



## MASTER THESIS

im Rahmen des

Universitätslehrganges »Geographical Information Science & Systems«  
(UNIGIS Msc) am Interfakultären Fachbereich für GeoInformatik (Z\_GIS)  
der Paris Lodron-Universität Salzburg

zum Thema

### **Integration von Expertenwissen in die reliefbezogene digitale Aktualisierung von Bodenkarten am Beispiel von Auenböden**

vorgelegt von

Dipl.-Geogr. Franziska Beau  
U1472, UNIGIS Msc Jahrgang 2010

Zur Erlangung des Grades  
»Master of Sciences (Geographical Information Science & Systems) - MSc (GIS)«

Gutachter: Ass. Prof. Dr. Hermann Klug

Dresden, den 30. Juni 2015

## **Erklärung**

Ich versichere, diese Master Thesis ohne fremde Hilfe und ohne Verwendung anderer als der angeführten Quellen angefertigt zu haben. Weiterhin versichere ich, dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat. Alle Ausführungen der Arbeit, die wörtlich oder sinngemäß übernommen wurden, sind entsprechend gekennzeichnet.

Dresden, den 30. Juni 2015

A handwritten signature in blue ink, appearing to read 'F. Beau', written over a horizontal line.

Franziska Beau

## **Danksagung**

Mein besonderer Dank gilt Dr. Markus Möller für die Anregungen und fachlichen Anstöße zum Thema dieser Arbeit. Seine große Begeisterung war mir eine starke Motivation. Weiterhin möchte ich Dr. Hermann Klug für die inhaltlichen und klärenden Rückmeldungen beim Erstellen der Arbeit danken und bei Dipl.-Ing. Sabrina Lott für die hilfreichen Korrekturen, Anmerkungen und Unterstützung. Jens Naumann danke ich für den Rückhalt, den er mir gegeben hat.

## **Zusammenfassung**

Anhaltende Bodendegradierung (bspw. durch Flächenversiegelung, Bodenversauerung und Bodenversalzung) führt zur Abnahme der als Pflanzenstandort verfügbaren Böden. Ein effektiver und schonender Umgang mit dem Boden durch präzisierete Bodennutzung gewinnt für verschiedene Nutzergruppen (Landwirtschaft, Planung, Naturschutz) an Bedeutung. Dafür werden groß- bis mittelmaßstäbige aktuelle Bodeninformationen benötigt. Ein zeit- und kostengünstiges Verfahren ist die Aktualisierung von vorhandenen Altdaten mittels digitaler Bodenkartierung. Digitale Bodenkartierung beruht auf geostatistischen Analysemethoden wie geomorphometrische Parameterableitung, Segmentierung und Clusterbildung. Vorliegende Altdaten sind durch Methoden der klassischen Bodenkartierung wie Bohrstockuntersuchungen, Profilsprachen, Feldversuche, pflanzensoziologische Aufnahmen entstanden. Bei der Verschneidung von wissensbasierten (qualitativen) Altdaten und statistikbasierten (quantitativen) Modellen kommt es aufgrund der Unterschiede in der methodischen Herangehensweise der Bodenkartierung zu Lageungenauigkeiten, welche bei größer werdendem Maßstab eine zunehmende Bedeutung erlangen. Das Ziel der Arbeit ist die Kombination von quantitativen und qualitativen Methoden der Bodenkartierung in einer Fachapplikation, um diese Lageungenauigkeiten zu identifizieren und die identifizierten Bereiche neu zu klassifizieren. Die Fachapplikation verbindet dabei quantitative und qualitative Methoden ohne zwischen unterschiedlichen Anwendungen und Benutzeroberflächen wechseln zu müssen. Ein 578,1 km<sup>2</sup> großes Untersuchungsgebiet im deutschen Bundesland Sachsen-Anhalt wurde gewählt. Als Datengrundlage dienen das digitale Geländemodell mit einer Auflösung von 25 m und die vorläufige Bodenkarte im Maßstab 1:50.000. Das digitale Geländemodell wurde hinsichtlich der Reliefparameter Massenbilanzindex, normalisierte Höhe und Neigung zur Charakterisierung fluvialer Prozessbereiche abgeleitet. Mittels einer regionenbasierten Segmentierung wurden Boden-Relief-Einheiten erstellt, die dann mit der k-mean-Methode geclustert wurden. Das erzielte Ergebnis stellt die Übertragung der Geländesituation in die digitale Umgebung dar. Für die Umsetzung des Verfahrensablaufes in einer Fachapplikation wurden Kombinationsvarianten von Algorithmen aus unterschiedlichen Geoinformationssystemen (GIS) (QGIS, Orfeo Toolbox, SAGA, GRASS, R,

GeoDMA) geprüft. Für die Validierung wurde ein visueller Vergleich innerhalb von Referenzbereichen (im Vorfeld festgelegte Bereiche mit potenziellen Lageungenauigkeiten) und den dort statistisch berechneten Clustern vorgenommen. Im Ergebnis wird festgestellt, dass die Kombination von Algorithmen aus unterschiedlichen GIS in einer Fachapplikation möglich ist. Die Identifikation von Lageungenauigkeiten konnte nicht hinreichend gewährleistet werden. Es werden Möglichkeiten der Verbesserung der integrierten Algorithmen (z.B. Instabilität des Segmentierungsalgorithmus, angewendetes Clusterbildungsverfahren) diskutiert sowie das Maß der Parameterkonfiguration, welches dem Bodenexperten angeboten wird.

## **Abstract**

Continued soil degradation (e.g. in form of soil sealing, soil acidification and salinisation of soils) leads to a decrease of available soils for plant growth. An effective and careful handling of the soil by precise land use becomes more important for different user groups (agricultural, planning, nature conservation). As a result, large- to middle-scale current soil information is needed. A time- and cost-efficient approach is to upgrade existing legacy soil data through digital soil mapping. Digital soil mapping is based on geostatistical analysis as geomorphometric derivation, segmentation and clustering. This legacy soil data comes from methods of classical soil mapping like soil sampling, profile speeches, excavation, in situ tests, and phytosociological surveys. At the intersection of knowledge-based (qualitative) legacy data and statistic-based (quantitative) models, the methodological approach of soil mapping results in position inaccuracies, which gain ever more importance as the scale increases. The aim of this study is the combination of quantitative and qualitative methods of soil mapping in an expert application to identify these position inaccuracies and reclassify the identified areas. The developed expert application combines quantitative and qualitative methods without having to switch between different systems and user interfaces. An exemplary study area of 578.1 km<sup>2</sup> in the German federal state of Saxony-Anhalt was elected. The digital elevation model with a resolution of 25 m and the preliminary soil map at a scale of 1:50.000 are used as a data source. The digital elevation model was derived using the relief parameters, mass balance index, normalised height and slope to characterize fluvial process areas. Soil-relief-units have been created by a region-growing segmentation, which were then clustered using the k-mean method. The obtained result represents the transfer of the land situation in the digital environment. For the implementation of the workflow in an expert application, variants of combinations of algorithms from different geographic information systems (GIS) (QGIS, Orfeo Toolbox, SAGA, GRASS, R, GeoDMA) were tested. For validation, a visual comparison within reference areas (prior defined areas with potential positional inaccuracies) and the statistically calculated clusters was performed. As a result, it is determined that the combination of algorithms from different GIS in an expert application is possible. The identification of positional inaccuracies could not be sufficiently guaranteed. Possibilities to improve the

integrated algorithms (e.g. instability of the segmentation algorithm, applied clustering methods) and the dimension of the parameter configuration, which is presented to the soil experts, are discussed.

# Inhaltsverzeichnis

Erklärung.....	II
Danksagung.....	III
Zusammenfassung.....	IV
Abstract.....	VI
Inhaltsverzeichnis.....	VIII
Abbildungsverzeichnis.....	X
Tabellenverzeichnis.....	XI
Abkürzungsverzeichnis.....	XI
1. Einleitung.....	1
1.1 Motivation.....	1
1.2 Stand der Wissenschaft.....	3
1.2.1 Eingangsdaten in der Digitalen Bodenkartierung.....	5
1.2.2 Methodische Ansätze in der digitalen Bodenkartierung.....	11
1.2.3 Vergleich von digitaler und konventioneller Bodenkartierung.....	12
1.2.4 Lageungenauigkeiten: Herausforderungen in der Digitalen Bodenkartierung.....	15
1.2.5 Digitale Bodenkartierung in Deutschland.....	16
1.3 Zielsetzung und Forschungsfrage der Arbeit.....	18
2. Material und Methoden.....	19
2.1 Das Untersuchungsgebiet.....	19
2.2 Datengrundlage.....	20
2.3 Verfahren zur Integration von Methoden der qualitativen Bodenkartierung in die digitale Bodenkartierung.....	24
2.3.1 Ableiten der Reliefparameter zur Charakterisierung fluvialer Prozessbereiche.....	26
2.3.2 Segmentierung zu Boden-Relief-Einheiten.....	27
2.3.3 Clusterbildung.....	29
2.3.4 Expertenbasierte Referenzclusterauswahl.....	31
2.3.5 Statistischer Vergleich der Referenzcluster mit übrigen Clustern.....	31
2.3.6 Neubewertung und abschließende Fixierung der Cluster.....	32



2.4 Integration des Bodenexpertenwissens: Erarbeitung einer Fachapplikation als Rahmenstruktur.....	32
2.5 Validierung der Ergebnisse / Identifizierung von Lageungenauigkeiten.....	36
3. Ergebnisse.....	39
3.1 Ergebnisse der quantitativen Analyse.....	39
3.1.1 Ableiten der Reliefparameter zur Charakterisierung fluvialer Prozessbereiche.....	39
3.1.2 Segmentierung.....	44
3.1.3 Clusterbildung.....	46
3.2 Validierung der Ergebnisse.....	51
3.3 Fachapplikation zur Integration des Fachwissen von Bodenexperten (qualitative Methode).....	52
4. Diskussion der Ergebnisse.....	55
4.1 Quantitative Analyse / Identifizierung von Lageungenauigkeiten.....	55
4.2 Validierung der quantitativen Methode.....	55
4.3 Fachapplikation zur Integration des Fachwissen von Bodenexperten (qualitative Methode).....	56
5. Schlussfolgerung und Ausblick.....	57
Literaturverzeichnis.....	59
Anhang.....	65

# Abbildungsverzeichnis

Abbildung 1: Einordnung der Arbeit (eigene Darstellung).....	3
Abbildung 2: Datenverarbeitung in der digitalen Bodenkartierung (eigene Darstellung in Anlehnung an Carré 2007, Grunwald 2009, Kempen et al. 2012, McBratney et al. 2003).....	5
Abbildung 3: Gegenüberstellung der Generalisierten Abläufe der klassischen Bodenerkundung in Anlehnung an Bui (2004), Kempen et al. (2012), Nauman und Thompson (2014) und eines DSM-Prozesses in Anlehnung an Dobos und Hengl (2009), Minasny et al. (2013).....	13
Abbildung 4: Lage des Untersuchungsgebietes in Deutschland, Untersuchungsgebiet in der LANDSAT8 Szene, rot umrandet (eigene Darstellung).....	19
Abbildung 5: Digitales Geländemodell des Untersuchungsgebietes (eigene Darstellung).....	21
Abbildung 6: Vorläufige Bodenkarte 1:50.000 mit Farbwertbelegung Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (eigene Darstellung).....	23
Abbildung 7: Verfahrensablauf zur Integration von Bodenexpertenwissen in die digitale Bodenkartierung (eigene Darstellung).....	25
Abbildung 8: Zusammensetzung der Bodengesellschaften in Anlehnung an Möller 2012, Farbwertbelegung Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie.....	30
Abbildung 9: Versuchsablauf (eigene Darstellung).....	36
Abbildung 10: Kontrollbereiche der Bodengesellschaft Auenböden.....	38
Abbildung 11: Massenbilanzindex.....	39
Abbildung 12: Histogramm und Boxplot des Massenbilanzindex für den Bereich der Auenböden.....	40
Abbildung 13: Normalisierte Höhe.....	41
Abbildung 14: Histogramm und Boxplot der normalisierten Höhe für den Bereich der Auenböden.....	42
Abbildung 15: Neigung.....	43
Abbildung 16: Histogramm und Boxplot der Neigung für den Bereich der	

Auenböden.....	44
Abbildung 17: Segmentierungsergebnis.....	45
Abbildung 18: Histogramm der Segmentverteilung.....	46
Abbildung 19: Clusterbildung mit der Orfeo Toolbox.....	47
Abbildung 20: Clusterbildung mit R.....	48
Abbildung 21: Clusterbildung mit SAGA.....	49
Abbildung 22: Design der Fachapplikation.....	53

## Tabellenverzeichnis

Tabelle 1: Eingangsdatenformat der Bodenbildungsfaktoren in die digitale Bodenkartierung (eigene Zusammenstellung in Anlehnung an Grunwald 2009, Jenny 1941, McBratney et al. 2003, Mulder et al. 2011, Minasny et al. 2013).....	6
Tabelle 2: Technische Hintergründe der vorausgewählten GI-Systeme.....	35
Tabelle 3: Für den Verfahrensverlauf vorhandene Algorithmen in der ausgewählten GI-Software.....	35
Tabelle 4: Histogramm der Clusterverteilung bei OTB, R, SAGA.....	50
Tabelle 5: Übereinstimmung in Prozent zwischen der visuell ermittelten und algorithmisch gemessenen Lageungenauigkeit.....	51

## Abkürzungsverzeichnis

CRAN	Comprehensive R Archive Network (deutsch Umfassendes Netzwerkarchiv für R)
CSM	Conventional Soil Mapping (deutsch Klassische Bodenkartierung)
DEM	Digital Elevation Model (deutsch Digitales Geländemodell)
DGM	Digitales Geländemodell
DHM	Digitales Höhenmodell
DLR	Deutsches Zentrum für Luft und Raumfahrt
DSM	Digital Soil Mapping (deutsch Digitale Bodenkartierung)
DSMo	Digital Soil Modelling (deutsch Digitale Bodenmodellierung)
DSMM	Digital Soil Mapping and Modelling (deutsch Digitale Bodenkartierung und Modellierung)
EU	Europäische Union
EUSIS	European Soil Information System (deutsch Europäisches Bodeninformationssystem)
FAO	Food and Agriculture Organization of the United Nations (deutsch Ernährungs- und Landwirtschaftsorganisation der Vereinten Nationen)
FE	Fernerkundung
GeoDMA	Geographic Data Mining Analyst
GRASS	Geographic Resources Analysis Support System
GUI	Graphic User Interface (deutsch Benutzeroberfläche)
INSPIRE	Infrastructure for Spatial Information in the European Community (deutsch Infrastruktur für räumliche Information in der Europäischen Gemeinschaft)
LiDAR	Light Detection and Ranging
NH	Normalisierte Höhe
MBI	Mass Balance Index (deutsch Massenbilanzindex)
MMK	Mittelmaßstäbige Landwirtschaftliche Standortkarte
OTB	Orfeo Toolbox

PSM	Predictive Soil Mapping (deutsch Prognostizierende Bodenkartierung)
SAGA	System for Automated Geoscientific Analyses
SAR	Synthetic-Aperture-Radar
SEIS	Shared Environmental Information System (deutsch Gemeinsames Europäisches Umweltinformationsnetzwerk)
scorpan-SSPFe	soil spatial prediction function with spatially autocorrelated errors
SRTM	Shuttle Radar Topography Mission
UG	Untersuchungsgebiet
UNO	United Nations Organization (deutsch Organisation der Vereinten Nationen)
VBK50	Vorläufige Bodenkarte im Maßstab 1:50.000

# 1. Einleitung

## *1.1 Motivation*

Der Boden ist das Umweltmedium, welches an der Grenze zwischen Luft, verwittertem Gestein und Wasser Lebensraum für Flora und Fauna bietet. Im Vergleich zu Luft und Wasser ist der Boden das im öffentlichen Bewusstsein am wenigsten beachtete Umweltmedium. Ein Indikator hierfür ist der Rückzug aller Entwürfe und Vorschläge zur Einführung einer Bodenrahmenrichtlinie auf Ebene der Europäischen Union (EU) im Mai 2014 durch die Europäische Kommission. Die EU-Bodenrahmenrichtlinie hatte den Erhalt der Funktionen des Bodens, die Verhinderung der Bodenverschlechterung und die Regeneration geschädigter Böden auf Ebene der Europäischen Union zum Ziel (COUNCIL OF THE EU 2014). Allein in Deutschland wurden im Vierjahresdurchschnitt 2007 bis 2010 täglich 87 Hektar Boden im Rahmen des Infrastruktur- und Gebäudebaus in Anspruch genommen (DEUTSCHER BUNDESTAG, 17. WAHLPERIODE 2012). Neben der Zerstörung der Bodenfunktionen und der natürlichen Bodenschichtung durch Versiegelung, Materialauf- und -abtrag, ist der Boden außerdem von Überdüngung, Verdichtung, Versauerung, Eintrag von Schadstoffen und Abtrag durch Luft- und Wassererosion gefährdet (COUNCIL OF THE EU 2014). Dabei ist die uneingeschränkte Wiederherstellung zerstörter Bodeneigenschaften an einem Standort nicht möglich (BRANDHUBER et al. 2008). Aufgrund der geringen Beachtung bei gleichzeitiger hoher Gefährdung des Schutzgutes Boden hat die United Nations Organization (UNO, Organisation der Vereinten Nationen) das Jahr 2015 zum internationalen Jahr des Bodens ausgerufen (VEREINTE NATIONEN 2013).

Durch die Abnahme der als Pflanzenstandort nutzbaren Bodenflächen nimmt die Notwendigkeit zu, über präzise und detaillierte Kenntnisse der vorhandenen Bodenstandorte als Grundlage für land- und forstwirtschaftliche Nutzung, Naturschutz und Planung zu verfügen (SENTHIL-KUMAR SELVARADJOU 2007). Aktuelle und digital vorliegende Kenntnisse über Bodentypen im groß- bis mittelmaßstäbigen Bereich bieten die Möglichkeit Rückschlüsse auf die Nährstoff- und Wasserverfügbarkeit des Standortes zu ziehen und ergänzende Pflanzenversorgung durch Wasser, Nährstoffe

und Pestizide standortgerecht und optimal dosiert auszubringen. Weiterhin können Bodeneigenschaften wie Körnung, Lagerungsdichte und Wasserspannung abgeleitet werden und somit Gewichts- und Zugkraft von Landmaschinen an den Bodenstandort angepasst werden, um Bodenverdichtung z.B. durch hohe Radlast auf feuchtem Untergrund zu vermeiden (BRANDHUBER et al. 2008). Dadurch wird das Bodengefüge, die Durchwurzelbarkeit und die Infiltrationsfähigkeit des Standortes erhalten und der Oberflächenabfluss verringert mit dem Ergebnis der Minderung der Bodenerosion durch Wasser, des Ausgleichs des Landschaftswasserhaushaltes (z.B. mit Hinblick auf Überschwemmung) und der Verbesserung der Gesamterträge (BRANDHUBER et al. 2008).

Knappe Zeit- und Geldressourcen schränken konventionelle Bodenuntersuchungen und -erkundungen im Sinne von Gelände- und Laborarbeiten ein. Um schnelle und genaue Methoden zur räumlichen Vorhersage von Bodentypen und -eigenschaften zu ermöglichen, entwickelt sich die digitale Bodenkartierung (GRIMM et al. 2009) als Werkzeug zum Generieren räumlicher Bodeninformationen. Mit Hilfe der digitalen Bodenkartierung kann der wachsende weltweite Bedarf an hochaufgelösten Bodenkarten gedeckt werden. Auch in Deutschland wird die digitale Bodenkartierung zu einem wichtigen Bestandteil bei der Beschaffung von Bodeninformationen, da die Geländearbeit durch die deutschen Bundesländer nicht länger finanziell tragbar ist (BEHRENS und SCHOLTEN 2006).

Seit Ende der 1990er Jahre erfolgt eine Methodenentwicklung in der Bodenkartierung vom Schwerpunkt der klassischen Labor- und Geländearbeit hin zu digitalen geoanalytischen Prozessierungen (MCBRATNEY et al. 2003, MINASNY et al. 2013). Um vorliegende Bodenaltdata in die digitale Bodenkartierung zu integrieren ist ein Wissenstransfer zwischen vorhandenem Bodenexpertenwissen in die Konzepte der digitalen Bodenkartierung, wie ihn Abbildung 1 darstellt, notwendig (BUI 2004, CARRÉ et al. 2007, KEMPEN et al. 2009). Dabei werden sehr unterschiedliche Wissensfelder, einerseits das der Bodenkunde und andererseits das der Prozessmodellierung zusammengeführt.



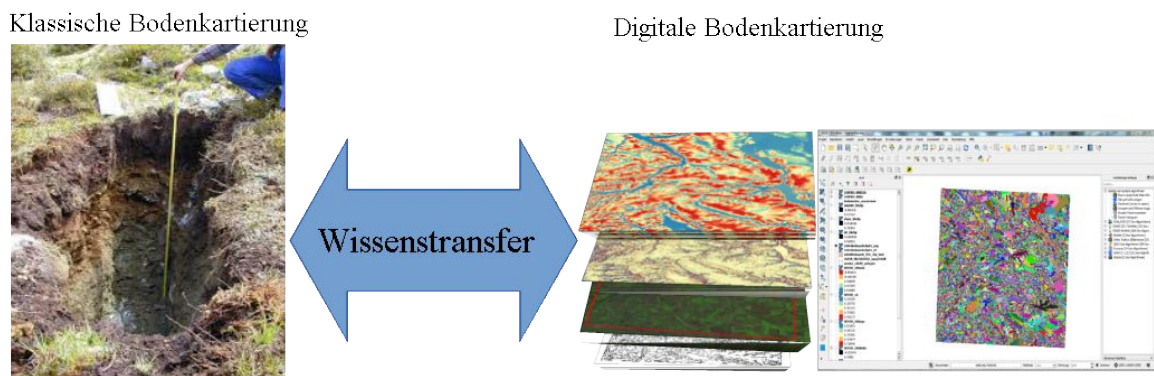


Abbildung 1: Einordnung der Arbeit (eigene Darstellung)

## 1.2 Stand der Wissenschaft

MCBRATNEY et al. legten 2003 mit „On digital soil mapping“ eine erste zusammenfassende Darstellung zur Entwicklung des Digital Soil Mappings (DSM) vor. MCBRATNEY et al. (2003) und CARRÉ et al. (2007) definieren die digitale Bodenkartierung als „die Erschaffung und Bestückung räumlicher Bodeninformation unter Anwendung von Feld- und Labormethoden im Zusammenspiel mit räumlichen und nicht räumlichen Bodeninferenzsystemen“. LAGACHERIE et al. (2007) erweitern diese Definition und bezeichnen digitale Bodenkartierung als „die Schaffung und Bestückung räumlicher Bodeninformationssysteme durch numerische Modelle abgeleitet aus räumlichen und zeitlichen Variationen der Bodentypen und Bodeneigenschaften sowie aus Bodenbeobachtung und Wissen sowie der in Bezug stehenden Umweltvariablen“. Somit beruht die digitale Bodenkartierung auf der statistischen Beziehung zwischen Messdaten aus Bodenerkundungen und räumlichen Kovariablen am Beprobungsort (GRIMM et al. 2009). Sie geht weit über das Digitalisieren vorhandener Bodenkarten hinaus und wird weiter zur kontinuierlichen Abbildung von Bodeneigenschaften ausgebaut (BEHRENS und SCHOLTEN 2006).

SCULL et al. (2003) verwendeten den Begriff prognostizierende Bodenkartierung (Predictive Soil Mapping PSM) und definieren damit die Entwicklung eines numerischen oder statistischen Modells der Beziehung zwischen Umweltvariablen und Bodeneigenschaften. Dieser Ansatz setzt nach SCULL et al. (2003) die

bodengeographische Perspektive in den Fokus, demnach die Untersuchung der Lage, Verteilung und räumlicher Muster der Böden in der Landschaft. GRUNWALD (2009) verwendet die Begriffe digitale Bodenkartierung und prognostizierende Bodenkartierung synonym. Sie führt den Begriff Digital Soil Mapping and Modelling (DSMM) als digitale Bodenkartierung und -modellierung ein. Dabei bezeichnet sie digitale Bodenmodellierung (DSMo) als umfassender im Vergleich zur digitalen Bodenkartierung. Denn die digitale Bodenmodellierung betont die Beschreibung und/oder Vorhersage von Böden in Raum und Zeit, in der Regel durch mechanistisch implementierte Simulation. Als Pedometrik bezeichnet GRUNWALD (2009) eine Rahmenstruktur, die die digitale Bodenkartierung und -modellierung mit den Umweltwissenschaften, der Geoinformatik, der Kartographie, der Statistik, der Geostatistik und der Mathematik zur Untersuchung der Verteilung und die Entstehung der Böden zusammenführt. In den vergangenen Jahren hat sich die Arbeit und Definition von MCBRATNEY et al. (2003) und LAGACHERIE et al. (2007) als häufig genutzte Definition durchgesetzt (ABDELFATTAH und PAIN 2012, BAGHERI BODAGHABADI et al. 2011, CAMBULE et al. 2013, CAVAZZI et al. 2013, KEMPEN et al. 2012, MINASNY et al. 2013, MÖLLER et al. 2015, SHI et al. 2012, VAN ZIJL et al. 2013). Aus diesem Grund wird in der vorliegenden Arbeit die Definition von MCBRATNEY et al. (2003) und LAGACHERIE et al. (2007) genutzt und der Begriff digitale Bodenkartierung verwendet.

Während zu Beginn der Entwicklung der digitalen Bodenkartierung die Herausforderung darin lag, eine spezifische räumliche Datenstruktur für die digitale Bodenkartierung in bestehende Kartierungsprojekte und bodenkundliche Wissenssysteme zu integrieren, ist die aktuelle Herausforderung Vergleichsdaten mittels Geländeaufnahmen zur Bewertung von Berechnungsmodellen der digitalen Bodenkartierung in Geoinformationsanwendungen zu integrieren (GRIMM et al. 2009). Abbildung 2 stellt die Datenverarbeitung in der digitalen Bodenkartierung dar. Die Kapitel 1.2.1 und 1.2.2 gehen auf die Punkte Eingangsdaten und Datenverarbeitung in der digitalen Bodenkartierung näher ein.

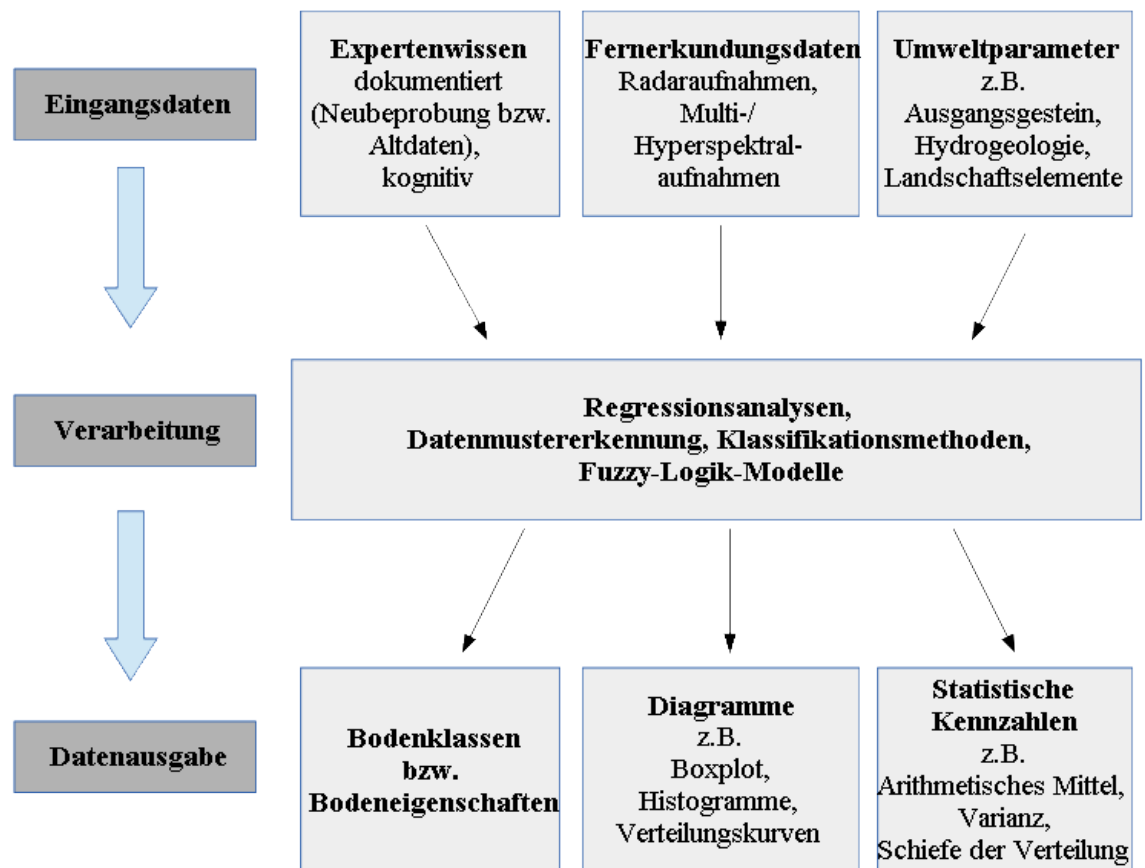


Abbildung 2: Datenverarbeitung in der digitalen Bodenkartierung (eigene Darstellung in Anlehnung an CARRÉ 2007, GRUNWALD 2009, KEMPEN et al. 2012, MCBRATNEY et al. 2003)

### 1.2.1 Eingangsdaten in der Digitalen Bodenkartierung

Für die digitale Bodenkartierung hat MCBRATNEY et al. (2003) die Gleichung zur Bodenentwicklung von JENNY (1942):  $s = f(c, o, r, p, t)$  mit den Bodenbildungsfaktoren Klima (c), Organismen (o), Relief (r), Ausgangsmaterial (p) und Zeit (t) erweitert und entwickelte die scorpan-SSPFe (soil spatial prediction function with spatially autocorrelated errors) Gleichung:

$$S = f(S, C, O, R, P, A, N) + e$$

Sie besagt, dass an einem Punkt in Raum und Zeit der Boden (entweder als Bodenklassen  $S_c$  oder Bodenattribut  $S_a$ ) eine empirische quantitative Funktion (f) des

Bodens als eine Klasse oder als ermittelte Eigenschaft (S), des Klimas (C), der Organismen, die Flora, Fauna und anthropogene Einflüsse umfassen (O), des Reliefs (R), des Ausgangsmaterials (P), des Alters (A) und der räumlichen Lage (N) ist. e sind räumlich korrelierte Fehler. Trotz standortabhängiger unterschiedlicher Gewichtung der Einflussfaktoren kommen alle Faktoren gleichzeitig zum Tragen (MCBRATNEY et al. 2003). Das scorpan-Modell stellt bis heute eine wichtige Grundlage vieler Studien des DSM dar (BALLABIO et al. 2012, CAMBULE et al. 2013, CARRÉ et al. 2007B, GRUNWALD 2009, KERRY et al. 2012, LACOSTE et al. 2014, LEMERCIER et al. 2012, MALONE et al. 2009, MINASNY et al. 2013, ODGERS et al. 2011, POGGIO et al. 2013, VAN ZIJL et al. 2013). Tabelle 1 ordnet die in der digitalen Bodenkartierung genutzten Eingangsdatenformate den Bodenbildungsfaktoren zu.

*Tabelle 1: Eingangsdatenformat der Bodenbildungsfaktoren in die digitale Bodenkartierung (eigene Zusammenstellung in Anlehnung an GRUNWALD 2009, JENNY 1941, MCBRATNEY et al. 2003, MULDER et al. 2011, MINASNY et al. 2013)*

<b>Bodenbildungs-faktor</b>	<b>typische Standorte (beispielhaft)</b>	<b>Anteil der Anwendung in DSM-Projekten</b> (nach MCBRATNEY et al. 2003)	<b>Datenformat</b>
Boden (S)		35 %	Expertenwissen in Form von Bodenalt Daten bzw. präsentem Expertenwissen
Klima (C)	Tropen, Permafrost	5 %	Messpunktaufzeichnungen, Klimakarten
Organismen (O)	Moore	25 %	hyper- und multispektrale Fernerkundungs(FE)daten, Landnutzungskarten, Vegetationskarten
Relief (R)	Hanglagen, stark reliefiertes Gelände	80 %	Digitales Geländemodell
Ausgangsmaterial (P)	Karstlandschaften	25 %	Geologische Karte, hyper- und multispektrale FEdaten
Zeit (A)	Moränenlandschaften, vulkanisch aktive Gebiete, Tropen	-	Bodenalt Daten, Bodenexpertenwissen
Position (N)		20 %	Geographische Koordinaten

*(a) Fernerkundungsdaten*

Die Rückstrahlung des Sonnenlichtes durch die Erdoberfläche und der auf ihr befindlichen Objekte wird bei passiven Aufnahmesystemen in der Fernerkundung zur Erstellung von multi- und hyperspektralen Fernerkundungsdaten genutzt. Diese Daten werden aufgrund der Rückstrahlungseigenschaften der Objekte zur Unterteilung von Elementen auf der Erdoberfläche durch Bildinterpretation genutzt. Vitale Vegetation hat bspw. eine besonders intensive Rückstrahlung im Wellenlängenbereich des Nahen Infrarotes, Gestein im Wellenlängenbereich des Mittleren Infrarotes. In der digitalen Bodenkartierung werden multi- und hyperspektrale Fernerkundungsdaten als Eingangsdaten für die Bestimmung der Bodenbildungsfaktoren Organismen und Ausgangsgestein genutzt (GRUNWALD 2009, MCBRATNEY et al. 2003, MULDER et al. 2011). BOCK et al. (2005), BUI et al. (2006), CAMBULE et al. (2013), CARRÉ et al. (2007b) und MALONE et al. (2011) nutzten Aufnahmen der LANDSAT-Satelliten, HANSEN et al. (2009) nutzen SPOT-Daten und BOCK et al. (2005) nutzten IKONOS-Daten.

Aktive Aufnahmesysteme in der Fernerkundung setzen Laser- oder Radiowellen ein und messen deren Rückstrahlung. Diese Aufnahmedaten dienen zur Ermittlung von Geländehöhen und damit zur Erstellung von digitalen Geländemodellen (DGM). Digitale Geländemodelle dienen in der digitalen Bodenkartierung als Eingangsdaten für den Bodenbildungsfaktor Relief (R). CAVAZZI et al. (2013) stellen fest, dass digitale Geländemodelle regelmäßig in der digitalen Bodenkartierung genutzt werden. Nach BISHOP et al. (2006) und MCBRATNEY et al. (2003) arbeiten 80% der Projekte der digitalen Bodenkartierung mit dem Bodenbildungsfaktor Relief (R), 25% nutzen jeweils die Bodenbildungsfaktoren Ausgangsgestein und Vegetation. BAGHERI BODAGHABADI et al. (2011) sowie HÄRING et al. (2012) wiesen den starken Zusammenhang zwischen der Verbreitung von Bodenserien und topographischen Eigenschaften nach. Dazu führten sie Korrelationsanalysen (BODAGHABADI et al. 2011) sowie entscheidungsbaumbasierte Klassifizierung (HÄRING et al. 2012) zwischen Geländeaufnahmen und aus digitalen Geländemodellen abgeleiteten Geländeparametern durch.

Die Auflösungen der DGM reichen von unter 1 m für Light detection and ranging (LiDAR) - Daten bis hin zu 1 km bei Datensätzen mit weltweiter Abdeckung. Der Einfluss der Auflösung (Rastergridgröße) der digitalen Geländemodelle auf die Landschaftsmodellierung wurde viel diskutiert (BEHRENS et al. 2010, SMITH et al., 2006, ZHU et al. 2012). Zur Wahl der Auflösung der DGM führen CAVAZZI et al. (2013) zwei Hauptprobleme auf:

1. bei feineren Auflösungen kommt es zum Detailüberschuss der Geländeattribute und es wird ein "Rauschen" erzeugt, was unweigerlich zum Genauigkeitsverlust der Vorhersage führt
2. bei größeren Auflösungen stellen die Geländeattribute nur verallgemeinerte Eigenschaften der Landfläche dar und verlieren so ihre Vorhersagekapazität.

CAVAZZI et al. (2013) haben ein DGM mit der Auflösung von 20 m in digitale Geländemodelle mit Auflösungen von 30, 40, 50, 60, 80, 100, 120, 140, 170, 200, 230, 260 m interpoliert. Im Ergebnis haben CAVAZZI et al. (2013) folgende Anwendungsvorschläge für die digitale Bodenkartierung formuliert:

1. Ebene homogene Flächen bevorzugen gröbere Auflösungen (über 140 m).
2. Morphologisch variierende Flächen mit charakteristischen Merkmalen wie Drummlins oder abrupte Änderungen in der Topographie z.B. in Steillagen benötigen feinere Auflösungen (30 m in dieser Studie).
3. In Bereichen, die bei optimaler Auflösung falsch eingestuft wurden, haben andere bodenbildende Faktoren wie z.B. das Ausgangsgestein einen größeren Einfluss auf die Bodenbildung als das Relief.

HÄRING et al. (2012) stellen ebenfalls fest, dass ein praktikabler Kompromiss zwischen räumlicher Genauigkeit und Modellgüte eine Rasterauflösung von 25 m ist. LACOSTE et al. (2014) wenden das Prinzip von CAVAZZI et al. (2013) an indem sie aufgrund vorhandener Störwerte ein LiDAR DGM mit der Auflösung von 2 m auf 5 m, 10 m und 20 m hochrechnen. LIU et al. (2012) kommen aufgrund ihrer Arbeit in ebenem Gelände wie CAVAZZI et al. (2013) zu dem Schluss, dass bei geringer

Reliefenergie andere Bodenbildungsfaktoren wirksam werden als das Relief und somit andere Eingangsdaten als digitale Geländemodelle genutzt werden müssen.

*(b) Bodenexpertenwissen*

BUI (2004), GRUNWALD (2009) sowie KEMPEN et al. (2012) setzen Bodenexperten und ihr Wissen u.a. in den Fokus ihrer Arbeit im Bereich der digitalen Bodenkartierung. Eine Definition des Bodenexperten wird dabei nicht gegeben. Der Begriff „soil surveyor expert knowledge“ wird häufig genutzt (BUI 2004, GRUNWALD 2009, KEMPEN et al. 2012). Daran angelehnt werden im Rahmen der vorliegenden Arbeit unter Bodenexperten Experten verstanden, die an der bodenkundlichen Datenaufnahme und/oder Interpretation von Feld- und/oder Labordaten durch ihre berufliche Tätigkeit aktiv und intensiv beteiligt sind.

Bodenexpertenwissen ist (a) in den Ergebnissen der klassischen Bodenkartierung (Bodenklassenkarten, Bodendatenbanken, Bodenprofilbeschreibungen, Beschreibungen von Bodeneinheiten) als Bodenaltdateien, (b) in neu erfassten projektbezogenen Daten gebündelt (AG BODEN 2005, BEAUDETTE und O'GEEN 2009, CARRÉ et al. 2007A, KEMPEN et al. 2012, SOIL SURVEY DIVISION STAFF 1993) sowie (c) aktuell präsenten Wissen der Bodenexperten aus Labor- und Geländearbeit (BUI 2004). Die Einbeziehung von Bodenexpertenwissen ist eine wichtige Säule in der digitalen Bodenkartierung schwer zugänglicher und entlegener Gebiete (HANSEN et al. 2009, GRIMM und BEHRENS 2010).

Bodenaltdateien sowie projektbezogene Bodendaten wurden und werden mittels Geländeaufnahmen erstellt. Die Methoden der klassischen Bodenkartierung sind nach AG BODEN (2005) und SOIL SURVEY DIVISION STAFF (1993):

1. Auswertung vorhandener Unterlagen (Topografische Karten, Historische Karten, Bodenkarten, Karten der Reichsbodenschätzung, Geologische Karten, Forstliche Standortskarten, Landwirtschaftliche Standortskarten, Hydrogeologische Karten, Ingenieurgeologische und Baugrundkarten, Luftbilder und Satellitenaufnahmen, punktuelle Messergebnisse, regionale Fachliteratur, Nutzungspläne, Gutachten u.s.w.)

2. Übersichtsbegehung/Feldbeobachtungen (Überblick über die allgemeinen bodenkundlichen Verhältnisse und deren Zusammenhänge mit Topografie, Geologie, Vegetation und Bodennutzung und Erfahrungen aus der Vergangenheit in ähnlichen Landschaften)
3. Profilaufnahme und Bohrstockuntersuchungen (Farbansprache, Bestimmung Bodenart, pH- und Redoxmessung, Carbonatbestimmung) für die Interpolation von Punkt auf die Fläche

Während der Geländeaufnahmen nutzt der Bodenexperte ein mentales konzeptionelles Boden-Landschaftsmodell (KEMPEN et al. 2012, SHI et al. 2009). Dazu steht in der Bodenkundlichen Kartieranleitung: „Werden alle pedologisch relevanten Kriterien bei der Geländebeobachtung berücksichtigt und in ein gedankliches Konzept eingebaut, so kann der Arbeitsaufwand durch eine gezielte Auswahl der Bohr- und Grabungspunkte optimiert werden“ (AG BODEN 2005).

BUI (2004) beschreibt zur Logik von Expertenwissen in der Bodenkunde die Beziehung zwischen dem Gedankenmodell des Bodenkundlers und der Bodenkarte nebst zugehöriger Legende als Wissenssystem (knowledge system):

„Eine Bodenkarte ist eine strukturierte Darstellung über die räumliche Verteilung von Böden, aus diesem Grund kann sie als logisches System bezeichnet werden. Die Logik kann von Expertenregeln erfasst werden und bei automatisierten Karten genutzt werden. Die Gedankenmodelle, welche genutzt werden, um Bodenkarten herzustellen, können in einer strukturierten Rahmenrichtlinie [...] repräsentiert werden.“

*(c) weitere Eingangsdaten*

In Tabelle 1 ist zu erkennen, dass der Bodenbildungsfaktor Klima durch Messpunktaufzeichnungen und Klimakarten in die digitale Bodenkartierung Eingang findet und eine Anwendung von 5 % bei digitalen Bodenkartierungsprojekten hat. Topographische Karten dienen bei BUI et al. (2006) und CAVAZZI et al. (2013) als Grundlage für die Berechnung von Geländemodellen aus Höhenlinien. Weiterhin werden Landnutzungskarten, Vegetationskarten, geographische Koordinaten,



geologische, hydrologische, hydrogeologische Karten als Eingangsdaten in der digitalen Bodenkartierung angewendet (MCBRATNEY et al. 2003).

### **1.2.2 Methodische Ansätze in der digitalen Bodenkartierung**

Die digitale Bodenkartierung verwendet Methoden, die die quantitative Beziehung zwischen Bodeneigenschaften und -klassen und ihrer Umwelt anpassen (KEMPEN et al. 2012, MCBRATNEY et al. 2003). Die Methode  $f()$  der SCORPAN-Gleichung kann von einem einfachen linearen Modell bis hin zu komplexen Werkzeugen der Datenmustererkennung wie Regressionsbäume und Entscheidungswälder reichen (MINASNY et al. 2013). Bei der systematischen Analyse von Veröffentlichungen im Themenfeld digitale Bodenkartierung kam GRUNWALD (2009) hinsichtlich der angewendeten Methoden zu dem Ergebnis, dass die häufigste verwendete Methode die Regressionsanalyse (41,1 %) ist, gefolgt von Klassifikationsmethoden (32,2 %), univariatem Kriging (18,9 %), Entscheidungsbaummethoden (13,3 %), Prozeßsimulierungsmodellen (14,4 %) und multivariate Krigingmethoden (7,85%). Andere Methoden wie neuronale Netzwerke, Fuzzy-Logik-Modelle und stochastische Simulationen sind sehr selten.

MCBRATNEY et al. (2003) resümieren, dass bei der Methodenauswahl unterschieden wird, ob Bodenklassen oder -eigenschaften berechnet werden sollen. Wenn Bodenklassen berechnet werden sollen, werden Klassifizierungsmethoden verwendet. NAUMAN und THOMPSON (2014) bestätigen, dass Entscheidungs- und Regressionsbäume eine angesehene Technik in der Klasseneinteilung sind. DEUMLICH et al. (2010), GRIMM und BEHRENS (2010), HÄRING et al. (2012), HANSEN et al. (2009), MORA-VALLEJO et al. (2008) sowie SCHMIDT et al. (2008) haben die Klassifikations- bzw. Regressionsbaumanalyse als Technik der Datenmustererkennung in ihren Arbeiten angewendet. Bodeneigenschaften werden hingegen mittels Regressionsanalysen bestimmt, wie z.B. von CAMBULE et al. (2013) sowie LACOSTE et al. (2014). MULDER et al. (2011) stellen fest, dass bis heute kein einheitliches Verfahren etabliert ist, in dem Ansätze der Segmentierung, Messung von Bodeneigenschaften und Interpolation unter Verwendung von Fernerkundungsdaten in einem Verfahrensablauf integriert sind.

Abschließend ist festzustellen, dass der Abgleich und die Einarbeitung von Bodenexpertenwissen in Form von Altdaten und neu erfassten projektbezogenen Daten gängige Praxis in der digitalen Bodenkartierung ist (HUANG et al. 2014, MINASNY und MCBRATNEY 2010, NAUMAN und THOMPSON 2014, SULAEMAN et al. 2013). Die interaktive Integration von Expertenwissen ist selten anzutreffen und nicht in einem Verfahrensablauf organisiert (DEUMLICH et al. 2010).

### **1.2.3 Vergleich von digitaler und konventioneller Bodenkartierung**

DOBOS und HENGL (2009) sowie KEMPEN et al. (2012) stellen heraus, dass trotz der technologischen Lücke zwischen traditioneller und digitaler Bodenkartierung keine bedeutenden Unterschiede in den beiden Ansätzen vorhanden sind. Beide Ansätze benötigen Eingangsdaten über den Boden und die Variablen, die die Umgebung beschreiben, in der die Bodengenese stattfindet. Um Vorhersagen zu treffen, nutzen beide Ansätze Boden-Landschafts-Modelle. Abbildung 3 stellt den generalisierten Prozessablauf der konventionellen Bodenkartierung der digitalen Bodenkartierung gegenüber. Sie verdeutlicht die Gemeinsamkeiten der beiden Prozessabläufe in den Bereichen Interpretation sowie Visualisierung der Ergebnisse in Form von Bodenkarten sowie Texten (Beschreibungen von Bodenprofilen und -einheiten bzw. wissenschaftliche Aufsätze).

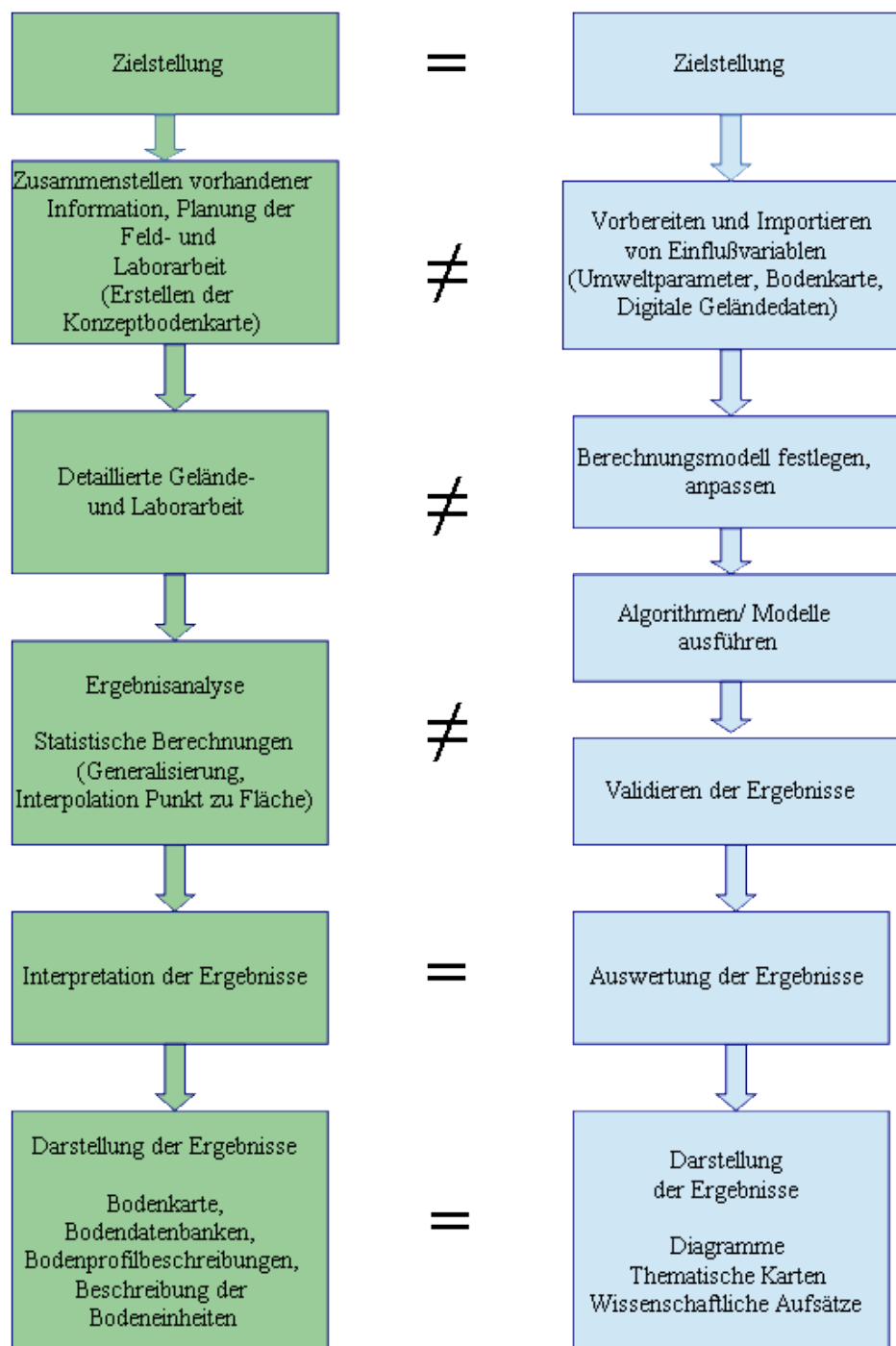


Abbildung 3: Gegenüberstellung der Generalisierten Abläufe der klassischen Bodenerkundung in Anlehnung an BUI (2004), KEMPEN et al. (2012), NAUMAN und THOMPSON (2014) (links) und eines DSM-Prozesses in Anlehnung an DOBOS und HENGL (2009), MINASNY et al. (2013) (rechts). (eigene Darstellung)

Der Unterschied zwischen konventioneller und digitaler Bodenkartierung ist laut DOBOS und HENGL (2009) sowie KEMPEN et al. (2012) in den Methoden der Ableitung der Bodeninformation aus den Eingangsdaten (z.B. Faktoren der Bodenbildung) zu finden. Die traditionellen Modelle basieren auf subjektiven, qualitativen Gedankenmodellen, die auf dem Expertenwissen des Bodenkundlers beruhen, während die digitale Bodenkartierung sich auf objektive, quantitative Boden-Landschafts-Modelle bezieht.

LAGACHERIE et al. (2008) stellen fest, dass die traditionelle Bodenerkundung explizit bodenkundliches Fachwissen in das Bodenkartierungsprodukt integriert. Jedoch ist sie im Vergleich zu den Ansätzen der digitalen Bodenkartierung zu teuer und zeitintensiv geworden. Weitere Einschränkungen der konventionellen Bodenkartierung betreffen die Nichtreproduzierbarkeit, die Sicht des Bodenkörpers als einzelne, homogene Einheiten und der Mangel an quantifizierten Bestimmungen von Unsicherheiten (BUI 2004, KEMPEN et al. 2012, NAUMAN und THOMPSON 2014). CARRÉ et al. (2007b) resümieren, dass es keine statistischen Kriterien für traditionelle Bodenaufnahmen gibt.

Demgegenüber stellen GRUNWALD (2009) sowie KEMPEN et al. (2012) fest, dass die digitale Bodenkartierung zum aktuellen Zeitpunkt keine Standardisierungen in Betracht zieht, während die konventionelle Bodenkartierung von Standardisierungen geprägt ist. GRUNWALD (2009) spricht in diesem Zusammenhang von einem „Wild West of DSMM“.

VAN ZIJL et al. (2013) haben in ihrer Arbeit die Genauigkeit von (1) klassischer Bodenkartierung, (2) digitaler Bodenkartierung und (3) einer Kombination aus klassischer und digitaler Bodenkartierung sowie Bodenexpertenwissen anhand des Prozentsatzes der Berechnungen, die korrekt vorhergesagt wurden auf Grundlage eines Vergleichsdatensatzes verglichen und kommen zu dem Ergebnis, dass die Genauigkeit der Kombination von klassischer und digitaler Bodenkartierung und Bodenexpertenwissen, der Genauigkeit von klassischer Bodenkartierung sehr nahe kommt (65% zu 67%). Allerdings ist bei klassischer Geländearbeit eine 3,7fache (kosten- und zeitintensive) Beprobungsdichte im Feld notwendig. Die ausschließliche digitale Bodenkartierung ohne Abgleich mit Geländedaten erreicht nach VAN ZIJL et al.

(2013) eine Genauigkeit von 35%. Da bei dieser Methode teilweise unlogische Bodenvergesellschaftungen geschaffen werden, kann sie nur dazu dienen, eine Idee von der Verteilung der Bodenklassen zu erhalten, wenn keine exakten/präzisen Bodeninformationen benötigt werden bzw. zur Vorinformation für weiterführende Arbeiten.

#### **1.2.4 Lageungenauigkeiten: Herausforderungen in der Digitalen Bodenkartierung**

Ein Grund für Fehler in digital abgeleiteten Boden-Landschafts-Modellen als Grundlage der digitalen Bodenkartierung rührt von der räumlichen Ungenauigkeit der Eingangsdaten. Sie können in Maßstab, Herkunft, Auflösung, Kartierschema und/oder Generalisierungsniveau variieren (GRIMM und BEHRENS 2010, BISHOP et al. 2006).

##### *(a) Maßstab*

In der digitalen Bodenkartierung ist der Umgang mit verschiedenen Maßstabsebenen von großer Bedeutung (BEHRENS et al. 2010, LAGACHERIE et al., 2008, MCBRATNEY et al. 2003). Auf Feld- bzw. Einzugsgebietsebene basieren die meisten Studien auf Datenerhebungen durch Bodenproben im Gelände (MINASNY et al. 2013). Auf regionaler und kontinentaler Ebene - außer in Frankreich, Großbritannien oder anderen Nationen, die ein nationales Monitoringsystem haben - beruhen die meisten Studien auf Bodenaltedaten. In der digitalen Bodenkartierung bewegen sich die meisten Studien bisher auf lokaler Maßstabsebene und nur wenige auf regionaler oder nationaler Maßstabsebene (MULDER et al. 2011).

Der Maßstab ist eine wichtige Überlegung bei der Erstellung von Bodenklassenkarten. Der Nutzer benötigt Informationen in einem bestimmten Maßstab und die einzelnen Parameter besitzen eine spezielle räumliche Auflösung. Probleme treten auf, wenn kleinmaßstäbige Karten als Eingangsdaten (z.B. geologische Karte mit Maßstäben unter 1:1.000.000) zur Berechnung von mittelmaßstäbigen Bodenkarten (mit Maßstäben von 1:25.000 bzw. 1:50.000) genutzt werden. Die nötige Detailschärfe und die Genauigkeit des Bodenattributes, in diesem

Fall Ausgangsgestein, ist nicht gegeben. Übertragbar ist die Problemlage, wenn ausgehend von Bodenalt Daten im Rahmen der digitalen Bodenkartierung Bodendaten aktualisiert werden (YANG et al. 2011).

*(b) Kartierschema, Abgrenzung von Bodeneinheiten*

Die den Bodenalt Daten zugrunde liegende Bodenkartierung ist in ihrer Intensität und Vorgehensweise abhängig vom Stand der Vorinformation, von der Diversität der Bodenverhältnisse und vom Darstellungsmaßstab (AG BODEN 2005). Da die Grenzen der Bodeneinheiten Übergangszonen sind (LAGACHERIE et al., 2007), erfolgt die Abgrenzung der Bodeneinheiten subjektiv. Es kommt zu Ungenauigkeiten bei der Überlagerung von auf subjektiven Gedankenmodellen beruhenden Bodenkarten und aktuellen durch Fernerkundung erfassten Bodeninformationen (BUI 2004, DOBOS und HENGL 2009, KEMPEN et al. 2012, MÖLLER et al. 2012). Dieser Sachverhalt beeinträchtigt die Modellierungsergebnissen in der Digitalen Bodenkartierung.

**1.2.5 Digitale Bodenkartierung in Deutschland**

KRASILNIKOV (2014) stellt heraus, dass Australien und Neuseeland hinsichtlich der Forschung in der digitalen Bodenkartierung eine Führungsposition inne haben. BEHRENS und SCHOLTEN veröffentlichten 2006 einen ersten Überblick über den Stand der Forschung zur digitalen Bodenkartierung in Deutschland. Sie resümieren, dass auch in Deutschland ein wachsender Bedarf an der Entwicklung von Ansätzen der digitalen Bodenkartierung besteht, da die traditionelle Bodenkartierung im Maßstab 1:25.000 und 1:50.000 zu teuer und zeitaufwendig ist. Forschungen mit diesen Zielmaßstäben führten DEUMLICH et al. (2010), HÄRING et al. (2012) und MÖLLER et al. (2012) durch. SCHMIDT et al. (2008) sowie MÖLLER et al. (2008) untersuchten den Zusammenhang von Zielklassen und variierenden Masstabebenen. Weiterhin stellen BEHRENS und SCHOLTEN (2006) fest, dass der Schwerpunkt der digitalen Bodenkartierung in Deutschland in der Bodenklassifizierung bzw. -gruppierung und nicht in der Prognose von Bodeneigenschaften liegt. Dies bestätigt sich in den

Arbeiten von DEUMLICH et al. (2010), HÄRING et al. (2012), MÖLLER et al. (2012), SCHMIDT et al. (2008).

Aufgrund der geomorphologischen und klimatologischen Gegebenheiten in Deutschland ist das Relief eine Haupteinflussgröße der Bodenentstehung. Das findet sich in der Forschung zur digitalen Bodenkartierung wieder, welche die große Bedeutung digitaler Geländemodelle in Kombination mit Informationen über das Ausgangsmaterial herausstellt (BEHRENS und SCHOLTEN 2006). Die Arbeit mit der Bodeneinflussgröße Relief abgeleitet aus digitalen Geländemodellen ist bspw. bei DEUMLICH et al. (2010), HÄRING et al. (2012), MÖLLER et al. (2012) und MÖLLER et al. (2015) wiederzufinden. Die Untersuchung von Lageungenauigkeiten führten GRIMM und BEHRENS (2010) anhand von manipulierten Datensätzen durch. MÖLLER (2004 und 2012) trennt Segmentierung und Klassifizierung, um Lageungenauigkeiten zu vermindern. MÖLLER et al. (2015) untersucht den Zusammenhang verschiedener Maßstabsebenen und Lageungenauigkeiten. Die Integration von Bodenexpertenwissen wird von DEUMLICH et al. (2010) und MÖLLER et al. (2012) angeregt. Es werden keine konkreten Standardisierungen beschrieben.

Auf institutionalisierter Ebene fördert die Arbeitsgruppe „Digital Soil Mapping“ der Deutschen Bodenkundlichen Gesellschaft die Diskussion neuer Methoden und Arbeitstechniken zur Erstellung von Boden(eigenschafts)karten. Schwerpunkt der Bestrebungen der Arbeitsgruppe ist eine möglichst effiziente und zielgerichtete Integration der Ansätze der digitalen und klassischen Bodenkartierung.

Eine Reihe von EU-Richtlinien und -Initiativen fordern in der nationalen Umsetzung die Bereitstellung von digitalen Umwelt- und somit auch Bodendaten. Im Wesentlichen handelt es sich dabei um die Umweltinformationsrichtlinie der Aarhus-Konvention, Annex III der Infrastructure for Spatial Information in the European Community (INSPIRE) - Richtlinie und das Shared Environmental Information System (SEIS) (COUNCIL OF THE EU 2014).

### ***1.3 Zielsetzung und Forschungsfrage der Arbeit***

In Deutschland existiert ein Bedarf an digitaler Bodeninformationen mit groß- bis mittelmaßstäbigem Detailgrad (BEHRENS und SCHOLTEN 2006). Ein zeit- und kostengünstiges Verfahren ist die Aktualisierung von vorhandenen Altdaten mittels digitaler Bodenkartierung. Bei der Verschneidung von wissensbasierten Altdaten und statistikbasierten Modellen kommt es aufgrund methodischer Unterschiede zwischen qualitativen (expertenwissenbasierten, konventionellen) und quantitativen (statistikbasierten) Methoden der Bodenkartierung zu Lageungenauigkeiten, welche bei größer werdendem Maßstab eine zunehmende Bedeutung haben. Die Identifizierung und Beseitigung von Lageungenauigkeiten erfordert eine Kombination von konventioneller und digitaler Bodenkartierung. Die Forschungsfrage lautet demnach: Wie können quantitative und qualitative Methoden der Bodenkartierung in einem Fachverfahren kombiniert werden, um mittels quantitativer Methoden Lageungenauigkeiten in vorhandenen Bodenkarten zu identifizieren und unter Einbeziehung des Fachwissen von Bodenexperten (qualitative Methode) im Geoinformationssystem diese Bereiche neu zu klassifizieren? Methoden der digitalen Bodenkartierung wie Parametrisierung, Segmentierung und Klassifizierung werden für die Kombination in einem Verfahrensablauf mit qualitativen wissensbasierten Methoden der Bodenkartierung analysiert. Als erstes Teilergebnis soll ein Verfahrensablauf erarbeitet werden, in dem das Expertenwissen so eingesetzt werden kann, dass der Transfer von qualitativem Wissen der klassischen Bodenkartierung in das quantitative Wissensumfeld der digitalen Bodenkartierung erfolgt. Anschließend wird die Umsetzung des Verfahrensablaufes in einer Plattform so vorgenommen, dass dem Bodenexperten ohne informationstechnische Fachkenntnisse das Einbringen seines Wissen ermöglicht wird. In einem letzten Schritt werden die mit dem entwickelten Prozessablauf identifizierten Lageungenauigkeiten durch einen visuellen Vergleich zwischen Schätzwerten und definierten Realwerten validiert.



## 2. Material und Methoden

### 2.1 Das Untersuchungsgebiet

Das Untersuchungsgebiet (UG) „Wipper-/Unteres Saaletal“ im deutschen Bundesland Sachsen-Anhalt befindet sich ca. 20 km nördlich der Stadt Halle/Saale und umfasst 578,1 km<sup>2</sup>. Aufbauend auf die Arbeiten von MÖLLER (2004 – 2015) lagen für das UG Eingangsdaten und Ortskenntnis vor. Abbildung 4 zeigt die topographische Lage des UG im Südosten Deutschlands und eine Satellitenbildaufnahme des LANDSAT 8 mit einer Echtfarbkanalbelegung. In der LANDSAT 8 Aufnahme vom 01. November 2014 sind dunkelbraune und grüne meist eckige Flurformen in regelloser Verteilung ohne Wechsel dieses Patterns zu erkennen. Diese Farb- und Formmerkmale lassen auf Ackerschläge schließen (jahreszeitbedingt: braun für Brache, grün für Gründünger) und geben die intensive landwirtschaftliche Nutzung in dem Gebiet wieder.

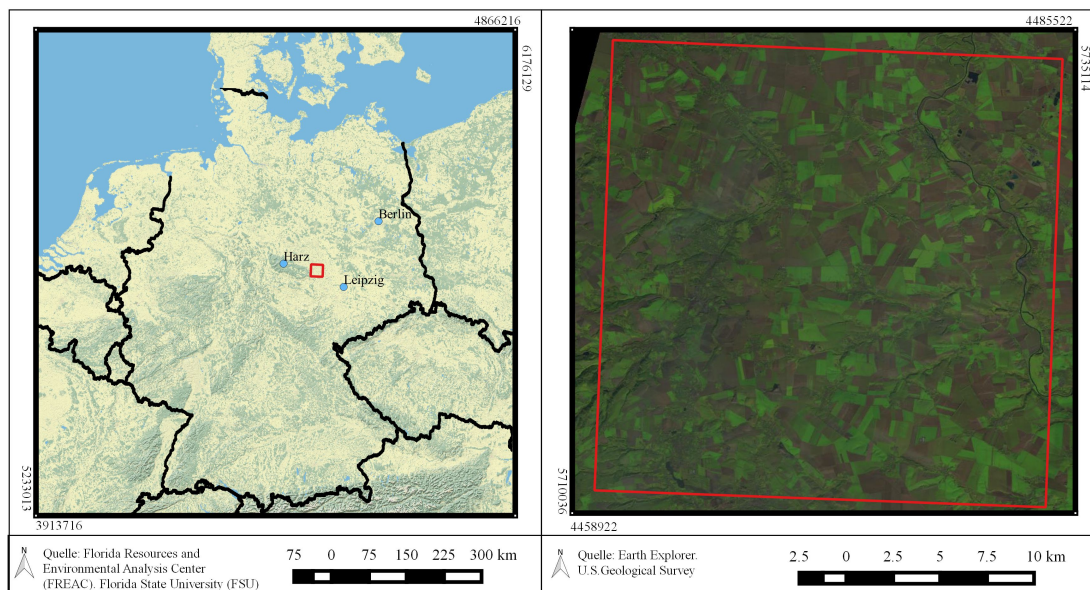


Abbildung 4: links: Lage des Untersuchungsgebietes in Deutschland, rechts: Untersuchungsgebiet in der LANDSAT 8 Szene, rot umrandet (eigene Darstellung)

Das Klima im UG ist kontinental geprägt. Die im bundesweiten Vergleich geringen mittleren jährlichen Niederschläge (450 bis 470 mm, gemessen 1901 bis 2000, WENDLING 2001) sind auf die Lage des UG im Regenschatten des Harzes

zurückzuführen. Die jährliche Verdunstungshöhe bewegt sich zwischen 600 bis 650 mm (MÖLLER 2004, TREFFLICH 1997).

Ausgangsgestein und Relief des UG sind durch glaziale und periglaziale Bedingungen während der Saale- und Weichselkaltzeit geprägt. Die bestimmende Reliefeinheit ist eine schwach reliefierte Hochfläche. Im Osten befindet sich das Durchbruchstal der Saale, im südlichen Bereich das Schlenzetal und im Westen das Wippertal. Die Hochflächenränder sind von Kerb- und Flachmuldentälern zerschnitten (KUNERT 1970, MÖLLER 2004).

Die Verteilung der Ausgangssubstrate und Böden ist eng an die Reliefverhältnisse gebunden. Die Hochflächen und deren Ränder sind bedeckt von Löß der Weichselkaltzeit und Geschiebemergel der Saalekaltzeit. Hier dominieren Ah/C Böden wie Pararendzina und Schwarzerden. In Bereichen in denen Sande, Tone und Lehme des Perm und Karbon vorherrschen, treten Ah/C-Böden aus Silikatgesteinen wie Braunerde und Ranker auf. Im Bereich der Flußauen haben sich grund- und oberflächenwasserbeeinflusste Böden wie Vega, Tschernitza, Kalkpaternia und Gley herausgebildet. Aufgrund intensiver Landwirtschaft und sommerlichen Starkniederschlagsereignissen sind kolluviale Böden (Kolluvisols) entstanden (ALTERMANN 1970, KAINZ et al. 1997, MÖLLER et al. 2012).

### **2.2 Datengrundlage**

Für die Untersuchungen stehen die nach BEHRENS et al. (2010) in Deutschland typischen Eingangsdaten für die digitale Bodenkartierung zur Verfügung: einerseits das digitale Geländemodell (mit einer Auflösung von horizontal 25 m und vertikal 1 m) im Rasterdatenformat sowie andererseits eine Bodenkarte im Vektorformat mit einem zugrundeliegenden Maßstab von 1: 50.000.

Das für wissenschaftliche Zwecke vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) kostenfrei zur Verfügung gestellte digitale Höhenmodell (DHM) ist aus SRTM-Daten abgeleitet wurden. Die vom DLR zur Verfügung gestellten SRTM-Daten sind bei der STS-99 Shuttle Radar Topography Mission (SRTM) mit dem deutsch-italienischen Radarinstrumentes Synthetic-Aperture-Radar (SAR) im X-Band

aufgezeichnet worden. Die Flugbahn des für die SRTM ausgerüsteten Space Shuttle Endeavour OV-105 befand sich zwischen 57° S und 60° N. Die Aufnahmen erfolgten vom 11. bis 22. Februar 2000. Die Auflösung beträgt horizontal 25 m und vertikal 1 m. Die Genauigkeit beträgt horizontal 20 m absolut und 15 m relativ sowie vertikal 20 m absolut und 15 m relativ. Das Radarinstrument X-SAR hat eine deutlich schmalere Bodenspur als das zweite, US-amerikanische Radargerät der SRTM, was zu Lücken zwischen den einzelnen Bahnen des X-SAR führte. Weitere Lücken entstanden dadurch, dass bei den Bahnmanövern das Shuttle gedreht wurde - die Radargeräte haben in der Zeit nicht aufnehmen können. Im Unterschied zum digitalen Geländemodell gibt das digitale Höhenmodell die Höhe des Geländes inklusive aller auf ihr befindlichen Objekte wie z.B. Vegetation und Bebauung wieder (WAGNER 2003). Da es sich bei dem UG um ein landwirtschaftlich genutztes Gebiet handelt und sich der Zeitpunkt der Datenaufnahme in der vegetationsfreien Zeit befindet, wird angenommen, dass die Radarstrahlen bis zur Erdoberfläche vorgedrungen sind und somit die Rückstrahlung von der Erdoberfläche erfolgte. Abbildung 5 zeigt die Abdeckung des Untersuchungsgebietes mit den DHM-Daten.

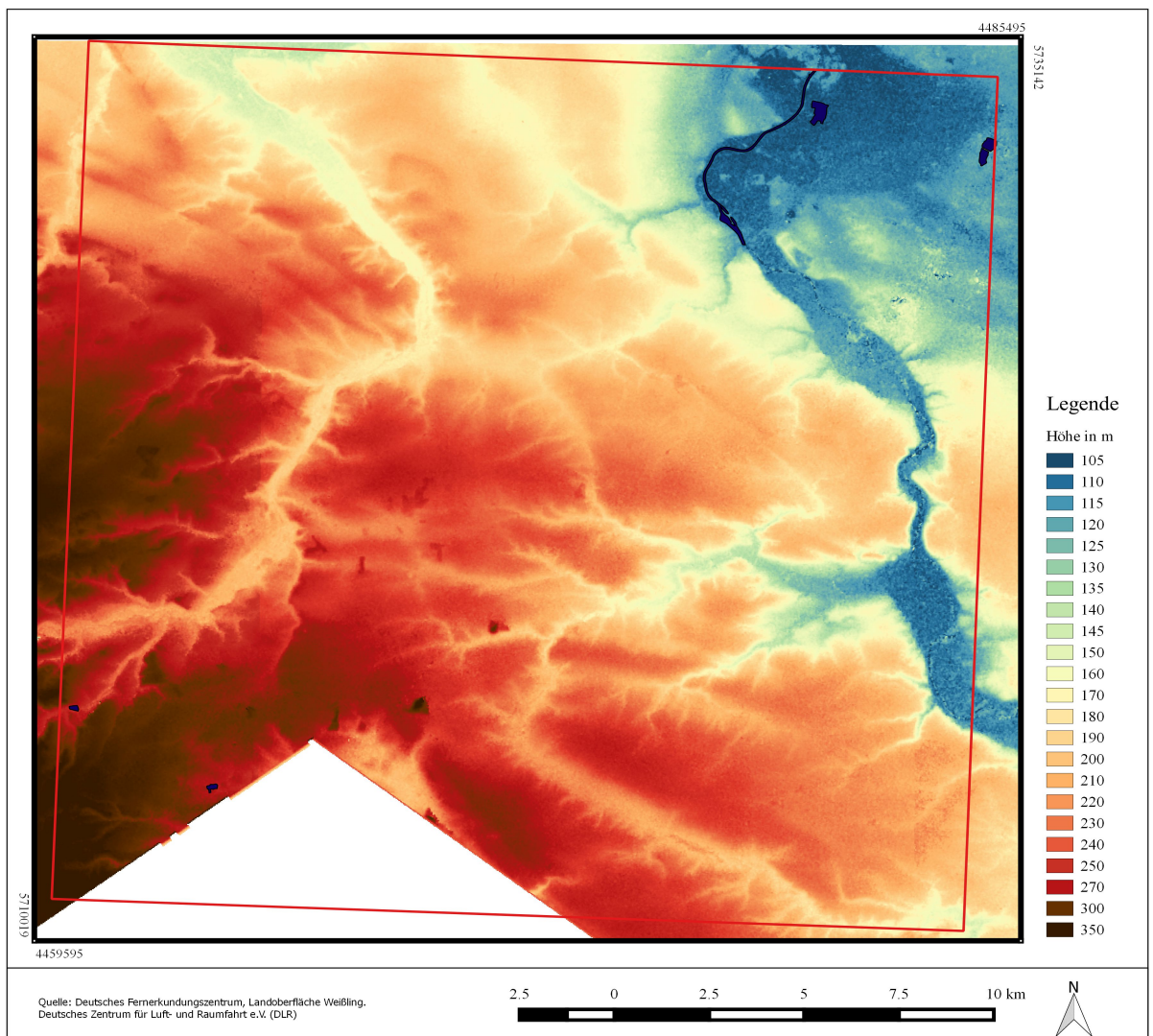


Abbildung 5: Digitales Höhenmodell des Untersuchungsgebietes (eigene Darstellung)

Mit der vorläufigen Bodenkarte im Maßstab 1:50.000 (VBK50) (Abbildung 6) werden im Bundesland Sachsen-Anhalt bodenkundliche Informationen für Flächen bereitgestellt, in denen zurzeit noch keine bodenkundliche Neukartierung im Maßstab 1:50.000 vorliegt. Grundlage der VBK50 ist die digitale Aufarbeitung und inhaltliche Vereinheitlichung von Altunterlagen. Dabei umfasst das Ausgangsmaterial im Wesentlichen drei Quellen (HARTMANN 2005, MÖLLER 2004, SCHMIDT und DIEMANN 1981):

- Altunterlagen aus Kartierungsberichten der Jahre 1970 – 1975

- Forstliche Standortkartierung
- Arbeitskarten im Maßstab 1:25.000 der Mittelmaßstäbigen Landwirtschaftlichen Standortkarte (MMK).

Die Mittelmaßstäbige landwirtschaftliche Standortkartierung (MMK) wurde für die gesamte landwirtschaftlich genutzte Fläche der ehemaligen DDR zwischen 1976 und 1982 durchgeführt, wobei der Schwerpunkt weniger in der Neukartierung als vielmehr in der Integration vorhandener Informationsquellen lag (z.B. Reichsbodenschätzung, Geologische oder Topographische Karte).

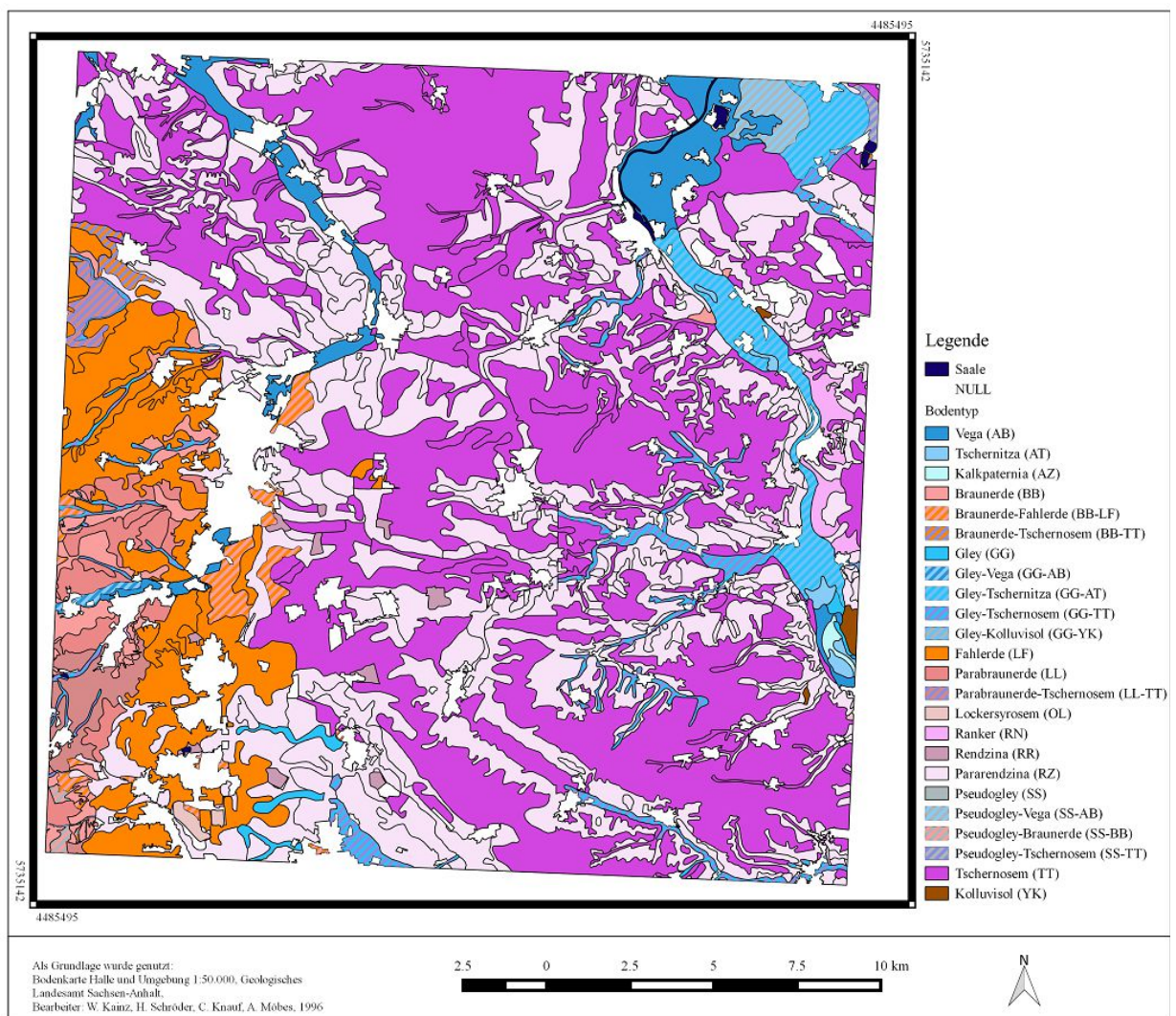


Abbildung 6: Vorläufige Bodenkarte 1:50.000 mit Farbwertbelegung Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (eigene Darstellung)

### ***2.3 Verfahren zur Integration von Methoden der qualitativen Bodenkartierung in die digitale Bodenkartierung***

Die Methoden der klassischen Bodenkartierung setzen im Gegensatz zu den Methoden der digitalen Bodenkartierung die direkte Präsenz im Gelände voraus (vgl. Kapitel 1.2.3). Die kognitive Leistung des Bodenexperten, d.h. der Aufbau der Gedankenmodelle zum Erkennen von Boden-Umwelt-Beziehungen sowie das Ziehen logischer Rückschlüsse ist ortsunabhängig. Auf Grundlage der Arbeit von MÖLLER et al. (2012) wurde ein Verfahrensablauf entwickelt, das dem Bodenexperten die Möglichkeit bietet, Boden-Umwelt-Beziehungen in der digitalen Geländesituation zu erkennen und sein Bodenexpertenwissen einfließen zu lassen.

Abbildung 7 stellt das angewendete Verfahren dar. In der Darstellung sind die Verfahrensschritte, die der quantitativen Bodenkartierung zuzuordnen sind, blau eingefärbt und die Verfahrensschritte, die der qualitativen Bodenkartierung zuzuordnen sind, grün eingefärbt.

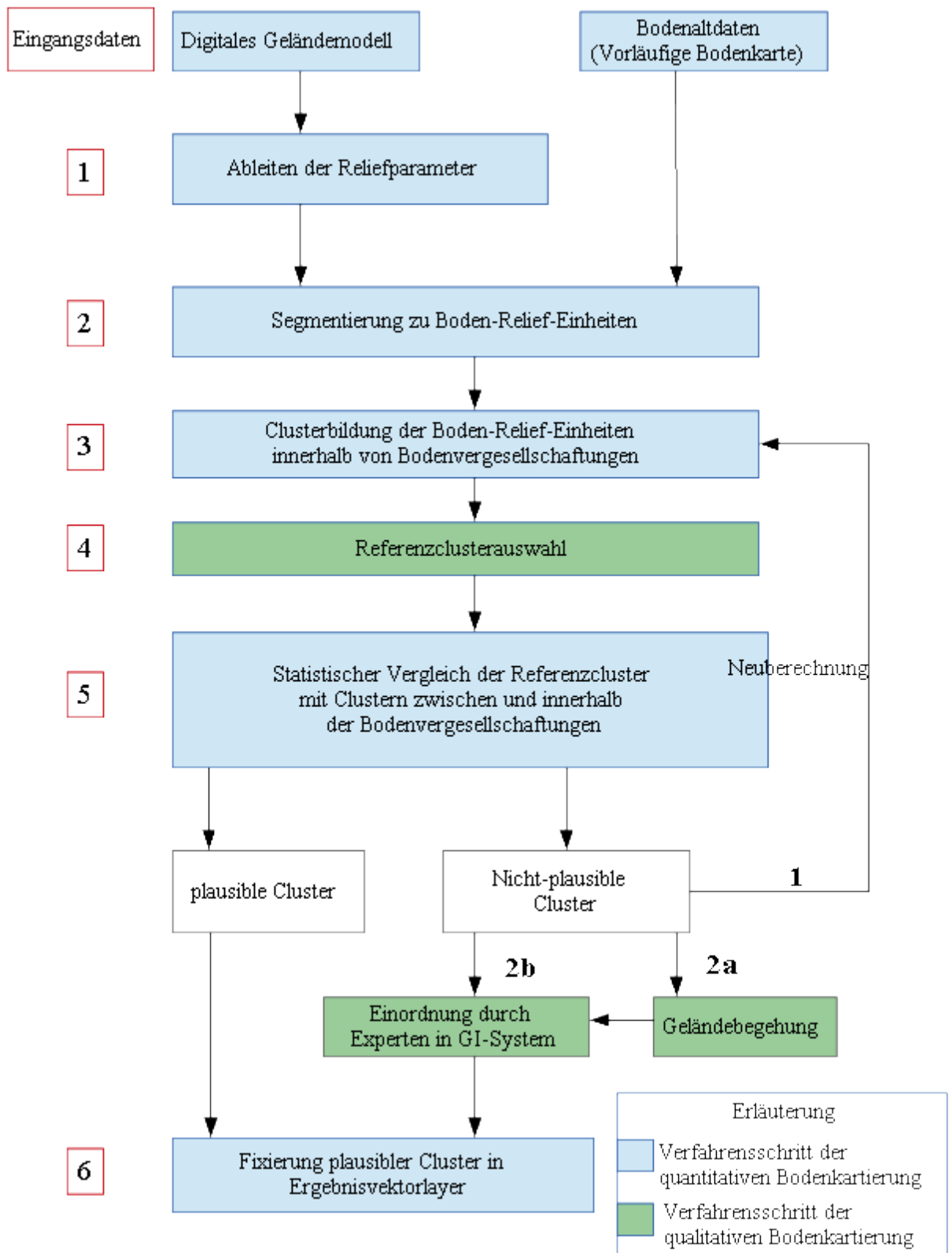


Abbildung 7: Verfahrensablauf zur Integration von Bodenexpertenwissen in die digitale Bodenkartierung (eigene Darstellung)

Der Verfahrensablauf unterteilt sich in (1) Ableiten der Geländeparameter aus dem DGM, (2) Verschneiden der Bodenalt Daten und Geländeparameter durch Segmentierung zu Boden-Relief-Einheiten und (3) Clusterbildung der Boden-Relief-Einheiten innerhalb von Bodengesellschaften. Die Schritte (1) bis (3) dienen der Datenbearbeitung zur Darstellung der Boden-Umwelt-Beziehung als Vorbereitung für die Interaktion des Bodenexperten. Die Darstellung der Boden-Umwelt-Beziehung wird am Beispiel des Bodenbildungsfaktors Relief vorgenommen. Die Arbeitsschritte (1) bis (3) werden im Rahmen der vorliegenden Masterarbeit durchgeführt und das Ergebnis auf seine Funktionsfähigkeit hinsichtlich der zuverlässigen Identifizierung von Lageungenauigkeiten überprüft (vgl. Kapitel 2.5). Ab Schritt (4) erfolgt die Integration des Bodenexpertenwissens, in (5) der statistische Vergleich zwischen Referenzclustern und übrigen Clustern und in (6) die Neubewertung und abschließende Fixierung der Cluster. Die Schritte (4) bis (6) sind dabei in Zusammenhang mit der in Kapitel 2.4 beschriebenen Erarbeitung einer Fachapplikation zu sehen, die die interaktive Integration des Bodenexpertenwissens ermöglicht. Die Durchführung der Arbeitsschritte (4) bis (6) ist nicht Teil der Masterarbeit. Die einzelnen Verfahrensschritte werden in den folgenden Kapiteln näher erläutert.

### **2.3.1 Ableiten der Reliefparameter zur Charakterisierung fluvialer Prozessbereiche**

Digitale Geländemodelle beinhalten je Rasterzelle eine spezifische Höheninformation des Geländes an diesem geographischen Punkt. Mittels der Nachbarschaftsbeziehungen der Rasterzellen untereinander können aus den DGM Rasteroberflächen mit Pixelwerten, die die spezifischen Geländeeigenschaften ausdrücken, bspw. die Neigung oder Wölbung des Geländes, abgeleitet werden (BÖHNER et al. 2006, HENGL und REUTER 2009, IWAHASHI und PIKE 2007). BÖHNER et al. (2006) unterscheiden dabei einfache und komplexe Reliefparameter. Einfache Reliefparameter können, ohne genaueres Wissen über die Erdoberfläche vorzusetzen, direkt aus einem digitalen Geländemodell abgeleitet werden. Die Berechnung erfolgt meist durch eine Filteroperation. Bekannte Beispiele für diese Art von Reliefparametern sind die Hangneigung (slope), die Ausrichtung (aspect) oder die



Wölbung (curvature). Komplexe Reliefparameter basieren auf zuvor berechneten einfachen Reliefparametern und stellen deren Kombination dar wie z.B. Feuchtigkeitsindex (terrain wetness index), Massenbilanzindex (mass balance index) oder der Flussintensitätsindex (stream power index).

Da sich fluviale Prozessbereiche aufgrund ihrer spezifischen Lage in ebenem Gelände in Nachbarschaft zu Flussläufen gut von anderen Reliefeinheiten abgrenzen lassen, werden in der vorliegenden Arbeit Referenzcluster im fluvialen Prozessbereich mit der Bodengesellschaft Auenböden ermittelt. Zur Identifizierung von Auenbereichen wird das digitale Geländemodell hinsichtlich der einfachen Reliefattribute Neigung (slope) und normalisierte Höhe (NH normalized height) sowie des komplexen Reliefattributes Massenbilanzindex (MBI) abgeleitet. In fluvialen Prozessbereichen weist NH die niedrigsten Werte im Vergleich zu ihrer Umgebung auf; die Neigung ist gleich null bis geringfügig positiv (MÖLLER 2004).

Der Massenbilanzindex wird zur Beurteilung von Sedimentationsprozessen und Abtragungsprozessen herangezogen und basiert auf Funktionen, welche die Hangneigung, die vertikale Distanz zum Tiefenliniennetzwerk und die Krümmung berücksichtigen (MÖLLER et al. 2012). Negative MBI-Werte repräsentieren Sedimentauffangbecken, also morphologische Depressionen und Täler, in welchen sich kolluviale wie auch alluviale Sedimente ansammeln. Erosionszonen (z.B. konvexe Hänge) sind durch positive MBI-Werte gekennzeichnet. MBI-Werte von 0 weisen auf ein Gleichgewicht zwischen Bodenabtrag und -akkumulation hin. Der MBI erreicht sein Minimum bei stark konkav gewölbten und schwach geneigten Reliefformen, die sich in minimaler Entfernung von der Tiefenlinie befinden. Das MBI-Maximum bzw. das Maximum der negativen Massenbilanz steht im Zusammenhang mit stark konvexen und stark geneigten Reliefformen bei maximaler Entfernung von der Tiefenlinie (MÖLLER et al. 2015). In Auenbereichen hat der Massenbilanzindex negative, aber nicht die höchsten Werte, da Auen nicht mit Senkenbereichen gleichzusetzen sind (MÖLLER 2004).

### 2.3.2 Segmentierung zu Boden-Relief-Einheiten

Ein notwendige Voraussetzung für die Bildung von Objektgruppen ist die Rasterbildsegmentierung (BAATZ und SCHÄPE 2000). Unter Segmentierung wird die räumliche Gliederung eines Zellenrasters in Polygone bzw. Segmente (FORTIN et al. 2000) verstanden. Boden-Relief-Einheiten sind räumliche Einheiten, die in Bezug auf den Bodentyp und das Relief homogen sind (HENGL et al. 2008, MCBRATNEY et al. 2003, MULDER et al. 2011). Um Boden-Relief-Einheiten zu bilden, werden die abgeleiteten Reliefparameter mit den digital aufbereiteten Daten der VBK50 verschnitten.

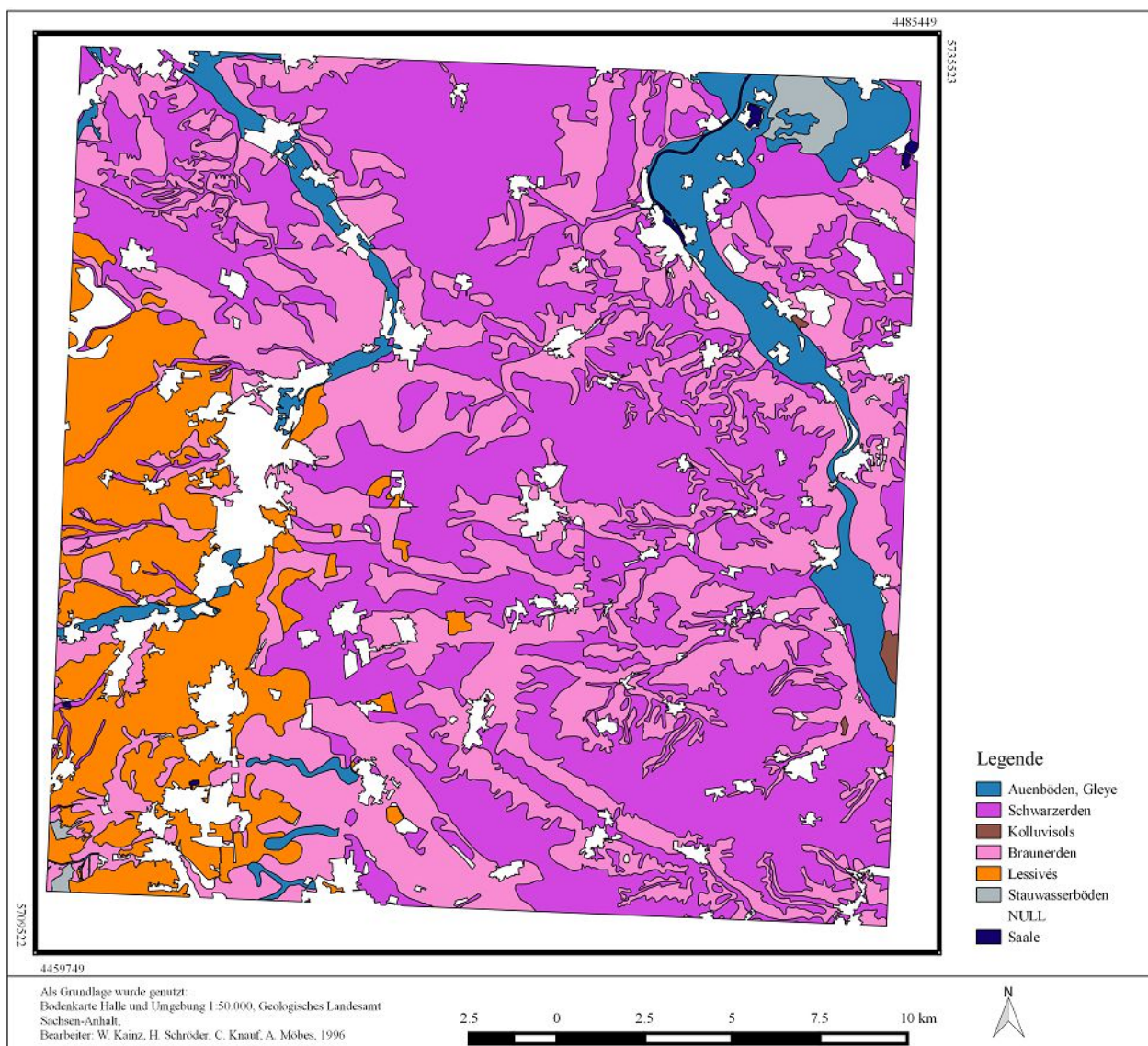
BLASCHKE (2010) unterscheidet vier Kategorien von Segmentierungstechniken: Punktorientierte Verfahren, Kanten-/Konturorientierte Verfahren, Regionorientierte Verfahren und Mischverfahren. Regionorientierte Techniken beruhen auf Ursprungspixeln (sog. seed points), die zu Beginn oder vor der Segmentierung algorithmisch oder durch einen Nutzer bestimmt werden. Ausgehend von diesen einzelnen Rasterzellen werden die Rasterzellen aufgrund statistischer (Vergleich von Mittelwerten) und räumlicher Ähnlichkeiten zu Segmenten zusammengefasst, bis bestimmte Abbruchbedingungen (Ähnlichkeitsmaß, minimaler Pixelwert) erfüllt sind. Kanten-/Konturorientierte Verfahren identifizieren Begrenzungselemente und versuchen dann diskontinuierliche Grenzen miteinander zu verbinden, um eine räumliche Unterteilung zu erreichen. Mischverfahren nutzen Ansätze aus dem Punktorientierten Verfahren, dem Regionwachstum und der Grenzerkennung (BÖHNER et al. 2006, FORTIN et al. 2000). In Anlehnung an HÄRING et al. (2012) und MÖLLER et al. (2012) wird ein regionenbasierter iterativer Segmentierungsalgorithmus angewendet, als Abbruchbedingungen werden die Grenzen der Bodenklassen der VBK50 dienen. Im Ergebnis dieses Arbeitsschrittes trägt jedes Segment, d.h. jede Boden-Relief-Einheit, einen Datensatz, der sich aus den Mittelwerten der zugehörigen Rasterzellen ergibt.

### 2.3.3 Clusterbildung

Die Clusterbildung ist die statistische Gruppierung eines Merkmalsraumes mit dem Ziel, ähnliche Objekte zusammenzufassen. Grundlegend für alle Cluster-Techniken ist die Wahl des Ähnlichkeitsmaßes zwischen zwei Objekten, die Clustermethode und die Clusteranzahl (HASTIE et al. 2011).

Nach HENGL und REUTER (2009) werden geomorphologische Klassen mit der k-mean-Methode erstellt. Von einer vom Nutzer definierten Anzahl von k-Zentren ausgehend, wird jeder Punkt iterativ mit dem nächsten Cluster-Zentrum verglichen. Als Ähnlichkeitsmaß wird die quadrierte euklidische Distanz genutzt. Die Wahl der geeigneten Clusteranzahl ist entscheidend für die Clusterbildung. Eine zu hohe Zielclusterzahl führt zu einer großen Heterogenität, d.h. es werden zu viele Cluster gebildet, die keine Rückschlüsse auf Bodentypen bzw. Lageungenauigkeiten selbiger zulassen. Eine zu geringe Clusteranzahl führt zur Ausbildung von zu homogenen Flächen, die Lageungenauigkeiten inkorporieren anstatt sie aufzudecken. Um einen Richtwert für die zu berechnende Clusteranzahl  $k$  zu ermitteln, wird die Gleichung von MARDIA et al. (1979) angewendet:  $k \simeq \sqrt{n/2}$ . Es wird eine Hauptkomponentenanalyse durchgeführt, die Korrelationen zwischen den Parametern prüft. Bei vorliegender Korrelation gehen die Parameter einfach in die Analyse ein.

In einem ersten Schritt werden Bodentypen aufgrund pedogenetischer Zugehörigkeiten zu den sechs Bodengesellschaften Auenböden und Gleye, Braunerden, Schwarzerden, Kolluvisols, Stauwasserböden und Lessivés (Abbildung 8) zusammengefasst. Die Cluster werden in jeder aggregierten Bodenvergesellschaftung unabhängig zu den übrigen bestimmt, da ansonsten bodengenetisch unlogische Cluster erzeugt werden.



Kurzzeichen Bodensystematische Einheit	Name Bodenvergesellschaftung
AB, AT, AZ, GG-AT, AA-YK	Auenböden, Gleye
BB, RN, RZ	Braunerden
BB-TT, GG-TT, SS-TT, TT	Schwarzerden
YK	Kolluvisols
SS, SS-AB, SS-BB	Stauwasserböden
BB-LF, LF, LL, OL, RR	Lessivés

Abbildung 8: Zusammensetzung der Bodengesellschaften in Anlehnung an MÖLLER 2012, Farbwertbelegung Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie

### **2.3.4 Expertenbasierte Referenzclusterauswahl**

Referenzcluster sind Cluster von Boden-Relief-Einheiten innerhalb einer aggregierten Bodenvergesellschaftung mit einer repräsentativen Werteverteilung, die dadurch als wahr definiert werden. Die Referenzcluster werden durch den Bodenexperten aufgrund seines Expertenwissens ausgewählt (durch z.B. Lage des Clusters im Gelände oder Kenntnis der lokalen Gegebenheiten des Experten). Im vorliegenden Anwendungsfall werden explizit Referenzcluster im Bereich der Bodengesellschaft Auenböden ausgewählt. Der Bodenexperte sucht die Cluster mit den Eigenschaften Auenzugehörigkeit, die er als plausibel bewertet im Rasterlayer aus.

### **2.3.5 Statistischer Vergleich der Referenzcluster mit übrigen Clustern**

Die durch den Bodenexperten als plausibel bewerteten Cluster werden mit den übrigen Clustern innerhalb und zwischen den Bodenvergesellschaftungen statistisch verglichen. Der Vergleich wird durch den Kolmogorov-Smirnov-Test (K-S-Test) realisiert. Der K-S-Test ist ein traditioneller Goodness-of-fit-Test. Der Hauptvorteil dieses nichtparametrischen zweiseitigen Test ist, dass jede Art von Verteilung miteinander verglichen werden können, ohne spezifische statistische Bedingungen zu erfüllen (THAS 2010). Aufgrund seiner statistischen Robustheit eignet sich der K-S-Test sehr gut im Umgang mit Umweltdaten. Basierend auf der empirischen Summenverteilungsfunktion überprüft der K-S-Test, ob zwei Verteilungen gleich sind (Nullhypothese) oder sich signifikant unterscheiden (MÖLLER et al. 2012). Cluster die statistisch gleich sind, werden als plausibel fixiert. Als Ergebnis liegen Cluster vor, die eindeutig plausibel sind bzw. nochmals durch den Bodenexperten geprüft werden müssen. Der Bodenexperte hat verschiedene Möglichkeiten mit den nicht plausiblen Clustern weiter zu verfahren.

### **2.3.6 Neubewertung und abschließende Fixierung der Cluster**

Die nicht plausiblen Cluster werden entweder neu berechnet (1) oder durch den Bodenexperten direkt im Geoinformationssystem (2a) oder unterstützt durch punktuelle Geländebegehung (2b) den Clustern zugeordnet (die Zahlen in den Klammern entsprechen den Prozessschritten in Abbildung 7 auf Seite 24). Für die Neuberechnung werden die Cluster in Segmente aufgelöst und ab Verfahrensschritt (3) der Clusterbildung neu berechnet. Im Ergebnis entsteht ein Vektorlayer.

## ***2.4 Integration des Bodenexpertenwissens: Erarbeitung einer Fachapplikation als Rahmenstruktur***

Damit die qualitativ-individuellen Erfahrungswerte der Bodenexperten in die objektive Clusterbildung des DSM integriert werden können, bedarf es eines Systems, welches den Bodenexperten an der Schnittstelle zwischen statistischer Datenanalyse und Ergebnisevaluierung interaktiv einbezieht. Hierfür eignet sich eine Fachapplikation, die verschiedene Geoinformationssysteme (GIS) in einen Rahmen fasst und so dem Bodenexperten die Möglichkeit bietet, sein Expertenwissen zielgerichtet und autark, d.h. ohne die Unterstützung eines Experten der Informationstechnologie (IT), einzubringen.

Die Fachapplikation soll den in Kapitel 2.3 erläuterten Verfahrensablauf in einer Plattform zusammenfassen und die Integration des Bodenexpertenwissens in den Verfahrensablauf umsetzen. Durch das Zusammenführen der für das Verfahren notwendigen GIS, sollen Informationsverluste vermindert werden, die ansonsten durch den Wechsel zwischen Geoinformations (GI)-Anwendungsumgebungen entstehen (z.B. durch Datenimport/-export, Datenkonvertierung).

Die QGIS-Verarbeitungswerkzeuge bieten die Möglichkeit, neben den QGIS-internen Geoalgorithmen auf weitere Geoanalysealgorithmen (GRASS GIS 153 Algorithmen, SAGA 235 Algorithmen, Orfeo Toolbox 83 Algorithmen, R 27 Algorithmen mit der Option eigene Skripte einzubinden) zuzugreifen. QGIS unterstützt Vektor- und Rasterformate. Das Programm ist in C++ entwickelt und

verwendet für die grafischen Benutzeroberflächen die Bibliothek Qt. Weiterhin bietet QGIS mit der Erweiterung Pluginbuilder die Möglichkeit, Fachapplikationen zu erstellen, die dann im Qt4-Designer bearbeitet werden, sowie mit einer Entwicklungsumgebung (bspw. Eric4, Eclipse) über die Programmiersprache Python die gewählten Algorithmen zu hinterlegen. Voraussetzung ist die Quelloffenheit des einzubindenden Algorithmus, idealerweise als Kombination mittels der Programmiersprache Python. Somit wurde die Vorauswahl für die Kombination der Werkzeuge hinsichtlich (1) der möglichen Integration in die QGIS-Verarbeitungswerkzeuge, (2) Quellcodeoffenheit, (3) Ansprache mittels Python und (4) Vorhandensein des benötigten Verarbeitungsprozesses getroffen. Folgende Geoinformationssysteme wurden für eine Integration in den Verfahrensablauf geprüft:

- Die **Orfeo Toolbox** ist eine quelloffene C++ Bibliothek für die Klassifikation, Segmentierung und Objektbildung von Fernerkundungsdaten (TINEL et al. 2012).
- **SAGA** (System for Automated Geoscientific Analyses) ist ein geografisches Informationssystem. Es unterstützt Raster-Daten wie z.B. digitale Höhendaten und Satellitendaten sowie Vektordaten und Tabellen. Der Schwerpunkt liegt in den Gebieten digitale Höhenanalyse, Geostatistik, Bildverarbeitung und Simulationen. SAGA ist in C++ und C entwickelt und unterstützt weiterhin Python (OLAYA und CONRAD 2009).
- **R** ist eine Programmiersprache und Softwareumgebung für statistische Berechnungen. Sie wird hauptsächlich zur Datenanalyse eingesetzt und bietet durch ihr matrizenorientiertes Datenmodell die Möglichkeit der Analyse von räumlichen Rasterdaten. Für diesen Zweck sind Pakete verfügbar, die unter der Aufgabenansicht *Spatial* im Archivverteiler (Comprehensive R Archive Network, CRAN) zusammengefasst sind. Über Rpy2 kann R mittels Python angesprochen werden (BIVAND et al. 2013).
- **GRASS** (Geographic Resources Analysis Support System) **GIS** ist ein modular aufgebautes, quelloffenes Geoinformationssystem. Es bietet Funktionen zur Verwaltung und Analyse räumlicher Vektor- und Rasterdaten mit Schwerpunkt

in der Bildverarbeitung. Seit 2005 kann GRASS GIS als Plugin und über die Verarbeitungswerkzeuge in QGIS integriert werden. GRASS ist in C entwickelt und bietet durch PyGRASS die Möglichkeit mittels Python GRASS GIS Module zu integrieren (NETELER et al. 2012). Aktuell liegt GRASS GIS in der stabilen Version 7 vor.

- **GeoDMA** (Geographic Data Mining Analyst) ist ein Geoprocessingwerkzeug für die Integration von Funktionalitäten der Bildsegmentierung und -klassifizierung sowie Validierung von Fernerkundungsdaten in das Desktop-GIS Terra View. GeoDMA ist in C++ entwickelt (KÖRTING et al. 2013).

Tabelle 2 stellt die nötigen Kriterien für eine Integration in die Fachapplikation und die möglichen GI-Systeme zusammen. Wenn ein GI-System die Kriterien überwiegend erfüllt (d.h. über 50 %), geht es in die weiteren Untersuchungen ein.

*Tabelle 2: Technische Hintergründe der vorausgewählten GI-Systeme*

GI-Software	Aktuelle Version	Quell-offen	In QGIS Verarbeitungswerkzeug integriert	Pythonanbindung	Weitere Betrachtung
Orfeo Toolbox	4.4	ja	ja	Otb-python	ja
SAGA	2.1.2	ja	ja	SAGA-Python	ja
R	3.1.2	ja	ja	Rpy2	ja
GRASS	7.0	ja	ja	pyGRASS	ja
GeoDMA	TerraView 4.2.2 mit GeoDMA 0.2.2	ja	nein	nein	nein
QGIS	2.8	ja	ja	PyQGIS	ja

Im Anschluss wurde untersucht, welche GI-Systeme für den Verfahrensverlauf benötigte Algorithmen aufweisen, unabhängig davon, ob der Algorithmus Bestandteil



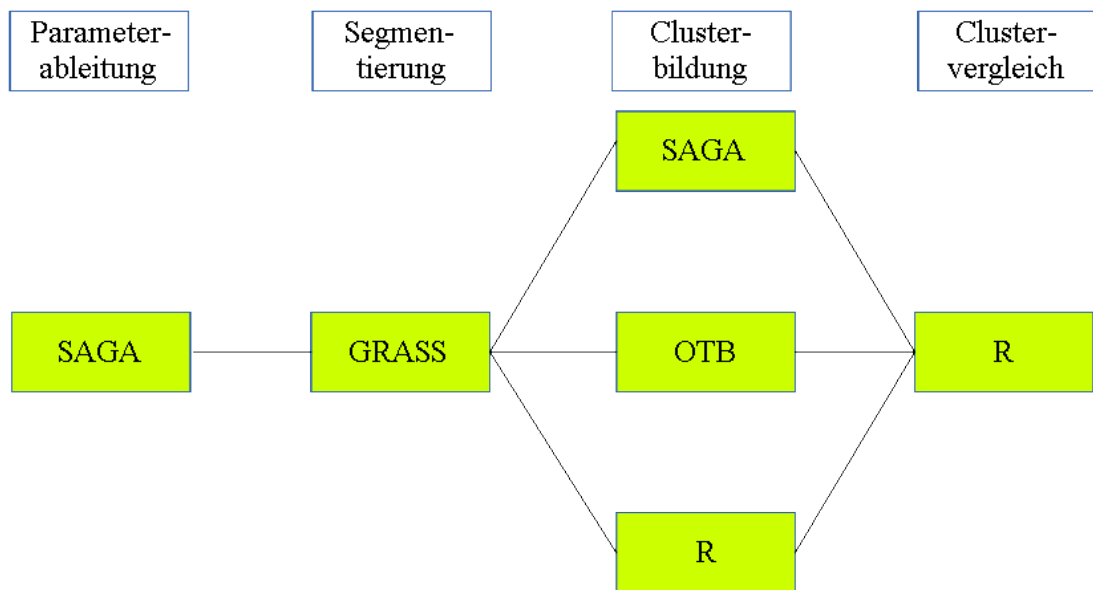
der QGIS-Verarbeitungswerkzeuge ist. In Tabelle 3 sind die für die Untersuchung ausgewählten GI-Systeme und die vorhandenen Algorithmen (grün hinterlegt) zusammengestellt.

*Tabelle 3: Für den Verfahrensverlauf vorhandene Algorithmen in der ausgewählten GI-Software (grün hinterlegt)*

GI-System	Ableiten mehrer Reliefattribute	regionenbasierte Segmentierung mit mehreren Eingangslayern	Clusterbildung (k-mean)	Clustervergleich (K-S-Test)
Orfeo Toolbox	nein	nein	ja	nein
SAGA	Mass balance index slope normalized height	Simple region growing*	ja	nein
R	nein	nein	ja	ja
GRASS	r.slope.aspect	i.segment	nein	nein
QGIS	nein	nein	nein	nein

\* Grenzen der VBK50 können nicht als Abbruchkriterium eingebracht werden, keine Dokumentation vorhanden, kein aktives Forum

Abbildung 9 stellt die drei möglichen Varianten der Kombination von Algorithmen dar, die im Folgenden untersucht werden sollen.



Nr.	Parameter-ableitung	regionenbasierte Segmentierung mit mehreren Eingangslayern	Clusterbildung (k-mean)	Clustervergleich (K-S-Test)
1	SAGA	GRASS	SAGA	R
2	SAGA	GRASS	R	R
3	SAGA	GRASS	OTB	R

Abbildung 9: Versuchsablauf (eigene Darstellung)

### 2.5 Validierung der Ergebnisse / Identifizierung von Lageungenauigkeiten

Fundamentaler Bestandteil einer geostatistischen Analyse ist die Bewertung der Ergebnisse in Hinsicht auf ihre Realitätsnähe. Üblicherweise wird hierzu der Übereinstimmungsgrad zwischen Analyseergebnissen (Schätzwerten) und Referenzinformationen (Messwerten) bestimmt. Die Referenzinformationen leiten sich aus Zufallsstichproben ab, deren inhaltliche Belegung das Resultat von Felduntersuchungen sein sollte. Allerdings können oftmals aus Kapazitäts- und/oder Kostengründen die dafür notwendigen Erhebungen nicht durchgeführt werden (MÖLLER 2004).

Zur Erstellung der Referenzinformationen wurde das DHM mit der Bodengesellschaft Auenböden der VBK50 (entsprechend Kapitel 2.3.3, Abbildung 8) überlagert. Das DHM wurde im Bereich der Bodengesellschaft Auenböden hinsichtlich Unterschieden der Farbtöne zwischen den Pixeln visuell untersucht. Farbtonunterschiede stellen im DHM Höhenunterschiede dar, da Auenböden in ebenem bis leicht geneigtem Gelände verortet sind, sind diese über den Farbton der Pixel ermittelten Höhenunterschiede ein Anzeichen für potenzielle Lageungenauigkeiten der Bodengesellschaft Auenböden in der VBK50. Es wurden 25 Referenzstellen mit einer Mindestgröße von 10.000 m<sup>2</sup> definiert. Abbildung 10 zeigt die Lage der Referenzstellen im UG, Anhang 1 zeigt die Referenzstellen detailliert. Diese Gebiete werden als Vergleichsgrundlage für die statistisch berechneten Werte genommen.

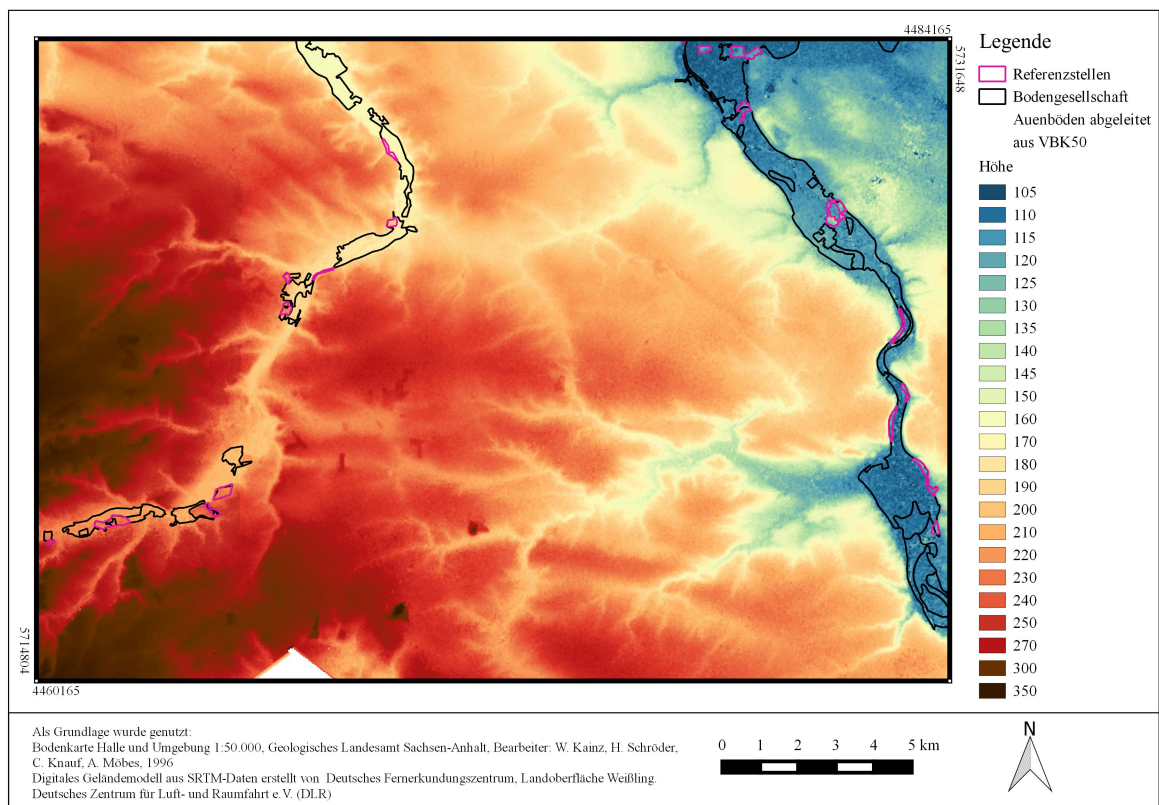


Abbildung 10: Lage der Referenzstellen im Untersuchungsgebiet

Die in Prozessschritt (3) berechneten Cluster werden validiert. Die Ergebnisse der drei Prozessvarianten (vgl. Abbildung 9) werden hinsichtlich der räumlichen Übereinstimmung zwischen den ermittelten Clustern und den Referenzstellen untersucht. Die prozentuale Abdeckung von separat abgegrenzten Clustern an den vorab visuell ermittelten Referenzstellen wird bestimmt. Anschließend wird der Durchschnitt der prozentualen Abdeckung an den 25 Referenzstellen ermittelt.

## 3. Ergebnisse

### 3.1 Ergebnisse der quantitativen Analyse

#### 3.1.1 Ableiten der Reliefparameter zur Charakterisierung fluvialer Prozessbereiche

Abbildung 11 zeigt die Rasteroberfläche des abgeleiteten Reliefparameters Massenbilanzindex. Schwarz begrenzt ist die Bodengesellschaft Auenböden und rosa begrenzt sind die Referenzstellen zur Evaluierung der ermittelten Lagegenauigkeiten.

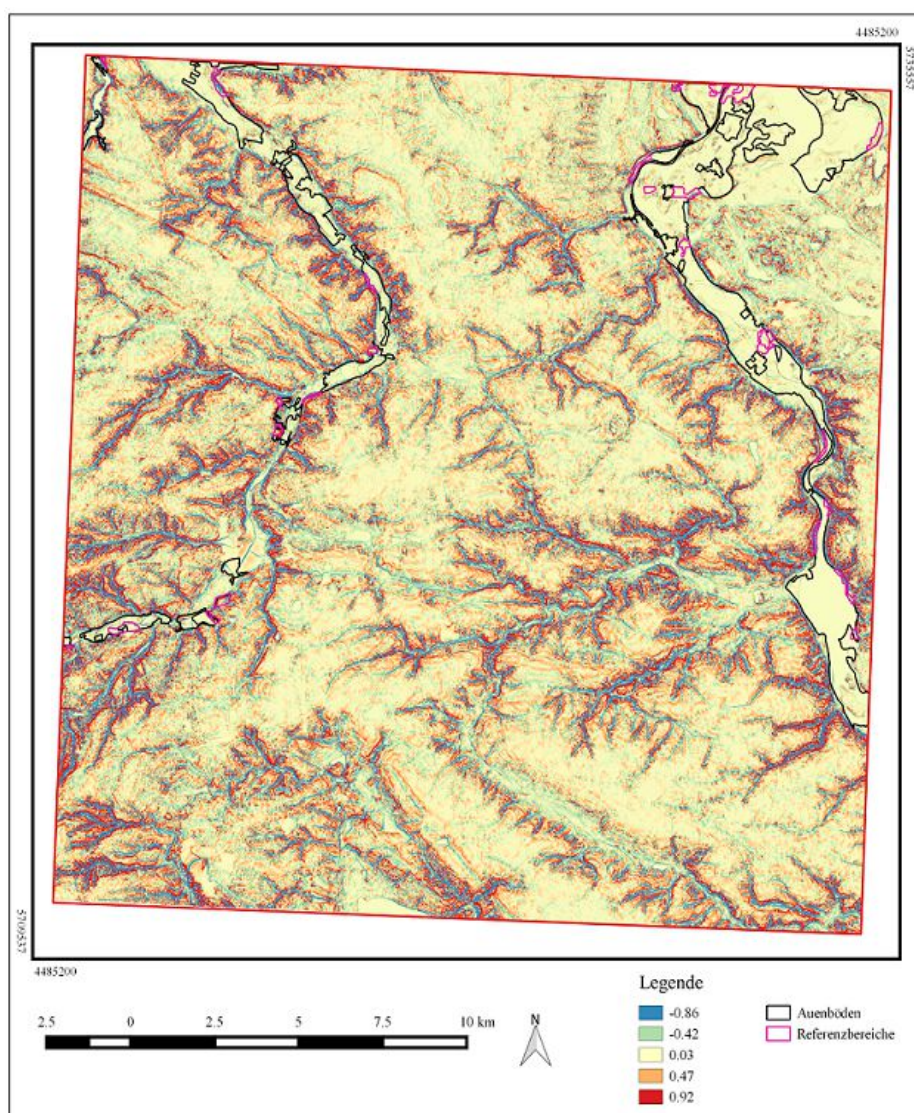
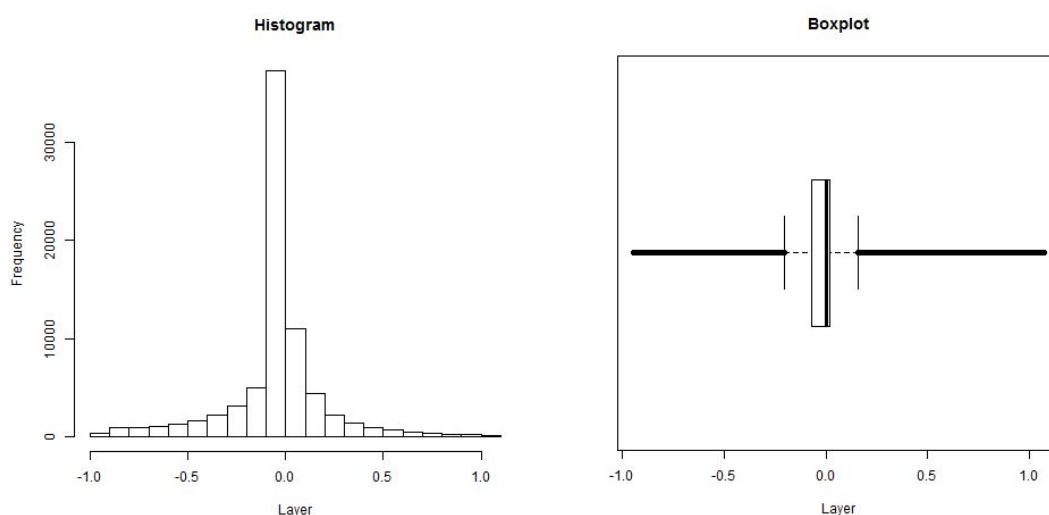


Abbildung 11: Massenbilanzindex

Zur visuellen Darstellung wurde das Rasterbild mit einem spektralen Farbverlauf in 5 Klassen belegt. MBI-Werte von 0,1 bis - 0,1 in blassgelb stehen für fluviale Prozeßbereiche. Die fluvialen Prozessbereiche werden von der Auenbodengesellschaft überlagert. Ausnahmen bilden die Randzonen der Auenbodengesellschaft und Bereiche im westlichen Auenbodengebiet des UG. Hier sind orange bis rote Pixelfarben (Erosionszonen) bzw. grüne bis blaue Pixelfarben (Sedimentation) zu erkennen. Das statistische Maximum der MBI-Werte liegt im Bereich der Auenböden bei 1,06, das statistische Minimum bei -0,95, das statistische Mittel bei -0,02 und die Standardabweichung bei 0,26. Das Histogramm und das Boxplot (Abbildung 12) zeigen, dass sich die mittleren 50 % der Daten eng im Wertebereich um 0 befinden. Die Werte streuen über die gesamte Spannweite.



*Abbildung 12: Histogramm und Boxplot des Massenbilanzindex für den Bereich der Auenböden*

Abbildung 13 zeigt die Rasteroberfläche des abgeleiteten Reliefparameters Normalisierte Höhe (NH). Die Kartendarstellung erfolgte analog der Darstellung des MBI: Das Rasterbild wurde mit einem spektralen Farbverlauf in 5 Klassen belegt. Schwarz begrenzt ist die Bodengesellschaft Auenböden und rosa begrenzt sind die Referenzstellen zur Evaluierung der ermittelten Lageungenauigkeiten. Die niedrigsten NH-Werte 0 bis 0,1 (blau) stehen für fluviale Prozessbereiche. Die fluvialen Prozessbereiche werden von der Auenbodengesellschaft überlagert. Ausnahmen bilden

die Randzonen der Auenbodengesellschaft und Bereiche im westlichen Auenbodengebiet des UG. Hier sind grüne bis rote Pixelfarben zu erkennen.

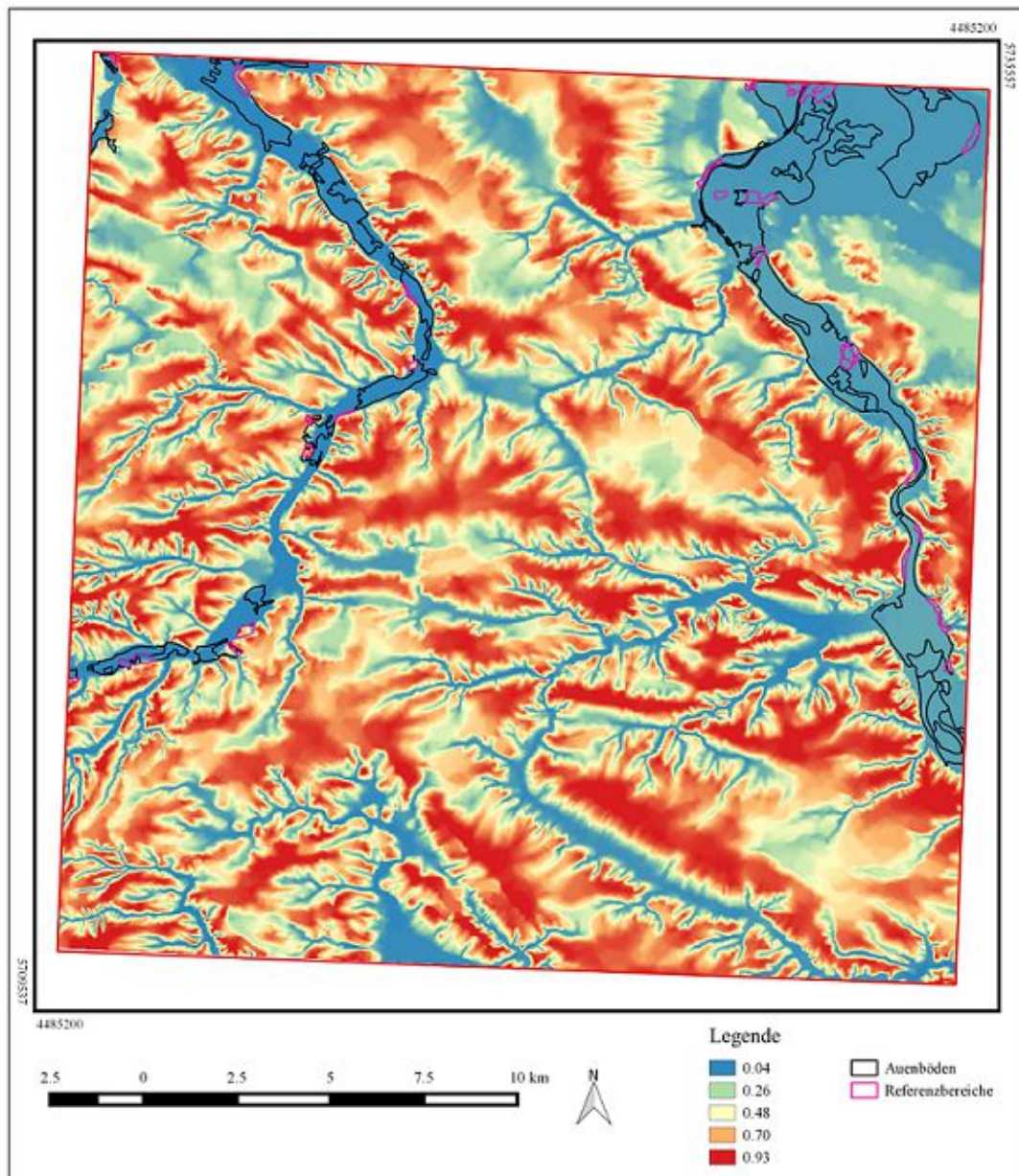


Abbildung 13: Normalisierte Höhe

Das statistische Maximum der NH-Werte liegt im Bereich der Auenböden bei 0,89, das statistische Minimum bei 0,02, das statistische Mittel bei 0,1 und die Standardabweichung bei 0,09. Das Histogramm und das Boxplot (Abbildung 14) der

NH-Werte im Auenbodenbereich zeigen, dass sich die mittleren 50 % der Daten eng im Wertebereich um 0,1 befinden. Die Werte streuen über die gesamte Spannweite.

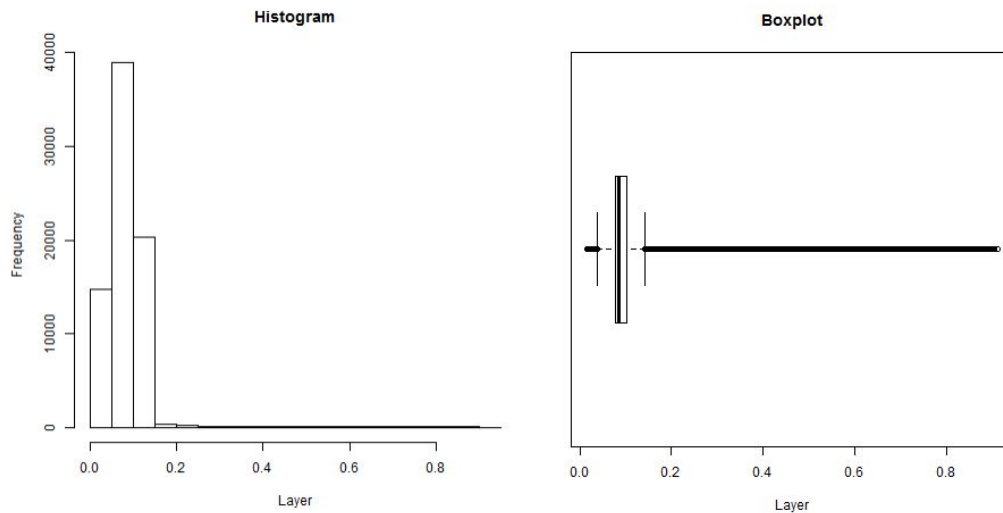


Abbildung 14: Histogramm und Boxplot der normalisierten Höhe für den Bereich der Auenböden

Abbildung 15 zeigt die Rasteroberfläche des abgeleiteten Reliefparameters Neigung. Die Farbwertbelegung der Kartendarstellung erfolgte analog der Darstellung des MBI und der normalisierten Höhe. Neigungswerte von null bis leicht positiv (blau bis leicht grün) stehen für fluviale Prozessbereiche. Die fluvialen Prozessbereiche werden von der Auenbodengesellschaft überlagert. Ausnahmen bilden die Randzonen der Auenbodengesellschaft. Hier sind grüngelbe bis rote Pixelfarben zu erkennen. Dies ist im westlichen Bereich des Untersuchungsgebietes auffällig.



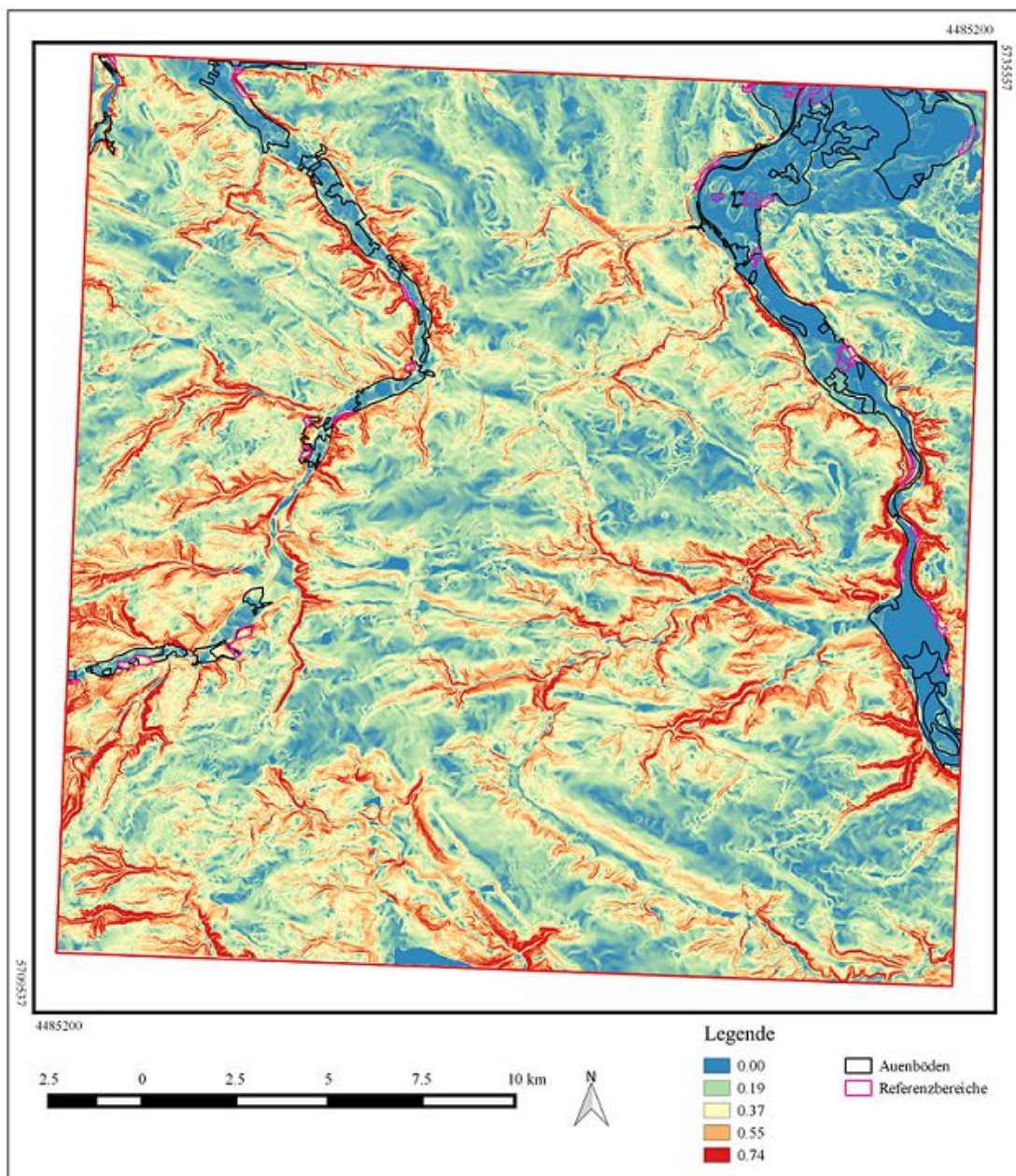


Abbildung 15: Neigung

Das statistische Maximum der Neigungswerte liegt im Bereich der Auenböden bei 0,83, das statistische Minimum bei 0,00, das statistische Mittel bei 0,09 und die Standardabweichung bei 0,16. Das Histogramm und das Boxplot (Abbildung 16) der Neigungswerte im Auenbodenbereich zeigen, dass sich die mittleren 50 % der Daten eng im Wertebereich um 0,1 befinden. Die Werte streuen über die gesamte Spannweite.

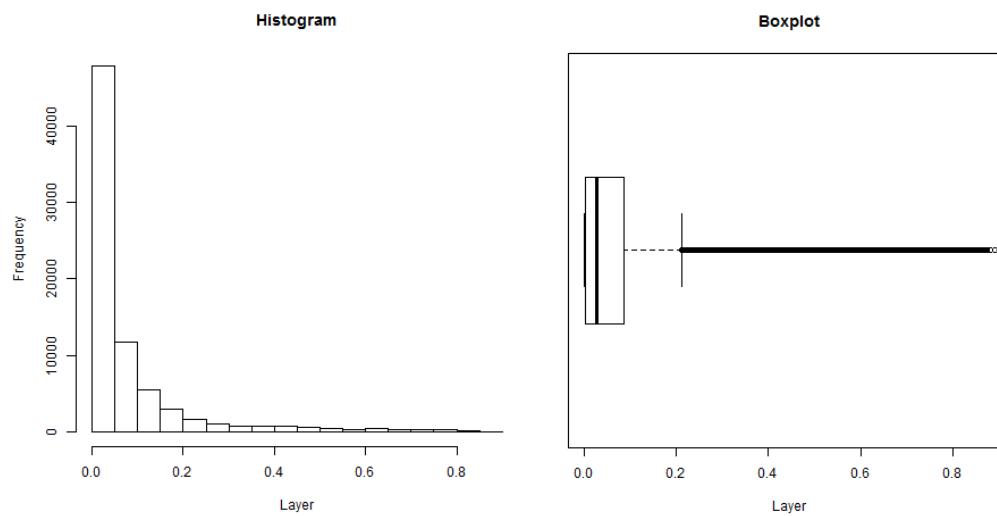


Abbildung 16: Histogramm und Boxplot der Neigung für den Bereich der Auenböden

### 3.1.2 Segmentierung

Das Ähnlichkeitsmaß und die minimale Anzahl von Pixeln pro Segment bestimmen die zu bildende Anzahl von Segmenten. Der Schwellenwert des Ähnlichkeitsmaßes bewegt sich im Bereich 0 (ausschließlich identische Pixel werden verbunden) und 1 (alle Pixel werden verbunden). 0,05, 0,1, 0,15, 0,2, 0,3 wurden als Schwellenwerte geprüft. Wenn die minimale Anzahl von Pixeln pro Segment zu gering gewählt wird, kommt es zu dem sogenannten Salz-und-Pfeffer-Effekt, d.h. die Segmente bestehen aus vielen einzelnen Pixeln (BLASCHKE 2010). Wenn die minimale Anzahl von Pixeln pro Segment zu hoch gewählt wird, verliert das Segmentierungsergebnis an Detailschärfe (CAVAZZI et al. 2013). 1, 2, 6, 10, 11, 13 und 15 wurden als minimale Pixelwerte, die ein Segment bilden, geprüft. Ein Schwellenwert des Ähnlichkeitsmaßes von 0,1 und ein minimaler Wert von 10 Pixeln pro Segment wurden für die Segmentierung angewendet. Abbildung 17 zeigt das Segmentierungsergebnis und Abbildung 18 das dazugehörige Histogramm. Dieses weist keine Normalverteilung auf.

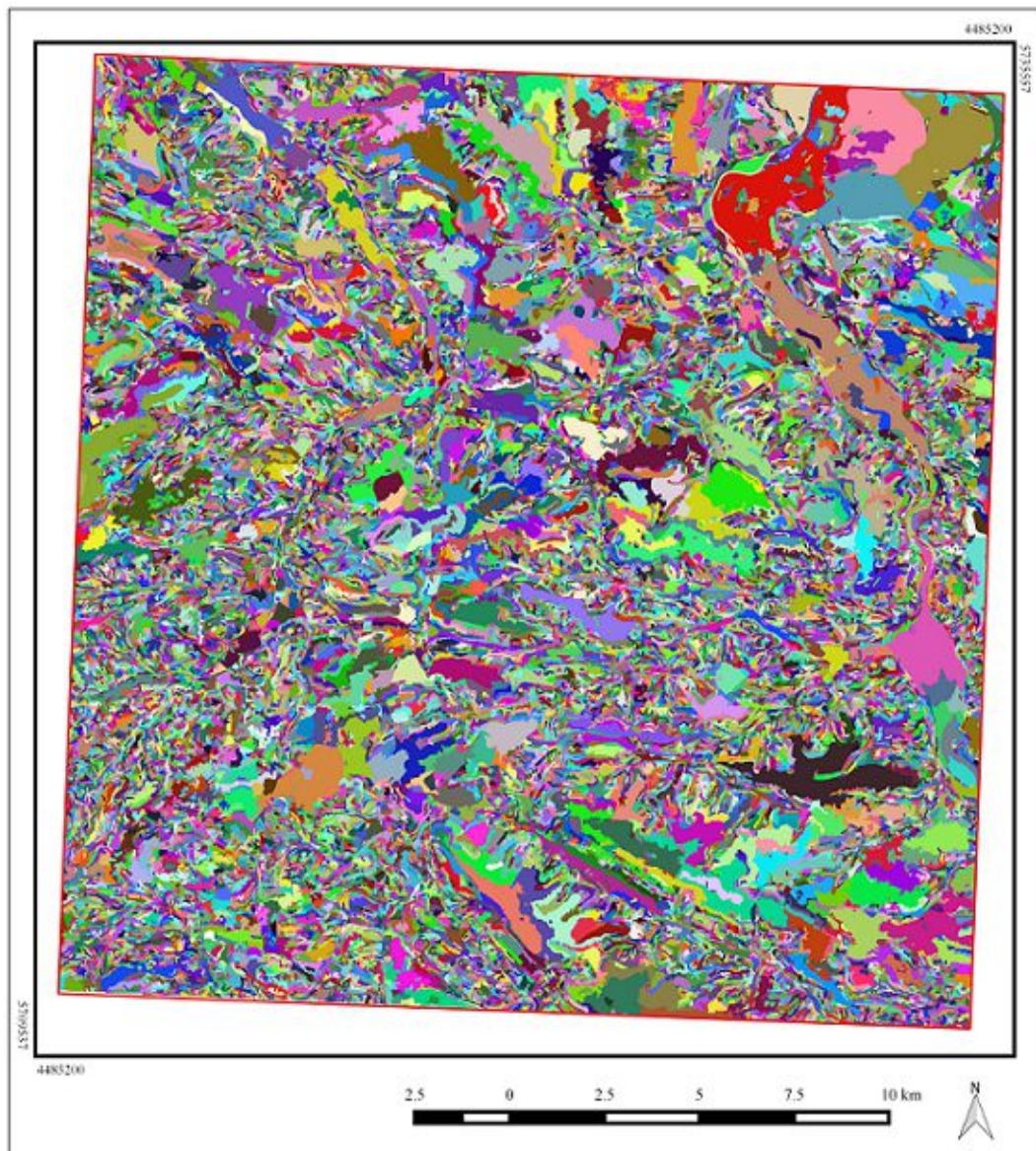


Abbildung 17: Segmentierungsergebnis

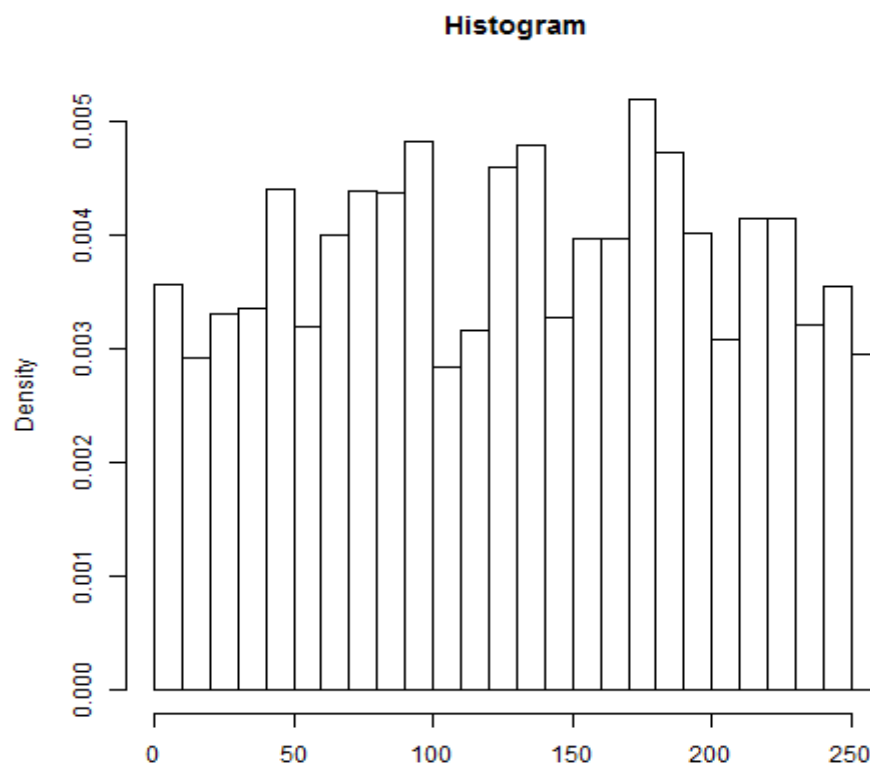


Abbildung 18: Histogramm der Segmentverteilung

### 3.1.3 Clusterbildung

Die Cluster wurden für den Bereich der Auenbodengesellschaft berechnet. Ausgehend von der Formel von MARDIA et al. (1979) zur Berechnung der Clusteranzahl:  $k \approx \sqrt{n/2}$  wurde bei 754 Segmenten im Bereich der Auenbodengesellschaft eine Clusteranzahl von ca. 19,4 ermittelt. Im Rahmen der Versuchsdurchläufe wurde eine Clusteranzahl von 20 genutzt. Abbildung 19, 20 und 21 zeigen die mit dem k-mean-Algorithmus der Orfeo Toolbox, R und SAGA ermittelten Cluster.

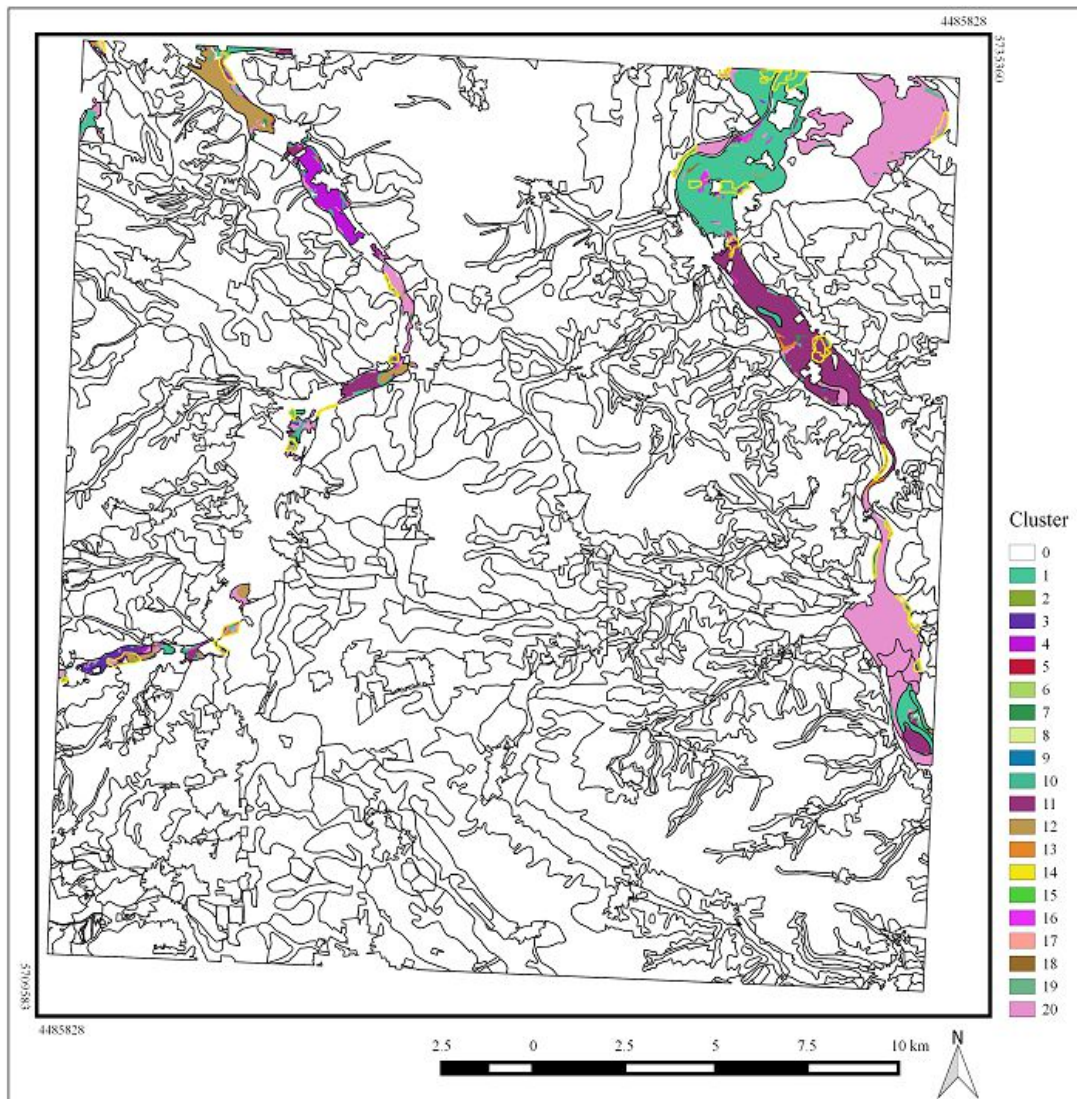


Abbildung 19: Clusterbildung mit der Orfeo Toolbox, Referenzstellen gelb umgrenzt

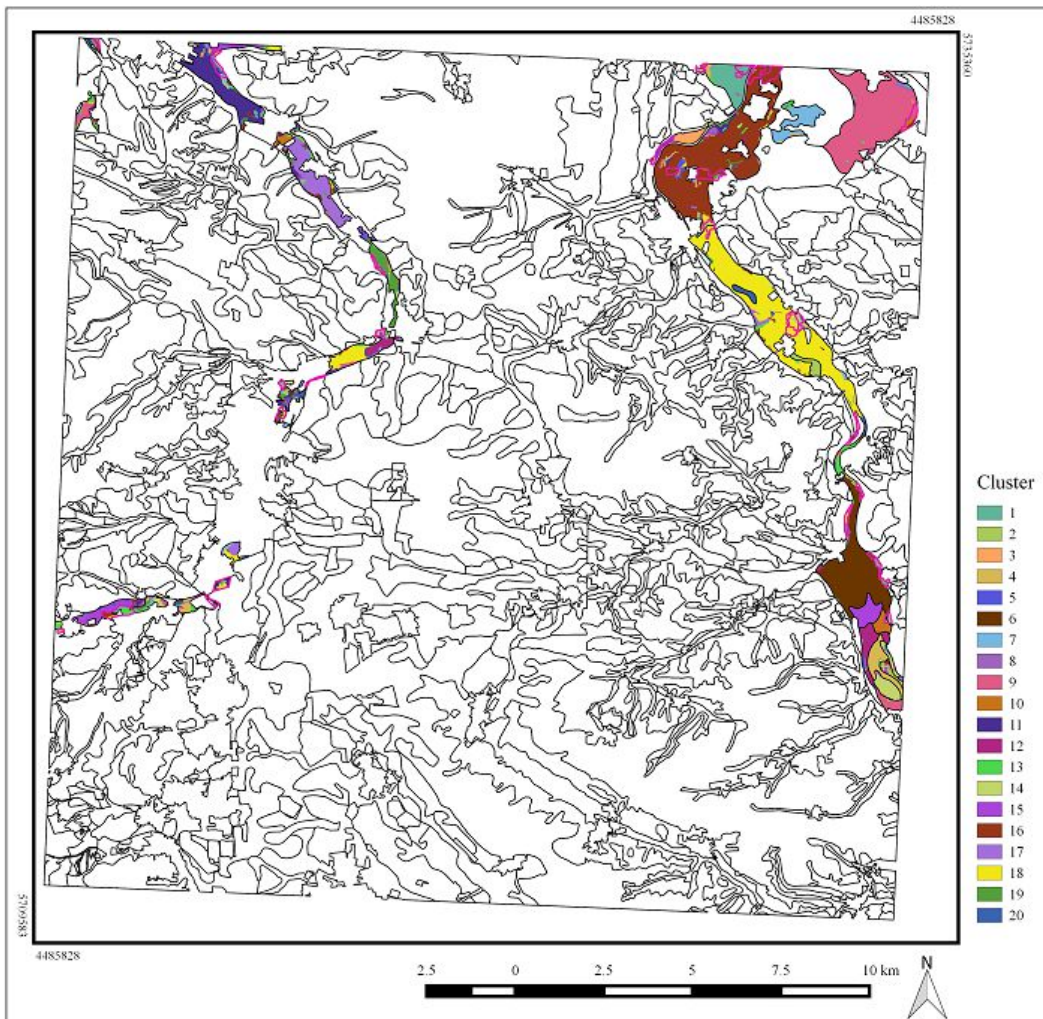


Abbildung 20: Clusterbildung mit R, Referenzstellen rosa umgrenzt

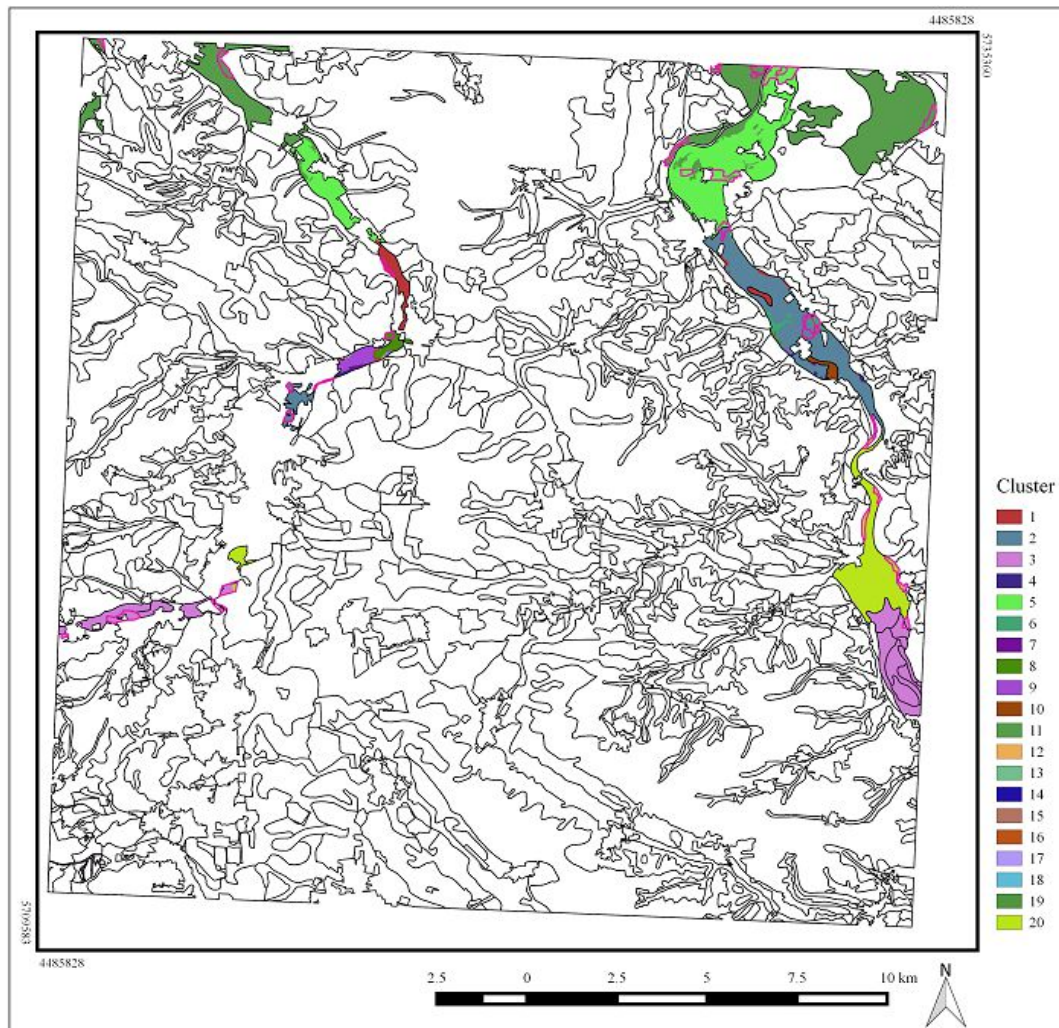
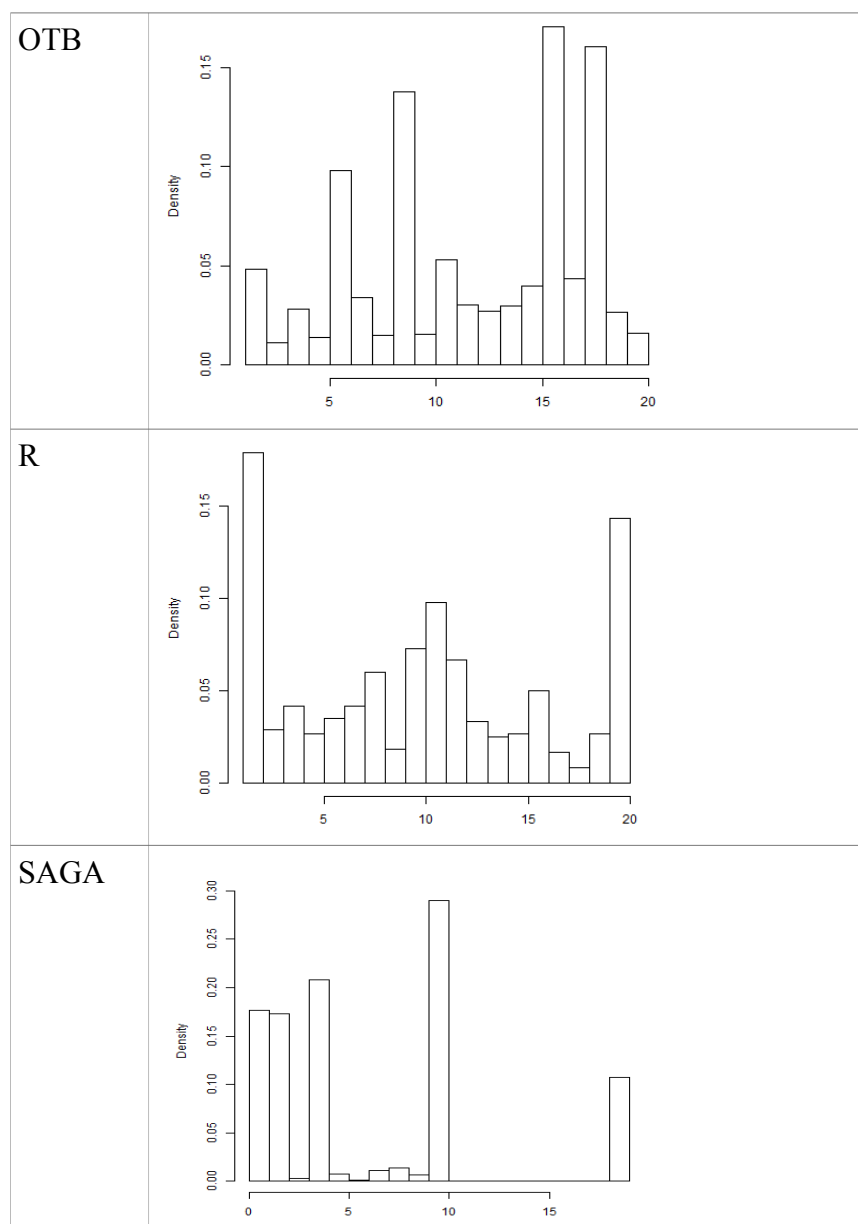


Abbildung 21: Clusterbildung mit SAGA, Referenzstellen rosa umgrenzt

Es ist zu erkennen, dass die mit der Orfeo Toolbox (Abbildung 19) und R (Abbildung 20) erstellten Cluster über die Bodengesellschaft kleinteiliger verteilt sind als die mit SAGA berechneten Cluster (Abbildung 21). Speziell die höher gelegenen Bereiche im Westen des UG weisen bei der Clusterbildung in R und OTB eine heterogenere Clusteroberfläche als die mit SAGA ermittelte Clusteroberfläche auf. In den Auenbödenbereichen im östlichen Teil des UG wurden in den Kernbereichen mit allen drei angewendeten Systemen homogene Cluster erstellt. Hier wurden in den Randzonen der Bodengesellschaft Auenböden speziell mit den Berechnungen in OTB und R kleinere Cluster abgegrenzt. Tabelle 4 stellt die Histogramme der Cluster-Berechnung mit einer relativen Häufigkeitsverteilung (engl. density) der einzelnen

Berechnungen gegenüber (Maximalwert =1). Dabei wird deutlich, dass das mit OTB berechnete Ergebnis 4 Cluster mit einer relativen Häufigkeit von größer 0,1 besitzt, das mit R berechnete Ergebnis 3 Cluster mit einer relativen Häufigkeit von größer 0,1 und das mit SAGA berechnete Ergebnis 5 Cluster mit einer relativen Häufigkeit von größer 0,1 besitzt. Die weiteren 16 mit OTB berechneten Cluster bewegen sich im Bereich zwischen 0,01 und 0,05. Die weiteren 17 mit R berechneten Cluster bewegen sich im Bereich zwischen 0,01 und 0,07. Die weiteren 15 mit SAGA berechneten Cluster bewegen sich zwischen 0,00 und 0,02 relativer Häufigkeit.

*Tabelle 4: Histogramme der Clusterverteilung bei OTB, R, SAGA*





### 3.2 Validierung der Ergebnisse

Anhang 1 beinhaltet die Darstellung des visuellen Vergleichs der räumlichen Übereinstimmung zwischen den Referenzstellen und den mit den Algorithmen in OTB, R und SAGA berechneten Clustern. Tabelle 5 gibt zusammenfassend die räumliche Übereinstimmung für jede Referenzstelle mit dem dazugehörigen Cluster an.

*Tabelle 5: Räumliche Übereinstimmung zwischen den Referenzstellen und den algorithmisch bestimmten Clustern (alle Angaben in Prozent)*

<b>ID</b>	<b>OTB</b>	<b>R</b>	<b>SAGA</b>
1	32,80	31,70	48,47
2	83,22	88,57	2,76
3	100	100	0
4	5,13	1,75	1,75
5	53,57	41,00	69,88
6	7,24	100	0
7	14,02	21,04	16,83
8	2,97	2,97	2,97
9	83,81	100	0
10	100	100	0
11	100	100	0
12	100	100	0
13	100	100	18,12
14	43,45	100	0
15	56,55	100	0
16	34,42	59,33	0
17	64,46	76,07	0
18	56,24	100	65,30
19	45,59	64,06	12,02
20	72,97	100	0
21	32,32	63,94	0
22	0	0	0
23	7,47	91,38	0
24	1,81	1,81	1,81
25	12,65	42,58	0
<b>Durchschnitt</b>	<b>48,43</b>	<b>67,45</b>	<b>9,60</b>

An 11 Referenzstellen konnte eine räumliche Übereinstimmung von 100 % zwischen den Referenzstellen und den mit dem k-mean-Algorithmus in R berechneten Clustern festgestellt werden. Für OTB war dies bei 5 Referenzstellen der Fall, wobei sich diese mit den 11 von R erkannten deckten. Mit dem k-mean-Algorithmus von SAGA konnte keine räumliche Übereinstimmung von 100 % zwischen den Referenzstellen und den berechneten Clustern festgestellt werden. An drei Referenzstellen (ID: 4, 8, 22) konnten von keinem der Algorithmen die Referenzstelle mit einer Abdeckung von über 10 % erkannt werden. Jenseits dieser maximalen und minimalen Werte bewegt sich die räumliche Übereinstimmung von Referenzstellen und der mit dem k-mean-Algorithmus in R berechneten Cluster im Bereich zwischen 21,04 % und 91,38 %. Im Durchschnitt erreichten die mit dem k-mean-Algorithmus in R berechneten Cluster eine Übereinstimmung von 67,45 % zu den Referenzstellen. Die räumliche Übereinstimmung von Referenzstellen und der mit dem k-mean-Algorithmus in OTB berechneten Cluster bewegt sich jenseits der maximalen und minimalen Werte im Bereich zwischen 12,65 % und 83,81 %. Im Durchschnitt erreichten die mit dem k-mean-Algorithmus in OTB berechneten Cluster eine Übereinstimmung von 48,43 % zu den Referenzstellen. Die räumliche Übereinstimmung von Referenzstellen und der mit dem k-mean-Algorithmus in SAGA berechneten Cluster bewegt sich jenseits der maximalen und minimalen Werte im Bereich zwischen 12,02 % und 69,99 %. Im Durchschnitt erreichten die mit dem k-mean-Algorithmus in SAGA berechneten Cluster eine Übereinstimmung von 9,6 % zu den Referenzstellen. 15 Referenzstellen hatten eine Übereinstimmung von 0 % zu den mit dem k-mean-Algorithmus in SAGA berechneten Cluster.

### ***3.3 Fachapplikation zur Integration des Fachwissens von Bodenexperten (qualitative Methode)***

Abbildung 21 zeigt die Benutzeroberfläche der Fachapplikation. Die Benutzeroberfläche der Fachapplikation wurde unter Anwendung der QGIS-Erweiterung Pluginbuilder und des Qt-Designers erzeugt. Grundlage ist der in Kapitel 2.3 Abbildung 7 vorgestellte Verfahrensablauf.

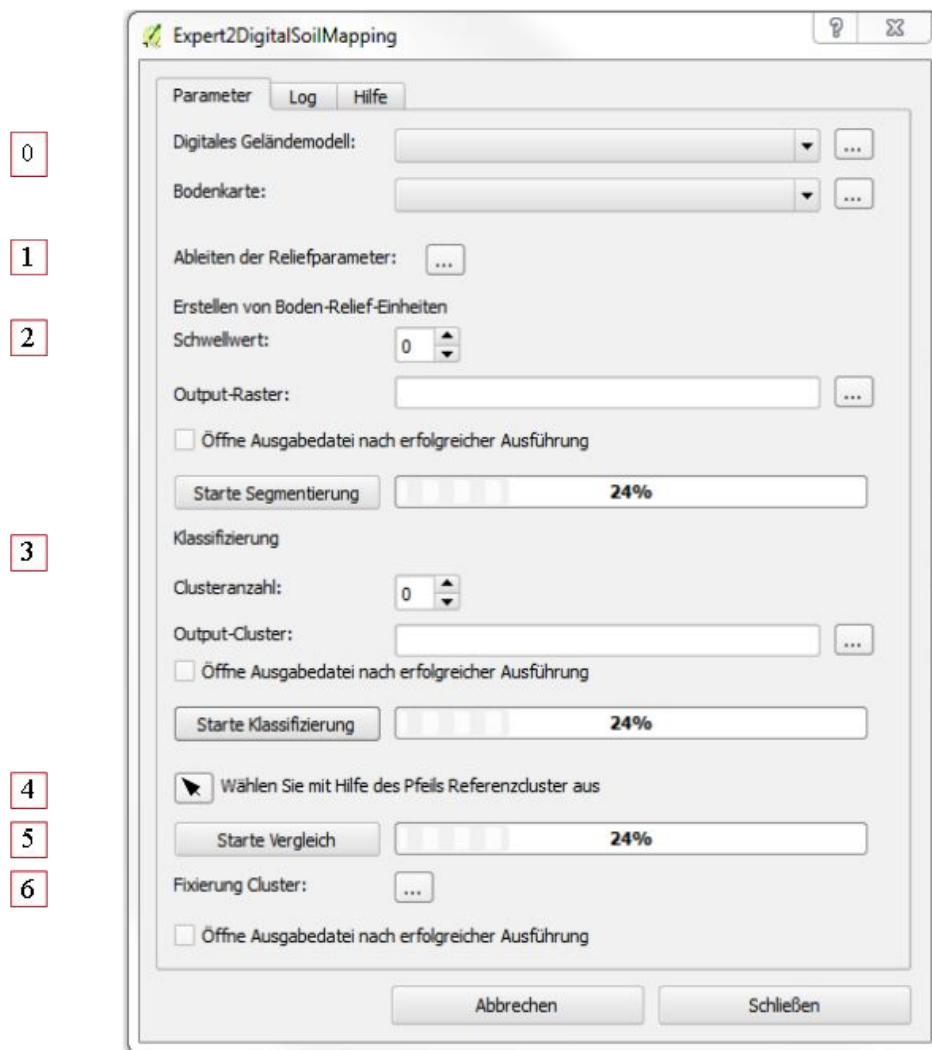


Abbildung 22: Design der Fachapplikation

Folgende Arbeitsschritte des Verfahrensablaufes werden hinterlegt:

- (0) Die Eingangsdaten werden geladen und die Bodenkarte wird vom Vektor- zum Rasterdatenformat umgewandelt.
- (1) Die Reliefparameter werden abgeleitet und in einer Gruppe i.group als Eingangsgruppe für die Segmentierung gespeichert. Eine Mehrfachauswahl ist hinterlegt.
- (2) Die Gruppe i.group wird mit dem GRASS-Algorithmus i.segment segmentiert. Die Grenzen der Bodenkarte werden als Abbruchgrenzen hinterlegt. Voreingestellt werden 10 Pixel als minimale Segmentgröße angenommen und

die gerasterte Bodenkarte als Abbruchgrenzen. Der Experte kann den Schwellenwert zwischen 0 und 1 festlegen.

- (3) Die Clusterbildung erfolgt mit dem k-mean-Algorithmus. Der Experte bestimmt die Anzahl der Zielcluster.
- (4) Der Experte legt die Referenzcluster, die er als plausibel bewertet, fest.
- (5) Der statistische Vergleich der Referenzcluster mit den übrigen Clustern erfolgt mit dem K-S-Test, welcher als R-Algorithmus hinterlegt ist.
- (6) Die plausiblen Cluster werden als Vektordaten gespeichert.

## 4. Diskussion der Ergebnisse

### *4.1 Quantitative Analyse / Identifizierung von Lageungenauigkeiten*

Bei der Analyse der abgeleiteten Parameter normalisierte Höhe, Neigung und Massenbilanzindex sind innerhalb der Grenzen der Bodengesellschaft Auenböden Pixelgruppen zu erkennen, die sich räumlich innerhalb der Bodengesellschaft befinden, deren Werte jedoch nicht in dem Bereich liegen, der die Bodengesellschaft Auenböden charakterisiert. Diese abweichenden Pixelgruppen sind in den Randbereichen der Bodengesellschaft Auenböden sowie im südwestlichen Teil der Bodengesellschaft Auenböden im UG lokalisiert. Sie stellen potenzielle Bereiche der Lageungenauigkeit der Bodengesellschaft Auenböden dar. Bei der Durchführung der Segmentierung der abgeleiteten Parameter wurde der Mittelweg zwischen nötiger Detailschärfe und zu hoher Detailliertheit (Salz-und-Pfeffer-Effekt) gesucht. Dazu wurden iterativ verschiedene Werte des Ähnlichkeitsmaßes sowie der minimalen Pixelanzahl pro Segment geprüft. Diese beiden Werte ließen sich für den konkreten Anwendungsfall ermitteln. Eine standardisierte Wertekombination konnte nicht ermittelt werden. Die Clusteranzahl wurde mit der Formel von MARDIA et al. (1979) bestimmt. Die so ermittelte Clusteranzahl führte für die Kernbereiche der Bodengesellschaft Auenböden mit allen angewendeten Systemen zu homogenen Ergebnissen, was als sehr günstig für die Weiterverarbeitung im Verfahrensablauf zu bewerten ist. Die Bereiche mit potenziellen Lageungenauigkeiten im westlichen Teil des UG sowie in den Randbereichen der Bodengesellschaft Auenböden wurden durch kleinteilige Cluster aufgedeckt. Ideal für die Weiterverarbeitung im Verfahrensablauf wäre, wenn diese Bereiche mit wenigen homogenen Clustern abgegrenzt werden würden und nicht mit kleinteiligen Clustern. Zur Clusterberechnung wurde die k-mean-Methode gewählt, da sie ein gängiges Verfahren der Clusterbildung in der Geomorphometrie ist (HENGL und REUTER 2009, MÖLLER 2012). Die k-mean-Methode setzt Normalverteilung der Daten voraus. Diese ist bei räumlichen Daten nicht immer gegeben. Dies belegen die Histogramme der berechneten Rasteroberflächen (MBI, NH, Neigung und Segmentierung). Es wäre zu prüfen, ob ein

hierarchisches Clusterverfahren, wie z.B. die Single-linkage, Average-linkage oder die WARD-Methode, geeigneter wäre.

#### ***4.2 Validierung der quantitativen Methode***

Aufgrund der Ergebnisse des Vergleiches von Referenzstellen und berechneten Clustern hat das mit R umgesetzte k-mean-Verfahren die besten Resultate in der Übereinstimmung von Referenzstellen und berechneten Clustern erzielt, gefolgt von der Orfeo Toolbox und dem mit SAGA umgesetzten k-mean-Verfahren. Da die gewählten Referenzstellen aufgrund ihres farblichen Unterschiedes der Pixel zu dem umgebenden Bereich der Bodengesellschaft Auenböden gewählt wurde und dies als Merkmal von Lageunterschieden definiert wurde, kann somit geschlussfolgert werden, dass der mit R umgesetzte k-mean-Algorithmus am geeignetsten für die Integration in den Verfahrensablauf ist.

#### ***4.3 Fachapplikation zur Integration des Fachwissen von Bodenexperten (qualitative Methode)***

Das Einbeziehen des Expertenwissens erfolgt durch die Umsetzung des gesamten Verfahrensablaufes in einer Plattform. Es kann festgestellt werden, dass die Kombination von qualitativen und quantitativen Methoden der Bodenkartierung in einem Verfahrensablauf möglich ist. Die Umsetzung dieses Verfahrensablaufes in einer GIS-Umgebung als Fachapplikation ist mittels der Programmierschnittstellen quelloffener Software ebenfalls möglich. Ein Wechsel zwischen Softwareprodukten ist somit nicht notwendig. Die Fachapplikation kann mit der QGIS-Erweiterung Pluginbuilder erzeugt werden. Über Python-Schnittstellen erfolgt die Kombination mit weiteren GIS (SAGA, GRASS) und R. Für die geomorphometrische Ableitung der Reliefparameter eignen sich die gleichnamigen SAGA-Algorithmen. Für die Segmentierung mittels pixelbasiertem Regionenwachstum erwies sich der GRASS-Algorithmus `i.segment` als geeignet. Der Algorithmus ist nicht in den QGIS-Verarbeitungswerkzeugen enthalten, kann aber aufgrund des frei verfügbaren Quellcodes von QGIS und GRASS in die QGIS-Verarbeitungswerkzeuge integriert

werden. Die Verarbeitung des Segmentierungsergebnisses ist aktuell durch ein GRASS-internes Exportproblem nicht reibungslos durchführbar. Dadurch können die Segmentierungsergebnisse nicht stabil weiterverarbeitet werden. Die Clusterbildung mit der k-mean-Methode wurde mit den Algorithmen der Systeme SAGA, Orfeo Toolbox und R durchgeführt. Dabei erwies sich die Umsetzung mit R am geeignetsten. Einerseits waren die Validierungsergebnisse im Vergleich zu den beiden weiteren überprüften Systemen die besten, andererseits kann das R-Skript durch eine QGIS-R-Schnittstelle in den QGIS-Verarbeitungswerkzeugen direkt angepasst werden. Somit ist eine höhere Flexibilität gegeben. Die Integration eines hierarchischen Clusterbildungsverfahrens in den Verfahrensablauf ist mit R ebenfalls umsetzbar. Inhaltlich konnten jedoch noch keine zuverlässigen Ergebnisse erzielt werden. Da der Verfahrensablauf komplexe Zusammenhänge beinhaltet, die eine Reihe von Parametrisierungen (Schwellenwerte, minimale Pixelzahl pro Segment, Anzahl der Iterationen, Zielclusteranzahl) benötigen. Die weitere Benutzerführung des Bodenexperten (Auswahl der Referenzcluster, Neubewertung, Fixierung der Cluster) lässt sich durch die Möglichkeiten des QGIS Pluginbuilders in Kombination mit dem Qt4-Designer sowie einer Entwicklungsumgebung (bspw. Eric4, Eclipse) umsetzen.

## 5. Schlussfolgerung und Ausblick

Qualitative und quantitative Methoden der Bodenkartierung können in einem Verfahrensablauf und in einer Fachapplikation kombiniert werden. Die Fachapplikation wird nach SHI et al. (2009) von den Nutzern/Bodenexperten akzeptiert, wenn sie einerseits aktiv in das System integriert werden und wenn andererseits Vertrauen in das System besteht. Die aktive Integration ist durch Kalibrierung der Parameter, Referenzclusterauswahl und Umklassifizierung gegeben. Das Vertrauen beruht auf Verständnis und Zuverlässigkeit des Systems. Das Verständnis ist durch die Komplexität der Fachapplikation nicht ohne weiteres gegeben. Dies kann durch eine sinnvolle Vorgabe von Richtwerten an den Stellen, an denen der Bodenexperte aufgefordert ist Parameter einzugeben, vermindert werden. Z.B. kann die mögliche Clusteranzahl algorithmisch berechnet und vorgegeben werden. Weiterhin ist eine gute Dokumentation zur Erläuterung der Algorithmen und Parameter notwendig. Da die Genauigkeit der Ergebnisse nicht sichergestellt werden konnte, ist die inhaltliche Zuverlässigkeit nicht gewährleistet. Die Integration verschiedener Algorithmen macht das System anfällig, insbesondere da der verwendete GRASS-Algorithmus nicht stabil einbezogen werden kann. Es wäre sinnvoll die Clusterberechnung und den Clustervergleich mit R oder direkt in Python zu programmieren, um Übersetzungsschritte, die bei der Einbindung fertiger Algorithmen entstehen, zu vermindern. Ein hierarchisches Clusterverfahren sollte geprüft werden. Neben dem Relief ist nach MÖLLER (2004) das Klima in Form von Starkniederschlägen ein wichtiger bodenbildender Faktor im UG. Für die Aktualisierung von Bodenaltdata, sollten daher zukünftig auch Niederschlagsdaten als Eingangsdaten Berücksichtigung finden und somit die Vielschichtigkeit der Boden-Umwelt-Beziehung abbilden.



## Literaturverzeichnis

- Abdelfattah, M.A., Pain, C., 2012. Unifying regional soil maps at different scales to generate a national soil map for the United Arab Emirates applying digital soil mapping techniques. *Journal of Maps* 8, 392-405.
- AG Boden, 2005. *Bodenkundliche Kartieranleitung*. 5. Aufl. E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart.
- Altermann, M., 1970. Böden, in: Kunert, R. (Hrsg.), *Erläuterungen Zur Geologischen Spezialkarte 1:25.000, Blatt Könnern*. Zentrales Geologisches Institut, Berlin.
- Baatz, M., Schäpe, A., 2000. Multiresolution Segmentation: An Optimization Approach for High Quality Multi-scale Image Segmentation, in: *Angewandte Geographische Informationsverarbeitung XII, Beiträge Zum AGIT-Symposium Salzburg, Heidelberg*.
- Bagheri Bodaghabadi, M., Salehi, M.H., Martínez-Casasnovas, J.A., Mohammadi, J., Toomanian, N., Esfandiarpour Borujeni, I., 2011. Using Canonical Correspondence Analysis (CCA) to identify the most important DEM attributes for digital soil mapping applications. *CATENA* 86, 66–74.
- Ballabio, C., Fava, F., Rosenmund, A., 2012. A plant ecology approach to digital soil mapping, improving the prediction of soil organic carbon content in alpine grasslands. *Geoderma* 187-188, 102–116.
- Beaudette, D.E., O'Geen, A.T., 2009. Soil-Web: An online soil survey for California, Arizona, and Nevada. *Computers & Geosciences* 35, 2119–2128.
- Behrens, T., Zhu, A.-X., Schmidt, K., Scholten, T., 2010. Multi-scale digital terrain analysis and feature selection for digital soil mapping. *Geoderma* 155, 175–185.
- Bishop, T.F.A., Minasny, B., Mcbratney, A.B., 2006. Uncertainty analysis for soil-terrain models. *International Journal of Geographical Information Science* 20, 117-134.
- Bivand, R.S., Pebesma, E., Gómez-Rubio, V., 2013. *Applied Spatial Data Analysis with R*, 2. Aufl. Springer, New York.
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 2–16.
- Bock, M., Rossner, G., Wissen, M., Remm, K., Langanke, T., Lang, S., Klug, H., Blaschke, T., Vrščaj, B., 2005. Spatial indicators for nature conservation from European to local scale. *Ecological Indicators* 5, 322–338.
- Böhner, J., Selige, Ringeler, 2006. Image Segmentation Using Representativeness Analysis and Region Growing. *Göttinger Geographische Abhandlungen* 115, 130.

- Brandhuber, R., Demmel, M., Koch, H.-J., Brunotte, J., 2008. DLG-Merkblatt 344: Bodenschonender Einsatz von Landmaschinen. Empfehlungen für die Praxis. Deutsche Landwirtschafts-Gesellschaft e.V.
- Bui, E.N., Henderson, B.L., Viergever, K., 2006. Knowledge discovery from models of soil properties developed through data mining. *Ecological Modelling* 191, 431–446.
- Bui, E.N., 2004. Soil survey as a knowledge system. *Geoderma* 120, 17–26.
- Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., 2013. A methodology for digital soil mapping in poorly-accessible areas. *Geoderma* 192, 341–353.
- Carré, F., McBratney, A.B., Mayr, T., Montanarella, L., 2007a. Digital soil assessments: Beyond DSM. *Geoderma* 142, 69–79.
- Carré, F., McBratney, A.B., Minasny, B., 2007b. Estimation and potential improvement of the quality of legacy soil samples for digital soil mapping. *Geoderma* 141, 1–14.
- Cavazzi, S., Corstanje, R., Mayr, T., Hannam, J., Fealy, R., 2013. Are fine resolution digital elevation models always the best choice in digital soil mapping? *Geoderma* 195-196, 111–121.
- Council of the EU, 2014. Soil protection in the Union - Background paper.
- Deumlich, D., Schmidt, R., Sommer, M., 2010. A multiscale soil-landform relationship in the glacial-drift area based on digital terrain analysis and soil attributes. *Journal of Plant Nutrition and Soil Science* 173.
- Deutscher Bundestag, 17. Wahlperiode, 2012. Drucksache 17/8478 - Aktivitäten der Bundesregierung zum konsequenten Schutz des Bodens auf nationaler, europäischer und internationaler Ebene.
- Dobos, E., Hengl, T., 2009. Soil Mapping Applications, in: *Geomorphometry: Concepts, Software, Applications*. Elsevier.
- Fortin, M.-J., Olson, R.J., Ferson, S., Iverson, L., Hunsaker, C., Levine, G., Butera, D., Klemas, K., Edwards, V., 2000. Issues related to the detection of boundaries. *Landscape Ecology* 15, 453–466.
- Grimm, R., Behrens, T., 2010. Uncertainty analysis of sample locations within digital soil mapping approaches. *Geoderma* 155, 154–163.
- Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma* 152, 195–207.
- Hansen, M.K., Brown, D.J., Dennison, P.E., Graves, S.A., Bricklemeyer, R.S., 2009. Inductively mapping expert-derived soil-landscape units within dambo wetland catenae using multispectral and topographic data. *Geoderma* 150, 72–84.

- Häring, T., Dietz, E., Osenstetter, S., Koschitzki, T., Schröder, B., 2012. Spatial disaggregation of complex soil map units: A decision-tree based approach in Bavarian forest soils. *Geoderma* 185-186, 37–47.
- Hartmann, K.J., 2005. Bereitstellung von Informationen der bodenkundlichen Landesaufnahme zur Bewertung von Bodenfunktionen, in: Möller, M., Helbig, H. (Hrsg.), GIS-Gestützte Bewertung von Bodenfunktionen - Datengrundlagen Und Lösungsansätze. Wichmann, Heidelberg, 27 – 34.
- Hastie, T., Tibshirani, R., Friedman, J., 2011. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2. Aufl. Springer, New York.
- Hengl, T., Bajat, B., Blagojević, D., Reuter, H.I., 2008. Geostatistical modeling of topography using auxiliary maps. *Computers & Geosciences* 34, 1886–1899.
- Hengl, T., Reuter, H.I., 2009. *Geomorphometry: Concepts, Software, Applications*. Elsevier.
- Iwahashi, J., Pike, R.J., 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology* 86, 409–440.
- Jenny, Hans, 1942. Geographical Publications. *Journal of Geography* 41.
- Kainz, W., Schröder, H., Knauf, C., Möbes, A., 1997. *Bodenkarte Halle und Umgebung: 1:50000*. Landesamt für Geologie und Bergwesen Sachsen-Anhalt, Halle/Saale.
- Kempen, B., Brus, D.J., Heuvelink, G.B.M., Stoorvogel, J.J., 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma* 151, 311–326.
- Kempen, B., Brus, D.J., Stoorvogel, J.J., Heuvelink, G.B.M., de Vries, F., 2012. Efficiency Comparison of Conventional and Digital Soil Mapping for Updating Soil Maps. *Soil Science Society of America Journal* 76, 2097.
- Kerry, R., Goovaerts, P., Rawlins, B.G., Marchant, B.P., 2012. Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale. *Geoderma* 170, 347–358.
- Krasilnikov, P., 2014. Book review. *Geoderma* 213, 131–132.
- Kunert, R. (Hrsg.), 1970. *Erläuterungen zur Geologischen Spezialkarte 1:25.000, Blatt Könnern*. Zentrales Geologisches Institut, Berlin.
- Lacoste, M., Minasny, B., McBratney, A., Michot, D., Viaud, V., Walter, C., 2014. High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. *Geoderma* 213, 296–311.
- Lagacherie, P., Baret, F., Feret, J.-B., Madeira Netto, J., Robbez-Masson, J.M., 2008. Estimation of soil clay and calcium carbonate using laboratory, field and

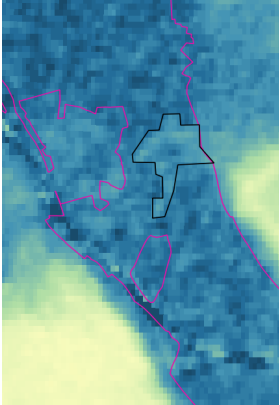
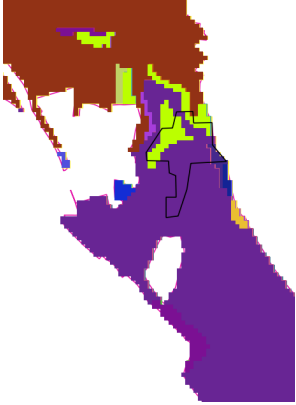
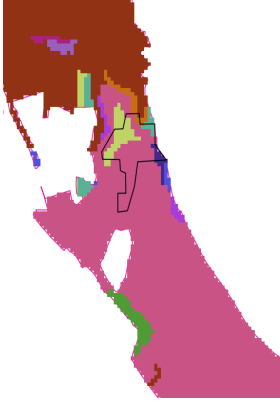
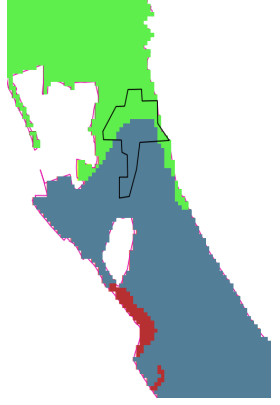
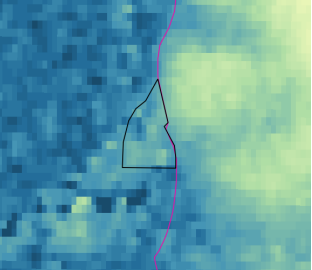
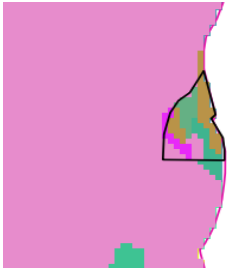
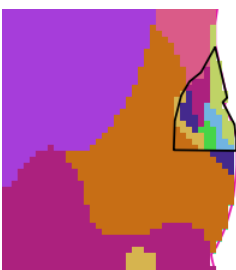
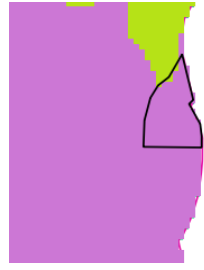
- airborne hyperspectral measurements. *Remote Sensing of Environment* 112, 825–835.
- Lagacherie, P., McBratney, A.B., Voltz, M. (Hrsg.), 2007. *Digital Soil Mapping: An Introductory Perspective*. Elsevier.
- Lemercier, B., Lacoste, M., Loum, M., Walter, C., 2012. Extrapolation at regional scale of local soil knowledge using boosted classification trees: A two-step approach. *Geoderma* 171-172, 75–84.
- Malone, B.P., McBratney, A.B., Minasny, B., 2011. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. *Geoderma* 160, 614–626.
- Malone, B.P., McBratney, A.B., Minasny, B., Laslett, G.M., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma* 154, 138–152.
- Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. *Multivariate analysis. Probability and mathematical statistics*. London u.a.
- McBratney, A.B., Mendonca Santos, M.L., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117, 3–52.
- Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I., 2013. Digital Mapping of Soil Carbon. *Advances in Agronomy* 118, 1-47.
- Möller, M., 2004. Disaggregation von Bodeninformationen auf der Grundlage digitaler Reliefdaten, in: Möller, M., Helbig, H. (Hrsg.), *GIS-gestützte Bewertung von Bodenfunktionen: Datengrundlagen und Lösungsansätze*. Wichmann, Herbert, Verlag in Verlagsgruppe Hüthig Jehle Rehm GmbH, Heidelberg.
- Möller, M., Koschitzki, T., Hartmann, K.J., Jahn, R., 2012. Plausibility test of conceptual soil maps using relief parameters. *Catena* 88, 57-67.
- Möller, M., Volk, M., Friedrich, K., Lymburner, L., 2008. Placing soil-genesis and transport processes into a landscape context: A multiscale terrain-analysis approach. *Journal of Plant Nutrition and Soil Science* 171, 419–430.
- Möller, M., Volk, M., Volk, M., Volk, M., 2015. Effective map scales for soil transport processes and related process domains – Statistical and spatial characterization of their scale-specific inaccuracies. *Geoderma* 247–248, 151–160.
- Mora-Vallejo, A., Claessens, L., Stoorvogel, J., Heuvelink, G.B.M., 2008. Small scale digital soil mapping in Southeastern Kenya. *Catena* 76, 44–53.
- Moran, C.J., Bui, E.N., 2002. Spatial data mining for enhanced soil map modelling. *International Journal of Geographical Information Science* 16, 533–549.
- Morvan, X., Saby, N.P.A., Arrouays, D., Le Bas, C., Jones, R.J.A., Verheijen, F.G.A., Bellamy, P.H., Stephens, M., Kibblewhite, M.G., 2008. Soil monitoring in



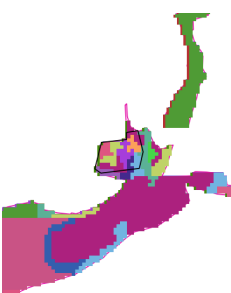
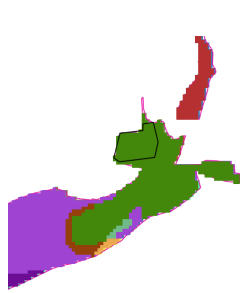
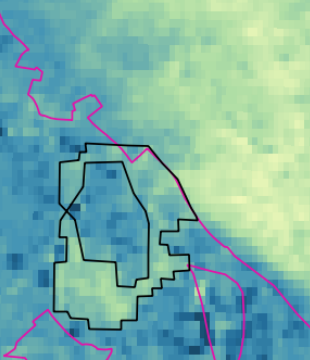
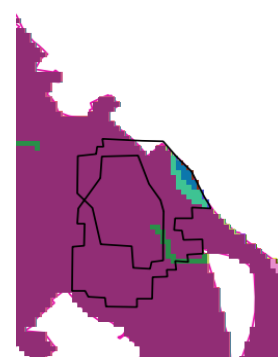
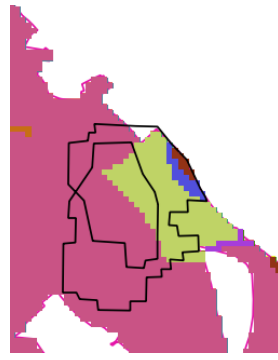
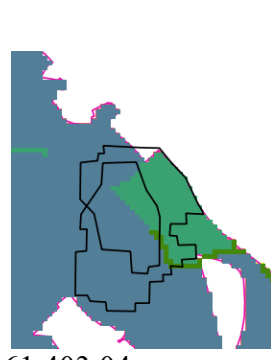
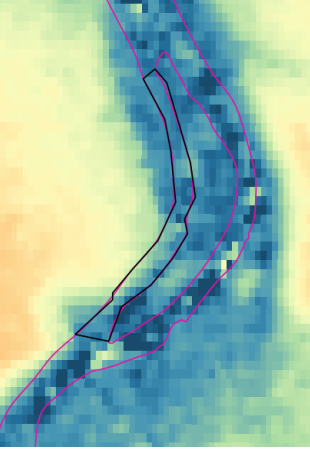
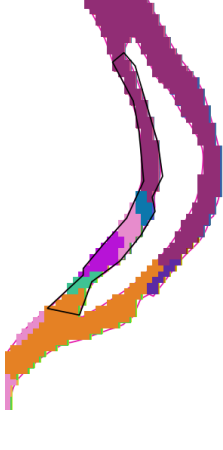
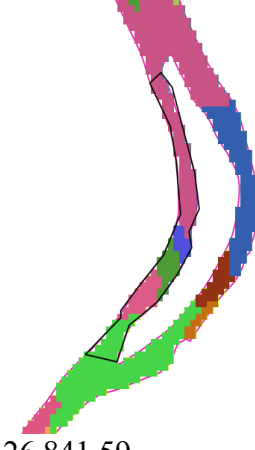
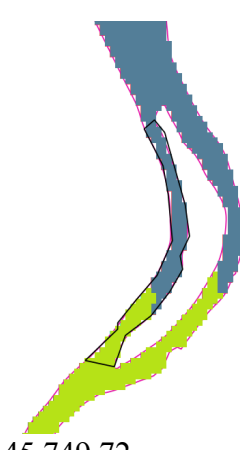
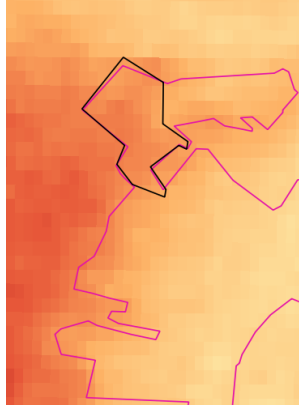
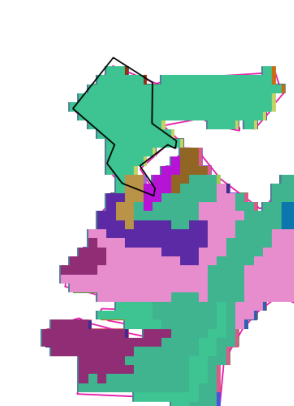
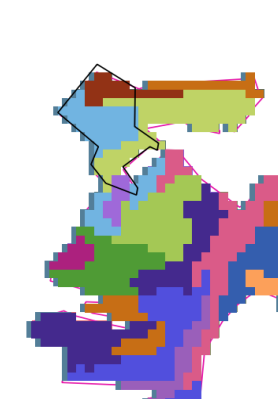
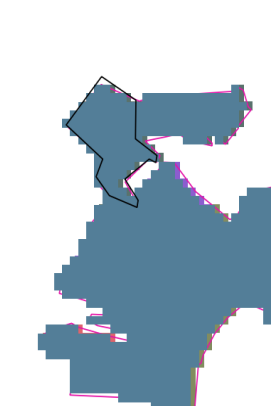
- Europe: A review of existing systems and requirements for harmonisation. *Science of The Total Environment* 391, 1–12.
- Mulder, V.L., de Bruin, S., Schaepman, M.E., Mayr, T.R., 2011. The use of remote sensing in soil and terrain mapping — A review. *Geoderma* 162, 1–19.
- Nauman, T.W., Thompson, J.A., 2014. Semi-automated disaggregation of conventional soil maps using knowledge driven data mining and classification trees. *Geoderma* 213, 385–399.
- Neteler, M., Bowman, M.H., Landa, M., Metz, M., 2012. GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling and Software* 31, 124–130.
- Ogders, N.P., McBratney, A.B., Minasny, B., 2011. Bottom-up digital soil mapping. II. Soil series classes. *Geoderma* 163.
- Olaya, V., Conrad, O., 2009. Geomorphometry in SAGA, in: Hengl, T., Reuter, H.I. (Hrsg.), *Geomorphometry: Concepts, Software, Applications*. Elsevier.
- Poggio, L., Gimona, A., Brewer, M.J., 2013. Regional scale mapping of soil properties and their uncertainty with a large number of satellite-derived covariates. *Geoderma* 209-210, 1–14.
- Schmidt, K., Behrens, T., Scholten, T., 2008. Instance selection and classification tree analysis for large spatial datasets in digital soil mapping. *Geoderma* 146, 138–146.
- Schmidt, R., Diemann, R., 1981. *Erläuterungen zur Mittelmaßstäbigen Landwirtschaftlichen Standortkartierung*. Akademie der Landwirtschaftswissenschaften der DDR, Institut für Bodenkunde, Eberswalde-Finow.
- Scull, P., Franklin, J., Chadwick, O.A., McArthur, D., 2003. Predictive soil mapping: A review. *Progress in Physical Geography* 27, 171-197.
- Shi, X., Girod, L., Long, R., DeKett, R., Philippe, J., Burke, T., 2012. A comparison of LiDAR-based DEMs and USGS-sourced DEMs in terrain analysis for knowledge-based digital soil mapping. *Geoderma* 170, 217–226.
- Shi, X., Long, R., Dekett, R., Philippe, J., 2009. Integrating Different Types of Knowledge for Digital Soil Mapping. *Soil Science Society of America Journal* 73, 1682.
- Smith, M.P., Zhu, A.-X., Burt, J.E., Stiles, C., 2006. The effects of DEM resolution and neighborhood size on digital soil survey. *Geoderma* 137, 58–69.
- Thas, O., 2010. *Comparing Distributions*, Springer Series in Statistics. Springer, New York.
- Tinel, C., Fontannaz, D., de Boissezon, H., Grizonnet, M., Michel, J., 2012. The ORFEO accompaniment program and ORFEO ToolBox, in: *Geoscience and*

Remote Sensing Symposium (IGARSS), 2012 IEEE International. Presented at the Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, 7102–7105.

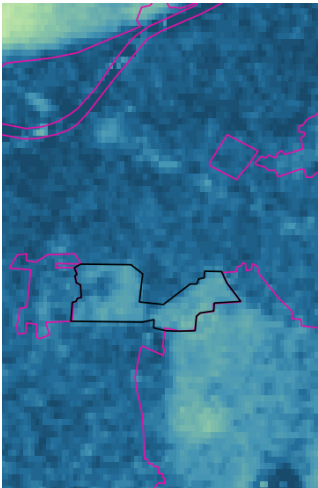
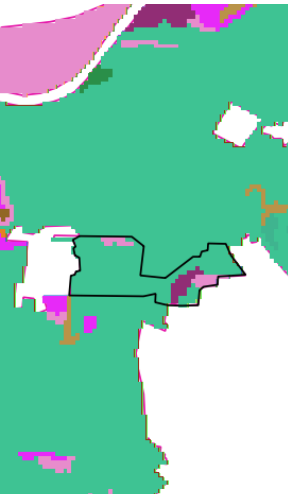

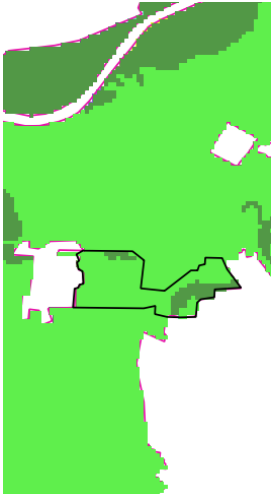
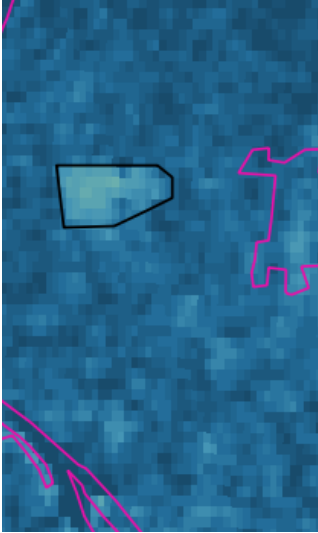


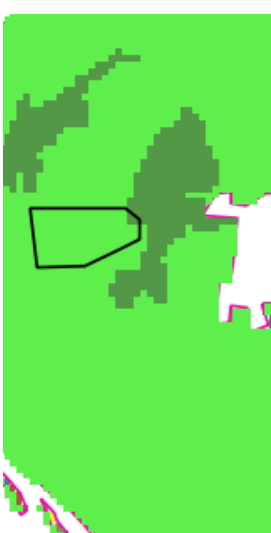
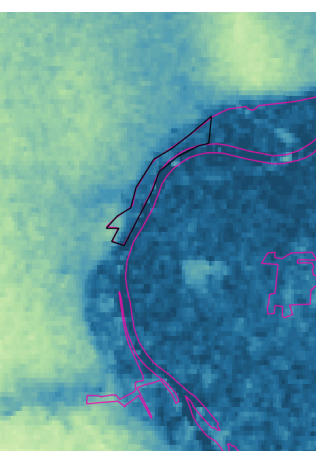
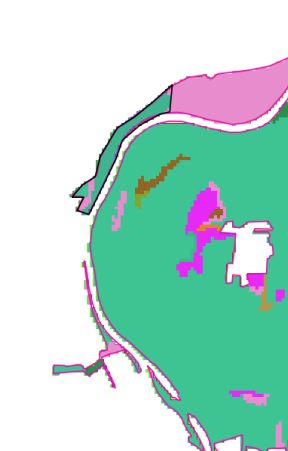


- Trefflich, A., 1997. Die Region Leipzig-Halle-Bitterfeld: Klimatologische Kennzeichnung des Gebietes, in: Feldmann, R., Henle, K., Auge, H., Flachowsky, J., Klotz, S., Krönert, R. (Hrsg.), *Regeneration und Nachhaltige Landnutzung - Konzepte für belastete Regionen*. Borntraeger, Berlin, Stuttgart.
- Van Zijl, G.M., Le Roux, P.A., Turner, D.P., 2013. Disaggregation of land types using terrain analysis, expert knowledge and GIS methods. *South African Journal of Plant and Soil* 30, 123–129.
- Vereinte Nationen, 2013. Resolution der 68. Generalversammlung der Vereinten Nationen - 68/232. Weltbodentag und Internationales Jahr der Böden.
- Wendling, U., 2001. Das Klima der Stadt Halle (Saale) nach den Wetterbeobachtungen 1901 bis 2000. *Klimastatusbericht 2001*. Deutscher Wetterdienst, Offenbach.
- Wagner, M., 2003. SRTM. DTED Format. Version 1.1. Deutsches Zentrum für Luft- und Raumfahrt, Köln.
- Zhu, A.-X., Lark, M., Minasny, B., Huang, Y., 2012. Entering the digital world (Pedometrics 2009). *Geoderma* 171-172, 1–2.

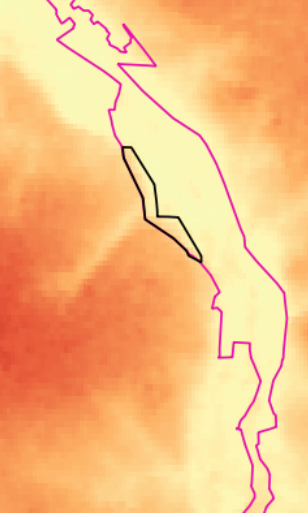


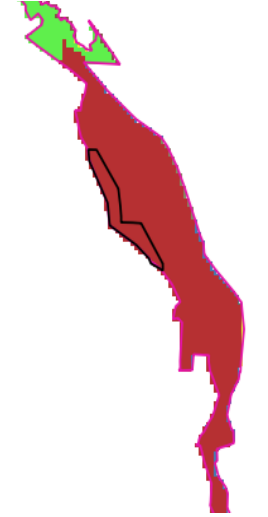
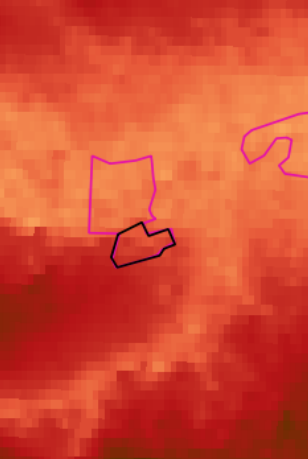
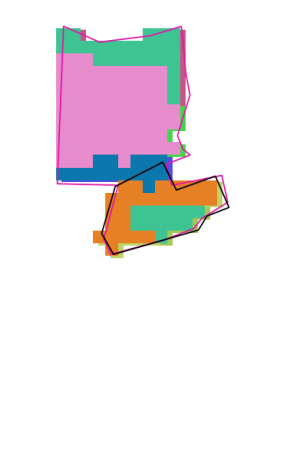
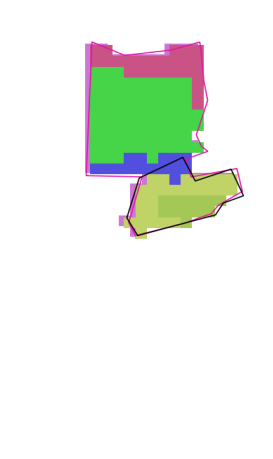
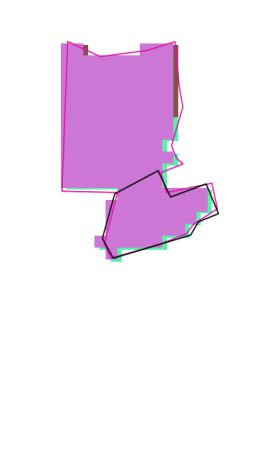
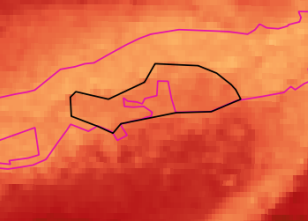
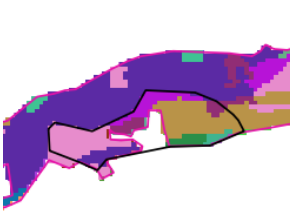
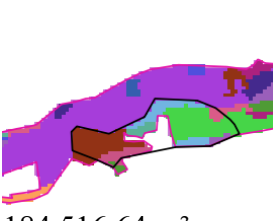
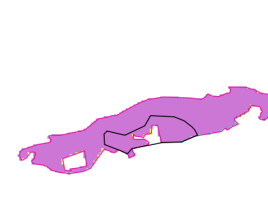
## Anhang: Visueller Vergleich der Referenzstellen (schwarz begrenzt) zwischen DHM und algorithmisch berechneten Clustern

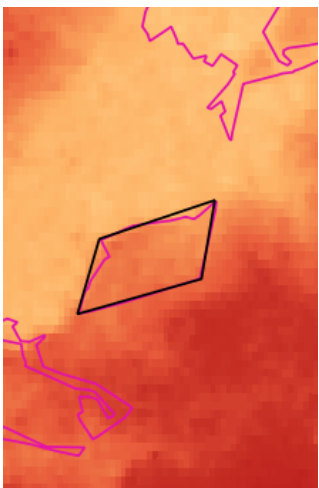
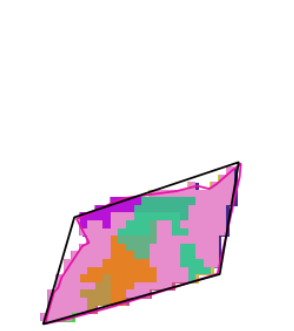
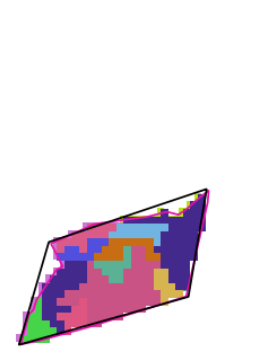
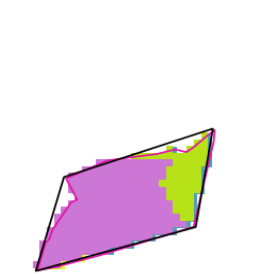
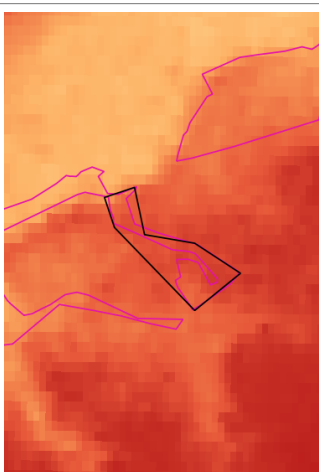
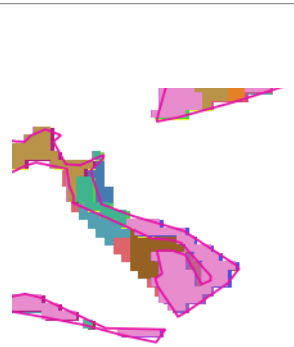

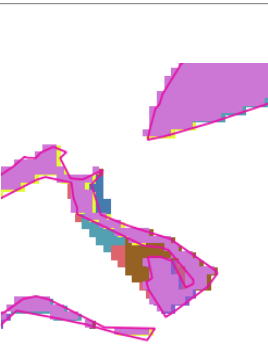
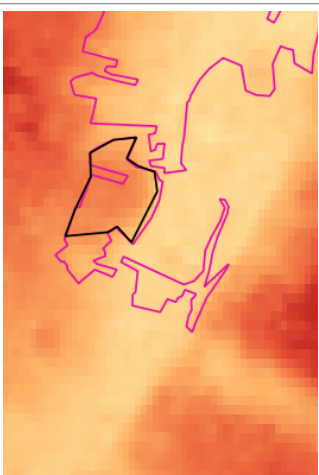
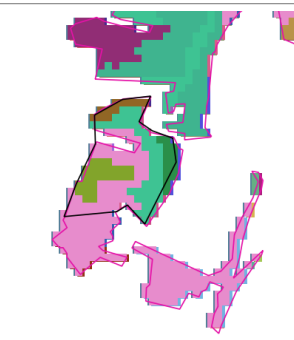
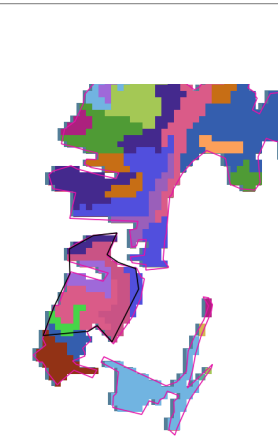
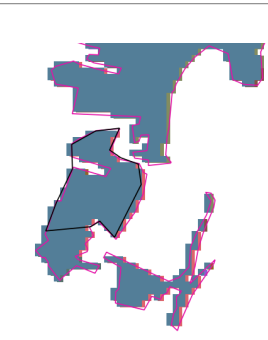
ID	visuell bestimmte Referenzstellen:  DHM  Bodengesellschaft Auenböden (rosa begrenzt)  Referenzstellen (schwarz begrenzt)  Größe der Referenzstelle in m <sup>2</sup>	DHM: Höhe in m <ul style="list-style-type: none"> <li>■ 105</li> <li>■ 110</li> <li>■ 115</li> <li>■ 120</li> <li>■ 125</li> <li>■ 130</li> <li>■ 135</li> <li>■ 140</li> <li>■ 145</li> <li>■ 150</li> <li>■ 160</li> <li>■ 170</li> <li>■ 180</li> <li>■ 190</li> <li>■ 200</li> <li>■ 210</li> <li>■ 220</li> <li>■ 230</li> <li>■ 240</li> <li>■ 250</li> <li>■ 270</li> <li>■ 300</li> <li>■ 350</li> </ul>	Räumliche Übereinstimmung zwischen den Referenzstellen (schwarz begrenzt) und an dieser Stelle separat berechneten Clustern (absolut in m <sup>2</sup> und relativ in %) der Bodengesellschaft Auenböden  berechnet mit den Algorithmen:		
			OTB	R	SAGA
1	 <p>81.008,34 m<sup>2</sup></p>	 <p>26.568,83 m<sup>2</sup> 32,80 %</p>	 <p>25.679,47 m<sup>2</sup> 31,70 %</p>	 <p>39.264,75 m<sup>2</sup> 48,47 %</p>	
2	 <p>45.218,09 m<sup>2</sup></p>	 <p>37.629,53 m<sup>2</sup> 83,22 %</p>	 <p>40.047,63 m<sup>2</sup> 88,57 %</p>	 <p>1.245,79 m<sup>2</sup> 2,76 %</p>	

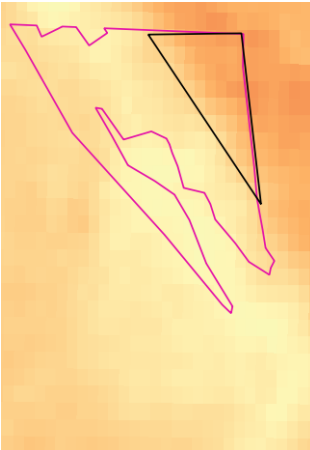

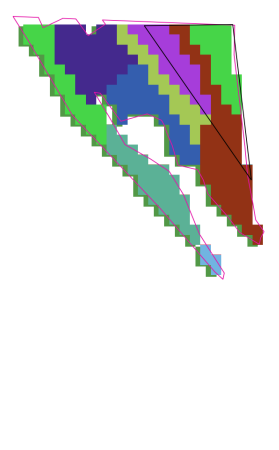
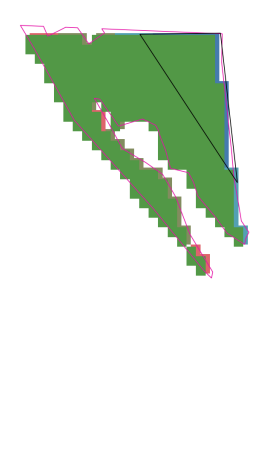

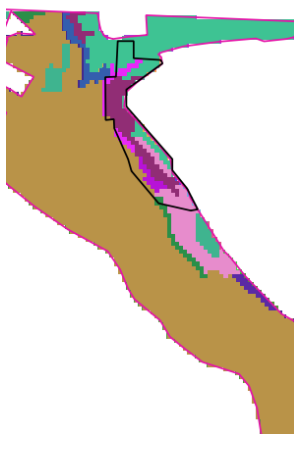
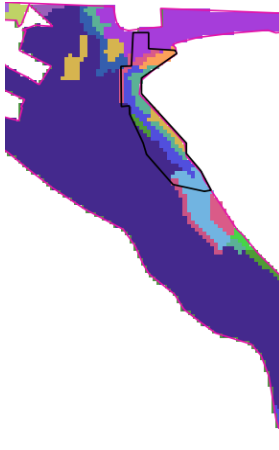
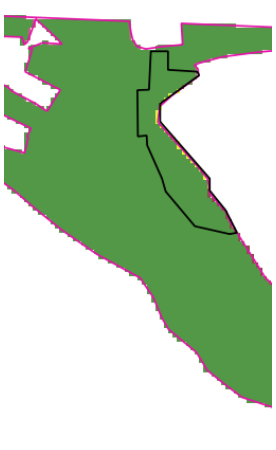
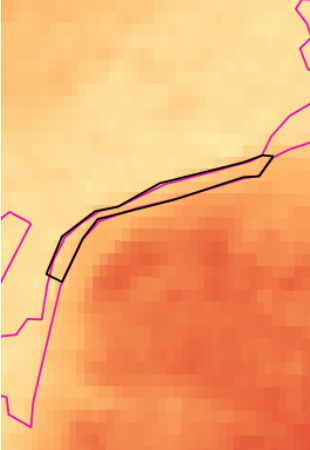
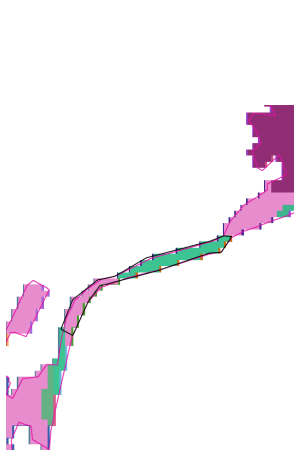
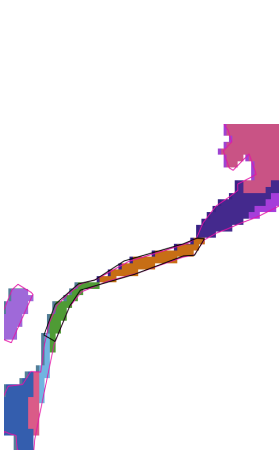
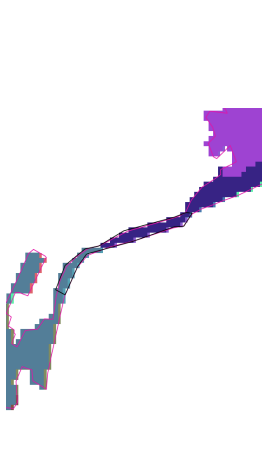
3	 <p>47.545,787 m<sup>2</sup></p>	 <p>47.545,787 m<sup>2</sup> 100 %</p>	 <p>47.545,787 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
4	 <p>350.507,94 m<sup>2</sup></p>	 <p>17.989,43 5,13 %</p>	 <p>61.403,04 1,75 %</p>	 <p>61.403,04 1,75 %</p>
5	 <p>65.470,15 m<sup>2</sup></p>	 <p>35.069,93 53,57 %</p>	 <p>26.841,59 41,00 %</p>	 <p>45.749,72 69,88 %</p>
6	 <p>31.728,06 m<sup>2</sup></p>	 <p>2.297,67 m<sup>2</sup> 7,24 %</p>	 <p>31.728,06 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>

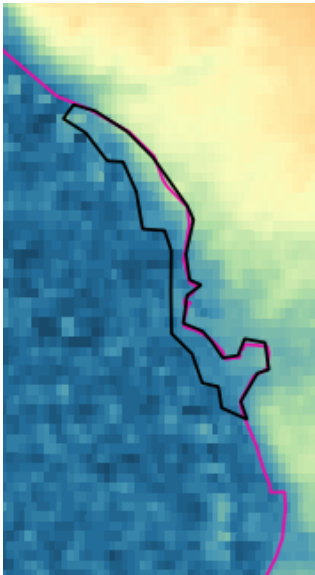
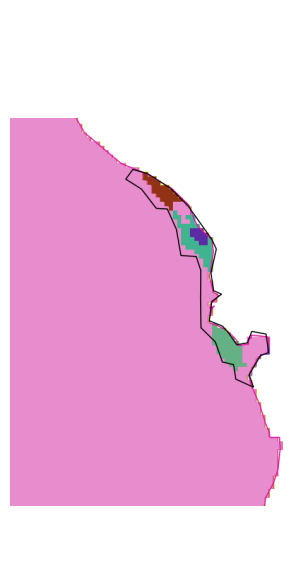
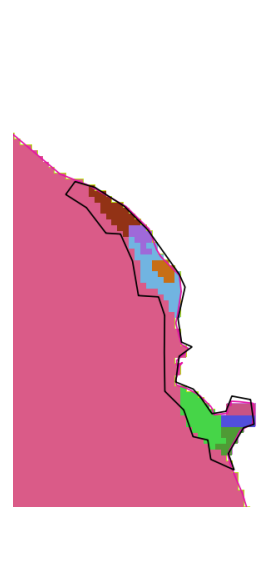
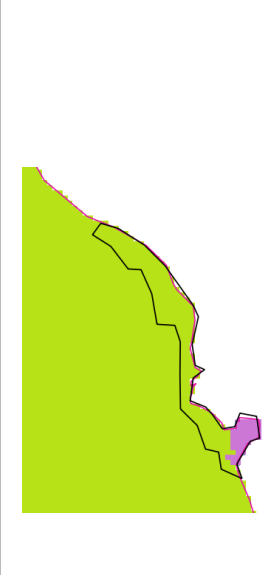
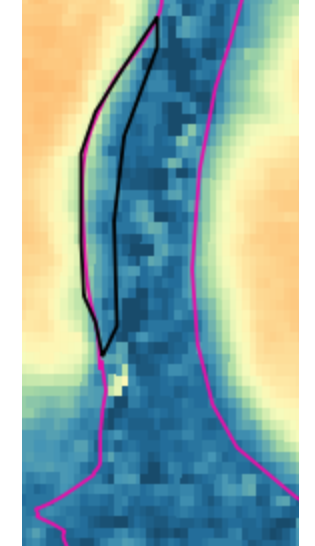
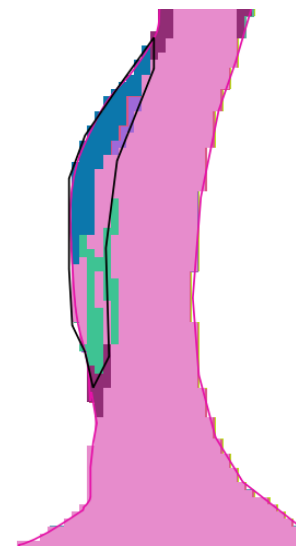

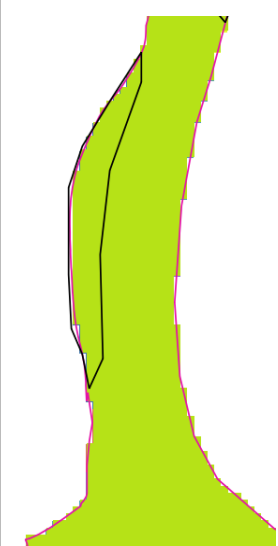
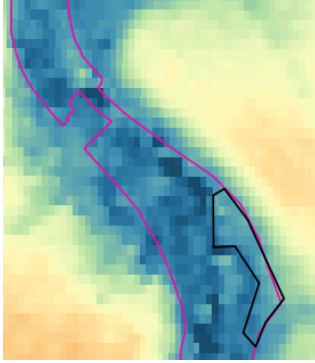
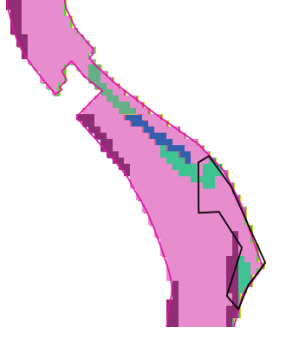
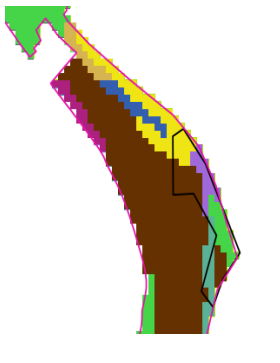
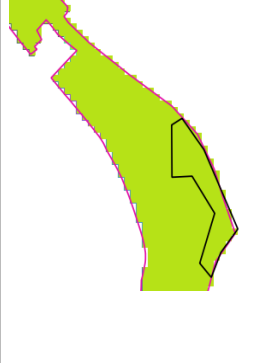


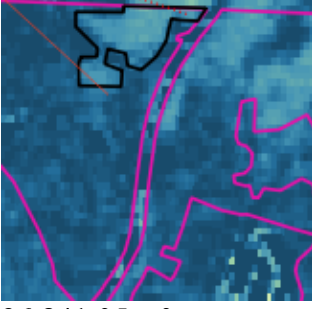
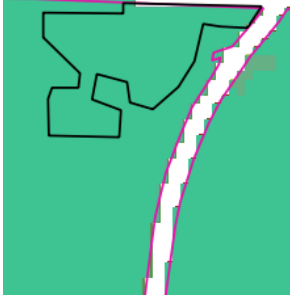
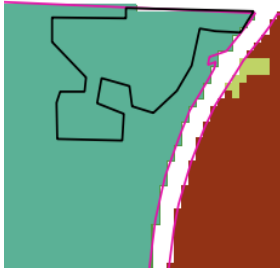
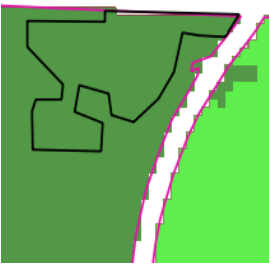
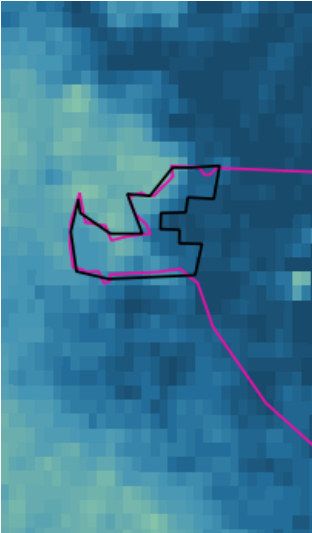
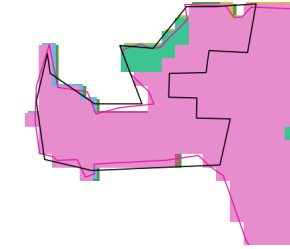
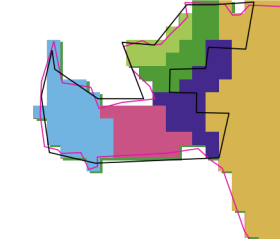
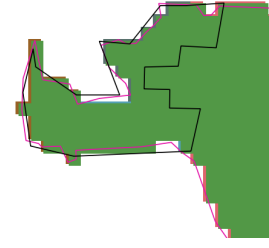
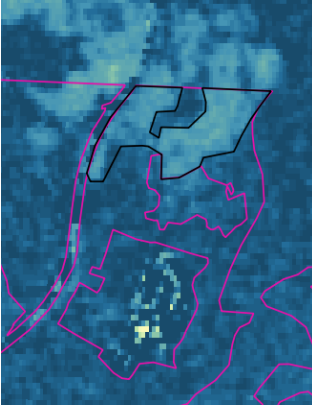
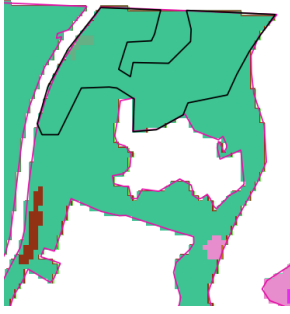
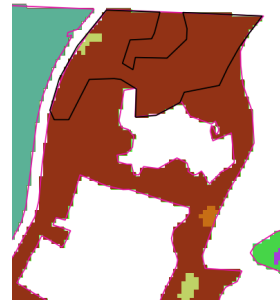
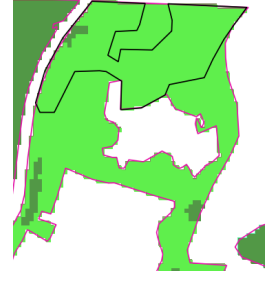
7	 <p>167.214,98 m<sup>2</sup></p>	 <p>23.437,44 14,02 %</p>	 <p>35.184,31 21,04 %</p>	 <p>28.140,16 % 16,83 %</p>
8	 <p>50.429,29 m<sup>2</sup></p>	 <p>1.506,23 2,97 %</p>	 <p>1.506,23 2,97 %</p>	 <p>1.506,23 2,97 %</p>
9	 <p>92.526,61 m<sup>2</sup></p>	 <p>77.550,79 83,81 %</p>	 <p>92.526,61 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>

10	 <p>46.201,50 m<sup>2</sup></p>	 <p>46.201,50 m<sup>2</sup> 100 %</p>	 <p>46.201,50 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
11	 <p>17.245,07 m<sup>2</sup></p>	 <p>17.245,07 m<sup>2</sup> 100 %</p>	 <p>17.245,07 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
12	 <p>184.516,64 m<sup>2</sup></p>	 <p>184.516,64 m<sup>2</sup> 100 %</p>	 <p>184.516,64 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>

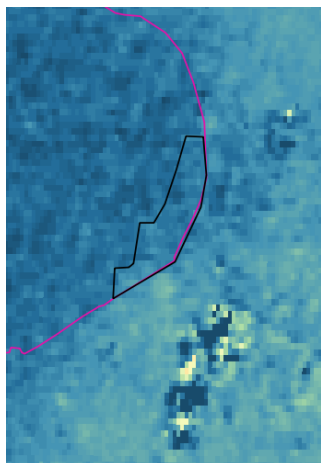
<p>13</p>  <p>116.493,10 m<sup>2</sup></p>	 <p>116.493,10 m<sup>2</sup> 100 %</p>	 <p>116.493,10 m<sup>2</sup> 100 %</p>	 <p>21.112,07 m<sup>2</sup> 18,12 %</p>
<p>14</p>  <p>52.459,48 m<sup>2</sup></p>	 <p>22.794,47 m<sup>2</sup> 43,45 %</p>	 <p>52.459,48 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
<p>15</p>  <p>61.851,52 m<sup>2</sup></p>	 <p>34.980,28 m<sup>2</sup> 56,55 %</p>	 <p>61.851,52 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>

16	 <p>26.505,31 m<sup>2</sup></p>	 <p>9.123,07 m<sup>2</sup> 34,42 %</p>	 <p>15.727,00 m<sup>2</sup> 59,33 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
17	 <p>132.656,47 m<sup>2</sup></p>	 <p>85.512,23 m<sup>2</sup> 64,46 %</p>	 <p>100.913,71 m<sup>2</sup> 76,07 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
18	 <p>25.408,03 m<sup>2</sup></p>	 <p>14.288,64 m<sup>2</sup> 56,24 %</p>	 <p>25.408,03 m<sup>2</sup> 100 %</p>	 <p>16.593,56 m<sup>2</sup> 65,30 %</p>

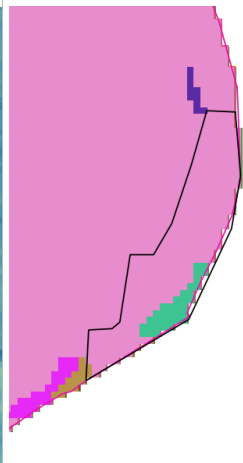
19	 <p>109.736,27 m<sup>2</sup></p>	 <p>50.033,57 m<sup>2</sup> 45,59 %</p>	 <p>70.299,37 m<sup>2</sup> 64,06 %</p>	 <p>13.191,73 m<sup>2</sup> 12,02 %</p>
20	 <p>79.350,45 m<sup>2</sup></p>	 <p>57.901,59 m<sup>2</sup> 72,97 %</p>	 <p>79.350,45 m<sup>2</sup> 100 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
21	 <p>37.615,92 m<sup>2</sup></p>	 <p>12.157,52 m<sup>2</sup> 32,32 %</p>	 <p>24.053,35 m<sup>2</sup> 63,94 %</p>	 <p>0 m<sup>2</sup> 0 %</p>

22	 <p>96.841,05 m<sup>2</sup></p>	 <p>0 m<sup>2</sup> 0 %</p>	 <p>0 m<sup>2</sup> 0 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
23	 <p>41.219,48 m<sup>2</sup></p>	 <p>3.080,39 m<sup>2</sup> 7,47 %</p>	 <p>37.666,11 m<sup>2</sup> 91,38 %</p>	 <p>0 m<sup>2</sup> 0 %</p>
24	 <p>270.275,28 m<sup>2</sup></p>	 <p>4.893,07 m<sup>2</sup> 1,81 %</p>	 <p>4.893,07 m<sup>2</sup> 1,81 %</p>	 <p>4.893,07 m<sup>2</sup> 1,81 %</p>

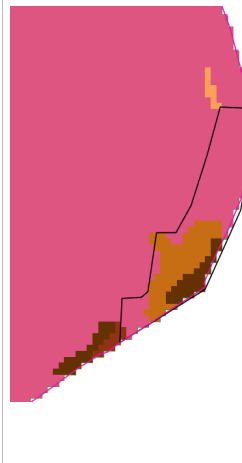
25



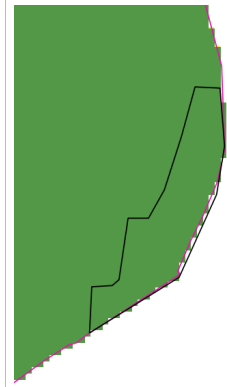
132.800,94 m<sup>2</sup>



16.801,10 m<sup>2</sup>  
12,65 %



56.553,33 m<sup>2</sup>  
42,58 %



0 m<sup>2</sup>  
0 %