

Datenvergleich und Integrationsmöglichkeiten von Landschaftsmodellen

Hubert Fünfer, Andreas Illert, Thomas Michaelis, Markus Müller, Kerstin Reinhold, Sonja Werhahn und Klaus-Peter Wodtke

Zusammenfassung

Aus einer Expertenrunde unter Leitung des Bundesministeriums des Innern entstand der Anstoß, die Potenziale zu analysieren, die sich aus dem Vergleich von Datensätzen mit räumlicher Überdeckung ergeben. Nutzer der Software von AED-SICAD haben sich zusammengefunden, um Verfahren zum Vergleich von Datensätzen an konkreten Anwendungsfällen zu testen. In drei Teilprojekten wurden unterschiedliche Aufgabenstellungen untersucht: die Erkennung von Unterschieden zwischen Daten des Basis-DLM zu zwei unterschiedlichen Zeitständen, der Abgleich von Daten des Landbedeckungsmodells für Deutschland (LBM-DE) mit dem Basis-DLM und die Zusammenführung von Daten der Open StreetMap (OSM) mit dem Basis-DLM.

Summary

A panel of experts headed by the Federal Ministry of the Interior provided the impetus to ascertain what could be learned from comparing different datasets covering the same areas. Users of AED-SICAD software collaborated to establish the best methodologies for comparing datasets based on concrete use cases. Investigations covered three different projects: the recognition of differences between base DLM data on two different dates, the comparison of the land cover model for Germany (LBM-DE) with the base DLM data and the merging of data from the OpenStreetMap (OSM) with the base DLM.

Schlüsselwörter: Vergleich von Datensätzen, räumliche Überdeckung, Basis-DLM, LBM-DE, OpenStreetMap (OSM)

1 Einleitung

Mit der fortschreitenden Digitalisierung nimmt die Sammlung von raumbezogenen Daten stetig zu. Die dabei entstehenden Datensätze sind in der Regel für spezifische Anwendungsfälle konzipiert und unterscheiden sich oft sowohl in ihrer inhaltlichen Ausprägung und räumlichen Ausdehnung als auch in weiteren qualitativen Eigenschaften wie Aktualität oder Präzision. Ein Ausschnitt der Erdoberfläche wird so von mehreren sich überdeckenden Datensätzen beschrieben. Aus einer Expertenrunde unter Leitung des Bundesministeriums des Innern entstand der Anstoß, die Potenziale zu analysieren, die sich aus dem Vergleich von Datensätzen mit räumlicher Überdeckung ergeben. Die Teilnehmer der Expertenrunde erwarten daraus Vorteile wie die Vermeidung von Doppelarbeit, die Zusammenführung von Information und die Optimie-

rung des Aktualisierungsprozesses der Daten. Um diese Vorteile näher auszuarbeiten, wurde ein Projekt initiiert mit dem Ziel, geeignete Verfahren zum Vergleich von Datensätzen an konkreten Anwendungsfällen zu testen.

Die Partner in dem Projekt »Vergleich von Datensätzen« leisten ihre Tätigkeiten im Rahmen des Vorhabens aus eigenen Mitteln. Der Mehrwert liegt für alle Beteiligten in den gemeinsam gewonnenen Erkenntnissen und Synergien für die eigenen Prozesse und Produkte. Als Partner im Projekt finden sich zusammen:

- Das Bundesamt für Kartographie und Geodäsie (BKG) als Fachbehörde für Geoinformation und Geodäsie des Bundes,
- das Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN) als Repräsentant für ein Flächenland,
- der Landesbetrieb Geoinformation und Vermessung (LGV) Hamburg als Repräsentant für einen Stadtstaat, in dem die Open Data Entwicklung bereits sehr fortgeschritten ist, sowie
- die AED-SICAD als Anbieter von Geoinformationssystemen für die öffentliche Verwaltung.

Die Motivation bei den Beteiligten deckt mit ihren unterschiedlichen Schwerpunkten eine Reihe der vermuteten Potenziale ab, die im Projekt am Beispiel verschiedener Anwendungsfälle exemplarisch untersucht werden:

- Das BKG sucht Methoden zur Ermittlung und Verbesserung der Qualität von Geodaten.
- LGLN und LGV wollen den Herstellungs- bzw. Aktualisierungsprozess effizienter gestalten und dazu Information aus einem Datensatz in einen anderen Datensatz übernehmen.
- Zudem sind alle Beteiligten an Verfahren zur Erkennung von Änderungen in einem Geodatenbestand zwischen zwei unterschiedlichen Bearbeitungsständen interessiert.

Zur Umsetzung der verschiedenen Teilaspekte ist das Projekt in drei Teilbereiche strukturiert. Entsprechend der Gliederung in diese Teilprojekte sind auch die Ergebnisse im vorliegenden Beitrag dargestellt:

- Im ersten Teilprojekt werden die Daten des Basis-DLM zu zwei unterschiedlichen Zeitständen verglichen, um daraus Hinweise für die Aktualisierung von weiteren abgeleiteten Datenbeständen wie z. B. dem DLM250 zu ziehen. Beteiligt sind das BKG und AED-SICAD.
- Im zweiten Teilprojekt werden die Daten des Landbedeckungsmodells für Deutschland LBM-DE und des

Basis-DLM verglichen, um Information zwischen beiden Datensätzen zu übertragen. Federführend ist das LGLN.

- Im dritten Teilprojekt werden Daten der OpenStreet-Map (OSM) mit dem Basis-DLM und weiteren anderen Datensätzen in Relation gesetzt, um Erkenntnisse zur Datenqualität zu gewinnen und Perspektiven für die gemeinsame Nutzung aufzuzeigen. Beteiligt sind AED-SICAD, LGV und BKG.

2 Vergleich von Daten des Basis-DLM zu zwei unterschiedlichen Zeitständen

Die digitalen Landschaftsmodelle des ATKIS werden in unterschiedlichen Auflösungsstufen vorgehalten. Das BKG bearbeitet die Auflösungsstufen DLM250 (entspricht dem Kartenmaßstab 1:250.000) und DLM1000 (entspricht dem Kartenmaßstab 1:1 Million). Die beiden Datenbestände werden jährlich aktualisiert. Die Fortführung erfolgt aus verschiedenen Quellen, wobei das Basis-DLM der Landesvermessungseinrichtungen die wichtigste Quelle für Veränderungsinformationen darstellt.

Im Idealfall stellen die Länder dem BKG die Fortführungsinformation in einer inkrementellen Form bereit. Nach der Erstausrüstung werden nur noch die Veränderungen der Objekte mitgeteilt (NBA-Verfahren). Diese Methode der Fortführung ist jedoch noch nicht flächendeckend durch alle Bundesländer implementiert. Die Berücksichtigung von Differenzfortführungen (NBA-Verfahren) im Rahmen des Prototyps wurde nicht weiter betrachtet. Im Falle einer Fortsetzung der Untersuchungen sollte dieser Punkt erneut aufgegriffen werden. Zudem ist nur eine Untermenge der Veränderung im Basis-DLM relevant für die Auflösungsstufen DLM250 und DLM1000. Dem BKG ist an einem Werkzeug gelegen, das Daten des Basis-DLM zu unterschiedlichen Zeitpunkten vergleicht, Änderungen erkennt und diese Änderungen nach den vom BKG vorgegebenen Kriterien filtert. Neben der Anwendung bei der Fortführung des DLM250 und DLM1000 könnte ein solches Werkzeug auch eingesetzt werden, um in einem GIS die Analyse nach zeitlichen Veränderungen zu unterstützen.

Bei diesem Anwendungsfall liegen die zu vergleichenden Datensätze im selben semantischen Modell vor. Ziel ist ein Werkzeug, das Differenzen erkennt und bewertet, wobei die Bewertung über Konfigurationsparameter wie geometrische Mindestabweichung und prägende Attributwerte gesteuert wird. Einfache Ansätze zur Erkennung von Veränderungen im Basis-DLM erweisen sich dabei schnell als nicht zielführend. So kommt man nach einer Veränderungsabfrage unter ausschließlicher Verwendung der Objekt-Identifikatoren zu der Erkenntnis, dass Bearbeiter im ATKIS nicht selten Objekte löschen und neu erfassen, auch wenn sich nur nachrangige Eigenschaften ändern. Die Nachverfolgung der Objekthistorie über den Identifikator ist dann unterbrochen. Für die Analyse der

Veränderung müssen die Objekte tatsächlich geometrisch verglichen werden.

Im Rahmen der gemeinsamen Arbeitsgruppe obliegt der AED-SICAD die Aufgabe, ein prototypisches Vergleichswerkzeug zwischen zwei Datenbeständen des Basis-DLM herzustellen. Aufgrund der Ausgangssituation und den Anforderungen des BKG wird die vorhandene Produktkomponente FUSION Data Service (FDS) als Softwarebasis herangezogen, um die relevanten Datenänderungen möglichst regelbasiert, automatisiert und mit flexiblen Parametern zu ermitteln. Für das Vergleichsprojekt werden seitens des LGLN drei Ausschnitte aus dem Basis-DLM mit unterschiedlichen Zeitstempeln (30.12.2014, 30.12.2015 und 30.12.2016) sowie zugehörige Statistikinformationen zu Datenänderungen zugeliefert. Das BKG definiert die Vergleichskriterien exemplarisch für die Objektarten »Straßen« und »Industrieflächen«. In diese Kriterien fließen die spezifischen Parameter zum Übergang vom Basis-DLM in das DLM250 wie geometrische Mindestdimensionen und Wertebereiche von Attributen ein.

Zudem legen BKG und AED-SICAD gemeinsam die Anforderungen an die Ergebnisdatei fest. Diese soll nur die geänderten Objekte enthalten, die Art der Veränderung anzeigen und bei Änderungen von Objektarten und Attributwerten sowohl die alte als auch die neue Bedeutung angeben.

Das für den Datenvergleich verwendete Produkt FDS ist von seiner ursprünglichen Grundausrichtung für flexible Datentransformationen konzipiert, wobei Datenvergleiche bisher nicht vorgesehen waren. Durch die grundsätzliche Produkteigenschaft, an verschiedenen Prozesszeitpunkten individuelle »Plugins« (FME-Workbenches) zu integrieren, besteht die Möglichkeit, die Standard-Produkteigenschaften von FDS durch individuelle Erweiterungen zu optimieren und zu ergänzen. Aus diesem Grund wird für den Datenvergleich in erster Linie ein sogenannter »FDS-Schreibe-Plugin« implementiert (Abb. 1).

Um aus den o.g. Komplettdatenbeständen die erforderlichen Objektarten für den Datenvergleich auszufiltern, wird eine entsprechende FDS-Steuerdatei konfiguriert. Das »FDS-Schreibe-Plugin« verwendet variable Einstellungen aus einer Steuerdatei sowie weitere Filterbedingungen in derzeit noch fest codierter Form. Die alten und neuen Daten werden in drei Ströme gelenkt, um neue Elemente, gelöschte Elemente und geänderte Elemente zu ermitteln. Innerhalb dieser drei Datenströme werden durch den direkten Vergleich aller Objekte über deren Objektidentifikator (UUID) die Merkmale »neu«, »gelöscht« bzw. »geändert« ermittelt. Wesentliche Schritte im Bearbeitungsprozess sind der Vergleich der Geometrie auf Lageidentität und das Ausfiltern von Objekten, die eine Mindestlänge und Mindestfläche unterschreiten. Sollte an ein und demselben Feature mehr als eine Änderung erfolgt sein, wird dies im Prototyp des Systems noch nicht erkannt, da die Vergleiche sequenziell abgearbeitet werden.

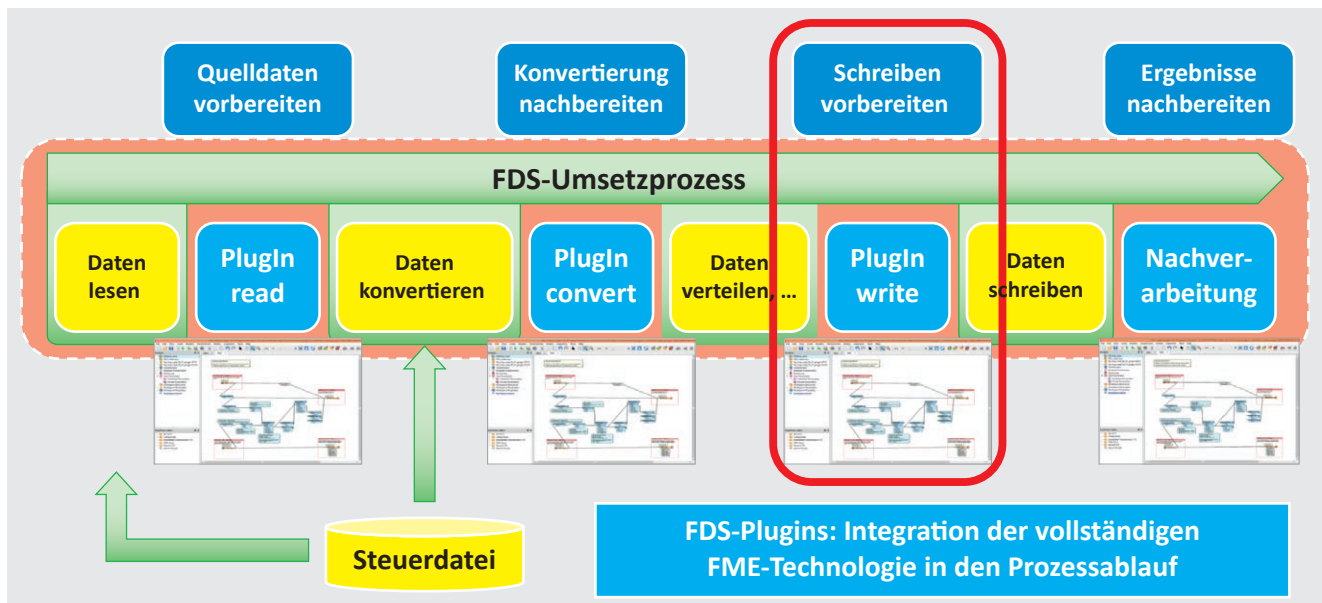


Abb. 1: Prozessablauf im FUSION Data Service

AED-SICAD hat die drei verfügbaren Basis-DLM-Datenbestände mit dem Werkzeug bearbeitet und jeweils in den Zeiträumen 2014–2015 und 2015–2016 verglichen. Für die Auswertung und Visualisierung der jeweiligen Vergleichsergebnisse wird zudem eine Kartendarstellung für ArcGIS for Desktop erstellt (Abb. 2).

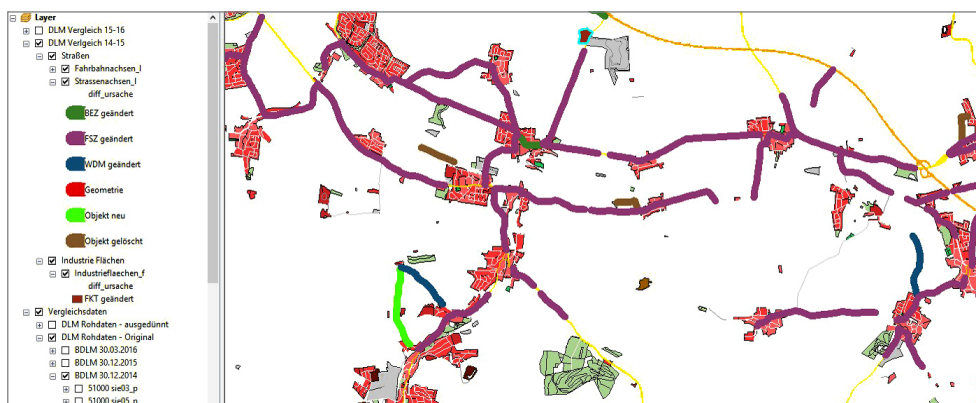
Obwohl das eingesetzte Werkzeug ursprünglich nicht für Geodatenvergleiche konzipiert ist, konnte im Rahmen der hier durchgeführten Tests nachgewiesen werden, dass gute Vergleichsergebnisse für Basis-DLM-Daten mit unterschiedlichen Zeitstempeln erzielt werden. Diese wiederum eignen sich als Grundlage für weitere Maßnahmen in anderen Nutzungsbereichen wie zum Beispiel für die Nachführung anderer Kartenwerke. Dies gilt insbesondere für die vom BKG für den Prototypen vorgegebenen Kriterien. Für eine umfängliche Einsetzbarkeit als allgemeines »Vergleichswerkzeug« sind aus Sicht der AED-SICAD noch weitere Feinspezifikationen und teilweise Implementierungen erforderlich. Dies gilt insbesondere für die Aspekte:

- Ausweitung auf alle ATKIS-Objektarten
- Flexibilisierung der vergleichbaren Attribute
- Mehrfache Änderungen an einem Feature erkennen und ausgeben

- Optionale Berücksichtigung von Änderungen an Relationen
- Informationsgewinnung auf der Grundlage von NBA-Daten

3 Vergleich von LBM-DE und Basis-DLM

Im zweiten Teilprojekt, bei dem die niedersächsische Landesvermessung und das Bundesamt für Kartographie und Geodäsie zusammenarbeiten, untersucht das LGLN den bidirektionalen Abgleich der Daten des ATKIS Basis-DLM mit denen des Landbedeckungsmodells für Deutschland (LBM-DE). Das LBM-DE beschreibt bundesweit flächendeckend die Landbedeckung mit einer räumlichen Auflösung von einem Hektar (Hovenbitzer et al. 2014, Arnold et al. 2011). Es wird vom BKG in einem dreijährigen Zyklus hergestellt und dient als Grundlage für die Erstellung des deutschen Beitrags im Projekt CORINE Landcover (CLC) der Europäischen Umweltagentur. Für die Erstausrüstung im Jahr 2009 leitete das BKG das LBM-DE aus den Daten des ATKIS Basis-DLM durch eine semantische Transformation in die CLC-Nomenklatur ab. Der so entstandene Datensatz wurde mit aktuellem Satellitenbildmaterial geprüft und ggf. korrigiert. Für das LBM-DE 2012 konnte erstmals fast vollständig auf die ATKIS-Daten im neuen AAA-Modell der AdV zurückgegriffen werden. Das



Satellitenbildmaterial geprüft und ggf. korrigiert. Für das LBM-DE 2012 konnte erstmals fast vollständig auf die ATKIS-Daten im neuen AAA-Modell der AdV zurückgegriffen werden. Das

Abb. 2: Visualisierung der Vergleichsergebnisse in ArcGIS for Desktop

in Bearbeitung befindliche LBM-DE 2015 wird dagegen nicht vollständig aus dem Basis-DLM neu abgeleitet, sondern durch Aktualisierung des LBM-DE 2012 mit Änderungsinformationen des Basis-DLM generiert. Im Anschluss erfolgt dann wiederum eine Kontrolle und Korrektur des Datensatzes mit Satellitenbildinformationen.

Ziel des Vergleiches von LBM-DE und Basis-DLM ist es nun, die Korrekturinformationen aus dem LBM-DE in das Basis-DLM zurückfließen zu lassen und damit einen bidirektionalen Austausch zwischen beiden Datenbeständen zu realisieren. Bei der Erstellung des LBM-DE fällt auf, dass die Daten des Basis-DLM bezüglich der Klassifizierung in Wald, Ackerland und Grünland besonders häufig anhand der Fernerkundungsdaten korrigiert werden müssen. Im Bereich des Waldes betrifft das die Differenzierung in Laub-, Nadel und Mischwald sowie die in CLC geforderte Objektklasse »Büsch, Sträucher, junge Bäume« (im Regelfall Wiederaufforstungsflächen). Da es sich bei der Waldklassifizierung im Gegensatz zur Differenzierung von Acker- und Grünland um relativ langlebige Informationen handelt, wurde in der niedersächsischen Landesvermessung frühzeitig an einem Konzept zur Übernahme der Änderungsdaten für die Objektart AX_Wald gearbeitet.

Der zu prüfende Objektkumfang konnte durch eine vom BKG bereitgestellte gezielte Auswertung der Waldklassifizierung für Niedersachsen und Bremen auf ca. 122.000 Waldflächen (ca. 47 % aller Waldflächen) abgeschätzt werden.

Abb. 3 zeigt für eine Bearbeitungseinheit 8×8 km (BE8) aus der südlichen Lüneburger Heide das Basis-DLM und die überlagerte Änderungsinformation aus dem LBM-DE 2012. In Magenta sind alle Waldflächen eingefärbt, die im LBM-DE ein abweichendes Vegetationsattribut aufweisen (beispielsweise Laubwald statt Mischwald).

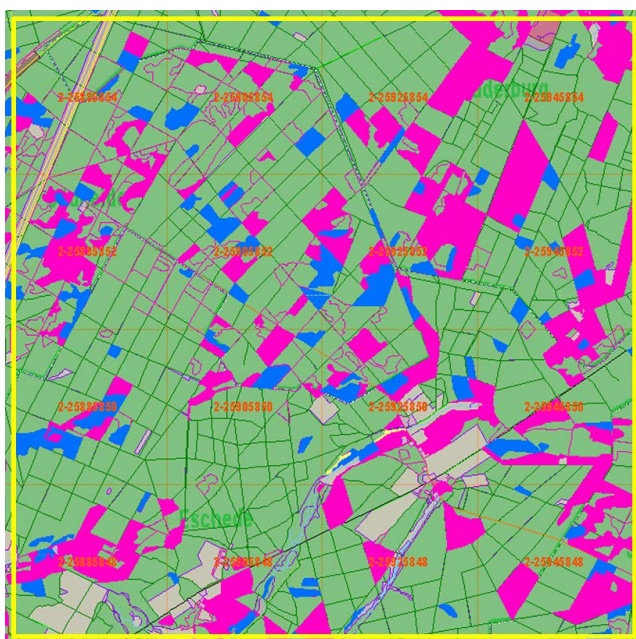


Abb. 3: Basis-DLM mit überlagelter Änderungsinformation aus dem LBM-DE

In Blau sind alle Flächen eingefärbt, die im LBM-DE mit der Objektklasse »Büsch, Sträucher, junge Bäume« belegt sind und damit im Basis-DLM als potenzielle Wiederaufforstungsflächen untersucht werden müssen.

Wegen des großen Umfangs betroffener Waldflächen wurde bereits von Beginn an nach einem automationsgestützten Ansatz für die Übernahme der LBM-DE-Information gesucht. In einer Bachelor-Arbeit in Zusammenarbeit mit der HTW Dresden (Lehmann 2013) konnte ein Weg für die automatisierte Übernahme der Waldklassifizierung aus dem LBM-DE (Datensätze 2009 und 2012) mit Hilfe der eingesetzten GIS-Software (3A Editor, AED-SICAD auf Basis von ArcGIS, Esri) aufgezeigt werden. Mit Hilfe des hier entwickelten Werkzeugs lassen sich einfache Fälle der Klassifizierungsänderung direkt in das Basis-DLM integrieren. Zu diesen einfachen Fällen zählen alle Waldobjekte, an denen im Basis-DLM seit der Ableitung des DLM-DE keine Änderungen vorgenommen wurden. Alle anderen Fälle erfordern die interaktive Bearbeitung.

Der automatisiert zu bearbeitende Anteil betrug zunächst 30 bis 40 % aller Anwendungsfälle. Mit zunehmender Bearbeitungsdauer im Übernahmeprozess verringerte sich jedoch der automatisierbare Anteil. Dies liegt einerseits an der größer werdenden Aktualitätslücke zwischen LBM-DE (Stand 2012) und Basis-DLM. Zum anderen wurden im Zuge der Nachmigration nach der Umstellung ins AAA-Datenmodell systematische Änderungen im Basis-DLM wie zum Beispiel eine Stützpunktreduktion der Umringsgeometrien durchgeführt, die den Vergleich geometrieidentischer Flächen zwischen LBM-DE (abgeleitet aus dem Basis-DLM 2012) und dem aktuellen Basis-DLM erschweren.

Abb. 4 zeigt eine Detailansicht einer Basis-DLM-Waldfläche mit überlagelter Änderungsinformation aus dem LBM-DE 2012. Die Waldfläche wurde im Basis-DLM als Mischwald- und im LBM-DE als Nadelwaldfläche klassifiziert. Ein zusätzliches Problem bei der Übernahme der LBM-DE-Daten stellt sich bei integrierter Führung von Landschaftsmodell und Kartendaten der DTK25, die in einigen Bundesländern (z. B. Hessen, Mecklenburg-Vorpommern, Niedersachsen und Sachsen) produktiv eingesetzt wird. Änderungen im Basis-DLM haben dann unmittelbar auch Einfluss auf die abgeleiteten Daten der DTK25. Im Falle der Objektart AX_Wald betrifft dies neben den ggf. durch Generalisierungsprozesse geänderten Objektgeometrien auch die als Präsentationsobjekte geführten Laub- und Nadelwaldsymbole. Die Änderung des Vegetationsmerkmals beispielsweise von Mischwald in Nadelwald erfordert das Nachziehen der Präsentationsobjekte in der Karte.

Abb. 5 zeigt die Umsetzung der Waldfläche in der DTK25. Hier wurden gemäß der bisherigen Klassifizierung im Basis-DLM zwei Präsentationsobjekte (Laub- und Nadelwaldsymbol) für den Mischwald platziert.

In Gebieten mit bereits gebildeten Kartengeometrie- und Präsentationsobjekten (ca. 50 % der Landesfläche) wird die Übernahme des LBM-DE daher fast ausschließ-

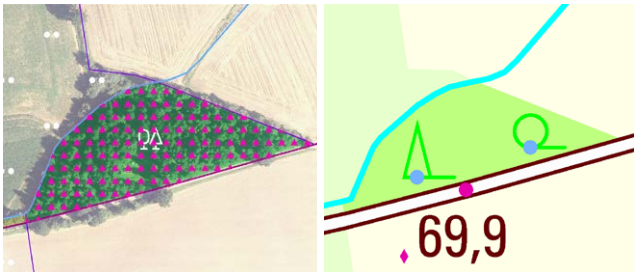


Abb. 4 [Links]: Basis-DLM mit überlagerter Änderungs-
information aus dem LBM-DE (Detail)

Abb. 5 [Rechts]: DTK25 (Detail)



Abb. 6: Basis-DLM und DTK25 nach der Aktualisierung

lich interaktiv durchgeführt. Abb. 6 zeigt den Zustand nach Einarbeitung der Änderungsinformationen aus dem LBM-DE 2012 im Basis-DLM und in der DTK25.

Die erstmalige Übernahme der Waldklassifizierung aus dem LBM-DE wird für Niedersachsen voraussichtlich im 2. Quartal 2018 abgeschlossen sein. Nach Bereitstellung des LBM-DE 2015 ist eine erneute Untersuchung der Differenzen in der Waldklassifizierung dieser Gebiete geplant, um einen zukünftigen Einarbeitungsaufwand abschätzen zu können.

Ein Grundproblem beim bidirektionalen Abgleich von Basis-DLM und LBM-DE bleibt – unabhängig von dem technischen Verfahren des Vergleichs – der unterschiedliche Zeitbezug der beiden Datenbestände. Das LBM-DE ist entsprechend den Anforderungen von CLC auf ein Stichjahr bezogen und wird alle drei Jahre auf ein neues Stichjahr aktualisiert. Das Basis-DLM wird dagegen kontinuierlich aktualisiert, sodass die Teile der Landesfläche im Rahmen des Grundaktualisierungszyklus einen unterschiedlichen Stand haben. Die Kapazitäten der Landesvermessungen werden deutschlandweit jedoch mittelfristig nicht für eine umfassende Grundaktualisierung aller Daten innerhalb eines Jahres ausreichen.

Bislang wird der bidirektionale Austausch von Informationen zwischen LBM-DE und Basis-DLM nur für die Waldklassifizierung betrachtet. Zukünftig könnten auch andere Informationen aus dem LBM-DE in das Basis-DLM zurückfließen, beispielsweise für Flächen von Solarkraftwerken oder für Salzwiesen. Mit der Einführung der Version 7 der GeoInfoDok werden eine ganze Reihe von Objektarten, Attributen und Attributwerten zum bundesweiten Grunddatenbestand des Basis-DLM aufgewertet bzw. neu aufgenommen, die aus dem LBM-DE übernommen und dann in den Aktualisierungsprozess des Landschaftsmodells einbezogen werden könnten.

4 Vergleich von OpenStreetMap und Basis-DLM/ALKIS

OpenStreetMap (OSM) und das Basis-DLM werden oft als konkurrierende Produkte wahrgenommen. Dies trifft in mancher Beziehung zu. Jedoch können sich beide Produkte auch ergänzen, wenn die gemeinsame Nutzung nicht durch unterschiedliche Lizenzbedingungen verhindert ist. Unter diesem Gesichtspunkt werden im entsprechenden Teilprojekt drei Aspekte untersucht:

- Vergleich der Gebäudenutzung von OSM mit dem ALKIS-Gebäudebestand zur möglichen Anreicherung von ALKIS
- Ergänzung von kommunalen Daten mit Objekten aus OSM am Beispiel der Kinderspielplätze in Hamburg
- Prüfung auf geometrische Vollständigkeit des Basis-DLM durch Vergleich mit OSM

4.1 Vergleich der Gebäude in OSM und ALKIS

Im Rahmen der gemeinsamen Arbeitsgruppe wird seitens der AED-SICAD die Einbeziehung von OSM-Daten bei der Laufendhaltung von ALKIS-Daten im urbanen Bereich untersucht. Es soll herausgefunden werden, welche Informationen aus OSM im Zuge der ALKIS-Fortführung sinnvoll und wirtschaftlich genutzt werden können. Voraussetzung für eine praktische Anwendung ist natürlich die Bereitstellung von beiden Datensätzen unter Open-Data-Lizenzen. Seitens der Beteiligten besteht Einvernehmen, dass sich die durchzuführenden Untersuchungen in erster Linie auf den Themenbereich »Gebäude« konzentrieren und ggf. bedarfsorientiert später auf andere Themenbereiche ausgeweitet werden.

Im Rahmen der Tests und der erforderlichen Prototypentwicklung wird ein ALKIS-Datensatz (ca. 5 × 5 km Ausdehnung) der Hansestadt Hamburg (LGV) aus dem Innenstadtbereich mit den Objektarten AX_Gebaeude sowie AX_Bauwerke und ein entsprechender Ausschnitt aus OpenStreetMap herangezogen. Nach Sichtung der Datensätze entscheiden sich die Projektpartner für die Bearbeitung der im Folgenden beschriebenen vier Fragestellungen. Die Durchführung der erforderlichen Vergleiche erfolgt ähnlich wie beim Teilprojekt Basis-DLM-Datenvergleich zu zwei unterschiedlichen Zeitständen.

1. Lassen sich die OSM-Gebäude den ALKIS-Gebäuden geometrisch zuordnen bzw. lassen sich mit Hilfe des geometrischen Vergleichs Lücken oder Veränderungen am Gebäudebestand in ALKIS feststellen?

Die geometrische Vergleichbarkeit zwischen ALKIS- und OSM-Gebäuden gestaltet sich schwierig, da hier neben den sehr unterschiedlichen Erfassungskriterien (Gebäude, Bauwerk, Einheiten, Bauteil, ...) zusätzlich noch die Transformation der OSM-Daten in das Koordinatenreferenzsystem der jeweiligen ALKIS-Daten erforderlich

ist und damit eine deutliche Lageungenaugkeit einhergeht. Für die geometrische Zuordnung wird deshalb die prozentuale Überlappung von Gebäudeflächen als Vergleichskriterium herangezogen. Die Bewertung der prozentualen Überlappung und deren Bedeutung hinsichtlich der wahrscheinlichen Identität der ALKIS- und OSM-Gebäude oder der Bewertung als Lücke bzw. Veränderung im Datenbestand obliegt dem jeweiligen Nutzer des Datenvergleichs. Eine objektive Bewertung der geometrischen Vergleiche ist nur sehr bedingt durchführbar und erfordert insbesondere die Durchführung von weiteren Vergleichstests mit unterschiedlichen Testgebieten sowie eine detailliert abgestimmte Selektion der zu vergleichenden Geometrien (Gebäude, Bauteile, Bauwerke, ...) aus dem jeweiligen Datenbestand. Sofern es gelingt, aus beiden Datensätzen die entsprechenden Geometrien zu identifizieren, bestehen durchaus sehr gute Möglichkeiten, Lücken oder Veränderungen am Gebäudebestand zu erkennen und ggf. gezielte Nachführungen zu veranlassen. Um aber eine realistische und wirtschaftliche Informationsgewinnung sicherzustellen, muss die Methodik der Geometrieidentifizierung und Zuordnung verfeinert und ggf. durch Hinzuziehung weiterer Geodatenquellen und Informationen optimiert werden. Dies gilt es in einem weiteren Schritt näher zu untersuchen.

2. Wie exakt passen die verfügbaren Adressangaben an den OSM-Gebäuden zu den Adressangaben in ALKIS (Lagebezeichnungen, Georeferenzierte Gebäudeadressen) und können aus dem Vergleichsergebnis die ALKIS-Objektarten *AX_GeoreferenzierteGebäudeadresse* oder *AX_Lagebezeichnung* ergänzt bzw. geprüft werden?

Im Umfeld der Adressinformationen kommt der Datenvergleich zum Ergebnis, dass ein Abgleich der Objektart *AX_GeoreferenzierteGebäudeadresse* mit OSM-Information durchaus als sinnvolle Ergänzung und Prüfmaßnahme Berücksichtigung finden kann, um sicherzustellen, dass in ALKIS vollständige Angaben vorliegen. Ein Abgleich zu den *AX_Lagebezeichnungen* gestaltet sich als tendenziell unbrauchbar, da in den OSM-Daten entsprechende Details nicht vollständig vorliegen bzw. die

Adressangaben in OSM lückenhaft sind, insbesondere für Lagebezeichnung mit Pseudonummer bzw. mit Hausnummer. Bei der Objektart *AX_LagebezeichnungOhneHausnummer* hingegen bestehen durchaus Möglichkeiten, Lücken in der Straßennamendarstellung (Präsentationsobjekte) in ALKIS zu entdecken und Ergänzungen zu veranlassen. Insgesamt scheinen die Adressangaben für den untersuchten Ausschnitt in ALKIS allerdings vollständiger vorzuliegen als in OSM (Abb. 7).

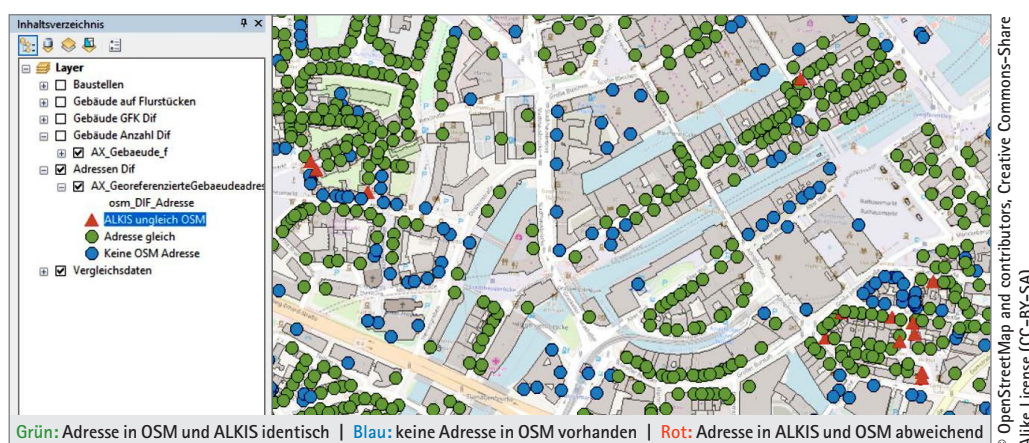


Abb. 7: Vergleich der Gebäudeadressen zwischen ALKIS und OSM

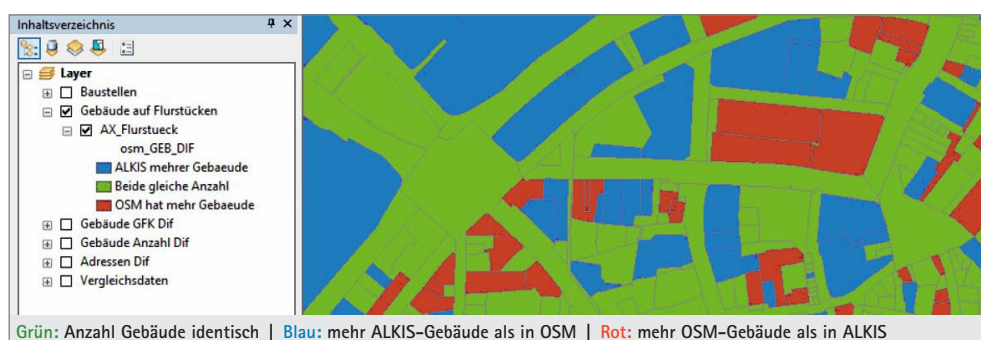


Abb. 8: Vergleich der Gebäudeanzahl pro Flurstück zwischen ALKIS und OSM

3. Wie viele ALKIS- bzw. OSM-Gebäude lassen sich dem jeweiligen Flurstück in ALKIS zuordnen und lässt sich hieraus Handlungsbedarf für die ALKIS-Fortführung ableiten?

Die Analyse, wie viele ALKIS- bzw. OSM-Gebäude pro Flurstück existieren, bestätigt in erster Linie die Ergebnisse des geometrischen Vergleichs. Die Auswertung der Ergebnisse zeigt sehr deutlich die unterschiedlichen Erfassungskriterien in ALKIS und OSM. Etwa ein Drittel der Flurstücke im Testgebiet (ca. 5 x 5 km) weist die gleiche Anzahl an Gebäuden aus, doch dies sind in erster Linie Verkehrsflächen und sonstige Freiflächen, bei denen die Gebäudeanzahl gleich Null entspricht. Je ein Drittel der Flurstücke weisen mehr ALKIS- bzw. mehr OSM-Gebäude auf (Abb. 8). Als Indikator für Fortführungsmaßnahmen in ALKIS scheint dieser Vergleich nicht zielführend zu sein.

Ein vielversprechender alternativer Ansatz ist ein Vergleich der überbauten Fläche (in OSM und ALKIS) im

Verhältnis zur Gesamtfläche des Flurstücks. Bei beiden Vergleichsansätzen zeigt sich aber, dass nur durch die fachgerechte Selektion der passenden OSM-Flächen in den verschiedenen Objektbereichen (building, landuse, ...) ein nutzbares Ergebnis erzielt werden kann. Die teilweise nicht eindeutige Zuordnung der relevanten (Gebäude-) Flächen in OSM zu unterschiedlichen Layern beeinflusst die Ergebnisse dieses Vergleichs sehr stark.

4. Lassen sich weitere Informationen (z.B. Nutzung) in OSM dazu heranziehen, um ggf. den Gebäudebestand in ALKIS aktuell zu halten oder mit weiteren Informationen (z.B. weitere Gebäudefunktion) anzureichern?

Der Abgleich der Gebäudeattribute in ALKIS und der Nutzung in OSM erscheint interessant. Durch diesen Vergleich lässt sich insbesondere das Attribut »WeitereGebäudefunktion« (z.B.: Bankfiliale, Hotel, Jugendherberge, Zoll, Post, usw.) in ALKIS anreichern bzw. die ALKIS-Gebäudefunktion überprüfen und die Aktualität erhöhen. Dazu ist aber im Vorfeld zwingend die Erstellung einer Vergleichs- und Zuordnungsmatrix zwischen den englischen Bezeichnungen in OSM und den festen Wertelisten in ALKIS erforderlich.

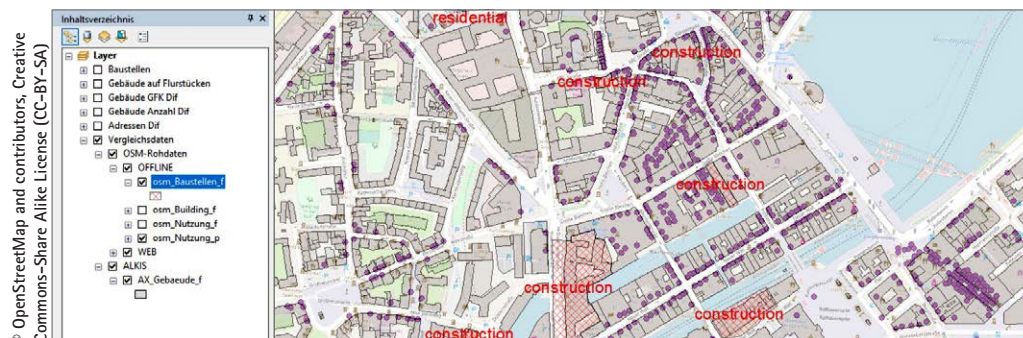


Abb. 9: Darstellung der Baustellen aus OSM im Stadtgebiet Hamburg

Als interessante zusätzliche Information für Bürger in Ergänzung zu kommunalen Datenquellen zeigen sich die aktuellen Angaben über Baustellen innerhalb des OSM-Datenbestandes (Abb. 9). An fast allen dieser Baustellen war im ALKIS noch die bisherige Bebauung dargestellt. Dies verdeutlicht den Fortführungsverzug in ALKIS, der sich aus dem Prozess der Amtlichen Vermessung ergibt.

4.2 Übernahme von Information aus OSM in Karten der Stadt Hamburg

Durch den Aufbau einer GeoDatenInfrastruktur (GDI) stellen die verschiedenen Dienststellen und Datenproduzenten ihre Informationen heute in Form eines Dienstes für die Darstellung in Karten zur Verfügung. Der topographische Kartenhintergrund reduziert sich von einer vollständigen Hintergrunddarstellung zu einer räumlichen Hintergrundkarte mit lediglich jenen Informationen, die nicht durch einen Dienst bereitgestellt werden. Dadurch profitiert die Kartographie von einer reduzierten Inhaltstiefe und einer erhöhten Aktualität. Sie orientiert sich an effektiveren modifizierten Workflows und am Gesichtspunkt der Datenkonsolidierung. Ziel ist es, den Paradigmenwechsel zum Web 2.0 zu schaffen, dabei aber das Printprodukt nicht aus dem Auge zu verlieren.

Zu berücksichtigen ist dabei, dass sich der Umfang und die Qualität der bereitgestellten Dienste räumlich stark unterscheiden können. Am Beispiel der großen Städte sind Hotspots mit einer sehr hohen und homogenen Bereitstellung von Diensten im dicht besiedelten Bereich und einer eher zögerlichen Umsetzung im ländlichen Umfeld erkennbar. Im Blickwinkel sind bei der kartographischen Nutzung nicht nur die amtlichen offenen Daten, sondern sicherlich auch die offenen Daten einer Community von Freiwilligen. Insbesondere die Übernahme oder der Abgleich von punktförmigen Daten stellt einen interessanten Aspekt dar. Aufgrund des Geometrietyps lässt sich ein räumlicher Abgleich bei punktförmigen Objekten leichter durchführen als bei linien- und flächenhaften Objekten.

phischen Nutzung nicht nur die amtlichen offenen Daten, sondern sicherlich auch die offenen Daten einer Community von Freiwilligen. Insbesondere die Übernahme oder der Abgleich von punktförmigen Daten stellt einen interessanten Aspekt dar. Aufgrund des Geometrietyps lässt sich ein räumlicher Abgleich bei punktförmigen Objekten leichter durchführen als bei linien- und flächenhaften Objekten.

Für das Projekt hat der LGV Hamburg die Übernahme der Spielplätze aus OSM in die Stadtkarte untersucht. Der Vergleich zwischen dem amtlichen Datenbestand und OSM zeigt dabei systematische Unterschiede (Abb. 10). Die OSM-Community erfasst sämtliche in der Örtlichkeit sichtbaren Spielplätze. Im amtlichen Datenbestand sind die als Spielplatz gewidmeten kommunalen Anlagen eingetragen. Eine Verifizierung der Eigentumsverhältnisse kann seitens der Community nicht erfolgen, da ihnen diese Informationen nicht zur Verfügung stehen. Im

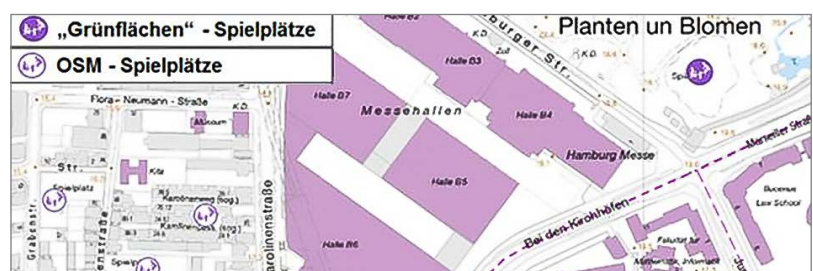


Abb. 10: Vergleich von Spielplätzen aus OSM und aus amtlichen Daten

Zusammenhang mit eigentumsrechtlichen Fragen haben daher die Dienste offizieller Dienststellen ein Alleinstellungsmerkmal.

Des Weiteren hat der LGV Hamburg die Darstellung von Kirchen zwischen den amtlichen Daten und OSM verglichen (Abb. 11). Am Beispiel der Kirchen kann man von recht statischen Informationen sprechen. Unterstellt man diesen Daten eine geringe Dynamik, so lassen sich diese in den Basisdatenbeständen gut vergleichen. Bezüglich der Lage lässt sich auch eine hohe Übereinstimmung erkennen. Die Attributtiefe hinsichtlich der Religion ist in OSM vorhanden. Die Daten einer GDI könnten damit angereichert bzw. verifiziert werden.

Datensätze mit sehr hoher Dynamik sind die Punktdaten zu Betrieben der Gastronomie und des Einzelhandels (Abb. 12). Ein Vergleich mit der Gebäudefunktion (GFK) aus ALKIS erscheint nützlich. Der Abgleich wird jedoch im Projekt nicht vertieft, da die Aspekte des Datenschemas und der Homogenität bei dieser Aufgabenstellung sehr komplex sind.

Objekte, die durch Dienste in der GDI bereitgestellt werden, sollen möglichst nicht mehr in die Basisdaten als originäre Präsentation einfließen. Sicherlich sind zur Nutzung solcher Informationen die Quellen zu verifizieren, auch wenn sie aus den öffentlichen Bereichen kommen. Bis hier der GDI-Gedanke voll umfänglich gelebt wird, sind noch einige Anstrengungen nötig. Aus diesem Grund ist es aber auch wichtig, die Quellen GDI-konform über Dienste bereitzustellen.

4.3 Vergleich von OSM und Basis-DLM zur Abschätzung der Vollständigkeit

Die Bundesverwaltung nutzt für zahlreiche Aufgaben die Geobasisdaten der Länder. Das BKG ist daher sehr an Aussagen über die Vollständigkeit und die Aktualität der Geobasisdaten interessiert. Diese Information kann durch einen Vergleich mit der Realität gewonnen werden oder mit etwas weniger Aussagekraft durch den Vergleich mit ähnlichen Datensätzen.

Im Teilprojekt des BKG soll gemeinsam mit der AED-SICAD ermittelt werden, ob durch einen Abgleich des Basis-DLM mit OSM Aussagen zur Vollständigkeit des Basis-DLM getroffen werden können. Für die Themen Straßen, Landnutzung/-bedeckung und Gebäude liegen bereits wissenschaftliche Untersuchungen vor (Dorn et al. 2015, Kunze et al. 2013, Schoof 2012). Gewässer wurden dahingehend bisher kaum untersucht. Daher hat das BKG exemplarisch die linien- und flächenhaften Gewässer auf

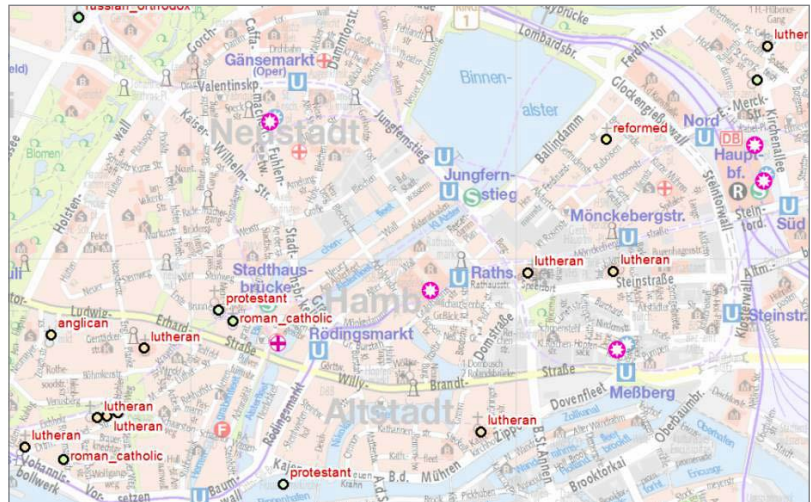


Abb. 11: Darstellung von Kirchen aus OSM auf der Stadtkarte Hamburg

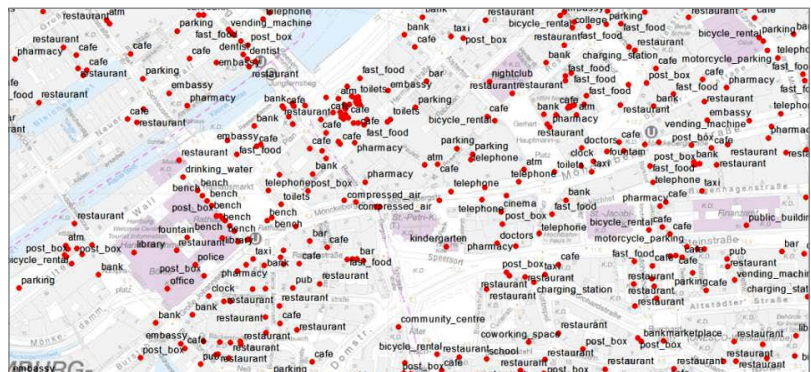


Abb. 12: Darstellung von Punktdaten zu Einzelhandel und Gastronomie aus OSM auf der Stadtkarte Hamburg

Vollständigkeit geprüft und als semantischen Test einen Vergleich der Namen durchgeführt. Die Lagegenauigkeit wird nicht geprüft.

In der Umsetzung des Vergleichs mit automationsgestützten Verfahren stellt sich die Abgrenzung von echten Unterschieden zwischen den Datensätzen (z.B. Objekt fehlt) gegenüber modellierungsbedingten Unterschieden (wie wird ein Objekt klassifiziert) als großes Problem heraus. Die »Tags« (Attribute) in OSM erweisen sich als sehr heterogen. Objekte, die im durchgeführten Vergleich als »in OSM fehlend« gefunden werden, sind nach genauerer Betrachtung meist mit ähnlichen oder verwandten Tags beschrieben, die aber im konkreten Vergleich bisher nicht berücksichtigt worden sind. Das Objekt liegt also in OSM vor, hat aber nicht die erwarteten Tags.

Die semantische Prüfung der Namen an einer kleinen, manuell ausgewerteten Stichprobe (knapp 2000 Objekte) hat ergeben, dass für größere Gewässer die Namen in den beiden Datensätzen weitgehend identisch sind. Bei den kleineren Gewässern sind Namen dagegen oft nur in einem der beiden Datensätze enthalten.

Die durchgeführten Vergleiche haben gezeigt, dass aufgrund der heterogenen Tags in OSM eine Prüfung auf Vollständigkeit für das Basis-DLM nicht sinnvoll durchgeführt werden kann. Hierbei liegen die Probleme weniger in der Entwicklung der Tools als vielmehr in der

grundsätzlich unterschiedlichen Herangehensweise bei der Erfassung und der Modellierung der Objekte. Für das Basis-DLM liegt ein bindender Objektartenkatalog vor, der festlegt, wie ein Objekt der realen Welt mit Objektart, Attributen und Geometrie für ATKIS zu erfassen ist. Für OSM hingegen gibt es zwar eine Liste möglicher Mapfeatures und Tags, die auch knapp erläutert werden. Grundsätzlich ist diese Liste aber für die freiwilligen Mapper nicht bindend und kann bei Bedarf von jedem einzelnen Mapper beliebig erweitert werden. Auch für die geometrische Erfassung liegen keine bindenden und damit streng einheitlichen Richtlinien vor. Insgesamt führt dies zu einer heterogenen Erfassung von ähnlichen oder gleichen Objekten und zu einer Vielzahl an Tags.

5 Fazit

Zur technischen Umsetzung der Aufgabenstellung sind im Projekt zwei Prototypen durch AED-SICAD bzw. LGLN erstellt worden. AED-SICAD hat den Prototyp aus bestehenden Softwaremodulen zusammengestellt und mit überschaubarem Aufwand angepasst. Das im LGLN eingesetzte Werkzeug entstand aus einer Bachelorarbeit. Beide Werkzeuge liefern gute Resultate bei einfachen Vergleichen. Für die Bearbeitung komplexer Sachverhalte und kartographischer Prozesse sind sie noch nicht geeignet.

Die aus dem Vergleich von Datensätzen abgeleitete Zuordnung erfolgt robust und sicher, wenn beide Datensätze im selben Datenmodell und nach gleichen Regeln erfasst sind, wie etwa beim Vergleich von Daten des Basis-DLM zu unterschiedlichen Zeitständen. Schwieriger wird der Vergleich, wenn systematische Unterschiede auftreten, wie zum Beispiel eine Stützpunktreduzierung der Geometrie. Problematisch ist die Zuordnung bei unterschiedlicher semantischer Ausprägung, wie am Beispiel der Spielplätze in Hamburg in OSM (Erfassung nach optischem Erscheinungsbild in der Landschaft) und in den kommunalen Daten (Erfassung aus der Widmung von städtischen Grundstücken) deutlich wird.

Die gemeinsame Nutzung von sich überdeckenden Daten ist einfach, wenn diese Daten über standardisierte Schnittstellen bereitstehen und lediglich grafisch überlagert werden. Die Stadtkarte Hamburg ist hierzu ein gutes Beispiel. Schwierig wird die gemeinsame Nutzung, wenn ein Abgleich auf Datenebene oder gar die Zusammenführung in eine einzige redundanzfreie Datenbank erforderlich ist. Zudem sind die lizenzrechtlichen Aspekte zu klären.

Allerdings bestätigen die verschiedenen Teilprojekte auch das aus der Expertenrunde vermutete Potenzial. Wenn in Zukunft immer mehr Datensätze über standardisierte Schnittstellen der GDI verfügbar sind, wird der Wunsch nach einer Nutzung dieses Potenzials weiter wachsen. Mit der Open-Data-Politik des Bundes und einiger Länder fallen vermehrt die lizenzrechtlichen Hürden.

Es ist durchaus möglich, dass unter diesen Rahmenbedingungen die für eine weitere Automatisierung bei der Zusammenführung von Datensätzen notwendige Verfahrensentwicklung in Angriff genommen wird.

Literatur

- Arnold, S., Rückert, F., Wende, C., Hovenbitzer, M. (2011): Digitales Landschaftsmodell Deutschland (DLM-DE). In: Publikationen der DGPF, 31. Wissenschaftlich-Technische Jahrestagung der DGPF in Mainz, Band 20, 201–210.
- Dorn, H., Törnros, T., Zipf, A. (2015): Quality Evaluation of VGI Using Authoritative Data – A Comparison with Land Use Data in Southern Germany. *ISPRS Int. J. Geo-Inf.* 2015, 4, 1657–1671.
- Hovenbitzer, M., Emig, F., Wende, C., Arnold, S., Bock, M., Feigenspan, S. (2014): Digital Land Cover Model for Germany – DLM-DE. In: Manakos, I., Braun, M. (Eds.): *Land Use and Land Cover Mapping in Europe: Practices & Trends, Remote Sensing and Digital Image Processing* 18, Springer, Berlin/Heidelberg, 255–272.
- Immler, W. (2014): *Das Handbuch OpenStreetMap*. 1. Auflage, Franzis Verlag.
- Kunze, K., Hecht, R., Hahmann, S. (2013): Zur Vollständigkeit des Gebäudedatenbestandes von OpenStreetMap. *KN Kartographische Nachrichten – Journal of Cartography and Geographic Information*, 2013, Heft 2/3, 73–81.
- Lehmann, C. (2013): Untersuchungen zur Aktualisierung des ATKIS-DLM mit Hilfe von Daten des DLM-DE. Bachelorarbeit HTW Dresden.
- Ramm, F., Topf, J. (2010): *OpenStreetMap: Die freie Weltkarte nutzen und mitgestalten*. 3. Auflage, Lehmanns Media.
- Retat, A., Schaffert, M. (2018): OpenStreetMap im Kontext kulturlandschaftlicher Fragestellungen. Ein Qualitätsvergleich mit Blick auf Landnutzungen und Landbedeckungen. *zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, Heft 1/2018, 143. Jg., 36–45. DOI: 10.12902/zfv-0190-2017.
- Schoof, M. (2012): ATKIS-Basis-DLM und OpenStreetMap – Ein Datenvergleich anhand ausgewählter Gebiete in Niedersachsen. *KN Kartographische Nachrichten – Journal of Cartography and Geographic Information*, 2012, Heft 1, 20–26.

Kontakt

Dipl.-Ing. Hubert Fünfer
AED-SICAD GmbH
Carl-Wery-Straße 22, 81739 München
hubert.fuenfer@aed-sicad.de

Dr.-Ing. Andreas Illert | Dipl.-Ing. agr. Sonja Werhahn
Bundesamt für Kartographie und Geodäsie
Richard-Strauss-Allee 11, 60598 Frankfurt am Main
andreas.illert@bkg.bund.de | sonja.werhahn@bkg.bund.de

Dipl.-Ing. (FH) Thomas Michaelis
Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung, Neuenfelder Straße 19, 21109 Hamburg
thomas.michaelis@gv.hamburg.de

Dipl.-Math. Markus Müller
AED-SICAD GmbH
Mallwitzstraße 1–3, 53177 Bonn
markus.mueller@aed-sicad.de

Dipl.-Ing. Klaus-Peter Wodtke
Landesamt für Geoinformation und Landesvermessung Niedersachsen
Podbielskistraße 331, 30659 Hannover
klaus-peter.wodtke@lgl.niedersachsen.de

Dieser Beitrag ist auch digital verfügbar unter www.geodaesie.info.