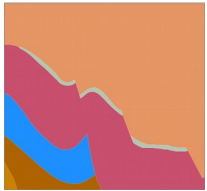
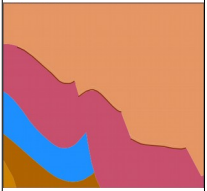




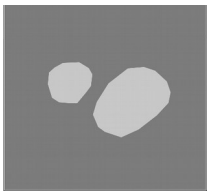




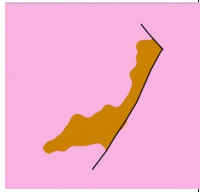
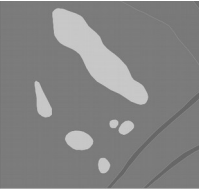

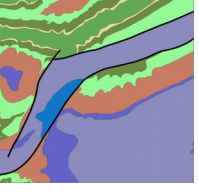

#### 4.5.2 Lokalisierung der unterschiedlichen Generalisierungssituationen/Konfliktfälle

Geologische Karten gehören zu den komplexesten thematischen Karten mit unterschiedlichsten z.T. diffizilen Formen und Strukturen, welche die Generalisierung verkomplizieren und entsprechend vorab eine eingehende Analyse solcher Strukturen erfordern (Sayidov & Weibel 2016). Gemäß der Einteilung in die hierarchischen Anforderungskategorien aus 1.4.1 zeigen sich entsprechend unterschiedliche Komplexitätsstufen der in den Ausgangsdaten auftretenden und zu lösenden Szenarien. In der nachfolgenden Zusammenstellung der Tab. 14 wurden vorhandene Generalisierungssituationen lokalisiert, kurz beschrieben, der zugrundeliegenden Anforderungsstufe<sup>16</sup> zugeordnet und mit entsprechenden Operatoren und Funktionen zur Lösungsfindung hinterlegt. Die in grau visualisierten Zeilen stellen dabei Generalisierungssituationen dar, die im Rahmen dieser Arbeit bzw. im Ergebnis-Workflow noch keine ausreichende Berücksichtigung fanden.

---

<sup>16</sup> Spalte Nr beinhaltet die zugrunde liegende Anforderungskategorie sowie eine fortlaufende Nummerierung.

Nr	Beschreibung	Vorgang / Funktion	Ausgangsbild	Ziel
B1	Einfache Zusammenfassung von Einheiten in neue Kategorien mit Grenzlinienauflösung → geometrisch-semantische Aggregation	aggregieren FME: dissolver		
B2	Null-Flächen ohne Kategorie → Zuordnung muss auf benachbarte Flächen aufgeteilt werden	Polygon-Partitionierung mit anschließender Aggregation PG: ST_StraightSkeleton FME: CenterLineReplacer		
C1	Objekt (Inselepolygon) unterhalb Mindestgröße → mit Beflagung: Geometrieübergang zu Punktsymbol ohne Beflagung: Auflösung	Punkt typisieren (kollabieren) FME: CenterPointReplacer		
C2a	Objekt unterhalb Mindestgröße ist als länglich identifiziert → mit Beflagung: Geometrieübergang zu Linie ohne Beflagung: Auflösung	Linie typisieren (kollabieren) FME: CenterlineReplacer		
C2b	Objekt unterhalb Mindestgröße, länglich, als geologische Bank identifiziert (jeweils Nachbar ober/unterhalb vorhanden) → Erhalt durch Beflagung	Linie typisieren (kollabieren) FME: CenterlineReplacer		
C3	Objekt unterhalb Mindestgröße mit mehreren Nachbarn unterschiedlicher Kategorien, soll anhand des lithologisch ähnlichsten Nachbarn aufgelöst werden (beinhaltet auch den Fall C2b)	Semantisch aggregieren FME: PythonCaller		
C4	Komplexe Objekte auf Makroebene → Formvereinfachung im Gesamtflächenkontext (starke Reduktion von Verästelungen und schmalen Fortsätzen)	Vereinfachen FME: Area Amalgamator AG: SmoothPolygon AG: GeoScaler		
C5	Disjunkte Objekte gleicher Kategorie innerhalb Mindestabstand, kumulierte Mindestgröße	Verschmelzen FME: Neighbour Finder, FME: Area Amalgamator AG: GeoScaler(?)		

Nr	Beschreibung	Vorgang / Funktion	Ausgangsbild	Ziel
D1	<i>Einbezug der tektonischen Linien als invariante Bereiche → Erhalt topologischer Bedingungen</i>	<i>AG: GeoScaler</i>		
D2	<i>Objektgruppe (Inselpolygone) mit bzw. unterhalb Mindestgröße, als Struktur ausgewiesen → Clustering</i>	<i>Verschmelzen</i>		
D3	<i>Grabenstrukturen - Abhängigkeit Bruchzone von tektonischen Linien als Abgrenzung Vereinfachung der Grabenzone anhand der Grabenränder</i>	<i>FME: Bufferer und CenterLineReplacer AG: GeoScaler</i>		

Tab. 14: Vorliegende Generalisierungssituationen

### 4.5.3 Generische Problemfelder

Zusätzlich zu den in der vorangegangenen Tabelle aufgeführten Generalisierungssituationen ergeben sich bei der praktischen Ausführung einige grundlegende und spezifische Herausforderungen, die es im Vorfeld genauer zu betrachten gilt.

#### a) Wegfallende Kategorien bzw. Umgang mit der Thematik Datenanreicherung

Den verschiedenen geologischen Klassen liegen naturgemäß unterschiedliche Ausgangsgeometrien mit sehr individueller räumlicher Verbreitung zugrunde. Schwierig zu handhaben in diesem Zusammenhang sind Klassen mit isolierten Einzelflächen geringer Größe, die durch die Generalisierung komplett eliminiert werden würden. Da sie inhaltlich erhaltenswerte Objekte darstellen, bedarf es einer Typisierung der Ausgangsgeometrien in Punkt- oder Linienobjekte. Die damit einhergehende Auswahl von Objekten zur Typisierung kann nicht algorithmisch, sondern nur durch Datenanreicherung eines zusätzlichen Attributs erfolgen. Die Werte dieses Zusatzattributes markieren bzw. "beflaggen" die erhaltenswerten Objekte und gewährleisten damit die inhaltliche Aussagekraft des Generalisierungsergebnisses<sup>17</sup>. Die dem späteren Workflow zugrunde liegenden kleineren Datenausschnitte wurden vorab dahingehend überarbeitet, um die exemplarische Handhabung besonderer Objekte zu simulieren (s. Kap. 5.3.2).

#### b) Auflösung von Flächen außerhalb der semantischen Hierarchisierung

Siedlungsflächen, aber auch Rohstoffabbauflächen und Gewässerflächen sind im Basismodell eigens ausgewiesene Einheiten, bleiben aber bei der semantischen Generalisierung ohne Zuordnung zu übergeordneten Kategorien - die Geometrien müssen also einer (oder mehreren) im Zielmaßstab vorhandenen Klasse(n) zuge-

<sup>17</sup> Ein Ergebnis dieser Arbeit ist die Erfordernis einer systematischen Datenanreicherung des landesweiten Geologiedatensatzes im Hinblick auf die Typisierung erhaltenswerter Objekte (Singularitäten). Dieser Arbeitsschritt ist an den Fachbereich Landesgeologie des LGRB adressiert und wird zeitnah vollzogen werden.

ordnet werden. Daraus ergibt sich die Aufgabe, wie diese Auf- oder Zuteilung der Ausgangsgeometrien zu den mitunter unterschiedlichen Kategorien der direkten und indirekten Nachbarschaft vorgenommen werden kann. Dies ist explizit ein anderer Fall als die Auflösung von mindergroßen Flächenobjekten mit Kategoriezuordnung, bei welchen der semantische Bezug (z.B. durch chronostratigrafische Einordnung) zu den umgebenden Einheiten zu berücksichtigen ist. Das Ausgangsszenario ist in beiden Fällen aber gleich: entweder hat das aufzulösende Polygon nur einen umgebenden Nachbarn, ist also lediglich der innere Ring eines anderen Polygon oder es grenzt an mehrere andere Flächenobjekte unterschiedlicher Kategorien. Im zweiten Fall wäre die Auflösung der ursprünglichen Geometrie also mit einer Neukonstruktion der inneren Grenzlinien bzw. Kanten verbunden, für die nun wiederum unterschiedliche Lösungsstrategien herangezogen werden können. Kap. 5.1 widmet sich diesem Teilproblem in einem eigenständigen Lösungsansatz durch Skelettlinien.

### *c) Bestimmung von Kenngrößen für die Durchführung der Generalisierung*

Die einzelnen Generalisierungsoperationen werden auch durch die Verwendung von geometrischen Kennwerten und Messungen gesteuert (Kap. 2.2.5). Die konkrete Festlegung solcher Kennwerte als Parameter ist eine zusätzliche Herausforderung, da die Qualität der Generalisierungslösung aufgrund falsch oder unzureichend festgelegter Parameterwerte abnimmt. In der Regel erfolgt die Optimierung der Kennwerte empirisch, d.h. durch festgelegte Testreihen. In Bezug auf die kartografischen Anforderungen, also auf zu verwendende Größen-Kennwerte hinsichtlich der Lesbarkeit, gibt es einschlägige Vorgaben bzw. Vorschlagswerte - nicht jedoch im Zusammenhang mit Kenngrößen, welche strukturelle Gegebenheiten erfassen oder eingrenzen.

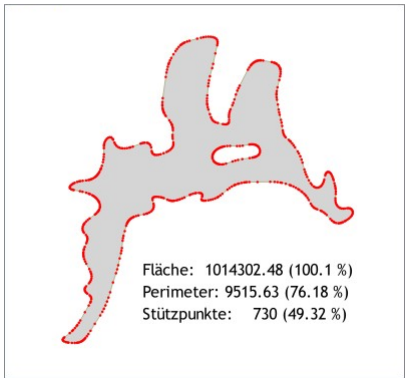
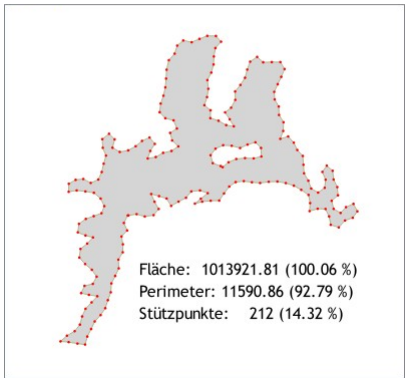
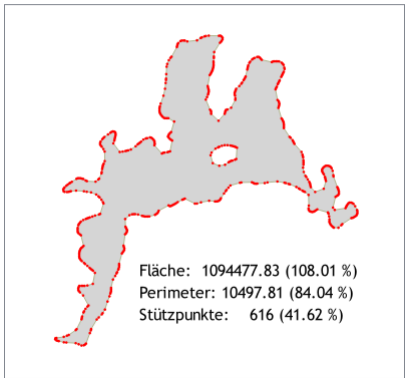
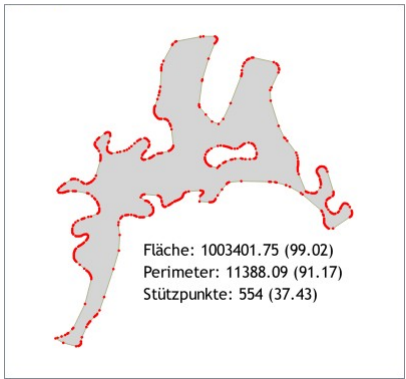
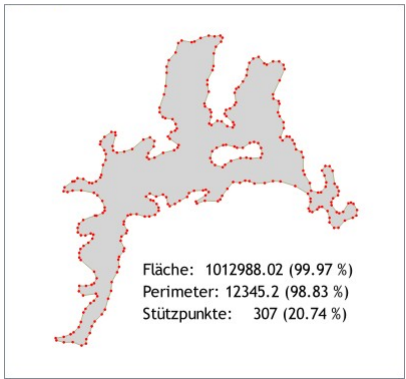
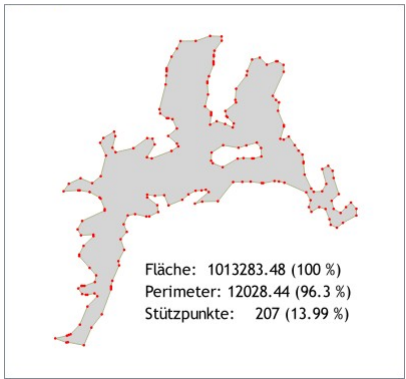
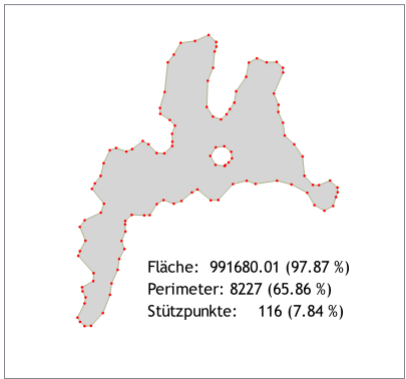
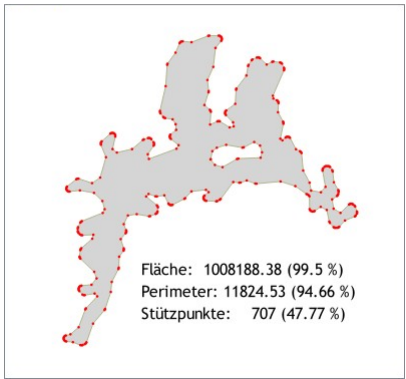
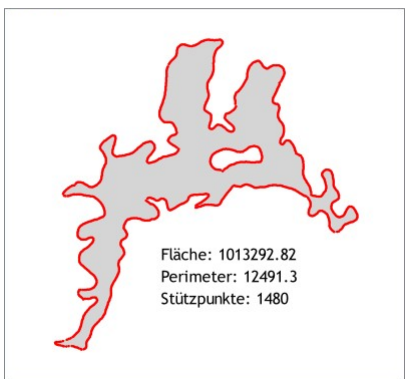
### *d) Simplifizierung versus Glättung (Generalisierung im Gesamtflächenkontext)*

Ein Generalisierungseffekt ergibt sich nicht nur in Bezug auf die Einzelgeometrie in ihrer Relation zur Nachbarschaft als Resultat entsprechend ausgeführter transformativer Operationen - ein wesentlicher Aspekt ist die Formvereinfachung der Objekte im Kartenkontext, also auf der Ebene des gesamten Polygonmosaiks. Hier zeigt sich das eigentliche Kernstück in der Findung einer gesamthaft wirkenden Vereinfachungsstrategie als Ersatz der ursprünglich manuellen oder teilautomatisierten kartografischen Wirkprozesse, welche ein möglichst homogenes und lesbares Generalisierungsergebnis unter Wahrung der topologischen Gegebenheiten zum Ziel haben. Die bloße Anwendung entsprechender Simplifizierungsalgorithmen ist hierfür nicht ausreichend.

Anhand der vorliegenden Datenthematik und der im Arbeitsumfeld zur Verfügung stehenden Softwarefunktionalität wurden Tests bezüglich der Eignung einschlägiger Simplifizierungswerkzeuge bzw. in der Praxis schon vorhandener Verfahren durchgeführt.

Bei den in Tab. 15 exemplarisch aufgeführten Verfahren handelt es sich um einzelne Algorithmen (B1, B2, C2), aus mehreren Einzelalgorithmen zusammengesetzten Methoden (B3), GIS-Funktionen (A1, A2, C1) sowie dem rasterbasierten CA-Verfahren aus der GeoScaler-Toolbox. Sie weisen hinsichtlich der Anzahl ihrer hinterlegten Eingabeparameter eine unterschiedliche Komplexität auf.

Tab. 15: Ausgewählte Simplifizierungsverfahren und ihre Wirkung bezüglich Stützpunktreduktion und Formvereinfachung

ArcGIS	PostGIS	FME
 <p>Fläche: 1014302.48 (100.1 %) Perimeter: 9515.63 (76.18 %) Stützpunkte: 730 (49.32 %)</p>	 <p>Fläche: 1013921.81 (100.06 %) Perimeter: 11590.86 (92.79 %) Stützpunkte: 212 (14.32 %)</p>	 <p>Fläche: 1094477.83 (108.01 %) Perimeter: 10497.81 (84.04 %) Stützpunkte: 616 (41.62 %)</p>
A1 SmoothPolygon	B1 RemoveRepeatedPoints	C1 AreaAmalgamator
 <p>Fläche: 1003401.75 (99.02 %) Perimeter: 11388.09 (91.17 %) Stützpunkte: 554 (37.43 %)</p>	 <p>Fläche: 1012988.02 (99.97 %) Perimeter: 12345.2 (98.83 %) Stützpunkte: 307 (20.74 %)</p>	 <p>Fläche: 1013283.48 (100 %) Perimeter: 12028.44 (96.3 %) Stützpunkte: 207 (13.99 %)</p>
A2 SimplifyPolygon	B2 SimplifyVW	C2 Generalizer (Deveau)
 <p>Fläche: 991680.01 (97.87 %) Perimeter: 8227 (65.86 %) Stützpunkte: 116 (7.84 %)</p>	 <p>Fläche: 1008188.38 (99.5 %) Perimeter: 11824.53 (94.66 %) Stützpunkte: 707 (47.77 %)</p>	 <p>Fläche: 1013292.82 Perimeter: 12491.3 Stützpunkte: 1480</p>
A3 GeoScaler-CA	B3 I/O-Buffer u. Simplify	Ausgangsdaten

Von den oben dargestellten Verfahren ermöglichen A1, C1, A3 und B3 eine über das Normalmaß (einer Stützpunktreduktion) hinausgehende Formvereinfachung, wobei lediglich die ArcGIS-Funktion SmoothPolygon (A1) und das ArcGIS-basierte GeoScaler-Verfahren (A3) zusätzlich auch die Möglichkeit zur Wahrung des topologisch-thematischen Bezugs im Flächenkontext beinhalten.

#### 4.5.4 Festlegung von Testbereichen

Aufgrund der Größe bzw. Datenmenge der Ausgangsdaten sind für die Implementierungsphase Trainingsbereiche bzw. kleinere, repräsentative Ausschnitte aus dem Gesamtdatenbestand zu definieren, die aber inhaltlich die vorhergehend identifizierten Generalisierungsszenarien und Problemstellungen weitestgehend enthalten. Darüber hinaus waren die Testbereiche solcherart zu wählen, dass sie auf den schon für die abgedeckte geologische Karte bearbeiteten Daten basieren. Zum Zeitpunkt der Tests war das bei ca. der Hälfte des Datenbestandes der Fall. Die exemplarische Anwendung entsprechender Operationen und Algorithmen auf die Testbereiche soll die Lösungsfindung erleichtern, ohne durch zusätzliche Besonderheiten später nicht mehr relevanter Konfliktsituationen beeinflusst zu werden. Von den insgesamt 29, nach vordergründig visuellen Kriterien festgelegten Testbereichen, waren 15 schon hinsichtlich der quartär-geologischen Abdeckung bearbeitet, von diesen wiederum wurden 5 aufgrund der geologischen Gegebenheiten als aussagekräftige Bereiche exemplarisch ausgewählt und für den weiteren Ablauf vorbereitet.

Box 1: ca. 144 km<sup>2</sup>

Box 2: ca. 128 km<sup>2</sup>

Box 3: ca. 188 km<sup>2</sup>

Box 4: ca. 117 km<sup>2</sup>

Box 5: ca. 64 km<sup>2</sup>

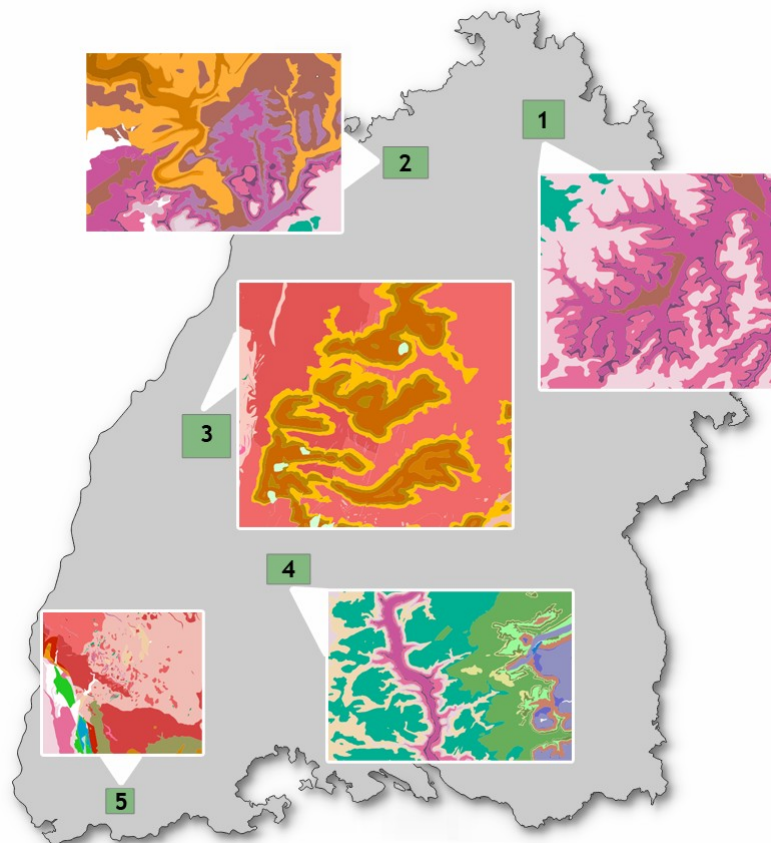


Abb. 13: Lage der Testbereiche

## 5 Methodenimplementierung

Die zuvor erzielten Erkenntnisse über Funktionsweise und Einsatz der verfügbaren Methoden und Funktionalitäten sowie die anhand der Tests gemachten Erfahrungen mit den vorliegenden Ausgangsdaten mündeten in unterschiedliche Überlegungen zur Umsetzung in einen Generalisierungsworkflow. Dabei spielten die unterschiedlichen Anforderungsszenarien aus Kap. 1.4.1 eine zentrale Rolle. In einem ersten testweisen Realisierungsversuch (s. 5.1) wurde die Anforderungskategorie B für eine Generalisierung von bodenkundlichen Gruppenkategorien ohne topologischen Bezug zueinander zugrunde gelegt. Der dort skizzierte einfache skriptbasierte Lösungsansatz diente für eine erste Auslotung der Möglichkeiten und Grenzen in einem relativ simplen Generalisierungsablauf und besitzt als pragmatischer Ansatz durchaus Weiterentwicklungspotenzial. Das Augenmerk im Ablauf lag neben der Anwendung des I/O-Buffer vor allem in der Implementierung der Polygon-Partitionierung auf Grundlage von Skelettlinien. Für eine topologisch konsistente Generalisierung der Flächen für die Anforderungskategorie C ist der Ansatz aber nicht ausreichend, insofern erfolgten die weiterführenden Tests anhand der geologischen Fachdaten mit dem schon in der Praxis vorhandenen GeoScaler-Verfahren bzw. dem zellulären Automaten (s. 5.2). Die Erkenntnisse führten zu einem gesamtheitlichen Lösungsansatz mit der Software FME (s. 5.3), welcher einen durchgehend automatisierten Prozessablauf in Art eines Workflows bereitstellt.

### 5.1 Generalisierung der bodenkundlichen Daten mittels Python/PostGIS-Geoprocessing

Beschrieben wird ein einfacher Generalisierungsansatz für die Prozessierung der bodenkundlichen Geofachdaten mittels PostGIS und Python. Die im GIS-Bereich weit verbreitete Skriptsprache Python bildet hierbei das Framework für die Steuerung der Prozesse, die anhand diverser Datenabfragen und -manipulationen SQL-basiert direkt in der PostGIS-DB mittels verfügbarer PostGIS-Funktionen durchgeführt werden.

Folgendes Ablaufprinzip kommt bei der Generalisierung der bodenkundlichen Einheiten zum Einsatz:

1. Zuweisung der designierten Zielkategorie anhand Zuordnungsschlüssel gemäß der hierarchischen Gliederung
2. Aggregation benachbarter Flächen gleicher Zielkategorie (Grenzauflösung)
3. Herauslösen und Auffüllen von Null-Flächen (Polygon-Partitionierung - s. Abb. 14)
4. Kategoriebezogene Zusammenfassung und Vereinfachung anhand verschiedener Operationen
  - a) Inward/Outward-Buffer
  - b) Stützpunktreduzierung und Liniensimplifizierung
5. Verebnung der in der Reihenfolge ihres Flächenanteils gestapelten Kategorielayer

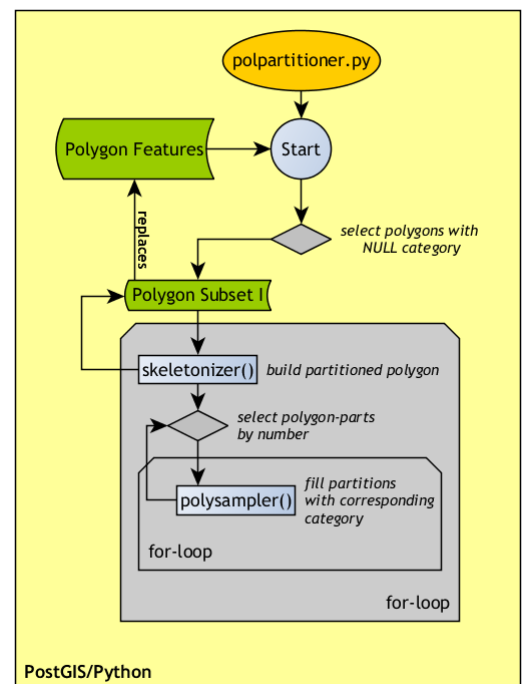


Abb. 14: Ablaufschema zum Teilprozess der Polygon-Partitionierung

### 5.1.1 Polygonaufteilung mittels Skelettlinien<sup>18</sup>

Als Lösungsansatz wird eine Flächenzuteilung auf Basis einer Polygonzerteilung mit Hilfe von Straight Skeletons implementiert. Straight Skeletons ist die Bezeichnung für einen Skelett-Algorithmus, der eine Polygonpartitionierung aus geraden Liniensegmenten ermöglicht. „Für die Konstruktion des Skeletts werden die Seiten eines Polygons schrittweise parallel nach innen versetzt. Jeder Schritt ist dabei so groß, dass der nächstfolgende Schnitt zweier Dachflächen erreicht wird. [...] Das Straight Skeleton liefert neben der linienförmigen Repräsentation des Objektes, die sich aus der Mittellinie des Skeletts, bzw. der Firstlinie ergibt, auch eine Unterteilung der ursprünglichen Flächen in Teilflächen. Jede dieser Teilflächen kann auf Grundlage der entsprechenden gemeinsamen Kante einem Nachbarn zugeteilt werden“ (Haurert 2004, 4). Auf dieser Basis lässt sich eine eindeutige Zuordnung einer Nullfläche zu unterschiedlichen Kategorien herstellen und eine inhaltlich sinnvolle Aufteilung der Teilflächen gewährleisten.

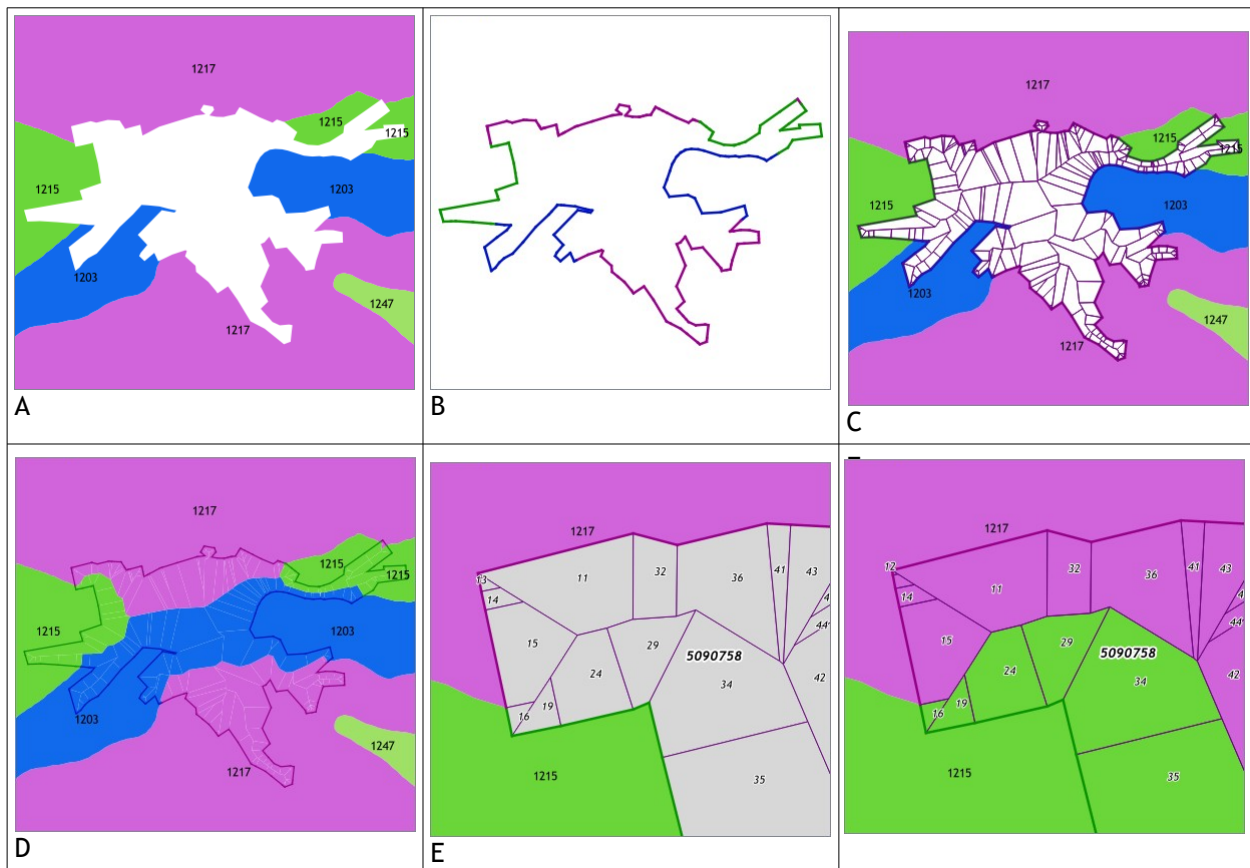


Abb. 15: Anwendung der Nullflächenzuteilung anhand des Straight Skeletons

Anhang C dokumentiert die Pythonskript-basierte praktische PostGIS/SQL-Implementierung.

Der Straight-Skeleton-Algorithmus begegnet uns an anderer Stelle noch einmal: Bei der Erzeugung linienhafter Repräsentationen gangartiger geologischer Flächen, also der geometrischen Generalisierungsoperation Kollaps, kann das Verfahren ebenfalls erfolgreich angewendet werden.

<sup>18</sup> Eingedeutscht aus dem englischen Terminus „straight skeleton“



## 5.2 Anwendung der GeoScaler-Toolbox

Im Zusammenhang mit der hier vorliegenden Aufgabenstellung und den unter Kapitel 4.5.2 aufgeführten Ausgangssituationen ist die Anwendung des rasterbasierten zellulären Automaten als allgemeine Generalisierungsstrategie auf Makro-Ebene bzw. im Gesamtflächenkontext von besonderem Interesse. Eine der grundlegenden Schwierigkeiten der Vereinfachung zusammenhängender Polygonmosaik liegt im Erhalt von topologischen Beziehungen bzw. der topologischen Konsistenz. Bei einer rein vektorbasierten Prozessierung sind die zu berücksichtigenden Abhängigkeiten und Komplexitäten deutlich höher als bei einer rasterbasierten Prozessierung, in welcher angrenzende Kategoriebereiche zueinander per se topologische Konsistenz aufweisen (Peter & Weibel 1999). Aufgrund dieser Tatsache bzw. auch den in der Praxis gemachten Erfahrungen im Umgang mit den vektoruell z.T. extrem komplexen Objektgeometrien hinsichtlich der Notwendigkeit der Formvereinfachung in der weiteren Prozessierung, bietet die Verwendung dieses Ansatzes bei der Entwicklung hin zu einer relativ einfachen Generalisierungslösung grundsätzliche Vorteile.

### 5.2.1 Funktionsweise des zellulären Automaten (CA)

Ein zellulärer Automat (im Folgenden mit der gängigen Bezeichnung „CA“ abgekürzt, nach „cellular automaton“) ist definiert als ein räumlich und zeitlich diskretes dynamisches System, bestehend aus einem regulären Raster von gleichförmigen identischen Zellen (Rommeney 2006, 11). Jede dieser Zellen ist in der Lage, einen der Zustände aus einer endlichen Zahl von Zuständen anzunehmen. Der Prozess beginnt mit dem Ausgangszustand der Zellen. Der neue Zustand jeder Zelle wird anhand der aktuellen Zustände benachbarter Zellen nach einer Regel berechnet. Die Anzahl und Position benachbarter Zellen, deren Zustände die betrachtete Zelle beeinflussen, wird als Nachbarschaft bezeichnet. Die Nachbarschaften können unterschiedlich groß und geometrisch sein. Sobald Zustände für alle Zellen berechnet sind, wird das gesamte Array gleichzeitig aktualisiert, was zu einer neuen Generation von Zellzuständen führt (iterativer Prozess). Die drei wichtigsten Modellparameter sind (1) die Regel, auf deren Grundlage sich der neue Zellzustand errechnet, (2) die Nachbarschaft, die den Zellzustand beeinflusst und (3) der Zustand der Zelle.

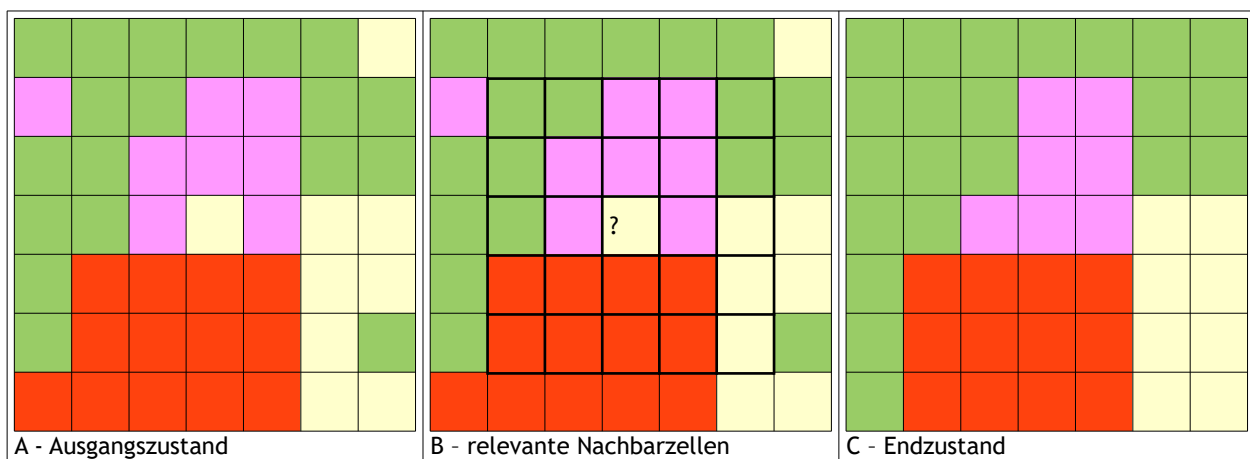


Abb. 16: Prinzip des zellulären Automaten (Smirnov et al. 2012)

Für die Polygon-Generalisierung ist das Array im Ausgangszustand ein Rasterbild der zu generalisierenden geologischen Karte (Abb. 57A). Die Zellen im Raster können eine endliche Anzahl von Zuständen annehmen, die der Anzahl der geologischen Einheiten entspricht, die auf der ursprünglichen Karte gefunden wurden und durch verschiedene Farben in der Abbildung dargestellt werden. Es wird die von Li, Wilkinson & Khaddaj (2001) vorgeschlagene Zustandsberechnungsregel verwendet, welche besagt, dass der Zustand unverändert

bleibt, wenn sich eine Zelle im gleichen Zustand wie die meisten Zellen in ihrer Nachbarschaft befindet - andernfalls wird er geändert, um den Zustand der Mehrheit wiederzugeben (am häufigsten vorkommender Zustand). Wie in Abb. 57B dargestellt, wird der Nachbarschaftsradius durch die Breite des Korridors beschrieben, der von den relevanten Nachbarzellen gebildet wird, die um die zu berechnende Zelle liegen (Moore's Nachbarschaft, d.h. 8 benachbarte Zellen um eine mittlere Zelle). Die letzte Generation ist definiert als diejenige, nach der keine weiteren Zustandsveränderungen mehr auftreten (Abb. 57C) (Smirnof et al. 2012).

### 5.2.2 Einbettung in der Toolbox

Die GeoScaler-Toolbox für ArcGIS ist untergliedert in die 5 Module Initialisierung (1), Verbreitungskarte (2a), Grundgebirgskarte (2b), Validierung (3) und Ergebnismanagement (4), die wiederum in weitere Teilprozesse aufgefächert sind. Das Herzstück der Flächengeneralisierung mit dem zellulären Automaten findet sich im jeweiligen Toolset „Geological Units“ als Teilprozess 02 in den Modulen 2a und 2b.

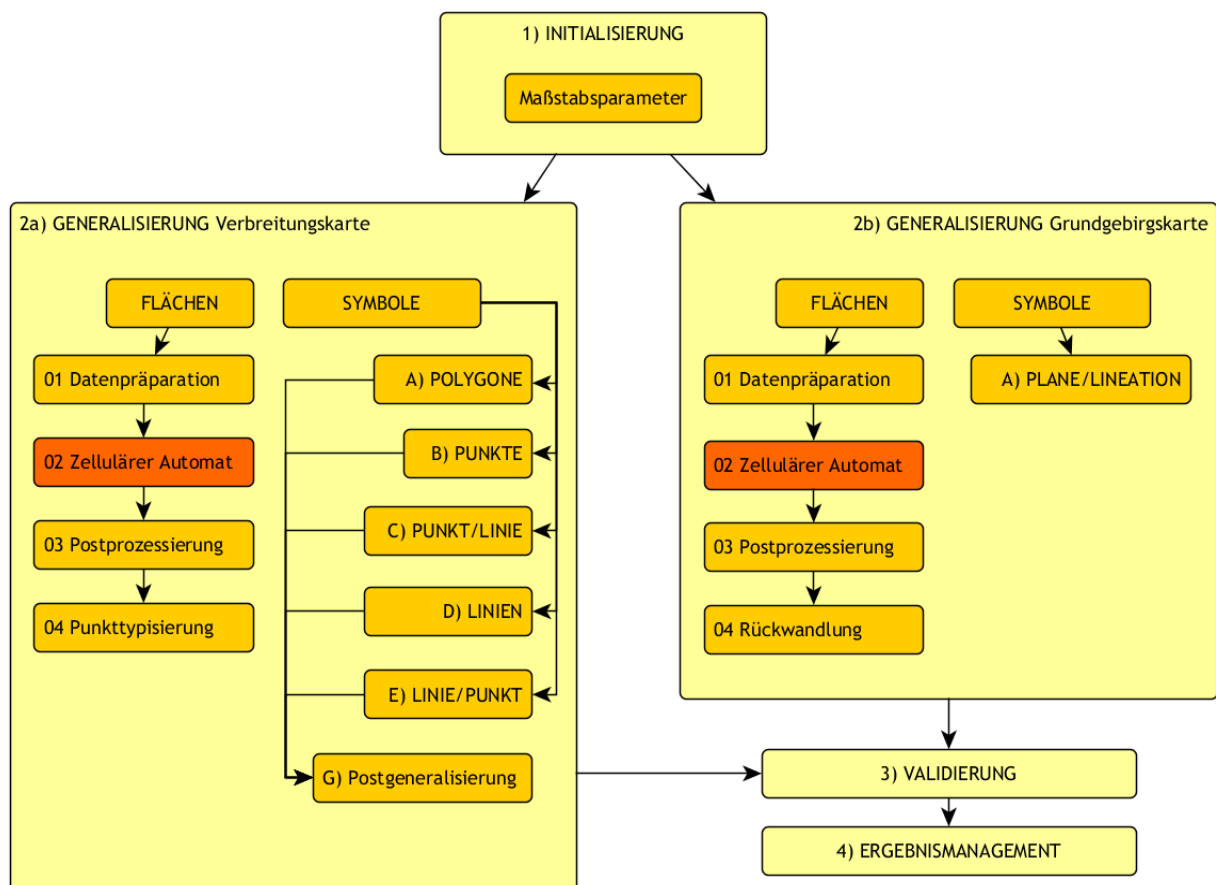


Abb. 17: Prozessmodule der GeoScaler-Toolbox

Diesem zentralen Teilprozess vorangestellt ist eine Vektor/Raster-Umwandlung der Eingangsdaten - also der Flächengeometrien mit/ohne hydrografischem Liniennetzwerk (bei der Verbreitungskarte) bzw. mit/ohne tektonischen Linienobjekten (bei der Grundgebirgskarte). Nach Ablauf des zellulären Automaten werden die Daten wieder in Vektorobjekte rückgewandelt. Die zusätzlich vorhandenen Toolbox-Elemente zur Verarbeitung von punkt- und linienhaften Kartenelementen bzw. deren grafischer Repräsentanz sind für unser Vorhaben von nachgeordnetem Interesse. Anders verhält es sich mit den Objekten, die aufgrund der Unterschreitung der Mindestgröße einer Typisierung (bzw. Kollabierung) vom flächenhaften hin zum linienhaften bzw. punkthaften Objekt unterliegen.

### 5.2.3 Generalisierungstests mit dem CA-Modul

Die Tests wurden auf Grundlage der vorab festgelegten Testbox 4 ausgeführt. Es handelt sich dabei um ein Gebiet, welches durch den tiefen Einschnitt des Neckarlaus in die umgebenden Muschelkalkplatten und Keuperschichten geprägt ist. Charakteristisch hierfür sind u.a. die Verästelungen der Seitentäler, die eine spezielle kartografische Relation im Bezug zur Referenztopografie durch das Gewässernetz beinhalten und im Falle einer höheren Anforderungsstufe (Kat. E) des expliziten Einbezugs und der gesonderten Beurteilung unterliegen müssten.

Bei der Testreihe im GeoScaler ergibt sich durch Anwendung der minimalen und maximalen Parameter bezüglich Radius und Iterationen der Rahmen des geringsten und größtmöglichen Veränderungsmusters (Abb. 18), in welchem sich die Formvereinfachung bewegt (s. Anhang D für gesamte Testreihe).

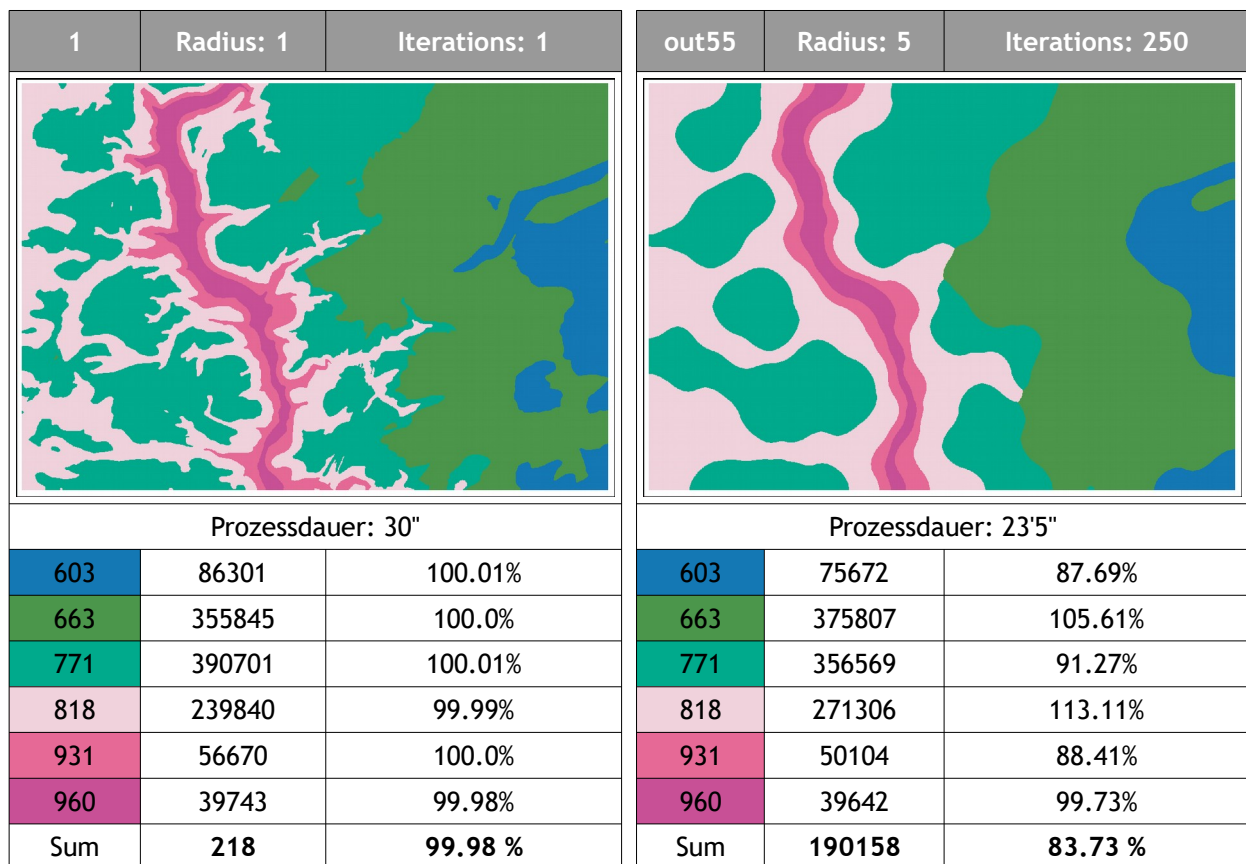


Abb. 18: Minimal- und Maximalwerte bei exemplarischer Durchführung des CA-Moduls im GeoScaler

(Anzahl der Rasterzellen je Kategorie sowie absoluter und prozentualer Summenwert der veränderten Rasterzellen zum Ausgangsdatensatz)

Ausgangsdaten Testbox 4: 1299x900 Pixel,  
Hierarchiestufe GE3  
Angewandte GeoScaler-Parameter:  
Zielmaßstab: 1 : 250 000  
Zellgröße: 10 m  
Mindestflächengröße: 62500 m<sup>2</sup>

Die Bandbreite der Flächengrößenveränderung über die vorkommenden Klasseneinheiten des Testbereichs in der Abbildung liegt dabei zwischen 0,02 und 16,27 %, die Formveränderung bzw. der Generalisierungseffekt äußert sich im Verschwinden von Klassenbereichen unterhalb der eingegebenen Mindestgröße und in der Verkürzung und Abrundung länglicher Strukturen, welches mit der Erhöhung von Radius und Anzahl der Durchläufe entsprechende Verstärkung findet (wobei letzteres den stärkeren Einfluss hat). Die Festlegung der Zellgröße und der abzubildende Extent sind zusätzlich zum Nachbarschaftsradius und der Anzahl der Iterationen weitere grundlegende Faktoren hinsichtlich des Generalisierungseffektes sowie auch der Prozessierungszeit,

die der zelluläre Automat in Anspruch nimmt. Ebenso wirkt sich der Einbezug der als invariant zu behandelnden Themenlayer der tektonischen Lineaturen als performanzmindernd aus.

	Ausgangsdaten	Radius 2 Iter 200	Radius 3 Iter 50	Radius 3 Iter 120	Radius 3 Iter 200	Radius 4 Iter 50
Zeit	-	15' 15"	3' 56"	10' 20"	15' 20"	4' 23"
<b>603</b>	86294	100.16%	99.23%	98.36%	98.07%	96.06%
<b>663</b>	355853	99.85%	100.19%	100.98%	100.49%	101.08%
<b>771</b>	390664	100.64%	100.18%	100.95%	99.02%	97.56%
<b>818</b>	239865	99.95%	100.47%	99.7%	102.3%	104.61%
<b>931</b>	56673	97.42%	98.78%	93.73%	98.57%	97.09%
<b>960</b>	39751	98.71%	97.21%	96.15%	97.63%	99.22%
<b>Sum</b>	1169100	<b>98.84 %</b>	<b>96.11 %</b>	<b>94.58 %</b>	<b>93.25 %</b>	<b>90.27 %</b>

Tab. 16: Verkürzte Testreihe mit quantitativen Aussagen zu Prozessdauer und Flächenanteile der geänderten Kategorien

Bezogen auf den Gesamtdatensatz mit der Landesfläche Baden-Württembergs, belaufen sich die Prozessierungszeiten auf mehrere Stunden, was eine Durchführung des Workflows über Nacht nahelegt. Nicht abgebildet in der Tab. 16 ist der Einfluss weiterer Gründe für die Verlangsamung im Prozess. Vor allem die Löschung temporärer Datensätze am Ende des Ablaufs nimmt je nach Anzahl der Durchläufe bis zu 20 % der Gesamtprozessdauer in Anspruch, ebenso sind die Werte nicht bereinigt von der Auswirkung, die eine jeweils zur Laufzeit gegebene Recherauslastung durch andere System- und Softwareprozesse bewirkt.

### 5.2.4 Entkapselung der CA-Prozessierung

Aufgrund der in den Tests erzeugten Ergebnisse mit dem zellulären Automaten ist dessen gezielte Implementierung innerhalb eines gesamtheitlich automatisierten Ablaufs eine erstrebenswerte Lösung. Für die anderweitige Verwendung der relevanten Geoprocessingsschritte ist eine Entkapselung und Herauslösung der entsprechenden Teilprozesse aus der Arc-Toolbox erforderlich. Es handelt sich dabei um Python-Skripte, die aus dem jeweiligen Toolscript exportiert werden können und durch einige wenige Anpassungen als Stand-Alone-Skripte fungieren können. Der detaillierte Blick auf die einzelnen Skripte, die jeweils benötigten Eingabedaten und -parameter sowie die Verzahnung der einzelnen Abläufe offenbart die Wirkungsweise und das Zusammenspiel der einzelnen Komponenten (Abb. 19). Die Abbildung nimmt Bezug auf das Modul zur Generalisierung von geologischen Grundflächen. Im Einzelnen handelt es sich um 4 Python-Skripte, die in sequentieller Abfolge die einzelnen Teilschritte ausführen.

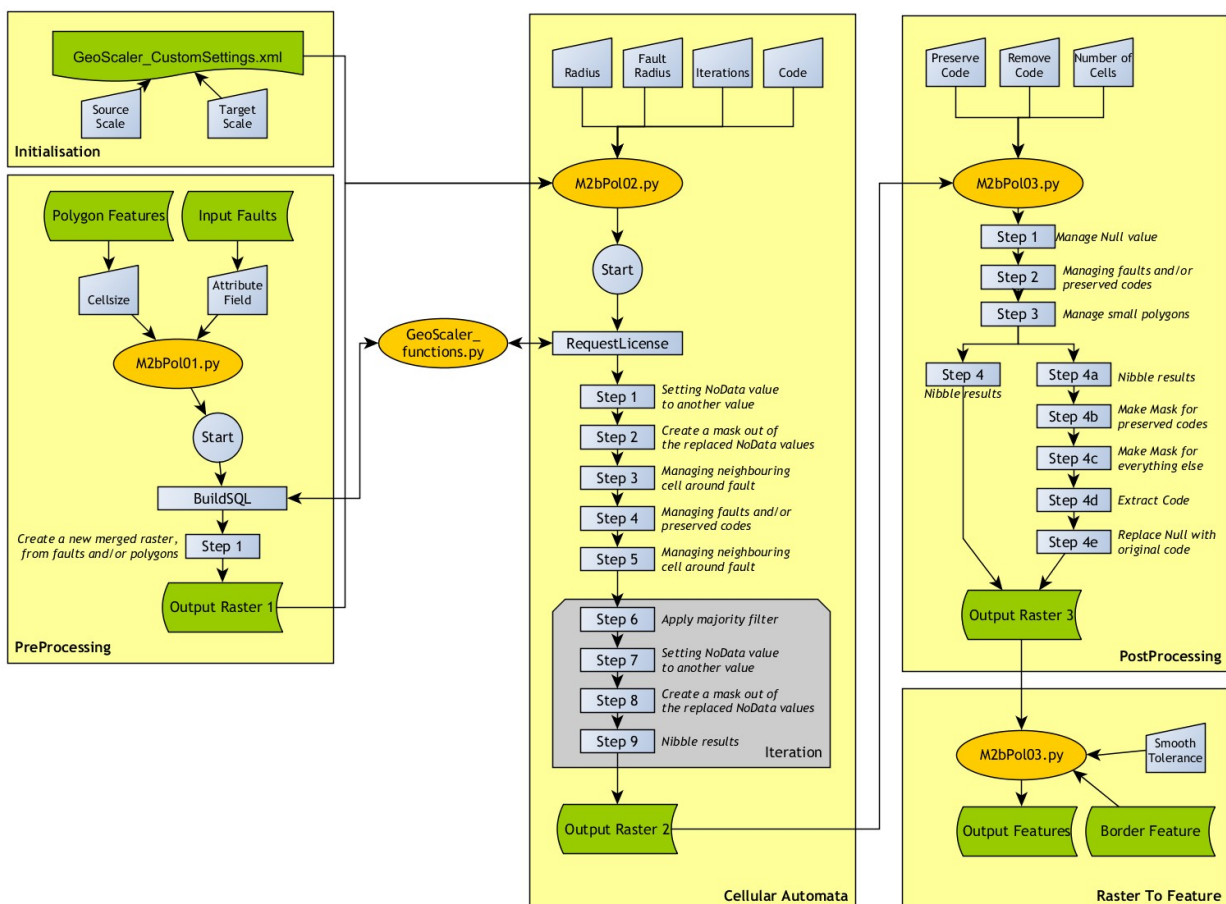


Abb. 19: Detailliertes Ablaufschema innerhalb des GeoScaler-Moduls 2b (s. Abb. 17)

Die semantisch aggregierten Eingangsdaten werden nach der Initialisierung im ersten Schritt (Skript M2b-Pol01) in ein Raster gewandelt (Abb. 20a + b). Im zweiten Skript (M2bPol02) erfolgen zuerst vorbereitende Schritte zum Management invarianter Zellbereiche (Step 1 bis 5). Danach folgt der iterative Durchlauf (Step 6 bis 9) in welchem die ArcGIS-SpatialAnalyst Funktionalitäten FocalStatistics mit angewandtem Majority-Filter und anschließendem Nibble<sup>19</sup> zur Anwendung kommen. Das resultierende Raster c) wird in einem weiteren Skript (M2bPol03) für die Rückumwandlung vorbereitet: den Barriere-Zellen werden wieder reguläre Ka-

<sup>19</sup> Nibble bedeutet übersetzt „knabbern“, was die Funktionsweise des CA sehr schön beschreibt: in jeder Schleife wird von den Ausgangsdaten ein Stückchen mehr „weggeknabbert“

tegorien zugewiesen, wahlweise kann auch die gesonderte Handhabung einzelner Kategorien und ihrer Restflächen in diesem Teilschritt Berücksichtigung finden. Das post-prozessierte Raster d) wird in einem abschließenden Skript (M2bPol04) wieder in ein Vektorfile zurückgewandelt (e). Bei Inselkarten wird das Raster anhand eines Grenz-Datensatzes auf den relevanten Bereich beschnitten.

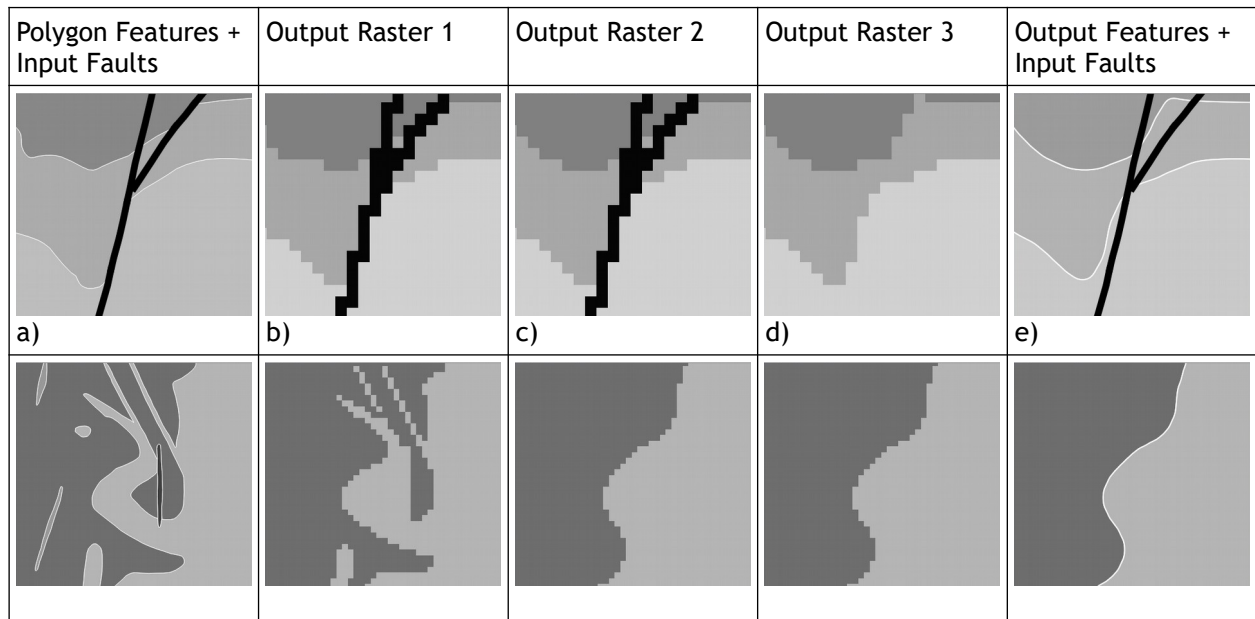


Abb. 20: Vektor - Raster - Vektor - Skriptdurchlauf im GeoScaler

Die obere Reihe (ca. M 1 : 15 000) in Abb. 20 zeigt die Datenzustände innerhalb des Prozessdurchlaufs in einem Bereich, in welchem tektonische Linien als Barrieren fungieren. Die untere Reihe (ca. M 1 : 30 000) dokumentiert die starke Formveränderung in Bereichen ohne einschränkenden Tektonik-Bezug.

### 5.3 Lösungsansatz durch FME-Workflow

Die Umsetzung des Gesamtprozesses erfolgte mit FME. Die Entscheidung hinsichtlich der Generalisierungsstrategie im Gesamtflächenkontext (zwischen dem GeoScaler-CA und dem FME-Transformer AreaAmalgamator) war zum Beginn der Umsetzungsarbeiten in FME noch nicht abschließend getroffen. Die Möglichkeit der Implementierung externer (ArcGIS-)Methoden sowie die anderweitigen Vorteile (gegenüber dem ModelBuilder) hinsichtlich der Eignung als Prozess-Framework gewährleistete letztlich die hierfür erforderliche Flexibilität. Technische Unterstützung bei der Umsetzung wurde durch die Firma Conterra geleistet<sup>20</sup>. Tests mit dem FME-Transformer AreaAmalgamator führten nicht zu den gewünschten Ergebnissen, vor allem der schwierig zu handhabende Aspekt der topologischen Konsistenz und der in einer künftigen Ausbaustufe gewünschte perspektivische Einbezug der tektonischen Relationen machten schließlich die CA-Methode zum zentralen Bestandteil im Prozessablauf. Die im vorhergehenden Kapitel beschriebene Generalisierungslösung durch die GeoScaler-Toolbox führt in Bezug auf den gesamtheitlichen Generalisierungsgrad zu durchaus guten Ergebnissen. Der Nachteil des GeoScaler, dass die Prozessablaufkette in ArcGIS quasi nur teilautomatisiert in aufeinander aufbauenden Einzelschritten durchführbar ist, wird durch die Implementierung in einen FME-Workflow aufgehoben. Das GeoScaler-Werkzeug wurde vordergründig unter dem Aspekt einer kartografischen Generalisierungslösung entwickelt, der Einbezug der Kartensymbolik und die damit verbundenen Ablaufschritte in Bezug auf Darstellung, Ausdünnung und Clusterung von punkt- und linienhaften geologischen Kartenelementen sind für die hier gesuchte Lösung von mindermem Interesse - gleichwohl sind die operativen Schritte zur Typisierung (Kollaps) bestimmter erhaltenswerter Flächengeometrien in Linien- oder Punktform auch bei der hier vorliegenden Zielsetzung und Fragestellung im weiteren Sinn perspektivisch relevant.



Abb. 21: FME-Workbench mit den schematisch hervorgehobenen Teilprozessen

Die resultierende Workbench ist modular strukturiert, entsprechende grau hinterlegte Teilbereiche in Abb. 21 (1-6) heben dies hervor. Die für die Prozessierung notwendigen Eingabeparameter (*Published Parameter*) werden vollumfänglich zum Start des Ablaufs eingegeben bzw. über Standardwerte definiert (s. Tab. 17). Die weitere Abfolge entspricht dem nachfolgenden Ablaufschema unter Verwendung der den entsprechenden Generalisierungsoperatoren gleichzusetzenden Transformern *Dissolver* (2 u. 5 - s. Abb. 22 A), *CenterPointReplacer* und *CenterLineReplacer* (4 - s. Abb. 22 C), ergänzt durch die wesentlichen Teilprozesse zur Handhabung der Objektauflösung von Einzelobjekten (5 - s. Abb. 22 D) und der geometrischen Vereinfachung durch die jeweilige Implementierung von Logik und externem CA-Verfahren mittels *PythonCallern* (6 - s. Abb. 22 F), worauf im nächsten Abschnitt noch einmal näher eingegangen wird. Die anzuwendenden Regeln und Bedingungen werden dabei an den entscheidenden (Schnitt)stellen durch sogenannte *Tester* gesteuert. Transformer aus der Gruppe der *Calculator* (3 - s. Abb. 22 E) sowie der *GeometryValidator* (1) und der *SpatialRelator* (s. Abb. 22 B) ermöglichen dabei die Überprüfung der Objekte auf die in diesem Zusammenhang erforderlichen Kennwerte, der geometrischen Validität und ihres topologischen Bezugs.

<sup>20</sup> LGRB-internes Projekt „Generalisierungsautomatisation“ unter Mitarbeit der FA. Conterra (Schuff, Hagemeister & Dupke 2018).

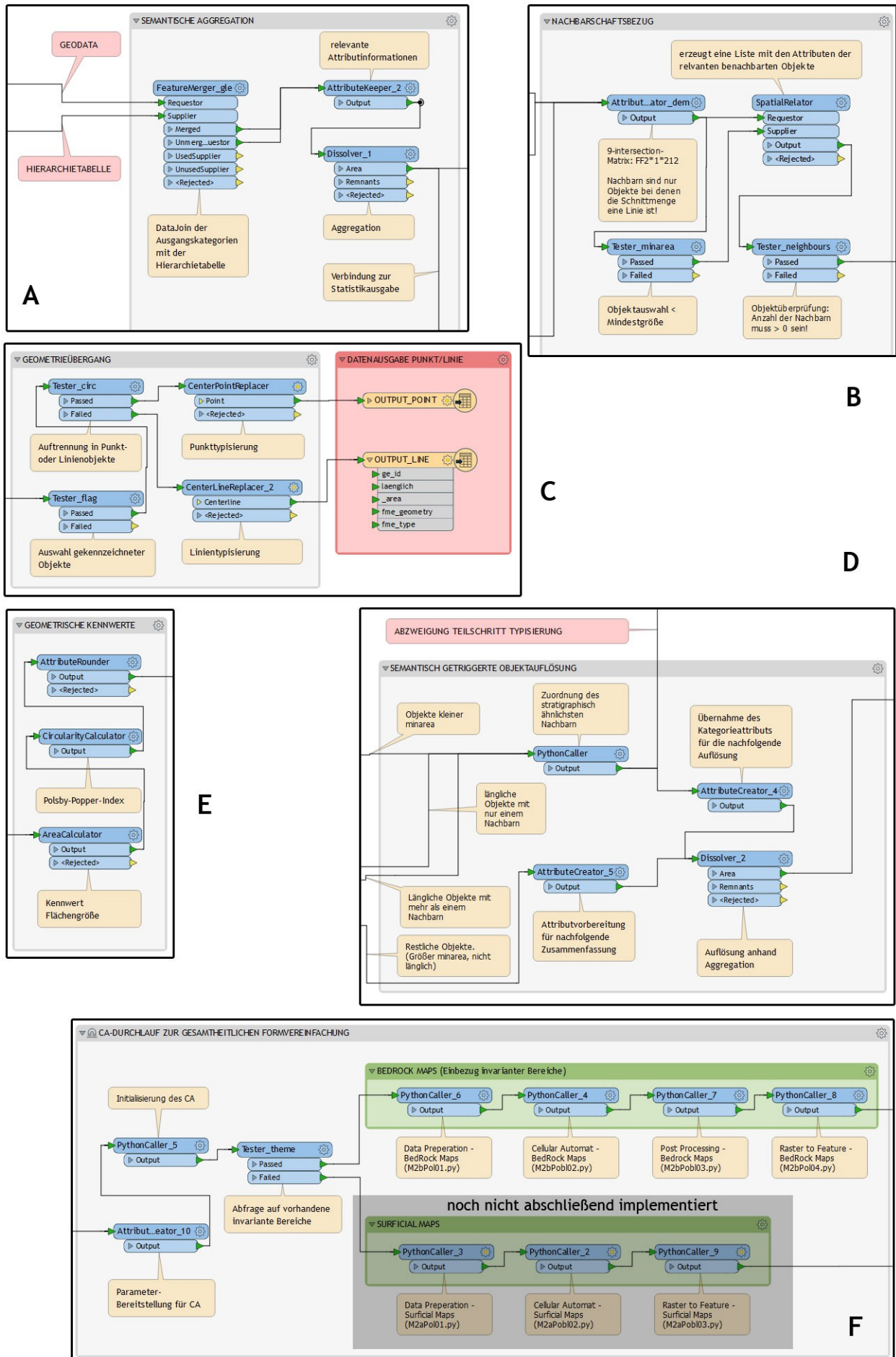


Abb. 22: verschiedene Teilmodule im Gesamtablauf der FME-Workbench



## 5.3.1 Ablaufschema

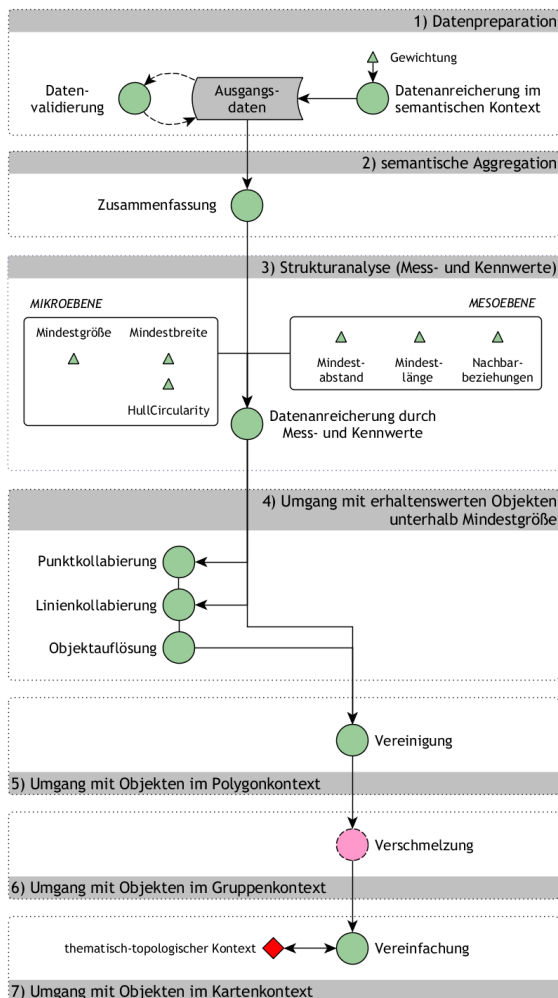


Abb. 23: Ablaufschema des designierten Workflows

Am Beispiel der flächenhaften geologischen Ausgangsdaten wurde der durchgehend automatisierte Ablauf folgendermaßen schematisch in Art einer Ereignis-Prozess-Kette (EPK) strukturiert:

1. Datenanreicherung der Ausgangsdaten durch Kennzeichnung fachspezifischer Objekte
2. vorgegebene semantische Aggregation mit nachfolgender geometrischer Grenzlinienauflösung  
→ Zusammenfassung
3. Strukturanalyse durch Mess- und Kennwerte (Größe, Circularity, Nachbarschaftsbeziehungen), Sortierung nach Flächengröße
4. Identifizierung und Durchlauf aller länglichen Geometrien
  - a) Objektgewichtung  $> 0$  oder  $>$  Mindestlänge  
→ Linienkollaps
  - b)  $<$  Mindestgröße und  $<$  Mindestabstand:  $\text{sum} >$  Mindestlänge  
→ Linienkollaps
5. Selektion von disjunkten Geometrien gleicher Kategorie  $<$  Mindestabstand
  - a) kumulierte Größe der Kandidaten  $>$  Mindestgröße  
→ Verschmelzung
6. Selektion und Durchlauf aller anderen Geometrien  $<$  Mindestgröße
  - a) Inseletpolygon
    - i. Objektgewichtung  $> 0$   
→ Punktollaps
    - ii. Auflösung nach umgebendem Nachbarn  
→ Zusammenfassung
  - b)  $> 2$  Nachbarn
    - i. Auflösung nach Kandidat mit lithologischer Ähnlichkeit/Abfolge
    - ii. Auflösung mit Polygon-Partitionierung durch Delauney-Triangulation
    - iii. Auflösung nach Kandidat mit größter Fläche
    - iv. Auflösung nach Kandidat mit längster Grenzlinie
7. Durchlauf Geometrien  $>$  Mindestgröße
  - a) ohne Objektrelation → Simplifizierung
  - b) Objektrelation vorhanden, erweiterter Kontext  
→ Glättung

### 5.3.2 Datenanreicherung

Die flächenhaften geologischen Ausgangsdaten wurden mit einem zusätzlichen Attribut „Flag“ versehen - erhaltenswerte Flächenobjekte mit dem Wert 1 werden somit unabhängig von ihrer Flächengröße einem gesonderten Verfahren unterzogen, d.h. bei Unterschreitung der Mindestgröße in Form eines Punkt- oder Linienobjekts weiter im Zieldatensatz erhalten bleiben.

Die tektonischen Lineationen erhalten in gleicher Weise das Attribut „Flag“. Hier werden die tektonischen Störungen anhand ihrer Bedeutung für die jeweilige Maßstabsstufe einer Kategorie von 0 bis 3 zugeordnet: 0 bedeutet höchste Relevanz bis hin zur größten Auflösungsstufe - es handelt sich um Hauptverwerfungslinien, d.h. diese Linienobjekte sind maßstabsunabhängig gegenwärtig; Linien der Kategorie 1 sind für die Hierarchieebenen GE4 bis GE2 maßgeblich, Kategorie 2 für GE4 bis GE3 und Kategorie 3 als minder-bedeutende Störungen nur in der detailliertesten Ebene GE4 präsent. Innerhalb des Prozessablaufs werden die Linienobjekte mittels einer Testklausel auf die zugrunde liegende Generalisierungsstufe und die in diesem Zusammenhang maßgeblichen Attributwerte (z.B. Generalisierungsstufe 1 => Hierarchieebene GE3 => „Flag“ in (0,1,2)) überprüft und entsprechend aus dem Gesamtdatensatz selektiert. Im weiteren Ablauf gehen diese Linienelemente als invariante Bereiche in die rasterbasierte Prozessierung durch den zellulären Automaten ein.

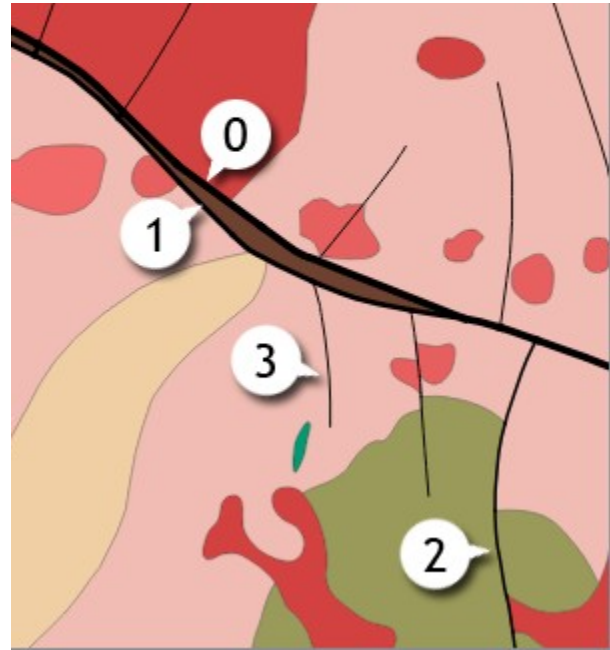


Abb. 24: Exemplarische Kategorien des Datenattributs „Flag“ bei tektonischen Linien

### 5.3.3 Objektauflösung

Eine Besonderheit im Prozessablauf ist der Einbezug des semantischen Kontexts bei der Flächenzuteilung der aufgrund Mindergröße aufzulösenden Objekte, die von mehr als einem angrenzenden Nachbarn umgeben sind. Die Zuteilung sollte idealerweise nicht nur nach topologischen oder geometrischen Gesichtspunkten, sondern in diesem Fall in erster Linie nach semantischer Ähnlichkeit durchgeführt werden (Ai & van Oosterom 2002). In der FME-Workbench wird dies durch einen sogenannten Python-Caller implementiert, der mittels Python-Funktion (s. Anhang E) die jeweiligen Kandidaten des aufzulösenden Objektes anhand der in einer Listenmatrix abgerufenen Werte bezüglich des a) Abstands in der Legendenhierarchie (inhaltliche Ähnlichkeit bzw. inhaltliche Nachbarschaft/Distanz), b) des Abstands zur übergeordneten Einheit und c) der Flächengröße untersucht und die Objektfläche dann dem lithostratigrafisch ähnlichsten Kandidaten zuweist.

Abb. 25 verdeutlicht die Wirkungsweise des Python-Algorithmus am Beispiel des als Requestor bezeichneten

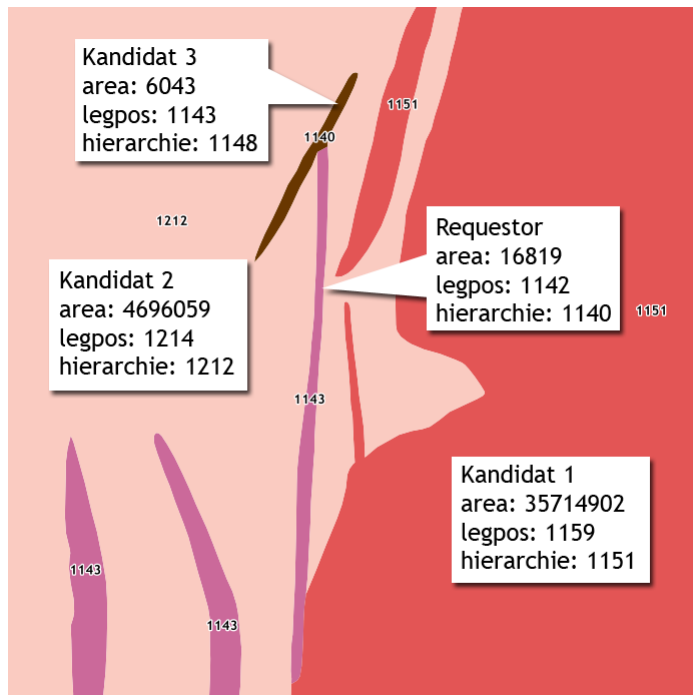
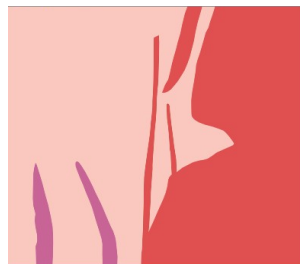


Abb. 25: Fallbeispiel für die Auflösung von Objekten nach ihrem semantischem Bezug

kl. Abb. Rechts: nach erfolgter Objektauflösung/-zuordnung



länglichen Polygons, welches aufgrund fehlender Beflagung und der Mindergröße < 20 000 m<sup>2</sup> zur Auflösung/ Eliminierung vorgesehen ist. Von den drei angrenzenden Nachbarn ist Kandidat 3 aufgrund der lithostratigraphischen Abfolge (Abstand zur Legendenposition *legpos*) am ähnlichsten, scheidet aber wegen der ebenfalls vorhandenen Mindergröße aus. Von den verbleibenden zwei Kandidaten ist Kandidat 1 in der Abfolge geringfügig ähnlicher und deutlich größer - somit wird die aufzulösende Fläche dieser Kategorie zugeschlagen<sup>21</sup>.

### 5.3.4 Generalisierungsstrategie für die Formvereinfachung auf Makroebene

Am zuvor aufgeführten Beispiel (Abb. 25 u.r.) offenbart sich das unter Kap. 4.5.3.1 d) schon angesprochene Problem einer weiterführenden Formvereinfachung der nun semantisch und geometrisch aggregierten Objekte. Erstrebenswert ist ein Generalisierungseffekt, der in seiner Wirkung im flächenhaften Kontext bei gleichzeitiger Wahrung der Topologie zu einem ansprechenden Resultat führt. Die gebräuchlichen Algorithmen zur Liniensimplifizierung sind hierfür nur bedingt geeignet, weil sie entweder Topologiebezüge nicht wahren oder den visuellen Ansprüchen nicht genügend entsprechen. Genau diese beiden Punkte aber sind es, deren Einhaltung auf vektorierter Ebene am meisten Schwierigkeiten machen.

Hier bietet die Übernahme des zellulären Automaten aus der GeoScaler-Toolbox eine pragmatische Lösung. Bewerkstelligt wird die Integration durch die Aneinanderreihung mehrerer Python-Caller, welche die zuvor entkapselten GeoScaler-Skripte inklusive der hierfür erforderlichen Modifikationen bezüglich Ein-/Ausgabepfaden, Parameter und Ausgabemeldungen beinhalten. Innerhalb des FME-Ablaufs erfolgt der Aufruf zur Laufzeit im Zugriff auf entsprechende ArcGIS-Funktionalität (u.a. PolygonToRaster, FocalStatistics, Con, Nibble,

21 Da beide Kandidaten in ihrer übergeordneten Hierarchie zu völlig unterschiedlichen Kategorien gehören und sich der Großteil des Polygons innerhalb Kandidat 2 befindet, sollte die Fläche eigentlich anhand des längsten Liniensegments aufgelöst und Kandidat 2 zugeschlagen werden. Allerdings ist diese Überlegung eher hypothetischer Natur, da im anschließenden CA-Verfahren je nach Zielmaßstab solche Flächen großzügig entfernt werden (s. Abb. 31 B).

RasterToPolygon, SmoothPolygon) - Voraussetzung ist also eine Systemumgebung, in der ArcGIS mit der erforderlichen Spatial Analyst-Erweiterung lauffähig ist.

### 5.3.5 Bestimmung und Handhabung variabler Werte (Parameter) im Ablauf

Im FME-Framework sind solche Variablen als „Published Parameter“ angelegt, d. h. die Werte müssen entweder frei oder anhand einer vorgegebenen Auswahl von Vorschlagswerten beim Starten des Workflows angegeben werden. Es handelt sich dabei einerseits um Angaben zur Festlegung von Rahmenbedingungen, die anhand von Auswahllisten bestimmt werden können und andererseits um maßstabsabhängige Größenangaben, welche konkret funktionale Abläufe mittels des angegebenen Wertes steuern.

Ausgehend von den designierten Zielmaßstäben der einzelnen Hierarchieebenen wurden entsprechende Kennwerte auf Grundlage eigener Messungen und vorhandener Vergleichswerte (McCabe 2008) festgelegt.

Die den GeoScaler betreffenden Referenzwerte wurden durch eigene Tests ermittelt, basieren aber auch auf der offiziellen Dokumentation (Huot-Vézina et al. 2012).

PARAMETER	WERTE (BEREICH)		
INITIAL_SCALE	50000		
BOX	[10,16,21,22,29]		
CA_MOORE_RADIUS	[1,2,3,4,5]		
CA_ITERATIONS	1-255		
FAULTS_NEIGHBOUR	1-5		
_max_rel_area_length	30		
KENNWERT	GE3 (100000) Bereich: 75000 - 150000	GE2 (350000) Bereich: 150000 - 500000	GE1 (750000) Bereich: 500000 - 1500000
_minarea	20 000 m <sup>2</sup>	200 000 m <sup>2</sup>	1 000 000 m <sup>2</sup>
GEN_STEP	1	2	3
CELLSIZE	10 m	25 m	50
TARGET_SCALE	100000	250000	1000000
SMALL_POLYGONS	200		
SMOOTH_TOLERANCE	50	150	500

Tab. 17: Parameter und (Default)-Kennwerte im Workflow für die Generalisierung der geologischen Daten

### 5.3.6 Modul für die statistische Ausgabe von Kennwerten

In der Workbench wurde ein Modul integriert, welches die Ausgabe von statistischen Kennwerten in Form eines tabellarischen Vorher/Nachher-Berichts ermöglicht. Der erste Ankerpunkt für die Übergabe von Mess- und Kennwerten liegt dabei nach Ausführung der semantischen Aggregation (dissolver), der zweite Ankerpunkt am Ende des Prozessablaufs. Zum jetzigen Zeitpunkt sind lediglich allgemein verfügbare Kenngrößenberechnungen (Anzahl der Objekte/Flächengrößen/Flächenanteile/Stützpunktzahl je Kategorie, Circularity) hinterlegt, die Hinzunahme weiterer Werte ist aber auf einfache Weise zu bewerkstelligen. Eine Beispieldatenbank findet sich in Anhang F.

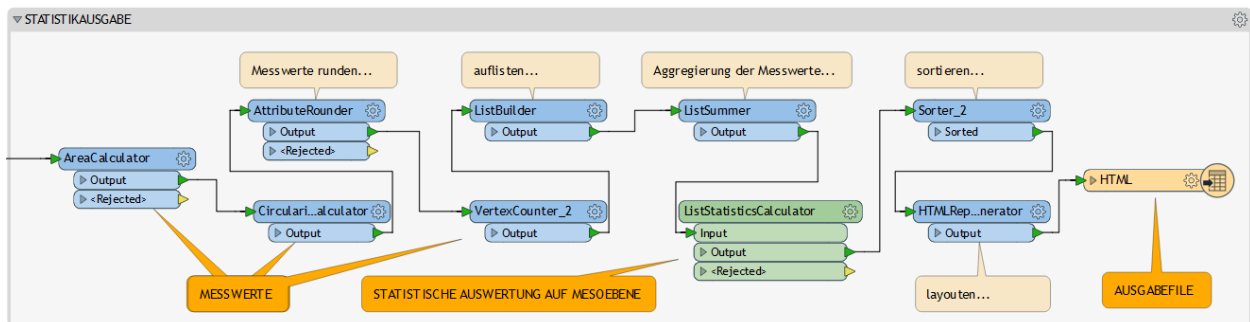


Abb. 26: Modul Statistikausgabe

### 5.3.7 Vorgehen bei der Anwendungssimulation des Workflows

Die soweit erzeugte Workbench wurde nun anhand der 5 ausgesuchten Datenbereiche für den ersten Maßstabsübergang GE4 (1 : 50 000) nach GE3 (1 : 100 000) getestet. Für die Erzeugung der nachfolgenden Generalisierungsstufen GE2 und GE1 wurden auf Basis der vorliegenden Workbench jeweils Kopien erstellt und mit notwendigen Detailanpassungen versehen. In einem weiteren Ablauftest wurde die Workbench hinsichtlich ihrer Verwendung auf einen flächenmäßig größeren Datenbereich modifiziert, um eine Aussage über die Durchführbarkeit in der Anwendung auf landesweite Daten sowie der damit einhergehenden Performanz treffen zu können.

## 6 Ergebnisse

### 6.1 Darstellung der Generalisierungsergebnisse

Anhand der sowohl nach Durchführung der semantischen Aggregation wie erneut am Ende der Prozesskette beim Generalisierungsergebnis vorgenommenen Messungen lassen sich quantitative Aussagen zur prozentualen Veränderung bezüglich Anzahl von Objekten und Flächenanteilen je Kategorie treffen. Die Aussagekraft dieser Werte in Bezug auf die fachliche Konsistenz der Resultate ist allerdings beschränkt, insofern gestaltet sich die quantitative Bewertung des vorliegenden Lösungsansatzes in der Praxis schwierig.

Für eine qualitative Bewertung wurden die Ergebnisse dem Fachbereich Geologie im LGRB vorgelegt. Die Experten haben eine gute Les- und Interpretierbarkeit der generalisierten Daten unter Wahrung des geologischen Kontextes bestätigt (Hagemeister & Franz 2018). Dies schließt aber nicht aus, dass in lokal begrenzten Bereichen die Geologen bei einer manuellen Generalisierung zu einer abweichenden Lösung kämen. Solche Einzelbereiche bleiben aber insgesamt an den hier vorliegenden Teilbereichen die Ausnahme und können als vernachlässigbar eingestuft werden.

Die folgenden Bildbeispiele zeigen ausgewählte Bereiche aus den prozessierten Testdaten mit den Generalisierungsergebnissen in ihrer maßstäblichen Abfolge auf Basis der jeweiligen semantischen Hierarchiestufe. In einer zweiten bildlichen Gegenüberstellung werden relevante Generalisierungssituationen aus Tab. 14 exemplarisch am entsprechend vergrößerten Ausschnitt dargestellt und erläutert.

### 6.1.1 Beispiel 1 - semantische Aggregation und Formvereinfachung (B1)

Abb. 27 verdeutlicht exemplarisch das Prinzip bei der Zusammenfassung der Einheiten in die jeweils nächsthöhere Hierarchiestufe sowie die durch die Anwendung des zellulären Automaten resultierende Formvereinfachung im Flächenkontext (C1 aus Tab. 14).

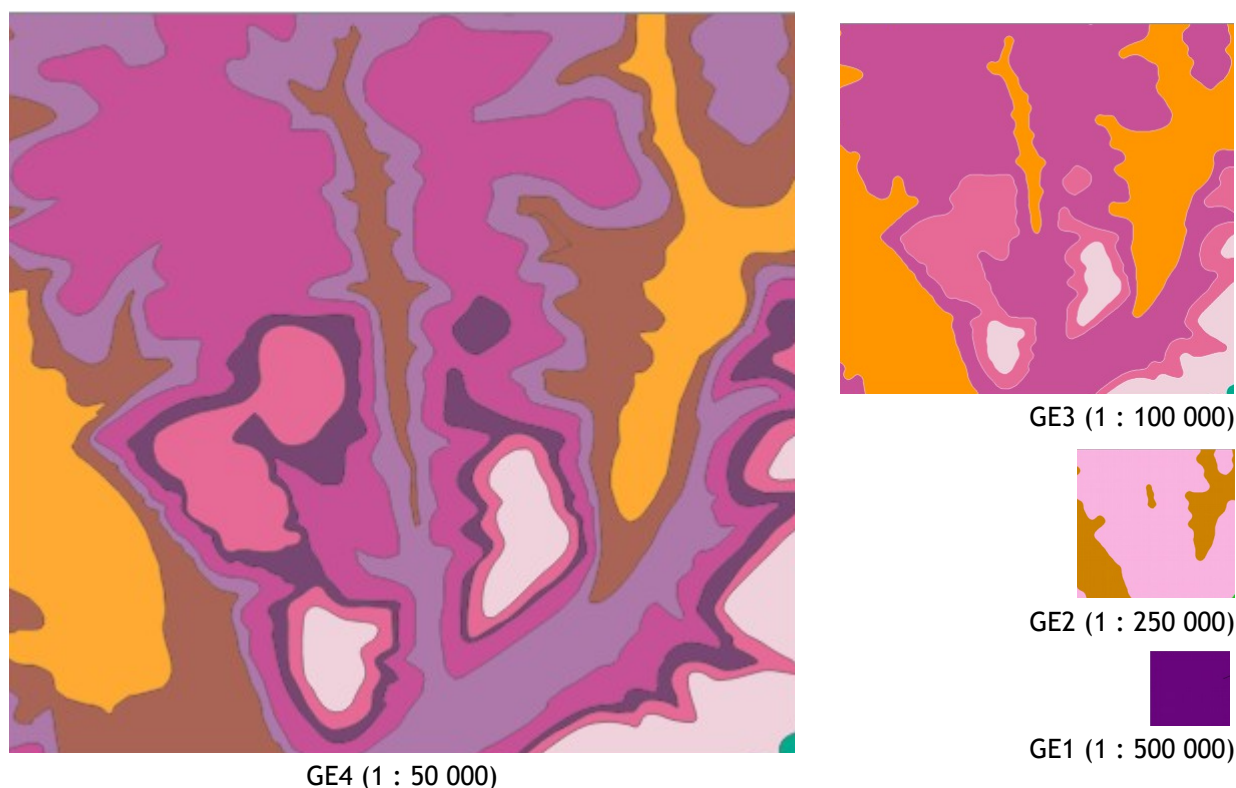


Abb. 27: Maßstabsreihe mit Ausgangssituation (GE4) und Generalisierungsergebnissen (GE3-GE1)

Die farbliche Hinterlegung der Kategorien entspricht der offiziellen Farbgebung der jeweils zugrundeliegenden geologischen Einheit<sup>22</sup>. Nachfolgende Tabelle belegt den semantischen Übergang der jeweiligen Einheiten zu den übergeordneten Kategorien.

GE4		GE3		GE2		GE1	
772	Erfurt-Formation (Lettenkeuper)	771	Unterkeuper	650	Keuper	649	Trias
818	Oberer Muschelkalk	818	Oberer Muschelkalk	817	Muschelkalk		
931	Mittlerer Muschelkalk	931	Mittlerer Muschelkalk				
954	Karlstadt-Formation	960	Unterer Muschelkalk				
961	Jena-Formation	1015	Oberer Buntsandstein	1014	Buntsandstein		
990	Freudenstadt-Formation						
1017	Obere Röttone						
1022	Plattensandstein-Formation						

Tab. 18: Überführungstabelle semantische Zusammenfassung

<sup>22</sup> Aufgrund der Neuordnung zu einer übergeordneten Kategorie wechselt die Grundfarbe der jeweiligen geologischen Einheit - durch die Hierarchieebenen hindurch ergibt sich deshalb kein homogener Farbeindruck in Kartenabbildungen.

## 6.1.2 Beispiel 2 - Punkt- und Linientypisierung (Kollaps - C1, C2)

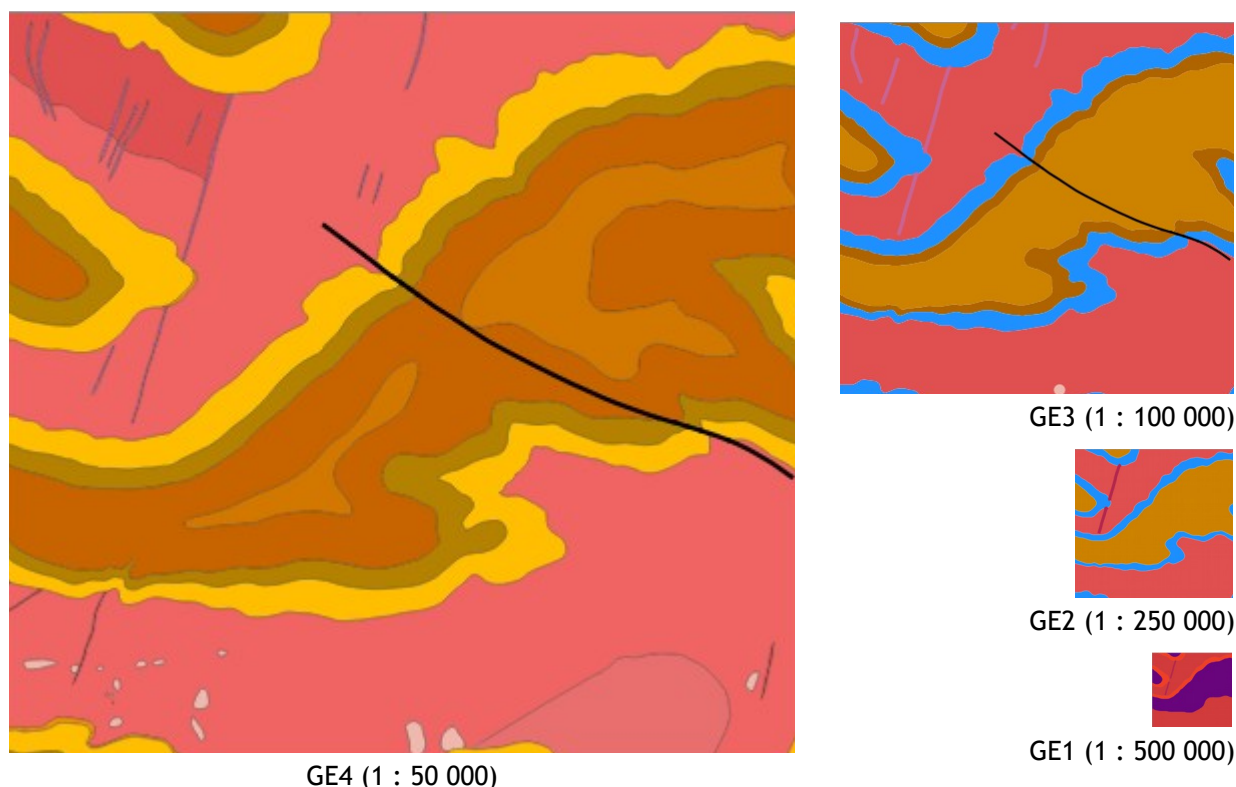


Abb. 28: Ausschnitte GE4 - GE1 für Beispiel 2

Im Beispielausschnitt finden sich auf Ausgangsebene GE4 eine Vielzahl länglicher, gangartiger geologischer Strukturen, die im Teilprozess 3 des FME-Workflows automatisch erkannt und mit einem temporären Attribut (laenglich: [ja|nein]) gekennzeichnet werden. Nur wenige dieser länglichen Gebilde weisen die erforderliche Mindestgröße auf. Durch die vorab durchgeführte Datenanreicherung (Kap. 5.3.2) sind aber einzelne ausgesuchte Objekte als erhaltenswert gekennzeichnet und werden nun anhand ihrer Gestalt (rundlich oder länglich) dem entsprechenden Ziel-Geometrietyp (Punkt oder Linie) zugeordnet. Die Anwendung des jeweiligen FME-Transformer (CenterPointReplacer oder CenterLineReplacer) sorgt dann im Ablauf für die operative Umsetzung. Die so generierten Punkt- und Liniendaten werden als neue FeatureClass in einer ESRI-Filegeodatabase abgelegt, während die weitere Prozessierung für die Objektauflösung und die anschließende Formvereinfachung durch den CA sorgt.

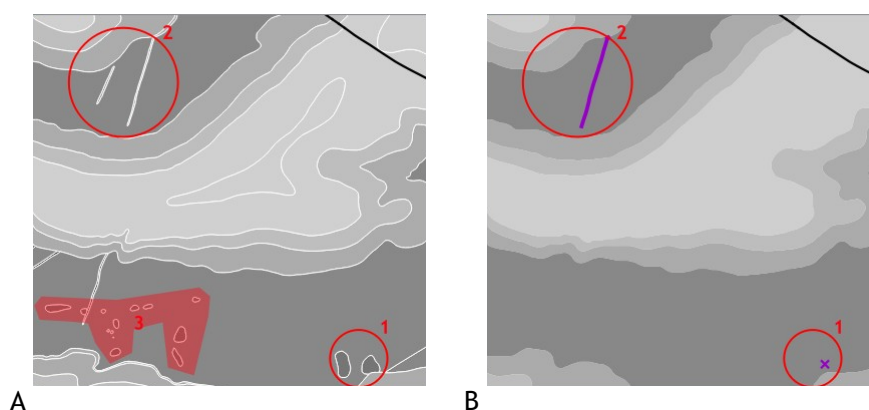


Abb. 29: Detailbeispiel für die Linien- und Punkttypisierung aufgrund objektbezogener Beflagung (Stufe GE3 - ca. M 1 : 30 000)

Beide Polygone in Abb. 29 innerhalb des Markierungskreises (1) unterschreiten die Mindestgröße, das linke Polygon weist eine Beflagung auf und bleibt im Generalisierungsergebnis als Punkt erhalten. Analog dazu wird in (2) von den beiden Gangpolygonen nur das entsprechend „beflaggte“ als Linie typisiert. Alle Flächen aus Markierungsbereich 3 werden aufgelöst. Die zugrunde liegende Generalisierungssituation D3 ist zum jetzigen Zeitpunkt noch nicht im Ablauf implementiert.

gen Zeitpunkt noch nicht im Ablauf implementiert.



## 6.1.3 Beispiel 3 - Auswirkung der Formvereinfachung auf Makro-Ebene (C4)

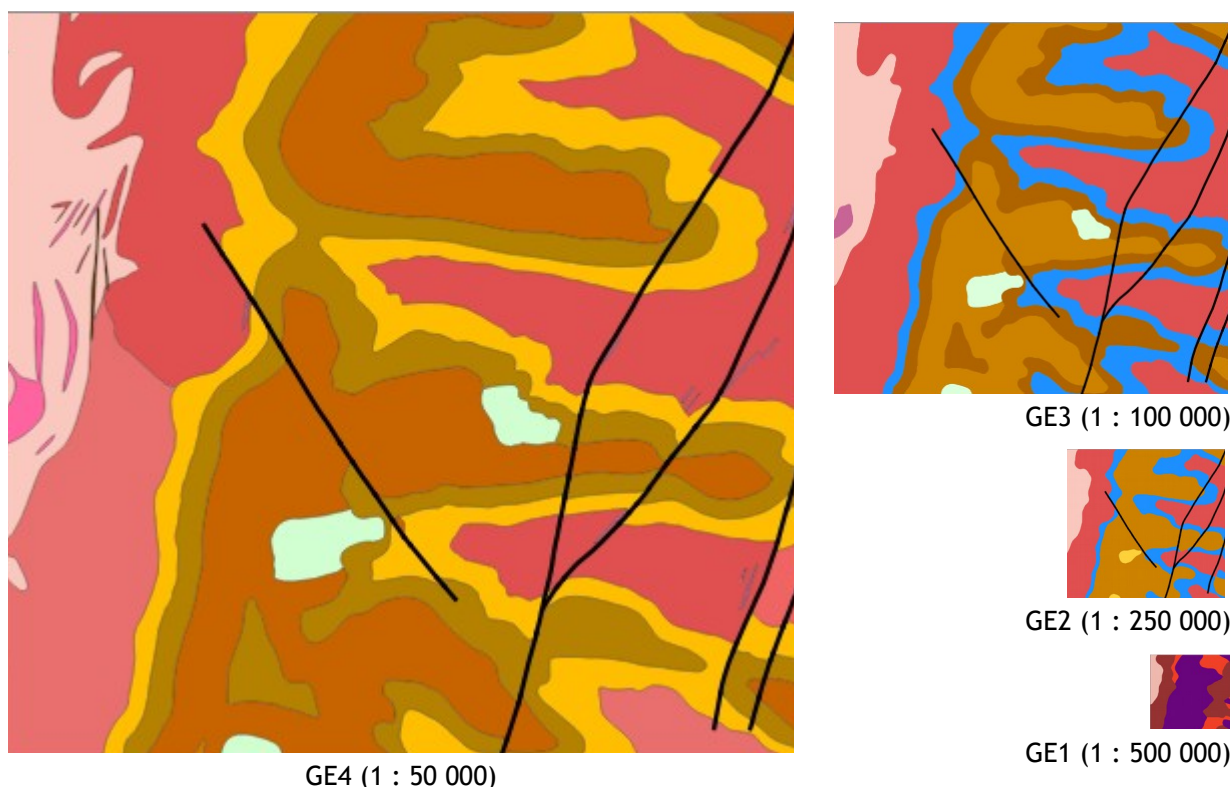


Abb. 30: Ausschnitte GE4 - GE1 für Beispiel 3

Der in Abb. 30 am linken Bildrand ersichtliche Datenausschnitt greift die Situation aus Abschnitt 5.3.3, Abb. 25 noch einmal auf und verdeutlicht hier die glättenden Auswirkungen des zellulären Automaten. So wird im Beispiel die längliche Einheit bei der Auflösung durch die lithostratigrafische Nähe gegebenenfalls einer un-günstigen Einheit zugeschlagen, durch die nachträgliche Generalisierung im zellulären Automaten findet aber quasi eine nachträgliche Bereinigung statt, so dass der vermeintliche Fehler im Endergebnis keine Relevanz hat.

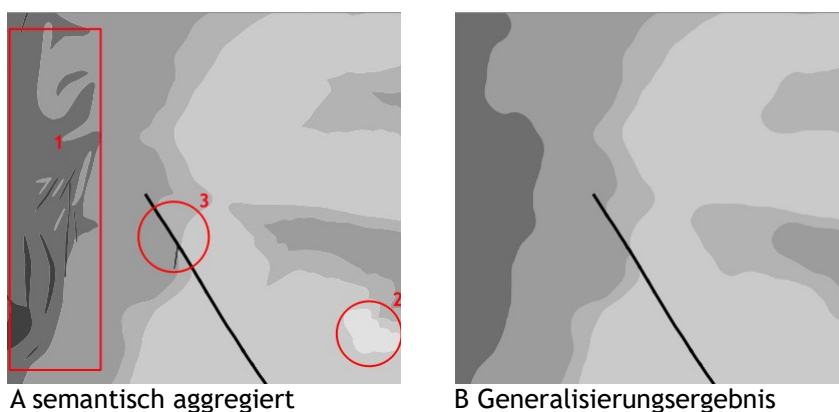


Abb. 31: Detailbeispiel für die CA-Wirkweise bei kleinteiligen Objekten (Stufe GE2)

Die im Markierungsrahmen 1 vorhandenen kleinteiligen Elemente unterliegen zum Teil anhand ihrer Mindergrößen schon im Vorfeld der Auflösung, einige der Objekte werden aber auch erst durch die Anwendung des zellulären Automaten weg generalisiert. Solche Objekte müssen - so sie denn als erhaltenswert definiert wurden - explizit attributiv gekennzeichnet werden.

#### 6.1.4 Beispiel 4 - Einbezug invarianter Bereiche durch tektonische Lineaturen (D1)

Eines der wesentlichen Merkmale bei der geologischen Karte ist die strukturelle Abhängigkeit der Ausstrichflächen von den tektonischen Verhältnissen in der direkten Umgebung. Eine entkoppelte Generalisierung ohne den Einbezug solcher invarianter Bereiche würde die fachliche Aussage in Frage stellen. Im folgenden Beispielausschnitt zeigen sich die Schwierigkeiten im Umgang mit dieser Randbedingung aus Tab. 14 (D1) besonders gut. Darüber hinaus lässt sich einmal mehr der Generalisierungseffekt des zellulären Automaten im Zusammenhang mit heterogenen, verstreuten Kleinstflächen belegen. Hier wäre über die vorab durchgeführte Datenanreicherung hinaus ein Verfahren zur Identifizierung und Behandlung von Mustern auf der Meso-Ebene (D2) erforderlich.

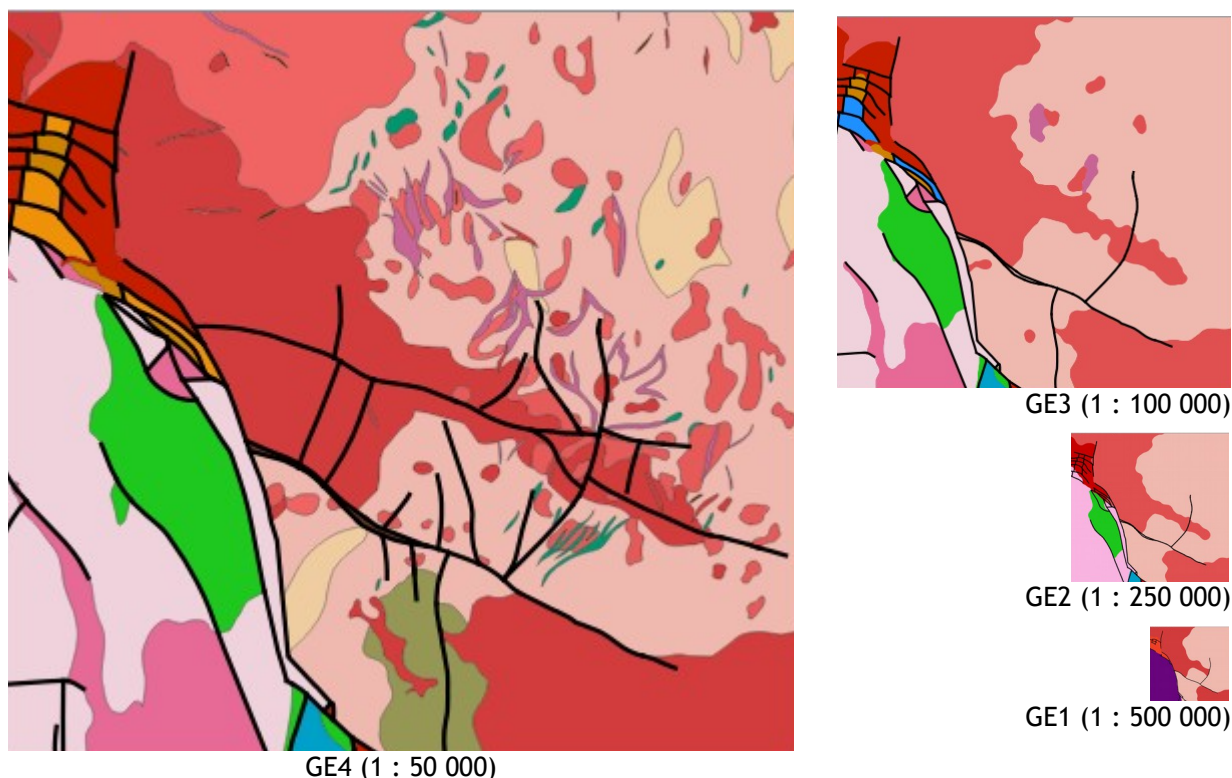
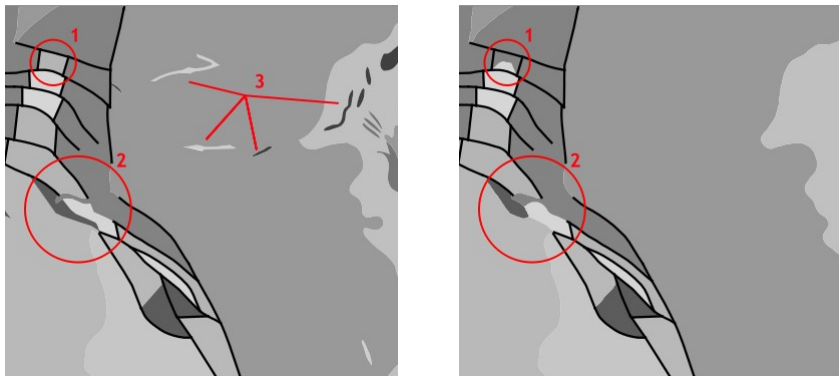


Abb. 32: Ausschnitte GE4 - GE1 für Beispiel 4

Am Beispiel in Abb. 32 zeigt sich ein bisher noch nicht ausreichend berücksichtigter Sachverhalt: Während die geologischen Flächen durch den zellulären Automaten und der anschließenden Glättung bei der Rückwandlung eine starke Formveränderung erfahren, bleiben die Geometrien der Tektonik-Lineationen unverändert. Daraus ergibt sich im Kartenbild an den Rändern der als invariant gehandhabten Bereiche eine Inkonsistenz zwischen den beiden Layern. Hier gilt es vorsichtig mit dem Eingangsparameter `SMALL_POLYGONS` im CA zu operieren. Andererseits zeigt sich hier als Bedingung, dass die attributive Datenanreicherung der tektonischen Ausgangsdaten nicht nur nach rein sachlich-formalen Kriterien, sondern auch im geometrisch-maßstäblichen Kontext durchzuführen ist. Ein Linienkonstrukt wie in Abb. 32 auf Stufe GE2 erscheint nach visuellen Kriterien als nicht zielführend<sup>23</sup>. Dennoch lässt sich im Workflow auch eine solche Situation bis zu einem gewissen Grad berücksichtigen (Abb. 33).

<sup>23</sup> Nach fachlicher Rücksprache erfolgt bei zu enger räumlicher Schärung von tektonischen Linienelementen eine besser auf den Zielmaßstab bezogene gröbere Ausdünnung derselben, auch wenn hiermit inhaltliche Verluste in Kauf zu nehmen sind.



A semantisch aggregiert

B Generalisierungsergebnis

Abb. 33: Detailbeispiel Stufe GE2 in tektonisch stark zergliederter Region

jekttaulösung für die Eliminierung aller mindergroßen Flächen.

Die Bereiche in direkter Umgebung der Tektoniklinien werden beim CA-Durchlauf als invariant behandelt. Nach dem CA-Durchlauf erfolgt vor der anschließenden Raster/Vektor-Rückwandlung noch eine gesonderte Behandlung von kleinen Zellregionen. In Randbereichen kann das zur nachträglichen Veränderung der im CA-Ergebnis eigentlich unverändert belassenen Flächen führen (1 und 2). Außerhalb des tektonischen Einflussbereichs (3) sorgt schon die vor dem CA durchgeführte Ob-

## 6.1.5 Beispiel 5 - Erhalt geologischer (Graben)Strukturen (D3)

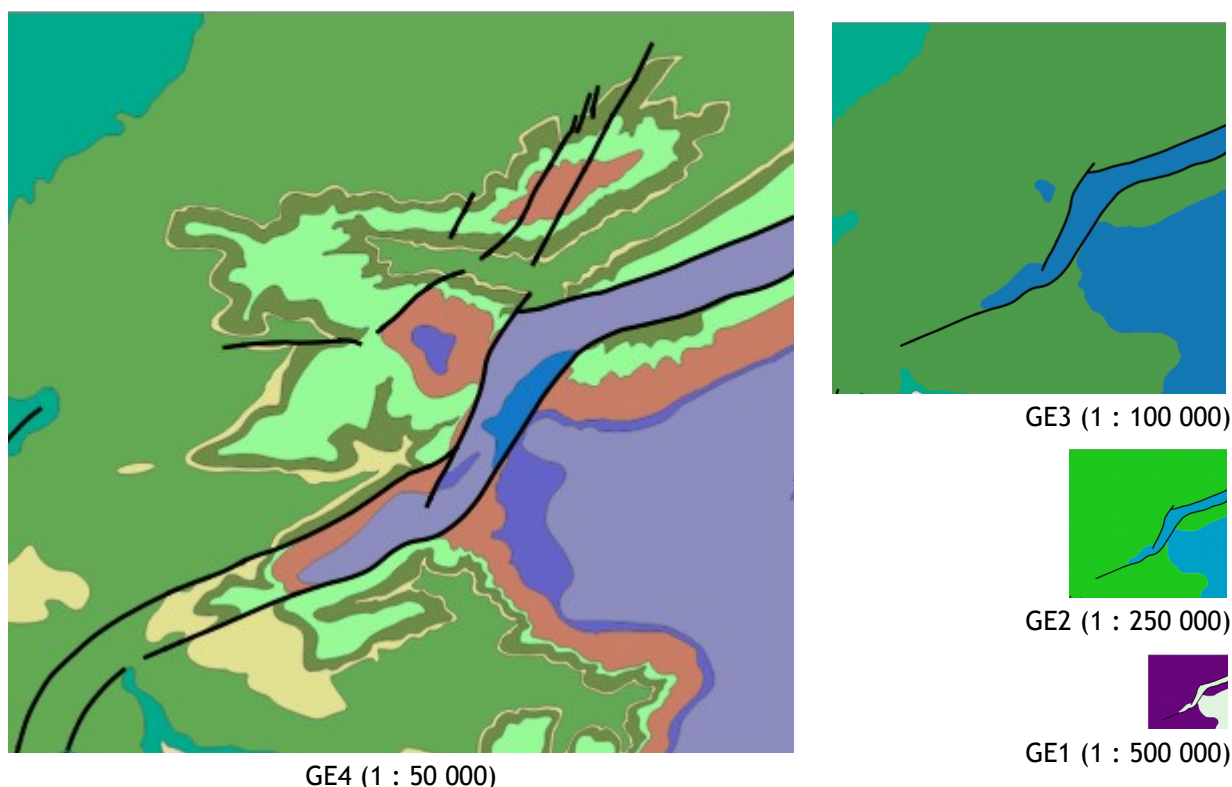


Abb. 34: Ausschnitte GE4 - GE1 für Beispiel 5

Das Beispiel in Abb. 34 zeigt als Ausgangssituation einen tektonischen Graben, der geprägt ist durch die unterschiedliche Zugehörigkeit der im Graben vorhandenen Kategorien zu den außerhalb vorkommenden geologischen Einheiten. Der Einbezug der invarianten Bereiche gewährleistet auch hier den Erhalt der ursprünglich gegebenen Situation durch die Hierarchiestufen hindurch im Rahmen der vorliegenden „Beflaggung“ der tektonischen Lineationen (Grabenränder).

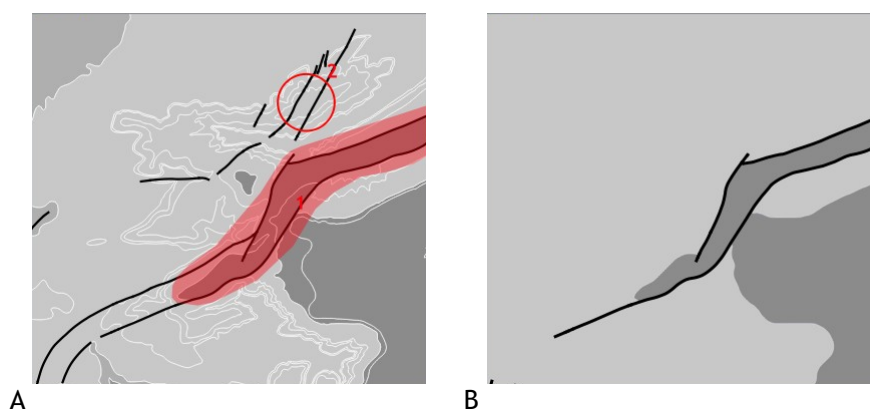


Abb. 35: Übergang von Hierarchiestufe GE4 in GE2

Während der Hauptgrabenverlauf (Abb. 35 - Bild A, Bereich 1) anhand der tektonischen Linien durch alle Hierarchiestufen hindurch erhalten bleibt, unterliegt die als minder wichtig definierte Bruchzone in Markierung 1 der Auflösung bzw. in diesem Fall der semantisch bedingten Aggregation.

## 7 Diskussion

### 7.1 Thematische Übertragbarkeit

#### 7.1.1 FME-Lösungsansatz in der Anwendbarkeit auf die bodenkundlichen Ausgangsdaten

Die Workbench in ihrer ursprünglichen Ausprägung weist nun erwartungsgemäß viele Geologie- bzw. allgemein Daten-spezifische Besonderheiten und Abhängigkeiten auf, dennoch ist eine Übertragung auf andere Fachdatenthemen möglich, sei es durch eine separat angepasste Workbench oder in Form eines parallelen Ablaufstrangs, der relevante Teilmodule bzw. einzelne Transformer wahlweise oder alternativ im Prozess ansteuert und so einen themenspezifisch angepassten Prozessdurchlauf ermöglicht. Für die integrierte Verwendung weiterer Fachthemen sind zusätzliche Parameter und Variablen einzuführen (bzgl. Thema, Objektartenname, Kategorisierungsattribut, Hierarchieebene, Ausgabepfade etc.) anhand derer die Durchführung des Ablaufs gesteuert werden kann. Ein konkreter Versuch mit den bodenkundlichen Flächendaten erfordert eine entsprechende Anpassung des Ablaufs bezüglich der genannten Variablen, wie auch die gezielte Steuerung und Handhabung des zellulären Automaten als gesamtheitliche Generalisierungsstrategie. Bei den geologischen Daten steht der Tektonikbezug mit seinen zu berücksichtigenden invarianten Bereichen und damit die Verwendung der entsprechenden GeoScaler-Methodik für die Generalisierung von Grundgebirgskarten im Fokus. Diese Abhängigkeit fehlt bei den bodenkundlichen Daten, dem entsprechend kann der etwas vereinfachte Ablauf für Verbreitungskarten angewendet werden. Beide Verfahren sind anhand entsprechender Python-Caller in der FME-Workbench hinterlegt und können so im Ablauf gezielt angesteuert werden.

Der Teilschritt mit der Identifizierung länglicher Objekte und deren Überführung in linienhafte Geometrien entfällt dabei komplett, während nun zusätzlich der Teilschritt der Polygonaufteilung von Nullflächen mit in den Ablauf integriert wurde. Die entsprechende Python-Skript-Funktionalität aus Kap. 5.1 wurde zu diesem Zweck in FME nachgebildet und als Modul in den Gesamtablauf übernommen (s. Anhang B).

Die unterschiedliche Methodik bei der semantischen Aggregation mit dem Übergang von detaillierten Verbreitungsflächen hin zu einer regional-räumlichen Zusammenfassung im Gegensatz zur stringenten lithostratigraphischen Hierarchisierung bei den geologischen Daten erschwert eine direkte Vergleichbarkeit bezüglich der Eignung des Verfahrens. Die in Abb. 36 dargestellten Ausschnitte C) und D) wurden bei gleicher CA-Parametrisierung mit unterschiedlicher Raster-Auflösung (25|50) und Zielmaßstab (250000|500000) prozessiert, die Ergebnisse zeigen die zu erwartenden Generalisierungseffekte und liefern hinsichtlich der Darstellung im Zielmaßstab ein (im Fall C) durchaus gutes Ergebnis. Der Vergleich mit einer durch einen Bodenkundler manuell erstellten Generalisierung (Bsp. E) zeigt dann aber schon an diesem kleinen Ausschnitt deutlich (v.a. anhand der gelblichen Einheit), dass die kartografischen Operationen des Amalgamieren und Überbetonen nicht adäquat simuliert werden können. Allerdings wurde die manuelle Überarbeitung für einen größeren Maßstab konzipiert, die entsprechend gegenübergestellte CA-Lösung D) zeigt sich anhand der gewählten Mindestgrößen als zu stark generalisiert.

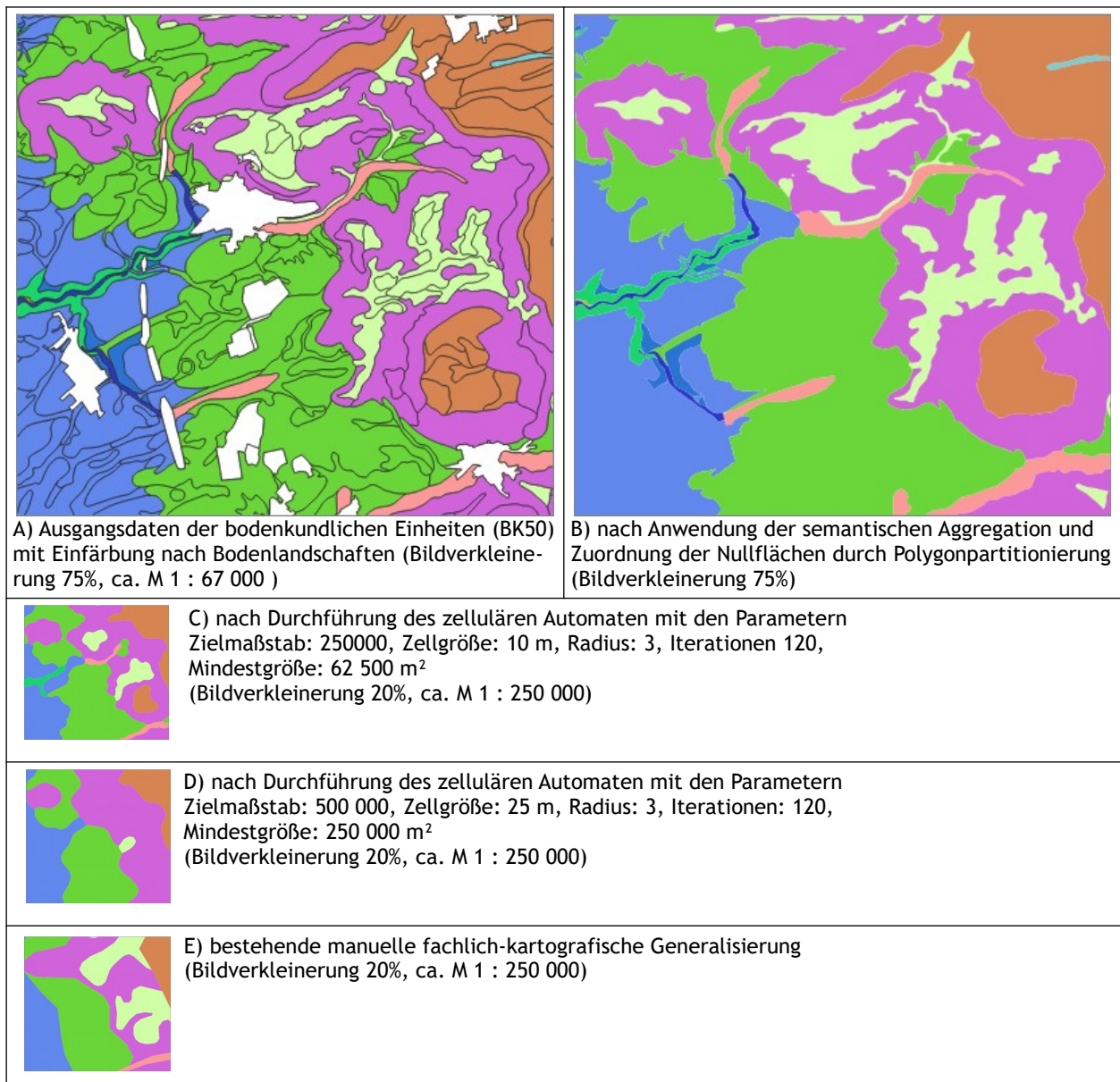


Abb. 36: Anwendung der CA-Methode auf die bodenkundlichen Daten im modifizierten Workflow

Grundsätzlich steht aber einer Verwendung der CA-basierten Methode auch bei den bodenkundlichen Verbreitungsflächen nichts entgegen, die Notwendigkeit zur Formvereinfachung ist unabhängig von den zugrundeliegenden Themendaten gleich. Inwiefern die durch die Anwendung des zellulären Automaten resultierenden Rundungseffekte bzw. die eventuell auftretenden Verluste kleinerer Kategoriebereiche problematisch für die Aussage im Generalisierungsergebnis sind, muss einer weitergehenden fachlichen Begutachtung unterzogen werden. Die Berücksichtigung solcher fachspezifischer Besonderheiten kann aber nach eingehender Strukturanalyse gegebenenfalls in gleicher Weise durch eine zusätzliche Attributierung in den Ausgangsdaten vorgenommen und im Ablauf eingebaut werden.

## 7.2 Weitere Fachthemen

Bezüglich der Anwendbarkeit des Workflows auf die hydrogeologischen Daten ist mehr oder weniger eine 1:1-Übertragung möglich. Strukturell unterscheiden sich die Daten - einschließlich der hier ebenfalls vorhandenen Tektonik - nicht von denen der Geologie. Die abgedeckte hydrogeologische Karte entsteht durch Zusammenfassung geologischer Einheiten zu hydrogeologischen Einheiten und Abdeckung der quartär-geologischen Deckschichten - hierbei besteht eine durchgehende geometrische Konsistenz der Flächen zwischen den beiden Themenbereichen. Anhand der Deckschichten lässt sich dadurch ein weiteres Generalisierungsszenario am neu entwickelten FME-Workflow testen.

### 7.2.1 Anwendbarkeit auf die Generalisierung von Deckschichten

Deckschichten treten als direkte gegenseitige Relation bei den Themen Hydrogeologie, Bodenkunde und Geologie auf (s. Kap. 4.3.2). Die jeweiligen Flächendaten gleichen sich bezüglich Art und Verbreitungsmuster. Als prägnante Eigenschaft der Geometrien ist ihre Verbreitung in der Art eines dendritischen Musters zu nennen, welches durch das hydrologische Abflusssystem definiert wird. Es weist also eine direkte topologische Relation mit dem topografischen Gewässernetz auf.

Bei der separaten Generalisierung von Deckschichten kann nun anhand eines fachlich definierten Abbruchkriteriums die gezielte Steuerung des Generalisierungsgrades vorgenommen werden, bis wohin Deckschichten in Seitentälern erhalten bleiben sollen. Es werden also lineare Zusatzdaten - ähnlich des einschränkenden Einbezugs der Tektonik bei der Generalisierung von geologischen Flächen - zur Definition invarianter Bereiche herangezogen. In diesem Fall handelt es sich um ein speziell auf die Datenbasis zugeschnittenes Gewässernetz aus den Geobasisdaten, welches durch eine „Beflagung“ in ähnlicher Art wie bei den tektonischen Linien (s. Kap. 5.3.2) für unterschiedliche Hierarchie-/Generalisierungstufen vorbereitet wurde.

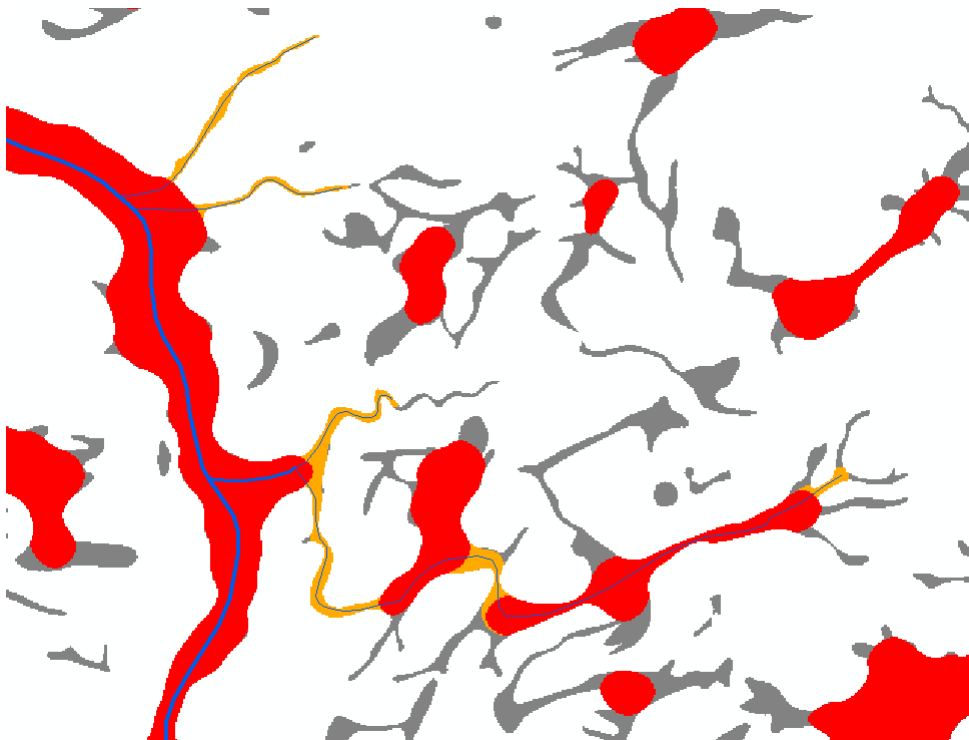


Abb. 37: Beispielausschnitt Generalisierung von Deckschichten mit und ohne Berücksichtigung des Gewässernetz

Die in Abb. 37 im Hintergrund erkennbaren grauen Flächen stellen die Ausgangssituation der vorliegenden Deckschichten dar. Die zuoberst liegenden roten Flächen verdeutlichen den Generalisierungseffekt durch den zellulären Automaten ohne Berücksichtigung des Gewässernetzes (in Blau), die orangen Flächen blieben durch Einbezug der invarianten Bereiche erhalten.

Eine optionale Berücksichtigung im CA-Modul kann analog zur Tektonik über einen zusätzlichen Parameter implementiert werden. Die Implementierung in den FME-Ablauf setzt eine vorherige Übernahme der inselhaften Flächen in ein flächendeckendes Polygonmosaik mit einem Umgebungspolygon und entsprechender Attributierung voraus sowie, ähnlich der Anwendbarkeit bodenkundlicher Daten, entsprechende datenspezifische Anpassungen innerhalb des Ablaufs.

### **7.2.2 Ingenieurgeologische und rohstoffgeologische Daten**

Analog zum vorhergehenden Generalisierungsszenario mit inselhaft verbreiteten Polygonen können auch die beiden angewandten Themenbereiche der ingenieurgeologischen und rohstoffgeologischen flächenhaften Datenlayer (s. 4.1) in ähnlicher Weise gehandhabt werden. Die Definition invarianter Datenbereiche und der Einbezug eines Liniennetzes bei der Generalisierung dieser Themen wäre dabei nur bei der unmittelbar gegebenen tektonischen Abhängigkeit von Flächenabgrenzungen erforderlich.



### 7.3 Fazit

Diese Arbeit zeigt, dass eine automatisierte Generalisierung von Geofachdaten in einem gesamtheitlich abgestimmten Ablaufprozess auch unter Nutzung gängiger GI-Software grundsätzlich möglich ist.

Dennoch bleiben selbst unter der größtmöglichen Reduktion der verwendeten Bedingungen und der Anforderungen an das Endergebnis eine Vielzahl relevanter fachspezifischer Einschränkungen und Abhängigkeiten, unvorhergesehener Schwierigkeiten und komplexer Prozessabläufe zu beachten, die eine Vollautomatisierung erschweren. Letztlich ist die Beschaffenheit der Ausgangsdaten und das avisierte Generalisierungsziel sowie die im thematischen Kontext zu berücksichtigenden fachlichen Parameter ausschlaggebend für den jeweils möglichen Automatisierungsgrad, den zu erzielenden Generalisierungseffekt und den dafür erforderlichen programmatischen Aufwand. Die Beurteilung von Qualität und Plausibilität eines solcherart erzielten Generalisierungsergebnisses liegt vor allem - und im besonderen Maße bei geologischen Karten/Daten - im Auge des bewertenden Experten und muss daher als relativ subjektiv empfunden werden.

In der vorliegenden Arbeit wurde schwerpunktmäßig am Beispiel geologischer Daten ein praxisgerecht einsetzbares Prozess-Framework für die automatisierte Generalisierung kategorischer Flächendaten entwickelt. Die Umsetzung des Ablaufs mit der Software FME ermöglichte dabei die software-übergreifende Nutzung und Implementierung unterschiedlicher Methoden. Das Zusammenspiel generischer Transformer wie auch die Zuhilfenahme und Integration der ursprünglich in ArcGIS gekapselten rasterbasierten Methode eines zellulären Automaten liefern die für die Generalisierung erforderlichen Teilschritte. Der Einbezug thematisch-topologischer Abhängigkeiten gewährleistet hierbei auch die weitestgehende Einhaltung und Erfüllung vorhandener fachlicher Ansprüche. Als Mischform aus vektoruellen und rasterbasierten Teilschritten kann der Workflow auch als hybrides Verfahren bezeichnet werden. Die während der Prozessierung vorgenommene Vektor-Raster-Vektor-Wandlung wird nicht als Bruch im Verfahrensablauf wahrgenommen, auch die technische Integration des CA aus ArcGIS findet nahtlos im Hintergrund statt. Eine themenübergreifende Übertragung des Workflows ist grundsätzlich gegeben, unter dem Gesichtspunkt unterschiedlicher semantischer und datenstruktureller Vorbedingungen sind aber eine Reihe weiterer Detailanpassungen erforderlich, die im Rahmen dieser Arbeit noch nicht durchgeführt werden konnten.

## 7.4 Ausblick - Erweiterbarkeit und Entwicklungspotenzial

Die soweit entwickelte prototypische Lösung ist nun in erster Linie auf die Bedürfnisse der Generalisierung geologischer Flächendaten abgestimmt, die Konzeption des Verfahrens als modular aufgebaute FME-Workbench ermöglicht aber die gezielte Hinzunahme oder Auslassung von Modulen bzw. eine themenbezogene Steuerung des Ablaufs.

Grundsätzlich läuft die Prozesskette stabil und fehlerfrei, bezüglich der Prozessierung des in naher Zukunft landesweit vorliegenden (abgedeckten) geologischen Datenbestands sind unter Umständen Modifikationen in der vorliegenden Workbench vorzunehmen, die eine Aufteilung der Ausgangsdaten in besser verarbeitbare Teilmengen liefert. Dies beträfe in erster Linie die vektorbasierten Teilschritte, das implementierte CA-Verfahren zeigte sich auch robust bei der Verarbeitung von landesweiten Rastern mit der maximalen Auflösung von 10 m.

Hinsichtlich der zu typisierenden Elemente wurden durch die vorhergehende attributive Datenanreicherung anhand der entsprechenden Beflagung von Einzelobjekten ein Großteil der Schwierigkeiten bezüglich nachfolgender Auswahl oder Clustering von Objekten aufgehoben. Hier bestünde die Möglichkeit vorhandene Ansätze (Downs & Mackaness 2002 bzw. Smirnoff et al. 2012) ebenfalls in den FME-Ablauf zu integrieren.

Die soweit bestehende Lösung bietet darüber hinaus weiteres Ausbaupotenzial hinsichtlich folgender Punkte:

- weitergehende Automatisierung bezüglich der Identifizierung besonderer (geologischer) Strukturen (wie Gänge, Bänke, etc.)
- Anwendung von Clusterungsverfahren für die Heraushebung zusammenfassender Objekte (wie z.B. Linienscharen)
- Verwendung des generischen Transformers AreaAmalgamator im Verbund mit weiteren semantisch-topologisch wirksamen Transformern als möglicher Ersatz für das CA-Verfahren
- Erweiterte automatisierte Auswertung der Vorher/Nachher-Ausgabereports bzw. Einbezug zusätzlicher statistischer Bezugswerte
- Steigerung des Handhabungskomforts durch mögliche Übernahme in FME-Server und Bereitstellung eines Web-Formulars zwecks Parameter-Eingabe und Steuerung

---

## Literaturverzeichnis

- Ai, T., Ke, S., Yang, M. & Li, J. 2017. "Envelope generation and simplification of polylines using Delaunay triangulation." *International Journal of Geographical Information Science* 31 (2): 297-319.
- Ai, T. & van Oosterom, P. 2002. "GAP-tree extensions based on skeletons." In *Advances in spatial data handling*, 501-13: Springer.
- Artioli, G. P., Ciardi, G., Garberi, M. L. & Martelli, L. 1995. "The digitised geological cartography of the Emilia Romagna Region to scales 10 000 and 1: 50 000: The cartographic generalization process." In *Proceedings of ICA Conference*, herausgegeben von ICA, 860-71.
- Bader, M. & Weibel, R. 1997. "Detecting and resolving size and proximity conflicts in the generalization of polygonal maps." In *Proceedings 18th International Cartographic Conference*. Stockholm, Schweden (23.-27.Juli), herausgegeben von Ottoson, L., 1525-32.
- Baig, S. U., Rahman, A. A. & Duncan, E. E. 2013. "A Review and Conceptual Framework for Generalization of Maps." In *Developments in Multidimensional Spatial Data Models*, herausgegeben von Abdul Rahman, A., Boguslawski, P., Gold, C. & Said, M. N., 191-206: Springer.
- Bard, S. 2004. "Quality Assessment of Cartographic Generalisation." *Transactions in GIS* 8 (1): 63-81.
- Barrault, M., Regnaud, N., Duchene, C., Haire, K., Baeijs, C., Demazeau, Y., Hardy, P., Mackaness, W., Ruas, A. & Weibel, R. 2001. "Integrating multi-agent, object-oriented and algorithmic techniques for improved automated map generalization." In *Proceedings of 20th International Cartography Conference*. Peking, China (6.-10. August), herausgegeben von ICA, Vol. 3, 2210-16.
- Bobzien, M. 2005. "Methodische Aspekte der Generalisierung von Geodaten." Dissertation, Hohe Landwirtschaftliche Fakultät, Rheinische Friedrich-Wilhelms-Universität.
- Bobzien, M., Burghardt, D., Petzold, I., Neun, M. & Weibel, R. 2008. "Multi-representation databases with explicitly modeled horizontal, vertical, and update relations." *Cartography and Geographic Information Science* 35 (1): 3-16.
- Boedecker, M. 2011. "Koordinierung thematischer Klassen unter Verwendung kartographischer Relationen." Master Thesis, Universität Salzburg.
- Boedecker, M. 2014. "Coordination of Different Map Themes using Cartographic Relations." *International Journal of Geoinformatics* 10 (1): 39-47.
- Bollmann, J. & Koch, W.G. 2001. *Lexikon der Kartographie und Geomatik-A bis Karti: kartographische Generalisierung*. 2 vols. Heidelberg: Spektrum, Akad. Verl.
- Bundesanstalt für Geowissenschaften und Rohstoffe. 2005. *Bodenkundliche Kartieranleitung: Mit 103 Tabellen und 31 Listen*. 5., verbesserte und erweiterte Auflage. Stuttgart: E. Schweizerbart'sche Verlagsbuchhandlung.
- Bundesanstalt für Geowissenschaften und Rohstoffe. 2012. "Bodenkundliche Karten und Datenbanken: Bodenregionen/Pedoregionale Gliederungen." Zuletzt geprüft am 17.01.2019. [https://www.robeha.de/DE/Themen/Boden/Informationsgrundlagen/Bodenkundliche\\_Karten\\_Datenbanken/Bodenregionen\\_Pedoregionale\\_Gliederungen.html](https://www.robeha.de/DE/Themen/Boden/Informationsgrundlagen/Bodenkundliche_Karten_Datenbanken/Bodenregionen_Pedoregionale_Gliederungen.html).
- Buttenfield, B. P., Stanislawski, L. V. & Brewer, C. A. 2011. "A comparison of star and ladder generalization strategies for intermediate scale processing of USGS national hydrography dataset." In *Proceedings of 14th ICA/ISPRS Workshop on Generalisation and Multiple Representation*. Paris, France (30.Juni-1.Juli), 13 S.
- Chen, Y., Zhu, D., Yang, J. & Niblett, T. J. 2014. "Constraints Framework for Automatic Quality Evaluation of Generalized Map." *Sensor Letters* 12 (3-4): 847-54.
- Corcoran, P., Mooney, P. & Winstanley, A. 2011. "Planar and non-planar topologically consistent vector map simplification." *International Journal of Geographical Information Science* 25 (10): 1659-80.
- Dick, T. & Retterath, A. 2012. "FOSSGIS 2012: NAS - BasisDLM Aufbereitung mit gdal/PostNAS." Zuletzt geprüft am 19.01.2019. <https://www.fossgis.de/konferenz/2012/programm/events/386.de.html>.

- 
- Downs, T. C. & Mackaness, W. A. 2002. "An Integrated Approach to the Generalization of Geological Maps." *The Cartographic Journal* 39 (2): 137-52.
- Duchêne, C., Baella, B., Brewer, C. A., Burghardt, D., Battenfield, B. P., Gaffuri, J., Käuferle, D., Lecordix, F., Maugeais, E. & Nijhuis, R. 2014. "Generalisation in practice within national mapping agencies." In *Abstracting Geographic Information in a Data Rich World*, 329-91: Springer.
- ESRI. 2018a. "Überblick über das Toolset "Generalisierung"—Hilfe | ArcGIS Desktop." Zuletzt geprüft am 08.12.2018. <http://desktop.arcgis.com/de/arcmap/latest/tools/data-management-toolbox/an-overview-of-the-generalization-toolset.htm>.
- ESRI. 2018b. "Überblick über das Toolset "Generalisierung"—Hilfe | ArcGIS for Desktop." Zuletzt geprüft am Dezember 08, 2018. <http://desktop.arcgis.com/de/arcmap/10.6/tools/cartography-toolbox/an-overview-of-the-generalization-toolset.htm>.
- ESRI. 2018c. "Automatisieren der Workflows für die Konfliktlösung und Generalisierung durch Geoverarbeitung—Hilfe | ArcGIS for Desktop." Zuletzt geprüft am 20.06.2018. <http://desktop.arcgis.com/de/arcmap/10.3/tools/cartography-toolbox/conflict-resolution-and-generalization-with-geoprocessing.htm>.
- Foerster, T., Stoter, J. E. & Köbben, B. 2007. "Towards a formal classification of generalization operators." In *Proceedings of the 23rd International Cartographic Conference*, Moskau, Russland (4.-10.August), 15 S.
- Fuchs, M. 2002. "Methoden zur objektiven Ableitung von Bodenkarten im Folgemaßstab." Dissertation, Freie Universität Berlin.
- Galanda, M. 2003. "Automated polygon generalization in a multi agent system." Dissertation, Universität Zürich.
- Gao, W., Gong, J., Yang, L., Jiang, X. & Wu, X. 2012. "Detecting Geometric Conflicts for Generalisation of Polygonal Maps." *The Cartographic Journal* 49 (1): 21-29.
- Gao, W., Stein, A., Yang, L., Wang, Y. & Fang, H. 2013. "Improving Representation of Land-use Maps Derived from Object-oriented Image Classification." *Transactions in GIS* 17 (3): 387-405.
- Gartner, G. & Schmidt, M. 2010. "Moderne Kartographie-Technologische Entwicklungen und Implikationen." *Kartographische Nachrichten* 6 (2010): 299-305.
- Hagemeyer, A. & Franz, M. 2018. mündliche Stellungnahme am 13.12.2018. Freiburg.
- Hake, G., Grünreich, D. & Meng, L. 2002. *Kartographie: Visualisierung raum-zeitlicher Informationen*. Berlin: Walter de Gruyter. 604 S.
- Hampe, M., Sester, M. & Harrie, L. 2004. "Multiple representation databases to support visualization on mobile devices." In *Proceedings of the XXth ISPRS Congress*. Istanbul, Türkei (12.-23.Juli), herausgegeben von Iltan, O., Vol. XXXV, 135-40.
- Hardy, P. 2014. "GIS tools for Generalization and Multi GIS-Tools for Generalization and Mapping." Präsentation NMA Symposium, Amsterdam, Niederlande (14.Dezember).
- Harrie, L. & Weibel, R. 2007. "Modelling the overall process of generalisation." *Generalisation of geographic information: cartographic modelling and applications*, herausgegeben von Mackaness, W., Ruas, A. & Sarjakoski, L. T., 67-87: Elsevier.
- Haurert, J. H. 2004. "Geometrietypwechsel in einer Multi-Resolution-Datenbank." *Mitteilungen des Bundesamtes für Kartographie und Geodäsie, Aga-Tagung*, 9 S.
- Haurert, J.-H. & Wolff, A. 2010. "Area aggregation in map generalisation by mixed-integer programming." *International Journal of Geographical Information Science* 24 (12): 1871-97.
- Huot-Vézina, G., Boivin, R., Smirnov, A. & Paradis S. J. 2012. *GeoScaler: Generalization tool (with a supplementary user guide in French)*. Geological Survey of Canada, Open File 6231, 2nd Edition.
- Lafay, S., Braun, A., Chandler, D., Michaud, M., Ricaud, L. & Mustière, S. 2015. "Automatic mapping and innovative on-demand mapping services at IGN France." *Cartography and Geographic Information Science* 42 (1): 54-68.

- 
- LGRB. 2015. Regierungspräsidium Freiburg - Landesamt für Geologie, Rohstoffe und Bergbau. "Integrierte geowissenschaftliche Landesaufnahme (GeoLa)." Zuletzt geprüft am 07.05.2018. [http://www.lgrb-bw.de/aufgaben\\_lgrb/geola](http://www.lgrb-bw.de/aufgaben_lgrb/geola).
- LGRB. 2017. Regierungspräsidium Freiburg - Landesamt für Geologie, Rohstoffe und Bergbau. "Ingenieurgeologische Gefahrenhinweiskarte von Baden-Württemberg." Zuletzt geprüft am 12.01.2019. <https://lgrb-bw.de/informationssysteme/geoanwendungen/geogefahren>.
- LGRB. 2018. Regierungspräsidium Freiburg - Landesamt für Geologie, Rohstoffe und Bergbau. "Geofachdaten BW - Hydrogeologie." Zuletzt geprüft am 06.01.2019. [https://produkte.lgrb-bw.de/docPool/readme\\_hk.pdf](https://produkte.lgrb-bw.de/docPool/readme_hk.pdf).
- Li, B., Wilkinson, G. G. & Khaddaj, S. 2001. "Cell-based Model For GIS Generalization." In *Proceedings of the 6th International Conference on GeoComputation*. Brisbane, Australien (24.-26. September), herausgegeben von Pullar, D. V., 6 S.
- Mackaness, W. A. & Ruas, A. 2007. "Evaluation in the map generalisation process." In *Generalisation of Geographic Information*, 89-111: Elsevier.
- McCabe, C. A. 2008. "Vector Approaches to Generalizing Faults and Polygons in 1: 24,000 Geologic Maps: Santa Rosa, California, Case Study." Department of Geography, The Pennsylvania State University.
- McMaster, R.B. & Shea, K. S. 1992. *Generalization in digital cartography*. Washington (DC): Association of American Geographers, 134 S.
- Neun, M. 2007. "Data enrichment for adaptive map generalization using web services." Dissertation, Mathematisch-naturwissenschaftliche Fakultät, Universität Zürich.
- Neun, M. & Steiniger, S. 2005. "Modelling cartographic relations for categorical maps." In *22nd International Cartographic Conference*. La Coruna, Spanien (11.-16. Juli), herausgegeben von ICC, 11 S.
- Peter. 2012. "Polygon Dissolve and generalize (with PostGis)." Zuletzt geprüft am 18.08.2018. <http://www.webspaces.net.nz/page.php?view=polygon-dissolve-and-generalise>.
- Peter, B. 2001. "Measures for the generalization of polygonal maps with categorical data." In *Fourth ICA Workshop on Progress in Automated Map Generalization*. Peking, China (2.-4. August), 21 S.
- Peter, B. & Weibel, R. 1999. "Using vector and raster-based techniques in categorical map generalization." In *Third ICA workshop on progress in automated map generalization*, Ottawa, Kanada (12.-14. August), 14 S.
- Petzold, I., Burghardt, D. & Bobzien, M. 2006. "Workflow management and generalisation services." In *Workshop of the ICA Commission on Map Generalisation and Multiple Representation*. Portland, USA (25. Juni), 12 S.
- Podolskaya, E. S., Anders, K.-H., Haurert, J.-H. & Sester, M. 2009. "Quality Assessment for Polygon Generalization." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Vol. 34 (Part XXX).
- PostGIS Project Steering Committee. 2018. "Postgis 2.4 Manual: Chapter 8. PostGIS Reference." Zuletzt geprüft am 19.01.2019. <http://postgis.net/docs/manual-2.4/reference.html>.
- Ramsey, P. 2010. "PostGIS - Tips for the Power Users." FOSS4G 2010: Selected Presentations. Zuletzt geprüft am 21.10.2018. [http://2010.foss4g.org/presentations\\_show.php?Fid=3369.html](http://2010.foss4g.org/presentations_show.php?Fid=3369.html).
- Rommeney, D. 2006. "Ein Zelluläres Automatenmodell zur Simulation urbaner Landnutzungsänderungen." Diplomarbeit, Mathematisch-Naturwissenschaftliche Fakultät, Universität Potsdam.
- Ross, K. 2009. *Ladder versus star: Comparing two approaches for generalizing hydrologic flowline data across multiple scales*. Pennsylvania State University. Seminar in Cartography: Multiscale Hydrography.
- Ruas, A. & Plazanet, C. 1996. "Strategies for automated Generalization." In *Proceedings of 7th International Symposium on Spatial Data Handling*. Delft, Niederlande (12.-16. August), herausgegeben von Kraak, M.J. & Molenaar, M., Vol. 1, 1-18 Nr. 6.
- Saalfeld, A. 1999. "Topologically consistent line simplification with the Douglas-Peucker algorithm." *Cartography and Geographic Information Science* 26 (1): 7-18.

- 
- SAFE Software. 2018. "FME Transformer Gallery." Zuletzt geprüft am 20.04.2018. <https://www.safe.com/transformers/#/>.
- Santilli, S. 2012. "Simplifying a map layer using PostGIS topology « Strk's Blog." Zuletzt geprüft am 24.11.2018. <http://strk.kbt.io/blog/2012/04/13/simplifying-a-map-layer-using-postgis-topology/>.
- Sayidov, A. & Weibel, R. 2016. "Constraint-based approach in geological map generalization." In *19th ICA Workshop on Generalisation and Multiple Representation*. Helsinki, Finnland (14. Juni), 8 S.
- Schuff, J, Hagemester, A. & Dupke, S. 2018. LGRB-internes Projekt "Automatisierte Generalisierung", Consulting der Conterra GmbH.
- Sester, M. 2000. "Maßstabsabhängige Darstellungen in digitalen räumlichen Datenbeständen." Habilitationsschrift, Fakultät für Bauingenieur- und Vermessungswesen, Universität Stuttgart.
- Šimková, K. & Talhofer, V. 2017. "The Creation of Cartography Model TM100 from DMU25 Data." In *2017 International Conference on Military Technologies (ICMT)*. Brünn, Tschechische Republik (31. Mai - 2. Juni), herausgegeben von Krivanek, V., 257-61.
- Smirnoff, A., Huot-Vézina, G., Paradis, S. J. & Boivin, R. 2012. "Generalizing geological maps with the Geocaler software: The case study approach." *Computers & Geosciences* 40: 66-86.
- Smirnoff, A., Paradis, S. J. & Boivin, R. 2008. "Generalizing surficial geological maps for scale change: ArcGIS tools vs. cellular automata model." *Computers & Geosciences* 34 (11): 1550-68.
- Steiniger, S. & Weibel, R. 2005a. "A conceptual framework for automated generalization and its application to geologic and soil maps." In *22nd International Cartographic Conference*. La Coruna, Spanien (11.-16. Juli), herausgegeben von ICC, 11 S.
- Steiniger, S. & Weibel, R. 2005b. "Relations and structures in categorical maps." In *The 9th ICA Workshop on Generalisation and Multiple Representation*. La Coruna, Spanien (7.-8. Juni), 18 S.
- Steiniger, S. & Weibel, R. 2007. "Relations among Map Objects in Cartographic Generalization." *Cartography and Geographic Information Science* 34 (3): 175-97.
- Stoter, J. E., Burghardt, D., Duchêne, C., Baella, B., Bakker, N., Blok, C., Pla, M., Regnaud, N., Touya, G. & Schmid, S. 2009. "Methodology for evaluating automated map generalization in commercial software." *Geographical Information Generalisation and Multiple Representation* 33 (5): 311-24.
- Stoter, J. E., Post, M., van Altena, V., Nijhuis, R. & Bruns, B. 2014a. "Fully automated generalization of a 1:50k map from 1:10k data." *Cartography and Geographic Information Science* 41 (1): 1-13.
- Stoter, J. E., van Altena, V., Post, M., Burghardt, D. & Duchêne, C. 2016. "Automated generalisation within NMAs in 2016." *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 41: 647.
- Stoter, J. E., Zhang, X., Stigmar, H. & Harrie, L. 2014b. "Evaluation in generalisation." In *Abstracting Geographic Information in a Data Rich World*, 259-97: Springer.
- Thiemann, F., Sester, M. & Bobrich, J. 2010. "Automatic derivation of land-use from topographic data." In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 38. 2 vols, 558-63 Part 2.
- Tong, X., Jin, Y., Li, L. & Ai, T. 2015. "Area-preservation Simplification of Polygonal Boundaries by the Use of the Structured Total Least Squares Method with Constraints." *Transactions in GIS* 19 (5): 780-99.
- Weibel, R. & Dutton, G. 1999. "Generalising spatial data and dealing with multiple representations." *Geographical information systems* 1: 125-55.
- Weibel, R. 2016. „GIS-Operationen in der Kartographie“. Vorlesungsskript. Geographisches Institut, Universität Zürich.
- Wikipedia. 2018. "Bodenkarte." Zuletzt geprüft am 09.12.2018. <https://de.wikipedia.org/w/index.php?oldid=176208157>.
- Wikipedia. 2019. "Geowissenschaften." Zuletzt geprüft am 25.12.2019. <https://de.wikipedia.org/w/index.php?oldid=184513726>.

---

Zhang, J., Wu, C. & Wang, L. 2016. "A conceptual framework for the automated generalization of geological maps based on multiple agents and workflow." *IEEE Access* 4: 6374-85.

Zimmermann, M. 2015. "PostGIS: Buffers, Intersections, Differences And Collections." Zuletzt geprüft am 02.12.2018. <https://www.zimmi.cz/posts/2015/postgis-buffers-intersections-differences-and-collections/>.

---

## Tabellenverzeichnis

Tab. 1: Zuordnung der Operatoren nach (Foerster, Stoter & Köbben 2007).....	11
Tab. 2: Kategorisierung von Mess- und Kennwerten (nach Peter 2001).....	13
Tab. 3: Bezugsmatrix kartografischer Relationen nach Boedecker 2014.....	17
Tab. 4: GeoLa-Fachdatenthemen.....	23
Tab. 5: Im LGRB genutzte Software mit Generalisierungsfunktionalität.....	26
Tab. 6: PostGIS-Funktionen mit Bezug zur Generalisierung (PostGIS Project Steering Committee 2018).....	29
Tab. 7: Für die Generalisierungsprozessierung relevante Funktionen (PostGIS Project Steering Committee 2018)...	30
Tab. 8: GIS-spezifische Generalisierungs-Funktionalität mit Zuordnung zu einschlägigen Operatoren.....	38
Tab. 9: Ausgangsdaten und Hierarchiestufen des Themas Bodenkunde.....	41
Tab. 10: Bezugsmatrix geologischer Themenebenen.....	44
Tab. 11: Kennwerte der Daten zur geologischen Karte in ihren einzelnen Hierarchien.....	45
Tab. 12: statistische Werte auf Aggregationsebene GE1.....	49
Tab. 13: Geometrien mit Stützpunktzahl und innerem/äußerem Topologiebezug.....	50
Tab. 14: Vorliegende Generalisierungssituationen.....	52
Tab. 15: Ausgewählte Simplifizierungsverfahren und ihre Wirkung bezüglich Stützpunktreduktion und Formvereinfachung.....	54
Tab. 16: Verkürzte Testreihe mit quantitativen Aussagen zu Prozessdauer und Flächenanteile der geänderten Kategorien.....	61
Tab. 17: Parameter und (Default)-Kennwerte im Workflow für die Generalisierung der geologischen Daten.....	69
Tab. 18: Überführungstabelle semantische Zusammenfassung.....	72



---

## Abbildungsverzeichnis

Abb. 1: Anforderungsstufen der Generalisierungsautomatisation von Geofachdaten im LGRB.....	5
Abb. 2: Überblick Gliederung der Masterarbeit.....	7
Abb. 3: Zeitleiste mit wissenschaftlichen Beiträgen zur Generalisierungsautomatisation geologischer Karten.....	9
Abb. 4: Konzeptioneller Rahmen der digitalen Generalisierung (nach McMaster & Shea 1992).....	11
Abb. 5: Interaktionsmodell zwischen Bedingungen, Messungen und Relationen (Steiniger & Weibel 2007).....	18
Abb. 6: Schema des Leiter- und des Sternansatz.....	20
Abb. 7: Aufbau der GeoScaler-Toolbox in ArcGIS Desktop.....	25
Abb. 8: In QGIS visualisierte verschachtelte räumliche SQL-Abfrage.....	29
Abb. 9: Beispiel einer FME-Workbench bei der erneuten Ausführung mit aktiviertem FeatureCaching.....	34
Abb. 10: Ausschnitt aus der GK50 (Arbeitsversion) im Randbereich des mittleren Schwarzwalds.....	42
Abb. 11: verkleinerter Ausschnitt einer Datenvisualisierung aus der GK-BW mit (l) und ohne (r) quartäre Deck- schichten.....	43
Abb. 12: Die drei Gliederungsebenen mit ihren geologischen Einheiten und deren Farbgebung.....	45
Abb. 13: Lage der Testbereiche.....	54
Abb. 14: Ablaufschema zum Teilprozess der Polygon-Partitionierung.....	55
Abb. 15: Anwendung der Nullflächenzuteilung anhand des Straight Skeletons.....	56
Abb. 16: Prinzip des zellulären Automaten (Smirnof et al. 2012).....	57
Abb. 17: Prozessmodule der GeoScaler-Toolbox.....	58
Abb. 18: Minimal- und Maximalwerte bei exemplarischer Durchführung des CA-Moduls im GeoScaler.....	59
Abb. 19: Detailliertes Ablaufschema innerhalb des GeoScaler-Moduls 2b (s. Abb. 17).....	61
Abb. 20: Vektor - Raster - Vektor - Skriptdurchlauf im GeoScaler.....	62
Abb. 21: FME-Workbench mit den schematisch hervorgehobenen Teilprozessen.....	63
Abb. 22: verschiedene Teilmodule im Gesamtablauf der FME-Workbench.....	64
Abb. 23: Ablaufschema des designierten Workflows.....	65
Abb. 24: Exemplarische Kategorien des Datenattributs "Flag" bei tektonischen Linien.....	66
Abb. 25: Fallbeispiel für die Auflösung von Objekten nach ihrem semantischem Bezug.....	67
Abb. 26: Modul Statistikausgabe.....	69
Abb. 27: Maßstabsreihe mit Ausgangssituation (GE4) und Generalisierungsergebnissen (GE3-GE1).....	71
Abb. 28: Ausschnitte GE4 - GE1 für Beispiel 2.....	72
Abb. 29: Detailbeispiel für die Linien- und Punkttypisierung aufgrund objektbezogener Beflagung (Stufe GE3 - ca. M 1 : 30 000).....	72
Abb. 30: Ausschnitte GE4 - GE1 für Beispiel 3.....	73
Abb. 31: Detailbeispiel für die CA-Wirkweise bei kleinteiligen Objekten (Stufe GE2).....	73
Abb. 32: Ausschnitte GE4 - GE1 für Beispiel 4.....	74
Abb. 33: Detailbeispiel Stufe GE2 in tektonisch stark zergliederter Region.....	75
Abb. 34: Ausschnitte GE4 - GE1 für Beispiel 5.....	76
Abb. 35: Übergang von Hierarchiestufe GE4 in GE2.....	76
Abb. 36: Anwendung der CA-Methode auf die bodenkundlichen Daten im modifizierten Workflow.....	78
Abb. 37: Beispielausschnitt Generalisierung von Deckschichten mit und ohne Berücksichtigung des Gewässernetz.....	79
Abb. 38: Beispiel für den Topologiefehler einer "self-intersection" in den Ausgangsdaten.....	91
Abb. 39: SQL-Abfrage auf Validität mit Lokalisierung der Inkonsistenzen.....	91

## Anhang

### Anhang A - Auflistung der Werkzeuge mit Generalisierungsfunktionalität in ArcGIS Desktop

TB	Werkzeug	Beschreibung und Kommentierung
Data Management	Zusammenführen (Dissolve)	Fasst Features auf der Grundlage angegebener Attribute zusammen. Zentrale Funktion im Generalisierungsprozess bei der Auflösung von Grenzen zwischen gleichartigen Einheiten höherer Gliederungsordnung.
	Entfernen (Eliminate)	Entfernt Polygone, indem sie mit benachbarten Polygonen zusammengeführt werden, die die größte Fläche oder die längste gemeinsame Grenze besitzen. Entfernen (Eliminate) wird häufig verwendet, um kleine Splitterpolygone zu entfernen, die das Ergebnis eines Überlagerungsvorgangs wie Überschneiden (Intersect) oder Vereinigen (Union) sind.
	Polygon-Teil entfernen (EliminatePolygonPart)	Erstellt eine neue Ausgabe-Feature-Class mit den Features von Eingabepolygonen, wobei einige Teile oder Löcher einer bestimmten Größe gelöscht werden.
Cartography	Punkte aggregieren (AggregatePoints)	Erstellt Polygon-Features um Cluster benachbarter Punkt-Features.
	Polygone aggregieren (AggregatePolygons)	Kombiniert Polygone innerhalb einer bestimmten Entfernung zueinander zu neuen Polygonen.
	Doppellinien zu Mittelachsen zusammenfassen (CollapseDualLinesToCenterline)	Leitet aus Features mit Doppellinien (z. B. Straßenumrisse) Mittelachsen (einzelne Linien) ab, die auf angegebenen Breitentoleranzen basieren.
	Straßendetails reduzieren (CollapseRoadDetail)	Reduziert kleine, offene Anordnungen von Straßensegmenten, die den allgemeinen Verlauf eines Straßennetzes unterbrechen, z. B. Kreisverkehre, und ersetzt sie durch eine vereinfachte Darstellung.
	Bebaute Flächen skizzieren (DelineateBuiltUpAreas)	Erstellt Polygone zur Darstellung bebauter Flächen, indem dicht konzentrierte Anordnungen von Gebäuden auf Karten mit kleinem Maßstab abgegrenzt werden.
	Kartografische Partitionen erstellen (CreateCartographicPartitions)	Erstellt ein Netz von Polygon-Features, die die Eingabe-Feature-Class abdecken, wobei jedes Polygon nicht mehr als eine angegebene Anzahl von Eingabe-Features umschließt, bestimmt durch die Dichte und Verteilung der Eingabe-Features.
	Getrennte Fahrbahnen zusammenführen (MergeDividedRoads)	Generiert Features mit einspurigen Straßen anstelle von zugeordneten Paaren von getrennten Fahrbahnen.
	Gebäude vereinfachen (SimplifyBuilding)	Mit diesem Werkzeug werden Grenzen oder Grundrisse von Gebäudepolygonen unter Beibehaltung der wesentlichen Form und Größe vereinfacht.
	Linie vereinfachen (SimplifyLine)	Hiermit werden Linien durch Entfernen relativ unwesentlicher Stützpunkte vereinfacht, behalten aber die wesentliche Form bei.
	Polygon vereinfachen (SimplifyPolygon)	Hiermit werden Polygon-Umrisse durch Entfernen relativ unwesentlicher Stützpunkte vereinfacht, behalten aber die wesentliche Form bei.
	Linie glätten (SmoothLine)	Glättet spitze Winkel in Linien, um die ästhetische oder kartografische Qualität zu verbessern.
	Polygon glätten (SmoothPolygon)	Glättet spitze Winkel in Umrisslinien von Polygonen, um die ästhetische oder kartografische Qualität zu verbessern.
Straßennetzwerk ausdünnen (ThinRoadNetwork)	Generiert ein vereinfachtes Straßennetzwerk, das die Konnektivität und allgemeine Merkmale zur Anzeige in einem kleineren Maßstab beibehält.	

TB	Werkzeug	Beschreibung und Kommentierung
Production Mapping <sup>24</sup>	Convert Polygons	Löscht Polygone unterhalb Mindestgröße.
	Delete Dangles	Entfernt Linienenden unterhalb einer angegebenen Mindestlänge, die an ihren Enden nicht mit anderen Features verbunden sind.
	Delete Polygons And Extend Lines	Entfernt Polygone unterhalb gegebener Mindestgröße, die und zwei oder mehr Linien verbunden sind. Sobald das Polygon gelöscht ist, werden die Linien so verlängert, dass sie sich am Mittelpunkt des Polygons treffen.
	Extend Polygon Sides	Verlängert die Kanten eines Polygons in Längen und Breite, bis sie einen bestimmten Mindestwert erreichen.
	Generalize Shared Features	Führt die ausgewählten Operationen Simplify and Smooth auf Linien- oder Polygon-Objekten aus. Innerhalb einer Topologie-FeatureClass hilft dieses Tool, die Koinzidenz zwischen gemeinsamen Kanten aufrechtzuerhalten.
	Increase Line Length	Verlängert Linien bis zu einer bestimmten Länge.
	Increase Polygon Area	Vergrößert die Polygonflächen bis sie zu einer Mindestgröße bzw. unter Einbezug einer als Barriere fungierenden anderen FeatureClass.
	Modify Underlying Polygon	Modifiziert Polygone, die unter einem Linienthema liegen hinsichtlich ihrer Lage bzw. ihres Mindestabstandes oder ihres Subtypes.
	Production Centerline	Erzeugt Mittellinien aus hydrographischen Polygonen.
	Prune Hydro	Erzeugt ein vereinfachtes hydrographisches Liniennetz für die Darstellung in kleinerem Maßstab. Das resultierende hydrographische Netzwerk erhält die Hauptarterien und schneidet weniger signifikante Objekte basierend auf Wichtigkeit, Länge und Abstand.
	Split Narrow Polygons	Identifiziert und trennt Bereiche von Polygonen, die schmaler als eine definierte Breite sind.
	Thin Hydro Features	Ausdünnung eines hydrographischen Liniennetzes anhand einer hierarchischen Klassenzuordnung basierend auf den Linienlängen der Nebenflüsse
Thin Spot Heights	Erzeugt Punktobjekte mit Maximal-/Minimalwerten aus einem Raster mit Höhenwerten.	

aus: ArcGIS-Desktop-Online-Hilfe (ESRI 2018b) und (ESRI 2018a)

<sup>24</sup> Das Toolset Production Mapping ist nur als kostenpflichtige Erweiterung verfügbar

## Anhang B - Prä-prozessierung der Ausgangsdaten

Bei der Arbeit mit den Ausgangsdaten zeigte sich ein Problem, das sich in jüngerer Zeit vor allem im Zusammenhang mit älteren bzw. historisch gewachsenen Geodatenbeständen in der Nutzung mit GI-Software neueren Datums zeigt. Die Geometrien der Ausgangsdatenbasis machen in der nativen Datenhaltung in der ArcGIS-SDE und bei der Bearbeitung mittels des ArcGIS-Desktop-Clients keine Schwierigkeiten und werden gemeinhin als valide angenommen, was sich aber bei genauerer Betrachtung als nicht gänzlich zutreffend zeigt. Bei der

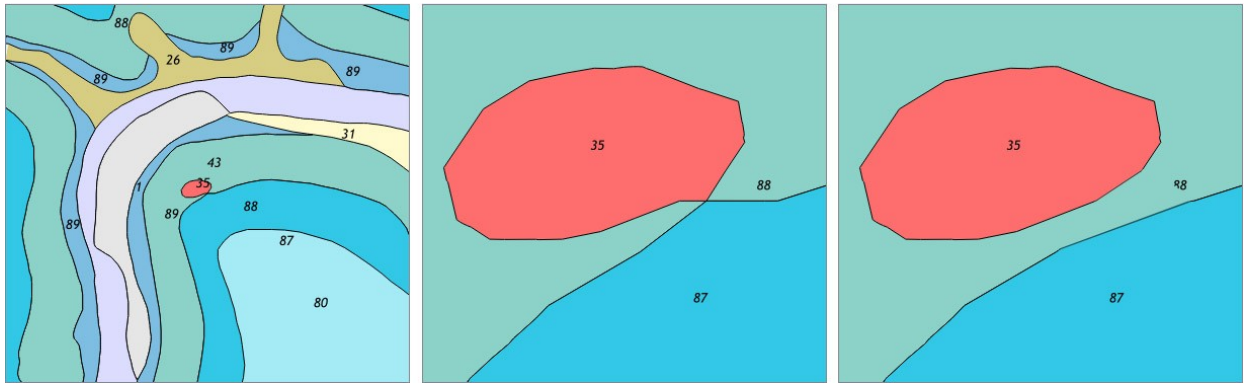


Abb. 38: Beispiel für den Topologiefehler einer "self-intersection" in den Ausgangsdaten

Übernahme in eine PostGIS-DB neuerer Versionierung und Analyse der Daten zeigen sich Unsauberkeiten von Geometrien, die sich in ArcGIS (zumindest in der im LGRB vorliegenden Version 10.4) für die weitere Bearbeitung als unproblematisch abbilden, in anderer GIS-Software neueren Datums aber durchaus als problematisch herausstellen bzw. ohne „Reparatur“ der Geometrien eine weitere Geoprozessierung z. T. nicht möglich ist. Bei den GeoLa-Fachdaten handelt es sich dabei um die topologische Inkonsistenz einer Ring-Selbstüberschneidung, d.h. in den Daten gibt es Polygone, die als innerer Ring genau an einem Knotenpunkt eine direkte Nachbarschaftsbeziehung zu einem eigentlich disjunkten anderen Polygon haben, somit also nicht in Gänze von einem äußeren Polygonring umschlossen sind. Die Ursache dieses in den Basisdaten der geologischen und bodenkundlichen Flächendaten vorhandenen topologischen Fehlers dürfte in einer Änderung der Snapping-Toleranzwert-Einstellung zu einem früheren Zeitpunkt der Bearbeitung liegen oder gegebenenfalls beim Umkopieren zwischen Geodatenbanken mit unterschiedlichen Umgebungswerten zustande gekommen sein. In beiden Fällen führte das dazu, dass Stützpunkte unterschiedlicher ursprünglich disjunkter Polygone aufeinander „gesnapp“ wurden.

In den landesweiten Daten wurden jeweils ca. über 200 Vorkommen solcher Inkonsistenzen identifiziert. Dafür wurde mittels SQL-Query eine direkte Abfrage auf die in der PostGIS-SDE abgelegten Basisdatenbestände der GK50 und BK50 durchgeführt. Diese liefert im Gegensatz zum in ArcGIS vorhandenen Werkzeug CheckGeometry auch eine Verortung der Ursache. Die Ausgangsdaten wurden in einem zusätzlichen aufwändigen Bearbeitungsschritt bereinigt.

id	st_invalidreason	geom
1	Ring Self-Intersection[3401600.0161 5310103.2173]	0106000020EB7...
2	Ring Self-Intersection[3490251.535 5362361]	0106000020EB7...
3	Ring Self-Intersection[3482240.4808 5369759.7588]	0106000020EB7...
4	Ring Self-Intersection[3517352.75 5408908.5]	0106000020EB7...
5	Ring Self-Intersection[3532752.5 5423193.5]	0106000020EB7...
6	Ring Self-Intersection[3521479.3022 5423974.4142]	0106000020EB7...
7	Ring Self-Intersection[3491304.6692 5452467.584]	0106000020EB7...
8	Ring Self-Intersection[3445892.25 5276489.5]	0106000020EB7...
9	Ring Self-Intersection[3489427.7632 5458168.7954]	0106000020EB7...
10	Ring Self-Intersection[3442460.851 5305307.527]	0106000020EB7...
11	Ring Self-Intersection[3424818.3412 5285113.7472]	0106000020EB7...

Abb. 39: SQL-Abfrage auf Validität mit Lokalisierung der Inkonsistenzen

## Anhang C - Python-Skript sowie alternative FME-Umsetzung zur Polygon-Partitionierung

```
# -*- coding: utf-8 -*-
#-----
# Name:          polpartitioner.py
# Purpose:       partition polygons with NULL-Values and
#               fill with category-values of adjacent polygons
# Created:       11.11.2018
# Copyright:     (c) Joachim Schuff 2018
#-----
import psycopg2,sys, types, time
start = time.clock()
conn = psycopg2.connect("dbname='geofachdaten' user='postgres' password='█' host='localhost' port='5433'")
cur = conn.cursor()

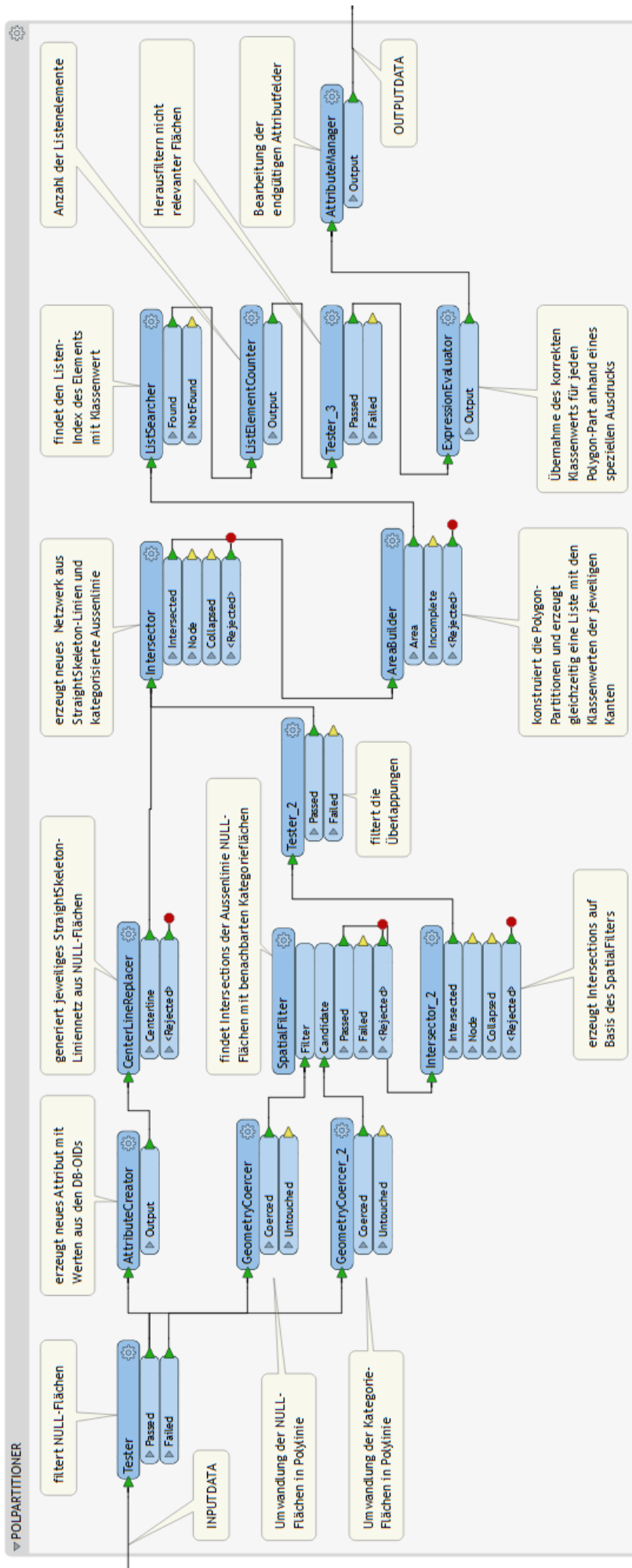
table = 'wissen.bod_boxes_29'

def Skeletonizer():
    # BUILD A PARTITIONED POLYGON OUT OF NULL_POLY-BOUNDARY AND IT'S STRAIGHTSKELETON
    cur.execute("drop table if exists skelet")
    cur.execute("""CREATE TEMP TABLE skelet AS
    WITH multi(geom) AS (
        SELECT st_node(st_collect(sub.geom)) FROM
        (SELECT st_exteriorring(geom) AS geom FROM """+table+"" WHERE oid = """+oid+""")
        union
        SELECT (st_dump((st_straightskeleton(geom))))).geom AS geom FROM """+table+"" WHERE oid =
    """+oid+""")
    sub)
    SELECT row_number() over () AS nr,0 AS reihe,subsub.geom FROM (SELECT (st_dump(st_polygonize(geom))).geom
    FROM multi) subsub""")
    conn.commit()
    cur.execute("SELECT max(nr) FROM skelet")
    row = cur.fetchone()
    return row[0]

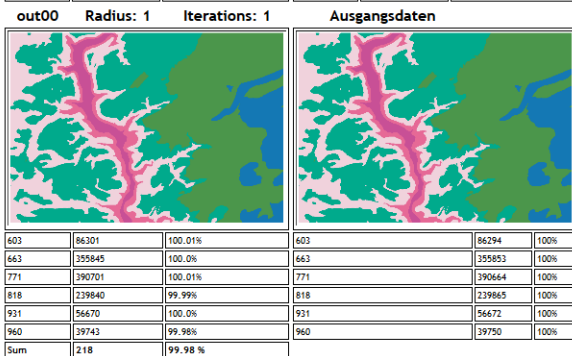
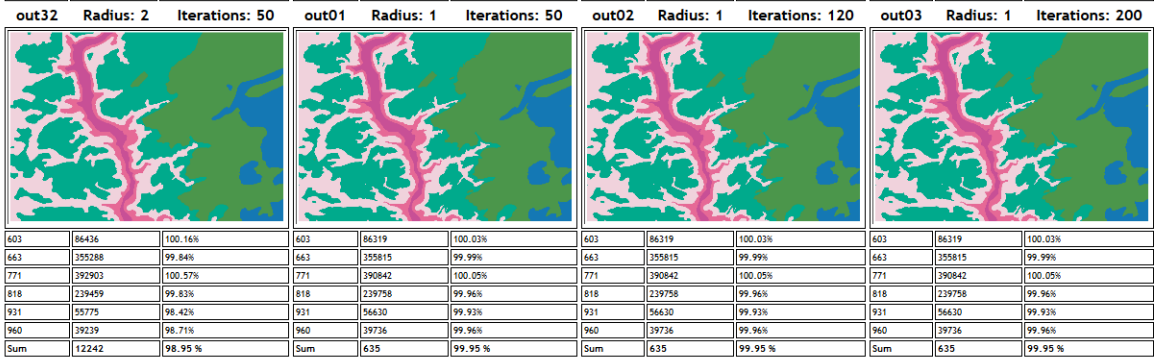
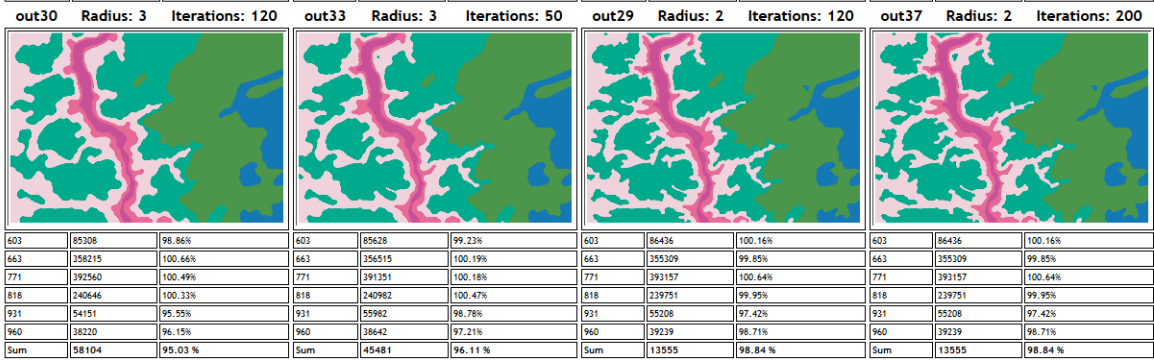
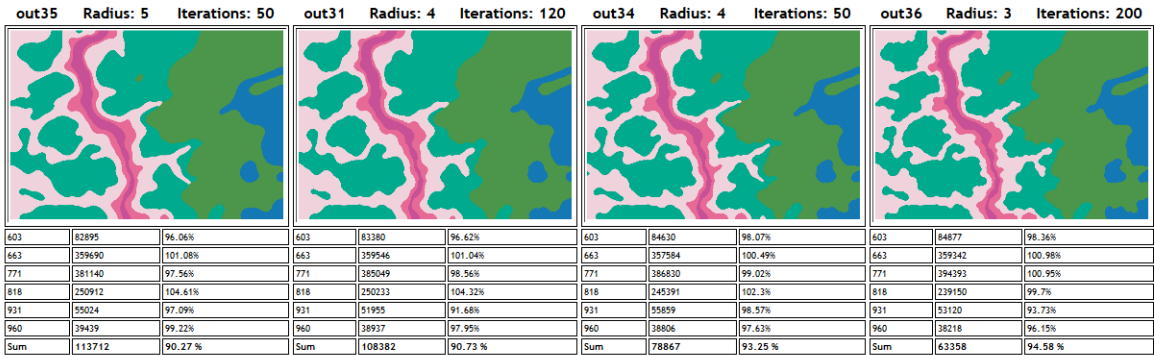
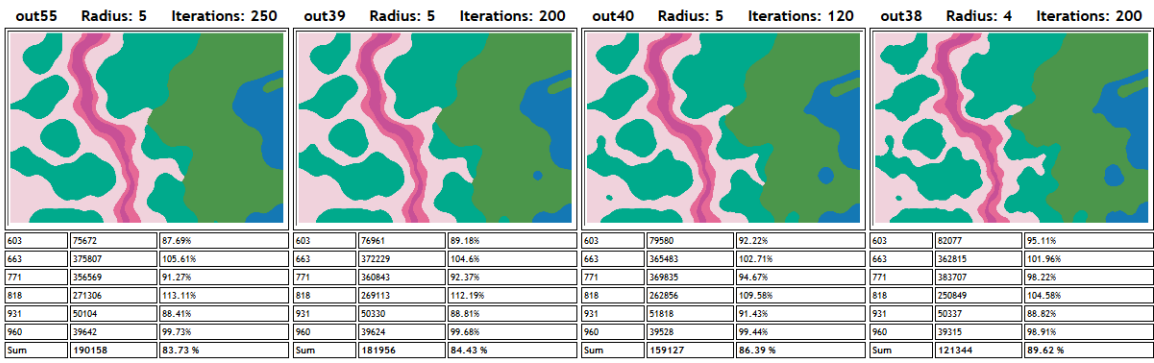
def PolySampler(mnr):
    # FILL THE PARTITIONS WITH THE CORRESPONDING CATEGORY
    cur.execute("""SELECT sub.reihe FROM (
        SELECT b.reihe,st_length(st_intersection(a.geom,b.geom)) AS lgth
        FROM skelet a,"+table+" b WHERE a.nr = "+mnr+" and st_intersects(st_exteriorring(a.geom),b.geom) and b.o-
    id != "+oid+") sub
    WHERE sub.lgth > 0""")
    row = cur.fetchone()
    ##print mnr,row
    if row:
        ##print("insert into "+table+" (select "+str(row[0])+ AS reihe,"+oid+" AS oid,geom FROM skelet where nr
    = "+mnr+"")")
        cur.execute("INSERT INTO "+table+"_ssk (SELECT "+str(row[0])+ AS reihe,"+oid+" AS oid,geom FROM skelet
    WHERE nr = "+mnr+"")")
        conn.commit()

# pre-process PostGIS-Output-Table
cur.execute("drop table if exists "+table+"_ssk")
cur.execute("create table "+table+"_ssk (reihe integer,oid bigint,geom geometry(geometry,31467)) with (OIDS=FAL-
SE)")
conn.commit()

# GET ALL NULL-POLYS...
cur.execute("select oid from "+table+" where reihe = 9999")
rows0 = cur.fetchall()
cnt = 1
# ...AND WALK THROUGH EACH NULL-POLYGON...
for row0 in rows0:
    # global variable
    oid = str(row0[0])
    ##print('...processing oid '+oid+' ('+str(cnt)+' of '+str(len(rows0))+')')
    # call function Skeletonizer to partition polygon
    outnr = Skeletonizer()
    # ...and walk through each Polygon-Part
    for nr in range(1,outnr):
        ##print(nr)
        # call function PolySampler
        PolySampler(str(nr))
    cnt+=1
print(time.clock() - start)
```



# Anhang D - CA-Testreihe



## Anhang E - PythonCaller-Funktion zur semantisch bedingten Objektauflösung

**Transformer**  
Transformer Name: PythonCaller

**Python Script**  
Class or Function to Process Features: FeatureProcessor

```

import fme
import fmeobjects
import numpy

class FeatureProcessor(object):
    def __init__(self):
        pass

    def input(self, feature):
        # listen mit legpos, nächster ebene, fläche
        ll = map(int, feature.getAttribute('_relationships{}.legpos_alles'))
        lh = map(int, feature.getAttribute('_relationships{}.hchy-1'))
        la = map(int, feature.getAttribute('_relationships{}.area'))

        # legpos und nächste ebene des features selbst
        legpos = int(feature.getAttribute('legpos_alles'))
        hchy-1 = int(feature.getAttribute('hchy-1'))

        z = numpy.column_stack((ll, lh))
        a = [list((i, abs(j[0]-legpos), abs(j[1]-hchy-1), la[i])) for i, j in enumerate(z))
        # gibt matrix mit index, abstand zu legpos, abstand zu hchy-1,
        # fläche zurück: [[0, 153, 162, 25731260], [1, 437, 421, 26834], ...]
        a = numpy.array(a)

        nearest_legpos = min(a[...], 1)

        if list(a[...], 1).count(nearest_legpos) == 1:
            feature.setAttribute("_index", list(a[...], 1).index(nearest_legpos))
            self.pyoutput(feature)
            pass
        else:
            # print "count: " + str(list(a[...], 1).count(nearest_legpos))
            indices = [i for i, x in enumerate(a[...], 1) if x == nearest_legpos]
            # print indices
            h = []
            for i in indices:
                h.append(a[i][2])
            nearest_hchy-1 = min(h)
            if list(a[...], 1).count(nearest_hchy-1) == 1:
                feature.setAttribute("_index", list(a[...], 2).index(nearest_hchy-1))
                self.pyoutput(feature)
                pass
            else:
                areas = []
                for i in indices:
                    areas.append(a[i][3])
                max_area = max(areas)
                feature.setAttribute("_index", list(a[...], 3).index(max_area))
                self.pyoutput(feature)
                pass
            pass

    def close(self):
        pass
    
```

**Output Attributes**

Attributes to Expose:

Attributes To Hide:

Lists to Hide:

Buttons: Help, Defaults, OK, Cancel



## Anhang F - Statistikausgabe im FME-Workflow

### Box 5 - Pre Stats

Klasse	Anzahl	Maximum	Minimum	Mean	Bereich	StandardDev	Summe	NPoints
456	1	108988.0	108988.0	108988.0	0.0		108988.0	63
575	1	637038.0	637038.0	637038.0	0.0		637038.0	201
650	6	1080098.0	6762.0	263813.833333	1073336.0	414175.26666	1582883.0	554
817	2	7245178.0	19074.0	3632126.0	7226104.0	5109627.13996	7264252.0	566
1014	7	129835.0	17327.0	62273.8571429	112508.0	39848.7725759	435917.0	412
1060	3	158564.0	22846.0	80497.3333333	135718.0	70124.4127058	241492.0	172
1071	3	852808.0	3907.0	464462.333333	848901.0	429032.520819	1393387.0	442
1140	7	2864.0	1168.0	2051.28571429	1696.0	714.431642236	14359.0	219
1143	30	96514.0	1850.0	16933.9666667	94664.0	22193.5632894	508019.0	2272
1150	24	57727.0	122.0	10221.2083333	57605.0	12113.8300301	245309.0	1026
1151	43	8503103.0	3.0	420408.023256	8503100.0	1803558.35364	18077545.0	5797
1186	1	4759.0	4759.0	4759.0	0.0		4759.0	40
1196	30	21444.0	1465.0	5543.1	19979.0	5326.67294329	166293.0	1203
1209	6	41709.0	1476.0	12105.6666667	40233.0	14768.0438334	72634.0	301
1233	17	30552481.0	1343.0	1938017.23529	30551138.0	7392163.78388	32946293.0	7950

### Box 5 - Post Stats

Klasse	Anzahl	Maximum	Minimum	Mean	Bereich	StandardDev	Summe	NPoints
456	1	97466.6017651	97466.6017651	97466.6017651	0.0		97466.6017651	370
575	1	700787.611597	700787.611597	700787.611597	0.0		700787.611597	1047
650	3	1051173.41492	154223.609499	493837.275547	896949.805418	486499.789211	1481511.82664	2188
817	1	7316503.72489	7316503.72489	7316503.72489	0.0		7316503.72489	2773
1014	1	115610.407701	115610.407701	115610.407701	0.0		115610.407701	205
1060	1	227793.874801	227793.874801	227793.874801	0.0		227793.874801	574
1071	2	937176.141223	578167.734069	757671.937646	359008.407154	253857.279202	1515343.87529	1904
1151	4	8727366.21723	201223.530383	4469651.61119	8526142.68685	4889154.18695	17878606.4448	6514
1233	2	32119484.2913	2296947.97856	17208216.1349	29822536.3127	21087717.6589	34416432.2698	6358