

Verfahren zur automatisierten Generalisierung flächenhafter Geofachdaten

A Workflow for the Automated Generalization of Polygonal Data

Joachim Schuff

Regierungspräsidium Freiburg · joachim.schuff@rpf.bwl.de

Zusammenfassung: Nach den weitreichenden Fortschritten bei der modellbasierten Generalisierungsautomatisierung topographischer Karten im letzten Jahrzehnt wächst nun die Nachfrage nach ganzheitlichen Lösungsansätzen für die Generalisierung flächenhafter Fachdaten. Die in jüngerer Zeit durchgeführte Harmonisierung von Basisdatenbeständen der geologischen und bodenkundlichen Landesaufnahme im Landesamt für Geologie, Rohstoffe und Bergbau (LGRB) in Baden-Württemberg ermöglichen nun in einem weiteren Schritt die modellbasierte Ableitung von Daten in generalisierter Form für verschiedene Zielmaßstäbe und unterschiedliche Zwecke. Die für die Umsetzung verfügbaren Verfahren und Methoden aus Standard-GIS wurden gesichtet und hinsichtlich eines praktikablen Lösungsansatz untersucht. Im vorliegenden Beitrag wurde am Beispiel geologischer Verbreitungsflächen ein praxisgerecht einsetzbares Prozess-Framework basierend auf der Software FME erstellt. Dem Verfahren wurde ein hybrider Lösungsansatz zugrunde gelegt, bestehend aus vektorieller und rasterbasierter Verarbeitung. Es liefert anhand der Implementierung sowohl generischer vektorieller Transformer wie auch unter der Zuhilfenahme und Integration der rasterbasierten Methode eines zellulären Automaten ein vollautomatisiertes Verfahren zur Generalisierung modellbasierter kategorischer Flächendaten. Vorausgesetzt wird diesem Ablauf ein hierarchisches Klassifizierungsschema bei den Ausgangsdaten. Eine fachspezifische Attributanreicherung ermöglicht die Berücksichtigung weiterer thematischer Aspekte.

Schlüsselwörter: Geologische Karte, Generalisierung, Automatisierung

Abstract: *After the progress in generalization-automation of topographic maps in the past decade, the demand for more comprehensive solutions for the generalization of categorical maps and data is growing. Recent harmonization of basic data sets of the Geological and Soil Survey of the Landesamt für Geologie, Rohstoffe und Bergbau Baden-Württemberg now enables a model-based derivation of data in generalized form for different target scales and purposes. A particular challenge is the implementation with existing functionality from available GI-software. The corresponding procedures and methods were reviewed and a solution approach has been developed. In this work, a practical process framework based on the FME-software was created using data from bedrock geology. The procedure was based on a hybrid solution approach, consisting of vector- and raster-based processing. A fully automated procedure for the generalization of model-based categorical data was developed, based on the implementation of generic vector transformers and the integration of the raster-based method of a cellular automaton. The process requires a hierarchical structure of the input data. That enables a semantic aggregation and includes the consideration of further thematic aspects by means of a data-specific attribute enrichment.*

Keywords: *Geological map, generalization, automation, workflow*

1 Einführung

Die geologische und bodenkundliche Landesaufnahme liefern wesentliche Grundlagen für die Kenntnis und die Nutzung der natürlichen Ressourcen eines Landes. Die Erstellung einer landesweiten, harmonisierten Datengrundlage für die Basisthemen Geologie und Bodenkunde war eine der vordringlichen Aufgaben des Landesamtes für Geologie, Rohstoffe und Bergbau (LGRB) im Rahmen der integrierten geowissenschaftlichen Landesaufnahme in Baden-Württemberg (GeoLa) im vergangenen Jahrzehnt (LGRB, 2015). Seit 2016 werden die Daten nun flächendeckend basierend auf landeseinheitlichen Taxonomien in einer Geodatenbank vorgehalten und unterliegen der ständigen weiteren Fortschreibung und Aktualisierung. Aus diesem Basisdatenmodell des Maßstabsbereiches 1:25.000 bzw. 1:50.000 soll nun eine generalisierte Ableitung der Daten für gröbere Maßstäbe erfolgen. Die bisher übliche Vorgehensweise einer zumeist nicht klar regelbasierten manuellen fachlich-kartografischen Generalisierung ist heutzutage nicht mehr zweckdienlich. Dennoch mangelt es nach wie vor – bei thematisch-kategorischen Flächendaten im Besonderen – an Lösungsansätzen zur ganzheitlichen automatisierten Generalisierung (Sayidov & Weibel, 2016). Der vorliegende Beitrag untersucht die Möglichkeiten zur praxisingerechten Umsetzung eines holistischen Verfahrens anhand verfügbarer Generalisierungsfunktionalität im Umfeld gebräuchlicher GIS-Software. Dabei soll auch die Methodenimplementierung aus schon bestehenden Lösungsansätzen in Betracht gezogen und auf ihre Eignung getestet werden. In diesem Fall handelt es sich um das rasterbasierte Verfahren eines zellulären Automaten, wie er in der ArcGIS-Toolbox GeoScaler (Huot-Vézina et al., 2012) Anwendung findet.

1.1 Rahmenbedingungen

Ausgangsbasis ist die Fragestellung, in wie fern die thematisch-kategorischen GeoLa-Daten mit den im LGRB zur Verfügung stehenden GIS-Funktionalitäten im Rahmen einer ganzheitlichen Generalisierungslösung für unterschiedliche Maßstäbe und Ausgabezwecke gehandhabt werden können. Das Ziel ist dabei die Umsetzung eines durchgehend automatisierten Prozessablaufs in der möglichst reproduzier- und übertragbaren Anwendung auf die Flächendaten der Bereiche Geologie und Bodenkunde sowie anhängiger angewandter Bereiche, wie z. B. die Hydrogeologie. Zentrales Element für die jeweilige themenbezogene inhaltliche Zusammenfassung (semantische Aggregation) ist ein vorgegebenes hierarchisches Gliederungsschema der im Land vorkommenden fachspezifischen Einheiten.

Diesem einführenden Kapitel folgt die Einordnung in den wissenschaftlichen Kontext mit Sichtung vorhandener einschlägiger Literatur. Als Grundlage für die weiteren Untersuchungen dient eine Kategorisierung der Anforderungsstufen an die automatisierte Generalisierung. Die folgende Situationsanalyse setzt sich mit den im GIS-Umfeld der Arbeit (ArcGIS, PostGIS, FME) verfügbaren Funktionen und Algorithmen sowie den vorliegenden Daten und konkret vorhandenen Generalisierungsfällen auseinander. Der Auseinandersetzung mit dem methodischen Konstrukt zur Bewältigung dieser Konflikte folgt dessen konkrete Implementierung in einem Lösungsansatz mittels FME sowie die Präsentation der Resultate. In der abschließenden Diskussion werden Aussagen zur räumlichen und thematischen Übertragbarkeit der Methode gemacht und in einem Ausblick noch auf Möglichkeiten zur Verfeinerung und Verbesserung des Ablaufs eingegangen.

2 Grundlagen

2.1 Abriss über den bisherigen Forschungsstand

Entgegen der mittlerweile fast unüberschaubaren Anzahl von Beiträgen im Bereich topographischer Karten und Datenbestände gibt es nur wenige Arbeiten, welche sich konkret mit der Generalisierungsautomatisation flächenhafter Verbreitungsdaten von thematischen Karten befassen. Als Themenfelder stehen dabei Landnutzung, Bodenkarte und Geologische Karte im Fokus. Alle drei beschäftigen sich mit dem gleichen Ausschnittsvolumen der sich an der Erdoberfläche abbildenden jeweiligen Einheiten. Im Zusammenhang mit der Landnutzung stehen allerdings Fernerkundungsdaten und dementsprechend rasterbasierte Methoden zur Analyse und Verarbeitung der Grundlagen im Fokus. Bei der Bodenkarte und geologischen Karte stehen vor allem semantische wie auch spezifische themenrelationale Aspekte im Vordergrund.

Bader & Weibel (1997) widmeten sich gegen Ende des 20. Jahrhunderts als Erste der Beschreibung von Konfliktsituationen bei der Polygongeneralisierung, während Peter (2001) eine vollständige Auflistung von wichtigen Mess- und Kenngrößen liefert. 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. 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. Nur wenige Autoren haben sich im speziellen und explizit mit den theoretischen und praktischen Problemen bei der Generalisierungsautomatisation von geologischen Karten beschäftigt. Neben Artioli et al. (1995) sind dies Downs & Mackaness (2002), die einen kontextuell sensitiven regelbasierten teilautomatisierten Prozessablauf zur Generalisierung eines geologischen Ausgangsdatensatzes im Maßstab 1:50.000 nach 1:250.000 vorstellten. McCabe (2008) entwickelte und beschrieb einen Generalisierungsworkflow, der basierend auf der Anwendung einer Reihe von ArcGIS-Werkzeugen eine geologische Karte im Gebiet Santa Rosa in drei aufeinanderfolgenden Maßstabsstufen (1:50.000, 1:100.000 und 1:250.000) 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 Smirnov, Paradis & Boivin (2008) die rasterbasierte Methode des zellulären Automaten im Umfeld der Generalisierung thematischer Karten angewandt und beschrieben, was zur Entwicklung einer umfangreichen ArcGIS-Toolbox führte (Smirnov et al., 2012).

Diese Software mit der Bezeichnung GeoScaler deckt weitestgehend die unterschiedlichen kartografischen Aspekte der Generalisierung geologischer Karten ab und ermöglicht mittels sequenzieller interaktiver Steuerung bei Verwendung gleicher Eingabeparameter reproduzierbare Generalisierungsergebnisse. In einer jüngeren Arbeit propagieren Zhang, Wu & Wang (2016) ein agentenbasiertes Rahmenkonzept unter Einbezug der semantisch-hierarchischen Belange geologischer Karten.

2.2 Anforderungskategorien

Auf der Grundlage des vorliegenden modellbasierten Fachdatenbestands lassen sich unterschiedlich differenzierte Zielsetzungen und Anforderungen an die Generalisierungsautomatisation ableiten (**Abb. 1**):

- A) Themenunabhängige Flächengeneralisierung ohne weitergehende geometrische Transformation: Mit der Zielsetzung einer einfachen Zusammenfassung und Reduktion vorhandener Einheiten durch Neuklassifizierung im Rahmen einer semantischen Aggregation.
- 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 Löschen von Flächen handeln. Der Erhalt des topologischen Bezugs ist nicht gefordert.
- C) Themenunabhängige Flächengeneralisierung unter Einbezug allgemeiner Bedingungen: Anwendung weiterer charakteristischer Generalisierungsoperationen wie geometrische Zusammenfassung, Auswahl oder Vereinfachung unter Wahrung topologisch-geometrischer Bezüge. Eignung für kartografische Ausgaben.
- D) Themenunabhängige Flächengeneralisierung unter Einbezug fachlicher Bedingungen: Einfluss zusätzlicher wertender und gewichtender fachlicher Faktoren unter Wahrung topologischer Abhängigkeiten zu Informationsebenen innerhalb des Datenmodells.
- E) Themenübergreifende Flächengeneralisierung unter größtmöglicher Berücksichtigung topologischer und fachlicher Bedingungen: Vollumfänglicher Einbezug von Topologie- und themenübergreifend-relationaler Abhängigkeiten innerhalb des Gesamtmodells.

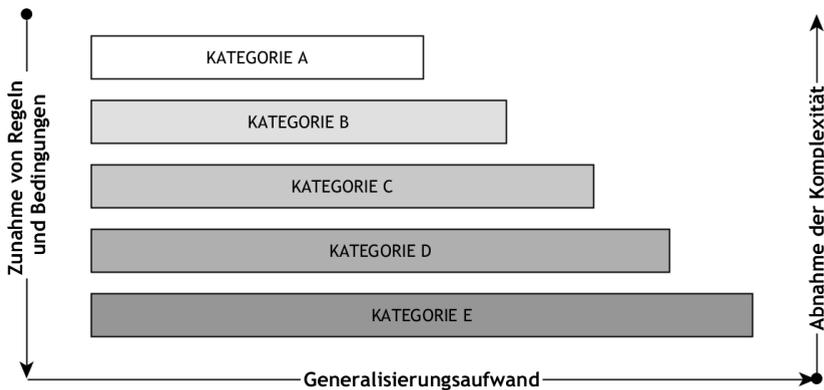


Abb. 1: Anforderungsstufen der Generalisierungsautomatisation von Geofachdaten im LGRB

3 Situationsanalyse

3.1 Vorhandene Funktionalität in verfügbaren GI-Systemen

Zahlreiche handelsübliche GIS stellen Standardfunktionalitäten zur Verfügung, die für eine Generalisierung genutzt werden können. Unterschiede bestehen in Art und Umfang der Funktionalität bzw. bei den hinterlegten Algorithmen. Die in der Geobranche weit verbreite-

ten Software-Systeme ArcGIS Desktop von ESRI, das RDBMS¹ PostgreSQL mit seiner räumlichen Erweiterung PostGIS sowie die ETL²-Software FME bilden sowohl einzeln wie auch im Zusammenspiel wesentliche Bausteine der Geodaten-Infrastruktur im LGRB. Die verfügbaren Funktionalitäten wurden hinsichtlich der nach Foerster, Stoter & Köbben (2007) identifizierten Operatoren Auswählen, Klassifizieren, Kollabieren, Aggregieren, Vereinfachen, Betonen, Eliminieren, Verdrängen und Verschmelzen zugeordnet und untersucht³.

Zusammenfassend lässt sich feststellen, dass ArcGIS eine Reihe explizit für die Generalisierung von topographischen Kartenwerken konzipierte Werkzeuge bereitstellt, wobei einige dieser Funktionen aber auch im Umfeld der Generalisierung von Polygonmosaiken angewendet werden können. PostGIS liefert keine vorkonfektionierten Funktionen, bietet aber ein überaus mächtiges Repertoire an Funktionalitäten, 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 liefert die Software FME zusammen mit einem ebenfalls äußerst breiten Spektrum an sogenannten „Transformern“, die gleichermaßen die wesentlichen Teilprozesse einer Generalisierungsautomatisation abdecken.

3.2 Datengrundlage

Die Geologie mit ihren lithostratigraphischen Verbreitungsflächen ist Bezugsgrundlage und Datenbasis für weitere Fachthemen wie die Hydrogeologie oder Ingenieurgeologie mit mehr oder weniger starken inhaltlich-geometrischen Abhängigkeiten. Die Komplexität der Generalisierungsaufgabe ist sowohl in semantischer wie auch geometrisch-topologischer Hinsicht im Vergleich zu den anderen GeoLa-Themen⁴ innerhalb des Themas „Geologie“ am größten, sodass hier der Ausgangspunkt für eine prototypische Generalisierungslösung anhand der Festlegung von Minimal- wie auch Maximalanforderungen zu sehen ist bzw. sowohl allgemeine wie auch speziell zu lösende Generalisierungsszenarien in der Mehrzahl aufzufinden sind. Die Unterschiede bei der Generalisierung geologischer Karten im Vergleich zu anderen Karten sind bezüglich des Inhalts, der Zielsetzung und des Zwecks sowie der Ausprägung auszumachen. 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 in der Zielsetzung Eingang finden (Zhang, Wu & Wang, 2016). Eine der Schlüsseleigenschaften einer geologischen Karte (sowie auch der bodenkundlichen und hydrogeologischen Kartengrundlagen) ist zudem die flächendeckende Verbreitung von Polygonen ohne Lücken und Überlappung – das Hauptaugenmerk gilt demnach in höherem Maße der Flächengeneralisierung mit der nachgeordneten Auswirkung bezüglich einer Typisierung (Kollaps) mindergroßer, herauszuhebender Flächenstrukturen in Linien- und Punktobjekte, womit sich zwangsläufig ein stärkerer Einbezug inhaltlich-semantischer Gegebenheiten zusätzlich zu den „normalen“ Generalisierungsoperationen ergibt.

¹ Relationales Datenbank-Management-System

² Extract – Transform – Load

³ s. Schuff (2019), p. 23 ff.

⁴ Bodenkunde, Hydrogeologie, Ingenieurgeologie, Rohstoffgeologie

Auch bezüglich der taxonomischen Vorgabe für die inhaltliche Gliederung weist das Thema „Geologie“ den differenziertesten Überbau auf: es ist aufgeteilt in eine 4-stufige Hierarchieabfolge, die ausgehend von der Geologischen Basiskarte 1:50.000 aus ursprünglich ± 450 flächenhaften und linienhaften Einheiten besteht. Diese Kartiereinheiten weisen einen jeweiligen 1:1-Bezug zu einer übergeordneten Klasseneinheit der nächstfolgenden Hierarchiestufe auf, welche sich wiederum direkt einer weiteren zusammenfassenden Kategorie höherwertigen Ranges zuordnen lässt. Die fachliche Klassenzuteilung basiert somit auf einer lithostratigraphischen Zusammenfassung in der Abfolge Formation \rightarrow Untergruppe \rightarrow Gruppe \rightarrow Hauptgruppe und führt zu den Kennwerten aus Tabelle 1, in welcher auch die für das Generalisierungsziel vorgesehenen Maßstabsbereiche hinterlegt sind. Die 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 Web-GIS bzw. Web Mapping-Anwendungen eine Rolle spielen.

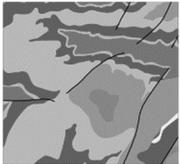
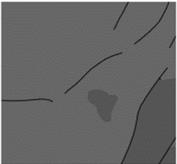
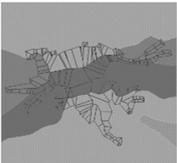
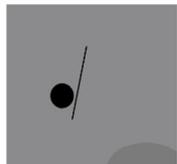
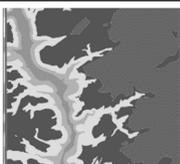
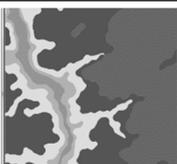
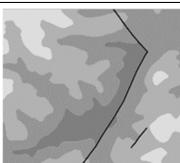
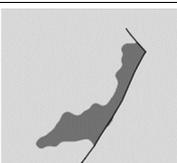
Als Ausgangsbasis für die spätere Generalisierung dienen die Daten der abgedeckten geologischen Karte. Durch das Fehlen der mitunter feingliedrigen dendritischen holozänen Deckschichten vereinfacht sich die Generalisierungsaufgabe. Da die abgedeckte geologische Karte zusammen mit den Deckschichten wieder die ursprüngliche geologische Karte ergibt, kann die Generalisierung an den beiden Datensätzen getrennt durchgeführt und die einzelnen Ergebnisse im Nachgang wieder zusammengefügt/überlagert werden.

Tabelle 1: Taxonomisch-hierarchische Kennwerte sowie Maßstabsrelevanz beim Geologie-Thema

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 – 1:3.500.000

Aus der vorliegenden Datensituation wurden typische Generalisierungssituationen detektiert. Anschließend fand eine Klassifizierung und Abstrahierung charakteristischer Konflikte nach den genannten Anforderungsstufen statt, für deren Lösung mögliche Operatoren gegenübergestellt wurden. In Tabelle 2 ist eine Auswahl an Ausgangsszenarien gelistet, welche als vorrangig für den späteren Lösungsansatz im Generalisierungsablauf identifiziert wurden.

Tabelle 2: Auswahl lokalisierter Generalisierungsszenarien als Ausgangssituation für die Umsetzung im Gesamttablauf (mit Kategorie-Zuordnung aus Abb. 1)

Kat	Beschreibung	Ausgangsbild	Ziel
B1	Einfache Zusammenfassung von Einheiten in neue Kategorien mit Grenzlinienauflösung → geometrisch-semantische Aggregation		
B2	Null-Flächen ohne Kategorie → Flächenzuordnung zu benachbarten Objekten: Polygon-Partitionierung basierend auf „straight skeletons“ mit anschließender Aggregation		
C1	Objekt unterhalb Mindestgröße → mit Beflagung: Geometrieübergang zu Punkt-symbol bzw. Linie ohne Beflagung: Auflösung		
C2	Objekt unterhalb Mindestgröße mit mehreren Nachbarn unterschiedlicher Kategorien, soll anhand des lithologisch ähnlichsten Nachbarn aufgelöst werden		
C3	Komplexe Objekte auf Makroebene → Formvereinfachung im Gesamtflächenkontext (starke Reduktion von Verästelungen und schmalen Fortsätzen)		
D1	Einbezug der tektonischen Linien als invariante Bereiche → Erhalt topologischer Bedingungen		

4 Vorgehensweise

4.1 Bewerkestellung der geometrischen Formvereinfachung

Anhand der vorliegenden Datenthematik und der im Arbeitsumfeld zur Verfügung stehenden Softwarefunktionalitäten wurden Tests bezüglich der Eignung einschlägiger Simplifizierungsalgorithmen bzw. in der Praxis schon vorhandener Verfahren durchgeführt (Schuff,

2019, pp. 52 f.). Der Generalisierungseffekt ergibt sich allerdings nicht nur in Bezug auf die Einzelgeometrie in ihrer Relation zur Nachbarschaft als Resultat entsprechender ausgeführter transformativer Operationen und der daraus abgeleiteten Messwerte. Ein wesentlicher Aspekt ist die Formvereinfachung der Objekte im Kartenkontext, also auf der Ebene des gesamten Polygonmosaiks. Die Einzelalgorithmen – vor allem unter Wahrung des topologischen Bezugs – können hier nicht alleine zum gewünschten Ergebnis führen, sie entsprechen auch nicht dem Vorgehen des Kartografen bei der Generalisierung (Ellsiepen, 2006, p. 124). Ziel führend ist vielmehr eine gesamthaft wirkende Vereinfachungsstrategie als Ersatz der ursprünglich manuellen oder teilautomatisierten kartografischen Wirkprozesse. Von den untersuchten Verfahren ermöglichen lediglich die ArcGIS-Funktion SmoothPolygon sowie das ArcGIS-basierte GeoScaler-Verfahren eine über das Normalmaß (einer Stützpunktreduktion) hinausgehende Formvereinfachung, die zusätzlich auch die Möglichkeit zur Wahrung des topologisch-thematischen Bezugs im Gesamtflächenkontext beinhaltet.

4.2 Zellulärer Automat als Kernkomponente zur Generalisierung im Gesamtflächenkontext

Ein Zellulärer Automat (im Folgenden mit der Bezeichnung „CA“ abgekürzt, nach „cellular automaton“) ist definiert als ein räumlich und zeitlich diskretes dynamisches System, bestehend aus einem regulären Raster (Array) von gleichförmigen identischen Zellen, deren Zustände sich in Abhängigkeit der sie umgebenden Nachbarzellen und der angewendeten Regel zu einer neuen Generation von Zellzuständen entwickeln (Rommeney, 2006, p. 11). Die vier wichtigsten Systemparameter sind (1) die Regel, auf deren Grundlage sich der neue Zellzustand errechnet, (2) die Nachbarschaft, die den Zellzustand beeinflusst, (3) der Zustand der Zelle sowie (4) die Anzahl der durchzuführenden Iterationen.

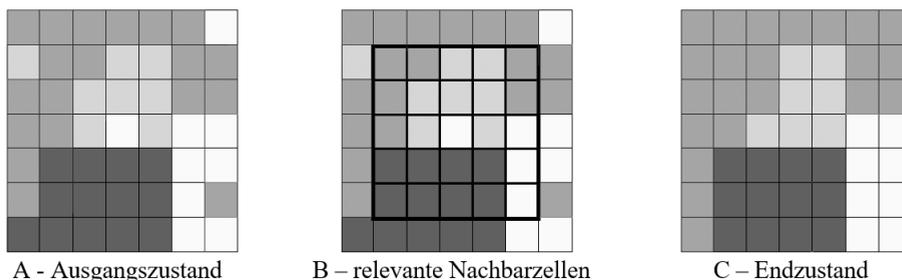
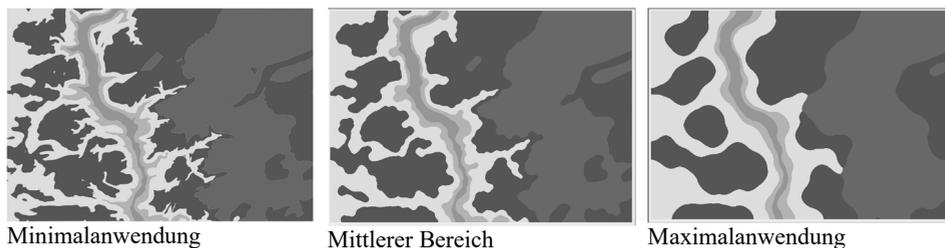


Abb. 2: Prinzip des zellulären Automaten (Smirnof et al., 2012, p. 69)

Für die Polygon-Generalisierung ist das Array im Ausgangszustand ein Rasterbild der zu generalisierenden geologischen Karte (**Abb. 2A**). 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. 2B** dargestellt, wird der Nachbarschafts-

radius durch die Breite des Korridors beschrieben, der von den relevanten Nachbarzellen gebildet wird, die um die zu berechnende Zelle liegen (Moore’s Nachbarschaft). Die letzte Generation ist definiert als diejenige, nach der keine weiteren Zustandsveränderungen (Abb. 2C) mehr auftreten (Smirnof et al., 2012).

Die geeignete Parametrisierung des CA wurde empirisch auf Basis einer Testreihe optimiert (Abb. 3). Der Veränderungsbereich der anteiligen Flächenkategorien bewegt sich dabei zwischen < 1 % im Minimum und > 13 % im Maximum, die Gesamtveränderungsrate der Rasterzellen kann bis zu 17 % betragen.



603	86301	100.01 %	603	85308	98.36 %	603	75672	87.69 %
663	355845	100.0 %	663	358215	100.98 %	663	375807	105.61 %
771	390701	100.01 %	771	392560	100.95 %	771	356569	91.27 %
818	239840	99.99 %	818	240646	99.7 %	818	271306	113.11 %
931	56670	100.0 %	931	54151	93.73 %	931	50104	88.41 %
960	39743	99.98 %	960	38220	96.15 %	960	39642	99.73 %
Diff	218	99.98 %	Diff	58104	94.58 %	Diff	190158	83.73 %

Radius: 1 Iterationen: 1

Radius: 3 Iterationen: 120

Radius: 5 Iterationen: 250

Abb. 3: Generalisierungseffekt des zellulären Automaten – Ausschnitt aus Testreihe

Das wesentliche Kriterium für die weitere Verwendung des zellulären Automaten als zentrale Methode im Generalisierungsablauf ist die Wahrung der topologischen Konsistenz einander angrenzender Polygone, was in einem rasterbasierten Verfahren per se gegeben ist. Zusätzlich besteht die Möglichkeit des Einbezugs invarianter Barriere-Bereiche (im Geologie-Kontext als tektonische Lineationen), in denen die Zellen von einer Veränderung ausgeschlossen sind. Ein weiterer durchaus gewünschter Nebeneffekt liegt in der durch die anschließende Raster-Vektor-Rückwandlung verursachten Geometrievereinfachung (Stützpunktreduktion).

5 Ergebnis

5.1 Generalisierungskonstrukt

Am Beispiel der flächenhaften geologischen Ausgangsdaten wurde der Ablauf (Abb. 4) folgendermaßen konzipiert:

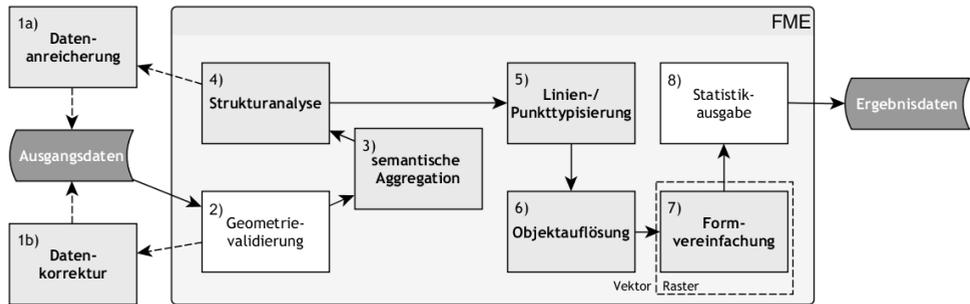


Abb. 4: Schematisierter Generalisierungsablauf

Voraus geht die Kennzeichnung (1a) speziell zu berücksichtigender Objekte im Ausgangsdatenmodell durch Datenanreicherung. Die Geometrievalidierung (1b) stellt sicher, dass der Workflow nicht aufgrund fehlerhafter Objekte vorzeitig abbricht. Die klassenbasierte Zusammenfassung der Objekte auf semantischer Ebene (2) beginnt den eigentlichen Generalisierungsprozess. Mittels Strukturanalyse werden verschiedene Mess- und Kennwerte zu länglichen, minder großen bzw. erhaltenswerten Objekten ermittelt (4). Es folgt die auf dieser Festlegung basierende Punkt- und/oder Linientypisierung (5). Daran schließt sich die Zuteilung der Flächenbereiche zu eliminierender Objekte an (6), gefolgt von einer allgemeinen Vereinfachung der Flächen im Kartenkontext mittels CA (7). Mit der Ausgabe der Ergebnisdaten ist auch die Erstellung eines tabellarischen Vorher/Nachher-Vergleichs verbunden (8). In den einzelnen Ablaufblöcken finden die Generalisierungsoperatoren Klassifizieren, Aggregieren, Eliminieren, Kollabieren (Typisieren) und Vereinfachen Anwendung.

5.2 Orchestrierung des Workflows mittels FME

Die Umsetzung des Gesamtprozesses erfolgte mit FME. Die Möglichkeit der Implementierung externer (auch ArcGIS-)Methoden gewährleistete die hierfür erforderliche Flexibilität. Die resultierende Workbench ist modular strukturiert. Die für die Prozessierung erforderlichen Eingabeparameter werden vollumfänglich zum Start des Ablaufs eingegeben bzw. über Standardwerte definiert. Als wesentliche Parameter zu nennen sind: Zielmaßstab, gewünschte Hierarchiestufe, Mindestgröße, Mindestlänge, Rasterzellgröße, Anzahl der Iterationen. Die weitere Abfolge entspricht dem Schema aus **Abb. 4** unter Verwendung der Transformer Disolver (2), CenterPointReplacer und CenterLineReplacer (5). Die wesentlichen Teilprozesse zur Handhabung der Objektauflösung (6) und der Flächenformvereinfachung (7) erfolgen anhand sogenannter PythonCaller, worauf im nächsten Abschnitt noch einmal näher eingegangen wird. Die anzuwendenden Regeln und Bedingungen innerhalb des gesamten Workflows werden durch Tester gesteuert. Transformer aus der Gruppe der Calculator (4, 8) ermitteln die erforderlichen Kennwerte, der GeometryValidator (2) ermöglicht die Überprüfung der Objekte auf ihre geometrische Gültigkeit und durch die Anwendung des SpatialRelator (4) werden topologische Bezüge ausgemacht.

5.3 Steuerung von Objektauflösung und Flächenvereinfachung

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 erster Linie nach semantischer Ähnlichkeit durchgeführt werden (Ai & van Oosterom, 2002). In der FME-Workbench wird dies durch die Implementierung einer Python-Funktion bewerkstelligt. Sie untersucht die jeweiligen Kandidaten des aufzulösenden Objektes bezüglich der inhaltlichen Nähe innerhalb der geologischen Legende, dem Abstand zur übergeordneten Einheit sowie der Flächengröße und weist die aufzulösende Fläche schließlich dem lithostratigraphisch ähnlichsten Kandidaten zu.

5.4 Ergebnisbeispiele

Aus den geologischen Grundlagendaten wurden fünf repräsentative Testbereiche vorbereitet und mit dem FME-Workflow für die vorgesehenen Hierarchiestufen prozessiert. Die folgenden drei Beispiele greifen Generalisierungsszenarien aus Tabelle 2 wieder auf und verdeutlichen die Generalisierungseffekte exemplarisch in bildlicher Vorher/Nachher-Gegenüberstellung.

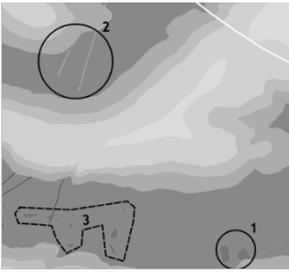
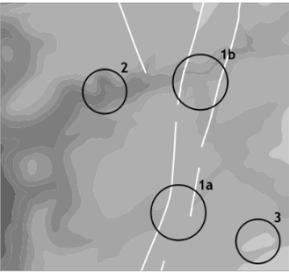
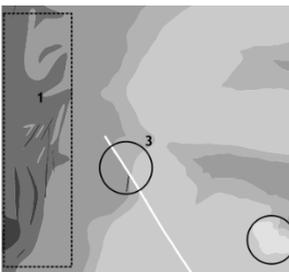
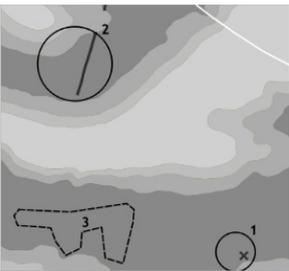
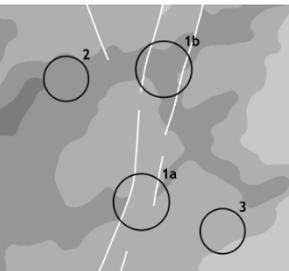
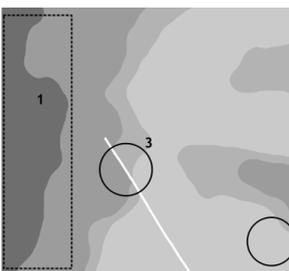
Beispiel 1	Beispiel 2	Beispiel 3
 <p>1a Ausgangssituation Stufe GE4</p>	 <p>2a Ausgangssituation Stufe GE4</p>	 <p>3a Ausgangssituation Stufe GE4</p>
 <p>1b generalisiert nach GE3</p>	 <p>2b generalisiert nach Stufe GE3</p>	 <p>3b generalisiert nach GE2</p>

Abb. 5: Detailbeispiele (mit Bezug auf designierte Hierarchiestufen aus Tab. 1)

Beide Polygone in Beispiel 1 innerhalb des Markierungskreises (1) unterschreiten die Mindestgröße, das linke Polygon weist eine Beflagung auf und bleibt somit im Generalisierungsergebnis als typisierter Punkt erhalten. Analog dazu wird in (2) von den beiden Gangpolygonen nur das entsprechend „beflaggte“ als Linie typisiert. Markierungsbereich (3) zeigt eine Gruppe von Polygonen, die der Auflösung unterliegt. Diese Generalisierungssituation

könnte in einer späteren Ausbaustufe in gesonderter Handhabung als amalgamiertes Polygon ausgewiesen werden.

Im Beispiel 2 ist ein Übergang von Hierarchiestufe GE4 nach GE3 dargestellt. Dabei wirken die tektonischen Lineationen gleichsam als Barrieren und einschränkende Bereiche für die Formvereinfachung (1a) und (1b), während im Bereich (2) der CA für die entsprechende Formvereinfachung sorgt. Das Polygon aus Bereich (3) wird aufgrund Mindergröße im Prozess aufgelöst.

Beispiel 3 verdeutlicht die glättende Wirkung des CA (1) sowie den Erhalt der geologischen Situation im Bereich der tektonischen Störung (3), während Flächen unterhalb der Mindestgröße bzw. außerhalb als invariant definierter Bereiche eliminiert werden (2).

Für eine qualitative Bewertung wurden die Ergebnisse dem Fachbereich Landesgeologie im LGRB vorgelegt. Die Experten haben eine gute Les- und Interpretierbarkeit der generalisierten Daten unter Wahrung des geologischen Kontextes bestätigt. 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 vorliegenden Teilbereichen die Ausnahme und können als vernachlässigbar eingestuft werden.

6 Diskussion

6.1 Übertragbarkeit und weiteres Potenzial

Die Workbench in ihrer ursprünglichen Ausprägung weist viele Geologie-spezifische Besonderheiten und Abhängigkeiten auf, dennoch ist eine themenübergreifende Übertragung des Prozessablaufes grundsätzlich gegeben. In Bezug auf hydrogeologische Daten ist eine 1:1-Übertragung möglich, da sich die hydrogeologischen Einheiten direkt aus den geologischen Einheiten ableiten und dadurch eine durchgehende geometrische Konsistenz – auch hinsichtlich der Deckschichten und tektonischen Verhältnisse – gegeben ist.

In einem Testdurchlauf mit den bodenkundlichen Daten wurde eine Generalisierung der ursprünglich in der Bodenkarte 1:50.000 vorhandenen Bodenformgruppen in semantisch zusammengefasste Bodenlandschaften vorgenommen. Die Workbench wurde hierfür an wenigen Stellen angepasst um relevante Teilmodule bzw. einzelne Transformer wahlweise oder alternativ im Prozess anzusteuern und so einen themenspezifisch angepassten Prozessdurchlauf zu ermöglichen. Die Generalisierungssituation aus Tabelle 2 B2) beinhaltet den themenspezifischen Sonderfall der Polygonpartitionierung (von Siedlungsflächen ohne hierarchische Zuordnung) durch StraightSkeletons mit nachfolgender semantisch bedingter Zuweisung. Dieser Teilschritt ist im Ablauf als separater Strang ausgewiesen.

Auch die Anwendung auf die gesonderte Generalisierung von Deckschichten wurde getestet. Hierbei stellt sich der Einbezug des hydrographischen Netzes als limitierender Faktor (Barrieren) bei der Flächenvereinfachung in ähnlicher Weise dar, wie dies bei geologischen Flächen und tektonischen Lineationen gegeben ist.

6.2 Fazit

In diesem Beitrag wurde schwerpunktmäßig am Beispiel geologischer Daten ein praxisgerecht einsetzbares Prozess-Framework für die automatisierte Generalisierung kategorischer Flächendaten entwickelt. Die Umsetzung des Ablaufs mittels FME ermöglichte dabei die software-übergreifende Nutzung und Implementierung unterschiedlicher Methoden. Das Zusammenspiel generischer Transformer, wie auch die Zuhilfenahme der rasterbasierten Methode eines zellulären Automaten, liefern die für die Generalisierung erforderlichen Teilschritte. Der Einbezug thematisch-topologischer Abhängigkeiten gewährleistet hierbei auch die weitest gehende Einhaltung und Erfüllung vorhandener fachlicher Ansprüche. Als Mischform aus vektoriellen und rasterbasierten Methoden kann der Workflow auch als hybrides Verfahren bezeichnet werden. Die während der Prozessierung vorgenommene Vektor-Raster-Vektor-Wandlung wird dabei nicht als Bruch im Prozessablauf wahrgenommen, auch die technische Integration der ArcGIS-Funktionalitäten findet nahtlos im Hintergrund statt. Das Verfahren zeigte sich auch stabil und praxistauglich im testweisen Durchlauf höherer Datenvolumen, vorausgesetzt wird dabei die absolute Validität der Eingangsgeometrien.

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

Literatur

- Ai, T., & van Oosterom, P. (2002). GAP-tree extensions based on skeletons. *Advances in spatial data handling* (pp. 501–513). Berlin/Heidelberg/New York: Springer.
- Artioli, G. P., Ciardi, G., Garberi, M. L., Martelli, L., 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. *Proceedings of ICA Conference* (pp. 860–871). ICA.
- Bader, M., & Weibel, R. (1997). Detecting and resolving size and proximity conflicts in the generalization of polygonal maps. *Proceedings 18th International Cartographic Conference*, 23.-27. Juli 1997, Stockholm, Schweden (pp. 1525–1532) (Ed. L. Ottoson).
- Bobzien, M. (2005). *Methodische Aspekte der Generalisierung von Geodaten* (Dissertation). Hohe Landwirtschaftliche Fakultät, Rheinische Friedrich-Wilhelms-Universität Bonn.
- Downs, T. C., & Mackaness, W. A. (2002). An Integrated Approach to the Generalization of Geological Maps. *The Cartographic Journal*, 39(2), 137–152.
- Ellsiepen, M. (2006). Nachhaltige Generalisierung topographischer Daten. *Zeitschrift für Vermessungswesen*, 131(3), 123–131.
- Foerster, T., Stoter, J. E., & Köbben, B. (2007). Towards a formal classification of generalization operators. *Proceedings of the 23rd International Cartographic Conference*, 4.-10.08.2007, Moskau, Russland.

- Galanda, M. (2003). *Automated polygon generalization in a multi agent system* (Dissertation). Universität Zürich.
- Harrie, L., & Weibel, R. (2007). Modelling the overall process of generalisation. In: W. Mackaness, A., Ruas, & L. T. Sarjakoski (Eds.), *Generalisation of geographic information: cartographic modelling and applications* (pp. 67–87). Elsevier.
- Huot-Vézina, G., Boivin, R., Smirnoff, A & Paradis, S. J. (2012). *GeoScaler: Generalization tool (with a supplementary user guide in French)*. Geological Survey of Canada, Open File 6231 (2nd Ed).
- LGRB – Regierungspräsidium Freiburg – Landesamt für Geologie, Rohstoffe und Bergbau (Ed.) (2015). *Integrierte geowissenschaftliche Landesaufnahme (GeoLa)*. Retrieved April 29, 2015, from https://www.lgrb-bw.de/aufgaben_lgrb/geola.
- Li, B., Wilkinson, G. G., & Khaddaj, S. (2001). Cell-based Model For GIS Generalization. *Proceedings of the 6th International Conference on GeoComputation*, 24.-26.09.2001, Brisbane, Australien (Ed. V. D. Pullar).
- 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.
- Peter, B. (2001). Measures for the generalization of polygonal maps with categorical data. *Fourth ICA Workshop on Progress in Automated Map Generalization*, 02.-04.08.2001, Peking, China.
- Rommeney, D. (2006). *Ein Zelluläres Automatenmodell zur Simulation urbaner Landnutzungsänderungen* (Diplomarbeit). Mathematisch-Naturwissenschaftliche Fakultät, Universität Potsdam.
- Sayidov, A., & Weibel, R. (2016). *Constraint-based approach in geological map generalization. 19th ICA Workshop on Generalisation and Multiple Representation*, 14.06.2016, Helsinki, Finnland.
- Schuff, J. (2019). *Verfahren zur automatisierten Generalisierung von flächenhaften Geofachdaten* (Masterarbeit). Universität Salzburg.
- Smirnoff, A., Huot-Vézina, G., Paradis, S. J., & Boivin, R. (2012). Generalizing geological maps with the GeoScaler 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), 1550–1568. doi:10.1016/j.cageo.2007.10.013.
- Steiniger, S., & Weibel, R. (2005a). A conceptual framework for automated generalization and its application to geologic and soil maps. *22nd International Cartographic Conference*, 11.-16.07.2005, La Coruna, Spanien (Ed. ICC).
- Steiniger, S., & Weibel, R. (2005b). Relations and structures in categorical maps. *The 8th ICA Workshop on Generalisation and Multiple Representation*, 07.-08.06.2005, La Coruna, Spanien.
- 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), 6374–6385.