

OpenStreetMap im Kontext kulturlandschaftlicher Fragestellungen – Ein Qualitätsvergleich mit Blick auf Landnutzungen und Landbedeckungen

Alessa Retat und Markus Schaffert

Zusammenfassung

Das Projekt OpenStreetMap bietet kostenfreie Geodaten, die aus dem Internet bequem – beispielsweise ohne vorherige Registrierung und längere Wartezeiten – bezogen werden können. Die Community von OpenStreetMap ist prinzipiell weltweit vertreten; abweichende staatlich-administrative Verantwortlichkeiten für die Datenerhebung (inklusive unterschiedlicher Lizenzierungsphilosophien) kennt das Projekt nicht. Diese Charakteristika machen OpenStreetMap attraktiv für Forschungsprojekte wie Regiobranding, das in mehreren Bundesländern Kulturlandschaften nach ihren Besonderheiten untersucht und dessen Ergebnisse auf andere Regionen (national wie international) übertragbar sein sollen.

Allerdings werden die Daten von freiwilligen »Mappern« erhoben, die i.d.R. zwar Erfassungsempfehlungen beachten, letztlich jedoch ohne verbindliche Vorschriften agieren. Die Datenqualität hängt folglich in starkem Maße von den Erfassern – ihrer Anzahl, ihren Fertigkeiten und ihren Kenntnissen – ab. Deshalb ist es für die Verwendung von OpenStreetMap-Daten in fachspezifischen Anwendungen von zentraler Bedeutung, ihre Qualität im jeweiligen Untersuchungsbereich mit Blick auf das relevante Fachthema einzuschätzen zu können.

Der Artikel beschreibt das Vorgehen und die Ergebnisse einer Qualitätsbestimmung für Daten aus OpenStreetMap im Gebiet der Hansestadt Lübeck. Die Qualität wird anhand eines Vergleichs mit dem amtlichen Datensatz ATKIS-Basis-DLM und mit Blick auf die Qualitätsparameter der Vollständigkeit sowie der thematischen Genauigkeit ermittelt. Inhaltlich legt der Vergleich einen Schwerpunkt auf das Themenfeld »Landnutzung und Landbedeckung«.

Summary

The OpenStreetMap project offers geo data that is free of charge and accessible via the internet without registration or considerable delivery time. Furthermore, the project's community operates worldwide and is not bound to (differing) data gathering responsibilities or licensing philosophies of public authorities. These aspects are attractive for a research project like Regiobranding that aims at investigating characteristics of cultural landscapes in different German federal states and tries to transfer findings and methods to other regions (in Germany and abroad).

However, OpenStreetMap data is gathered by volunteers. These voluntary mappers usually follow recommended mapping instructions. As following these instructions is not mandatory, data quality strongly relies on the mappers' skills and number. Consequently, assessing the data quality for a

certain study area is crucial for using OpenStreetMap in real world use cases.

This paper presents a quality assessment for OpenStreetMap data in the city of Lübeck in northern Germany, focusing on land use and land cover. The assessment is conducted by comparing OpenStreetMap with the authoritative data set ATKIS-Basis-DLM and by applying the quality parameters of completeness and thematic accuracy.

Schlüsselwörter: OpenStreetMap, Datenqualität, Landnutzung und Landbedeckung, ATKIS-Basis-DLM

1 Einleitung

Um Daten aus OpenStreetMap (OSM) in fachspezifischen Anwendungen zu nutzen, ist es essenziell, deren Qualität zu kennen. Diese kann gemäß der internationalen Organisation für Normung durch unterschiedliche Parameter, u.a. ihre Vollständigkeit und ihre thematische Genauigkeit, beschrieben werden (ISO 19157).

Für verschiedene Länder existieren Untersuchungen, die die Qualität von OSM-Daten im Vergleich zu anerkannten Referenzdatensätzen und anhand unterschiedlicher Qualitätsparameter beurteilen. Nicht wenige dieser Datenvergleiche konzentrieren sich auf die Analyse von Linienobjekten, insbesondere Straßennetze. Ein solcher Qualitätsvergleich wird z.B. von Mondzech und Sester (2011) durchgeführt, die das Straßennetz in London bezüglich der Parameter »Vollständigkeit« und »Positionsgenauigkeit« untersuchen. Allerdings werden in OSM auch Punkt- und Polygonobjekte erfasst, die zusätzliche Nutzungsmöglichkeiten eröffnen und weitere Anwendungsfelder erschließen. Entsprechend beschäftigen sich Qualitätsuntersuchungen auch mit OSM-Daten, die auf Landnutzungen und Landbedeckungen (LULC, von engl. Land Use/Land Cover) fokussieren. Landnutzungen in diesem Zusammenhang beziehen sich auf »man's activities on land which are directly related to the land« (Burley 1961), während Landbedeckungen »the vegetational and artifical constructions covering the land surface« beschreiben (Clawson und Stewart 1965).

In diesem Kontext gelangen Schoof et al. (2011) zu Aussagen über die Datenqualität, indem sie OSM und das ATKIS-Basis-DLM (Amtliches Topographisch-Kartographisches Informationssystem; Digitales Landschaftsmodell) in drei niedersächsischen Untersuchungsgebieten

mit je 25 km² verglichen. Während dabei Linienobjekte auf Vollständigkeit und Positionsgenauigkeit hin untersucht werden, werden LULC-Klassen nur auf ihre Vollständigkeit geprüft. Estima und Painho (2013) bewerten die Qualität von LULC-Daten aus OSM gegenüber CORINE Land Cover¹ für das portugiesische Festland. Als Bewertungskriterium nutzen die beiden Autoren neben der Vollständigkeit den Qualitätsparameter der thematischen Genauigkeit. Diese beiden Parameter stehen auch in weiteren Studien im Fokus: Dorn et al. (2015) untersuchen diesbezüglich die Qualität von OSM in der Rhein-Neckar-Region und nutzen das ATKIS-Basis-DLM als Referenz. Jokar Arsanjani und Vaz (2015) berechnen entsprechend OSM-Qualitäten für sieben europäische Stadtregionen gegenüber Daten des Urban Atlas-Projekts¹. Mit gleicher Methodik und Datengrundlage ermitteln Jokar Arsanjani et al. (2015) die Qualität von LULC-Daten für die Stadtregionen Berlin, Frankfurt, Hamburg und München, während Vaz und Jokar Arsanjani (2015) OSM mit amtlichen kanadischen Daten im Großraum Toronto vergleichen. Von Johnson und Lizuka (2016) schließlich werden thematische Genauigkeiten im Abgleich mit Luftbildern aus Google Earth für eine Region nahe der philippinischen Hauptstadt Manila bestimmt.

Landnutzungen und -bedeckungen können als Elemente einer Kulturlandschaft gesehen werden, die aufgrund ihrer ausgeprägten Flächenpräsenz das sichtbare Antlitz dieser Landschaft wesentlich mitbestimmen. Das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Projekt Regiobranding möchte deshalb Besonderheiten von regionalen LULC, neben weiteren landschaftlichen Spezifika, für ein umfassendes Verständnis norddeutscher Kulturlandschaften und schließlich für das Entwickeln von regionalen Marken nutzen (Schaffert et al. 2016). In Regiobranding arbeiten acht Einrichtungen aus Praxis und u.a. ingenieurs-, planungs- sowie wirtschaftswissenschaftlich orientierten Hochschuldisziplinen zusammen, die ihre Fachsicht auf das Thema Kulturlandschaft einbringen und gemeinsam in drei norddeutschen Regionen umsetzen. Im Zusammenspiel aller Partner möchte das Projekt mittels Branding² dazu beitragen, die Identifikation der Bewohner mit ihrer Region und deren Kulturlandschaft zu festigen, ihr Engagement für den Erhalt der Kulturlandschaft zu steigern und die Attraktivität der Region nach außen zu fördern (Kempa und Herrmann 2014). Die Untersuchung von LULC stellt insofern einen von mehreren Fachbeiträgen im Projektverbund dar.

¹ www.eea.europa.eu/data-and-maps

² Beim Branding (dt. Markenaufbau) wird eine neue Marke mit dem Ziel geformt, ihr einen eigenen und unverwechselbaren Charakter zu verleihen. Dieser Ansatz, der im Unternehmensumfeld etabliert ist, wird heute auch auf Städte und auf Regionen – beispielsweise mit der Zielrichtung touristischer Entwicklung – angewandt (u.a. Blain et al. 2005).

Um entsprechende Analysen auch nach Projektende durchführen zu können – ohne dabei an die Lizenzbedingungen für amtliche Daten mehrerer projektrelevanter Bundesländer gebunden zu sein –, wurde OSM als mögliche Alternative im Projektgeschehen diskutiert. Dies motivierte zu Qualitätsuntersuchungen von OSM in den Projektgebieten, wovon nachfolgend das Vorgehen und die Ergebnisse für die Hansestadt Lübeck dargestellt werden. Dabei orientiert sich die Methodik an früheren Studien. Dies erfolgt, indem die Qualität anhand von Vollständigkeit und thematischer Genauigkeit im Vergleich zum amtlichen Datensatz ATKIS-Basis-DLM bestimmt wird. Allerdings definieren kulturlandschaftsbezogene Fragestellungen in einer Stadt- bzw. Stadt-Land-Region eigene Schwerpunkte bei der Bearbeitung, die die Vergleichbarkeit mit vorangegangenen Untersuchungen beschränken.

Im Folgenden werden zunächst grundlegende Informationen, die zum Verständnis des Artikels beitragen, vermittelt. Anschließend wird die Vorgehensweise erläutert, anhand derer die Qualitätseinschätzung erfolgt. Diese umfasst neben der eigentlichen Berechnung von Qualitätsparametern vorbereitende Schritte der Datenaufbereitung und -harmonisierung. Darauf aufbauend werden Ergebnisse präsentiert und es wird ein abschließendes Fazit gezogen. Da die vorbereitenden Schritte die Ergebnisse wesentlich beeinflussen können, wird das dabei verfolgte Vorgehen ausführlich dargestellt.

2 Grundlagen

2.1 Untersuchungsgebiet

Die Hansestadt Lübeck liegt im Südosten von Schleswig-Holstein an der Grenze zu Mecklenburg-Vorpommern. Sie umfasst eine Fläche von ca. 21.174 ha, auf der Ende 2016 gut 220.000 Einwohner lebten (Hansestadt Lübeck 2017). Gemeinsam mit östlich angrenzenden Kommunen bildet sie eine von insgesamt drei Untersuchungsregionen des Projekts Regiobranding.

2.2 Daten

Die OSM-Daten wurden von <http://planet.openstreetmap.org> (Stand: Juni 2015) bezogen. Als Referenz dient das ATKIS-Basis-DLM des Landesamtes für Vermessung und Geoinformation Schleswig-Holstein; es stammt aus der ersten Hälfte des Jahres 2015.

2.2.1 OpenStreetMap

OSM wurde im Jahr 2004 initiiert, um eine offen zugängliche Weltkarte zu schaffen. Zu diesem Zweck sammeln an dem Projekt Teilnehmende (»Mapper«) seither auf

ehrenamtlicher Basis vektorielle Geodaten, die über das Internet »frei« (gemäß Open Database License, ODbL, 1.0) für kommerzielle und für nicht kommerzielle Zwecke zur Verfügung gestellt werden (OSM 2016). Die erfassten Objekte sind durch sogenannte Tags beschrieben, die die thematischen Eigenschaften der Objekte kennzeichnen. Tags werden als Schlüssel-Wert-Paare gespeichert, wobei der Schlüssel (z.B. *landuse*) die Art der beschriebenen Eigenschaft bestimmt und der Wert den Schlüssel genauer beschreibt (z.B. *landuse = residential*).

Es gibt mehrere Möglichkeiten, Objekte in OSM zu kartieren. Eine Variante besteht in der Erfassung vor Ort mit mobilen GPS-Empfängern, wie sie in Smartphones heute serienmäßig verbaut sind. Des Weiteren gibt es die Optionen, vektorielle Geodaten auf Basis bereitgestellter Luftbilder manuell zu erzeugen oder – vorbehaltlich lizenzrechtlicher Möglichkeiten – aus anderen Quellen zu importieren (Mooney et al. 2010). Jeder eingetragene Nutzer kann Daten erfassen und Änderungen vornehmen. Diese Freiheiten dürften wesentlich zur Beliebtheit des OSM-Projekts und seiner vergleichsweise großen Community beitragen. Andererseits kann durch das Fehlen von verbindlichen Erfassungsvorschriften keine Homogenität der Daten garantiert werden (Antoniou und Schlieder 2014). Zwar unterbreitet das OSM-Wiki (unter <http://wiki.openstreetmap.org>) konkrete Vorschläge für die Benennung von Tags und beschreibt diese i. d. R. eindeutig; allerdings sind sie nur als Empfehlungen und nicht als verpflichtende Richtlinien zu verstehen. Hinzu kommt, dass die Mapper ggf. keine in der Erfassung von Geodaten geschulten Experten sind. Häufig liegt ihr Fokus zudem auf eigenen Interessensgebieten, die sie ohne Koordination mit anderen Mappern bearbeiten (Mondzech und Sester 2011). Infolgedessen zeigen sich in räumlicher und zeitlicher Hinsicht mitunter deutliche Unterschiede in der Datenqualität, die gemeinhin mit Anzahl und Kenntnissen der aktiven OSM-Erfasser erklärt werden (Neis et al. 2013).

2.2.2 ATKIS-Basis-DLM

Das ATKIS-Basis-DLM ist ein amtlicher Datensatz, der die topographischen Objekte der Landschaft und Reliefformen der Erdoberfläche im Vektorformat beschreibt (Laggner 2016). Es strukturiert und modelliert die Landschaft systematisch nach Objektarten und zugehörigen beschreibenden Informationen. Das Basis-DLM wird bundesweit auf Grundlage der GeoInfoDok (aktuelle Version: 6.0.1; AdV 2015) und dem zugehörigen ATKIS-Objektartenkatalog Basis-DLM (aktuelle Version: 6.0) erfasst und bietet einen einheitlichen Grunddatenbestand als länderübergreifenden Mindestinhalt. Für die Arbeit in Regiobranding, das in drei Bundesländern kulturlandschaftliche Besonderheiten untersucht, stellt es deshalb eine zentrale Datengrundlage dar. Objekte werden im Basis-DLM vollzählig erfasst, »wenn ihre topographische Bedeutung dem Maßstabsbereich 1:10.000 / 1:25.000 an-

gemessen ist« (AdV 2016). Entsprechend arbeitet es mit Mindestfassungsgrößen, sodass Objekte von LULC-Klassen mehrheitlich erst ab einer Flächengröße von mindestens 1 ha kartiert werden (AdV 2015). Abgrenzungen zwischen benachbarten Nutzungsarten bieten eine Lagegenauigkeit von 15 Metern. Nach Angaben des Landesamtes für Vermessung und Geoinformation Schleswig-Holstein ist sie jedoch vielfach besser (LVermGeo SH 2017).

In der Modellsicht des Basis-DLM bilden die Objektartengruppen ›Gewässer‹, ›Siedlung‹, ›Verkehr‹ und ›Vegetation‹ die Oberfläche der Erde überschneidungsfrei und lückenlos ab. Diese werden ergänzt durch überlagernde topographische Objekte, wie beispielsweise Bauwerke. Eine strikte Trennung zwischen Landnutzungs- und Landbedeckungsklassen erfolgt im Basis-DLM, das die Erfassung der tatsächlichen Nutzung vorsieht, insofern nicht. Das Auflösen dieser Vermischung und die fachlich-semantische Trennung von Landnutzung und Landbedeckung im Basis-DLM (sowie im ALKIS, dem Amtlichen Liegenschaftskatasterinformationssystem), wird derzeit – u. a. aus Gründen der statistischen Vergleichbarkeit – diskutiert (Arnold et al. 2017).

Objektarten, Attribute und Attributwerte des Basis-DLM werden momentan mindestens innerhalb eines Aktualisierungszeitraumes von fünf Jahren überprüft und bei Veränderungen fortgeführt. Für die Objektarten des Objektbereichs ›Verkehr‹ und einige Objektarten des Objektbereichs ›Gebiete‹ werden darüber hinaus Spitzenaktualitäten von drei, sechs und zwölf Monaten realisiert (Kurstedt 2013, Meinel 2008).

2.3 Prüfung der Datenqualität

Von den in der ISO-Norm 19157 definierten Qualitätsparametern spielen nach Jokar Arsanjani et al. (2015) die Vollständigkeit und die thematische Genauigkeit eine zentrale Rolle für die Qualitätsbeurteilung von LULC-Daten aus OSM. Daher konzentriert sich die Untersuchung in Lübeck auf diese Parameter.

2.3.1 Vollständigkeit

Die Vollständigkeit »informs the user about the spatial, thematic, and temporal coverage capabilities of the data according to the predefined purposes (Caprioli et al. 2003)«. Die nachfolgend dargestellte Untersuchung bezieht sich, ebenso wie vorangegangene Vergleichsstudien zur LULC-Datenqualität von OSM, nur auf die Vollständigkeit nach Fläche (spatial coverage). Sie entspricht dem Verhältnis der in einem Testdatensatz (hier: OSM) kartierten Fläche (A_{OSM}) zur Gesamtfläche des untersuchten Gebiets (A_{ges}):

$$\text{Vollständigkeit [%]} = \frac{A_{OSM}}{A_{ges}} * 100 . \quad (1)$$

2.3.2 Thematische Genauigkeit

Die thematische Genauigkeit gibt Aufschluss über die Richtigkeit der Attribute im Testdatensatz im Vergleich zu den Referenzdaten (hier: Basis-DLM). Zur Ermittlung der thematischen Genauigkeit dienen Parameter wie die Gesamtgenauigkeit, der Kappa-Koeffizient und die Produzenten- bzw. Nutzergenauigkeit, die im Folgenden näher erläutert werden.

Die Gesamtgenauigkeit ist ein Genauigkeitsmaß, das eine allgemeine, klassenübergreifende Qualitätsbeurteilung eines Datensatzes erlaubt. Für ihre Berechnung wird eine Konfusionsmatrix erstellt, durch die sich Test- und Referenzdaten hinsichtlich der Übereinstimmung ihrer Attribute vergleichen lassen (siehe Tab. 1). Die Felder der Matrix enthalten die Häufigkeit korrespondierender LULC-Klassen. Auf der Hauptdiagonalen, die in Tab. 1 durch die Gesamtheit der gelb dargestellten Zellen repräsentiert ist, befinden sich die Übereinstimmungen und in den restlichen Zellen die Abweichungen zwischen den beiden Datensätzen (Foody 2002).

Tab. 1: Schematische Darstellung der Konfusionsmatrix

		ATKIS-Basis-DLM			
		Klasse 1	Klasse 2	...	Klasse n
OSM	Klasse 1	h_{11}	h_{12}	...	h_{1n}
	Klasse 2	h_{21}	h_{22}	...	h_{2n}
	:	:	:	:	:
	Klasse n	h_{n1}	h_{n2}	...	h_{nn}

Anhand der Konfusionsmatrix lassen sich die statistischen Maße True Positives (TP_i), False Positives (FP_i) und False Negatives (FN_i) für jede Klasse i (mit $i = 1, 2, \dots, n$) bestimmen. True Positives entsprechen LULC-Objekten, die in OSM und im Basis-DLM an gleicher Stelle liegen und zudem dieselbe Klasse aufweisen.³ In der Matrix werden TP_i durch die Werte auf der Hauptdiagonalen (in Tab. 1 also $h_{11}, h_{22}, \dots, h_{nn}$) dargestellt. FP_i stehen für Objekte, die zur betrachteten LULC-Klasse in OSM gehören, im DLM aber keine thematische Entsprechung an gleicher Position haben. Dies entspricht in Tab. 1 – bezogen auf eine einzelne LULC-Klasse in OSM – der Summe der horizontalen Werte abzüglich des Wertes auf der Hauptdiagonalen. Beispielhaft für die Klasse 1 aus OSM sind dies die Werte h_{12}, \dots, h_{1n} . Von FN_i spricht man dagegen, wenn an der Stelle, an der ein Objekt einer betrachteten Klasse im Basis-DLM liegt, kein Objekt der gleichen Klasse

in OSM vorhanden ist. In Tab. 1 entspricht dies – bezogen auf eine einzelne LULC-Klasse des DLM – der Summe der vertikalen Werte abzüglich des Wertes auf der Hauptdiagonalen. Dies ist beispielhaft für Klasse 1 aus dem Basis-DLM die Summe der vertikalen Werte h_{21}, \dots, h_{n1} .

Die Gesamtgenauigkeit (p_0) stellt den relativen Anteil der Übereinstimmungen (TP_i) zur Gesamtzahl an Stichproben (N_{ges}) dar (vgl. Story und Congalton 1986):

$$p_0 = \sum_{i=1}^n \frac{TP_i}{N_{ges}}. \quad (2)$$

Die Gesamtgenauigkeit ist nur bedingt aussagekräftig, da auch bei einer zufälligen Verteilung ein gewisses Maß an Übereinstimmung zu erwarten ist. Dem begegnet der Kappa-Koeffizient, indem er die zufällig zu erwartenden Häufigkeiten ermittelt und ihren Unterschied zu den tatsächlichen Übereinstimmungen wiedergibt. Seine Berechnung erfolgt nach Cohen (1960) durch folgende Gleichung:

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \quad (3)$$

$$\text{mit } p_e = \sum_{i=1}^n \frac{(TP_i + FP_i) * (TP_i + FN_i)}{(N_{ges})^2}.$$

Die zufällig zu erwartenden Häufigkeiten werden aus den Zeilen- und Spaltengesamtanzahlen der Konfusionsmatrix bestimmt, die über alle Klassen aufsummiert werden. In der Gleichung geben p_0 den Anteil der tatsächlich beobachteten Übereinstimmung und p_e den Anteil der zufallsbedingten Übereinstimmung an. Im Zähler wird somit die Abweichung der Beiträge durch die beobachtete Übereinstimmung und der durch Zufall prädizierten Übereinstimmung berechnet. Durch den Nenner erfolgt eine Standardisierung, sodass der Kappa-Koeffizient im Normalfall im Wertebereich von 0 bis 1 liegt. Ein Wert von 1 zeigt eine vollständige Übereinstimmung an, während ein Wert von 0 auf eine zufällige Übereinstimmung hinweist (Foody 2002). Theoretisch sind auch negative Werte möglich, die eine geringere Übereinstimmung, als die durch den Zufall bedingte, implizieren.

Gesamtgenauigkeit und Kappa-Koeffizient beziehen sich auf die untersuchten LULC-Klassen in ihrer Gesamtheit. Aussagen zur Qualität einzelner Klassen lassen sich dagegen durch die Berechnung von Produzenten- und Nutzer-Genauigkeit treffen. Die Produzenten-Genauigkeit gibt die Vollständigkeit an, mit der Objekte einer bestimmten LULC-Klasse im Testdatensatz (OSM) erfasst worden sind. Dazu wird die Anzahl der richtig klassifizierten Objekte dieser Klasse ins Verhältnis zu allen Objekten der entsprechenden Klasse im Referenzdatensatz (Basis-DLM) gesetzt. Die Nutzer-Genauigkeit ist ein Maß für die Zuverlässigkeit, mit der Objekte in den Testdaten der gleiche Klasse aufweisen wie im Referenzdatensatz.

³ Diese Stelle wird in der hier verfolgten Vorgehensweise durch eine Stichprobe, die Punkte in zufälliger Verteilung in jedem DLM-Polygon setzt, definiert.

Dafür wird die Anzahl von richtig klassifizierten Objekten im Verhältnis zu allen Objekten dieser Klasse in OSM angegeben. Die Produzenten-Genauigkeit und die Nutzer-Genauigkeit, die auch als completeness bzw. correctness bezeichnet werden, sind bedingte Häufigkeiten und lassen sich wie folgt bestimmen (Heipke et al. 1997):

$$\text{Produzenten-Genauigkeit}_i = \frac{TP_i}{TP_i + FN_i} \quad \text{mit } i = 1, 2, \dots, n, \quad (4)$$

$$\text{Nutzer-Genauigkeit}_i = \frac{TP_i}{TP_i + FP_i} \quad \text{mit } i = 1, 2, \dots, n. \quad (5)$$

3 Methodisches Vorgehen

Um die Qualität von OSM im Vergleich zum Basis-DLM bewerten zu können, sind vorbereitende Schritte notwendig. So besteht eine Voraussetzung darin, ein für beide Datensätze gemeinsames Klassifizierungsschema zu entwickeln. Zudem ist eine Topologieprüfung durchzuführen und durch Eliminierung sich überlappender Flächen sind eindeutig vergleichbare Datensätze zu schaffen. Dies ist regelmäßig eine Herausforderung bei entsprechenden Qualitätsvergleichen, da – insbesondere aus unterschiedlichen Erfassungsmaßstäben resultierende – Überlappungen in OSM häufig sind (Sester et al. 2014). Die Abb. 1 stellt die wesentlichen Schritte der verfolgten Vorgehensweise dar.

3.1 Datenvorbereitung und Schlüsselauswahl

Die Daten aus OSM werden zunächst auf das Gebiet von Lübeck zugeschnitten (»Clipping«) und in das Koordinatensystem des DLM (ETRS 1989 UTM Zone N32) überführt. Zudem erfolgt eine Auswahl auf flächenhafte Objekte mit Schlüsseln, die gemäß OSM-Wiki Informationen über die Landnutzung und Landbedeckung enthalten – namentlich *landuse*, *natural*, *leisure*, *sport*, *waterway*, *railway*, *aeroway*, *highway*, *building*, *parking*, *bicycle* und *historic*.

Die Schlüssel *bicycle* und *building* werden im Folgenden jedoch nicht weiter betrachtet. Dies erfolgt, da alle Objekte, die einen Wert für den Schlüssel *bicycle* aufweisen, im Untersuchungsgebiet ebenso einen Eintrag

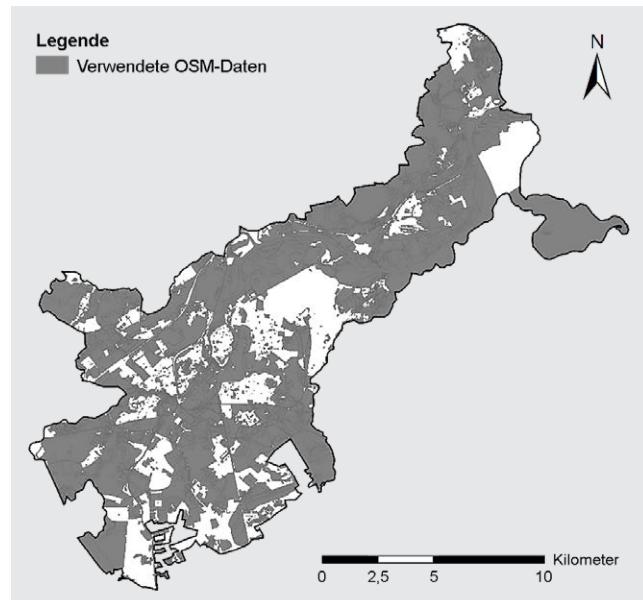


Abb. 2: OSM-Objekte mit LULC-Schlüsseln und mindestens einem Wert; diese gingen in die Untersuchung in Lübeck ein.

für den Schlüssel *highway* haben. Informationen über die Landnutzung – für die im gemeinsamen Klassifizierungsschema die Oberklasse »Straßenverkehr« eingeführt wird – sind für diese Objekte insofern redundant. Objekte mit einem Eintrag im Schlüssel *building* besitzen im Untersuchungsgebiet häufig keine Werte, die die Art des Gebäudes beschreiben. Stattdessen bieten sie i. d. R. aber Werte für andere Schlüssel, v. a. in der Kombination mit dem Tag *landuse = residential*. Objekte der übrigen Klassen, die keinerlei Werte aufweisen, werden ebenfalls nicht verwendet, da sie keinen weiterführenden thematischen Informationsgehalt bieten. Die Abb. 2 zeigt (in grau) diejenigen Objekte, die letztlich Eingang in die Berechnungen fanden.

3.2 Semantische Harmonisierung

Die semantische Harmonisierung führt die Nomenklaturen von OSM und vom Basis-DLM in einem gemeinsamen Klassifizierungsschema zusammen. Sie ist erforderlich, um die Vergleichbarkeit der beiden Datensätze zu gewährleisten. Die unterschiedliche thematische Modellierung von OSM und Basis-DLM macht sich u. a. durch verschiedene Bezeichnungen für LULC-Klassen und eine abweichende Klassenzahl bemerkbar.

Auf dem Weg zu einem gemeinsamen Klassifizierungsschema werden zunächst diejenigen Klassen in beiden Datensätzen identifiziert, die sich inhaltlich entsprechen und somit eindeutig zuzuordnen sind. Solche Klassen bilden jeweils eine Kategorie im

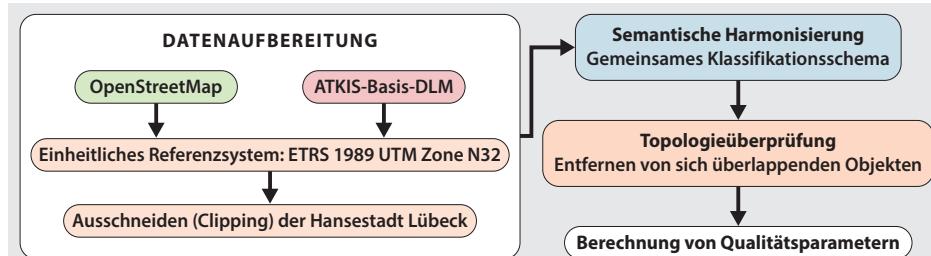


Abb. 1: Methodisches Vorgehen

Tab. 2: Gemeinsames Klassifikationsschema als Ergebnis der semantischen Harmonisierung

Klasse	Klasse
Bahnverkehr	Ackerland
Brücke	Baumschule
Festplatz	Gewächshaus
Flugverkehr	Grünland
Friedhof	Obstplantage
Golfplatz	Düne
Hafen	Fließgewässer
Industrie & Gewerbe	Gehölz
Kleingarten	Grasland
Parken	Heide
Schleuse	Insel
Schwimmbad	Moor
Sonstige urbane Fläche	Naturschutzgebiet
Sportanlage	Röhricht
Stadion	Schutt
Städtische Grünanlage	Stehendes Gewässer
Straßenverkehr	Sumpf
Tagebau	Wald
Wohnen	Unbekannt

gemeinsamen Klassifikationsschema. Anschließend werden OSM-Objekte, die mit einem Tag versehen sind, der im OSM-Wiki nicht vorgesehen ist, dem nächstähnlichen OSM-Wiki-Tag gleichgesetzt. So kennt das Wiki z.B. den Tag *leisure = park*, während die OSM-Daten in Lübeck zusätzlich den Eintrag *landuse = park* enthalten. Zudem werden Tags, die sich inhaltlich nur gering unterscheiden, zusammengefasst. So ist die Bedeutung der Tags *land-use = garden* und *landuse = park*, die verschiedene Arten von Gärten bzw. einen Park darstellen, vor dem Hintergrund des hier verfolgten Anwendungsfalls ähnlich. Beide Tags werden im gemeinsamen Klassifikationsschema zur Kategorie „Städtische Grünanlage“ zusammengeführt. Bei Unklarheiten hinsichtlich der Bedeutung von Tags wird durch Stichproben überprüft, welche DLM-Objekte an der entsprechenden Stelle liegen. So stellt sich z.B. beim Tag *landuse = farmland* die Frage, ob damit ausschließlich Ackerflächen oder auch landwirtschaftliches Grünland kartiert werden. Das Wiki ist an dieser Stelle nicht eindeutig: Während die deutsche Version des Wikis nur Acker als Entsprechung für *farmland* nennt, sieht die englische Version Ackerbau (*tillage*) und Weide (*pasture*) vor (Stand: 12.12.2016). Im gemeinsamen Klassifikationsschema wird zudem eine Restklasse „Unbekannt“ eingeführt. Diese enthält alle OSM-Tags, die nicht eindeutig zuzuordnen sind. Die im Basis-DLM vorhandenen Objektarten konnten in den meisten Fällen anhand der Definitionen im ATKIS-Objektartenkatalog einem gleichwer-

tigen OSM-Tag zugeordnet werden. Einzelne Objektarten, für die im OSM-Wiki kein entsprechender Tag zu finden ist, werden ebenfalls der Klasse „Unbekannt“ zugeordnet.

Die semantische Harmonisierung führt zu dem in Tab. 2 dargestellten gemeinsamen Klassifikationsschema.

3.3 Datenaufbereitung nach topologischen Kriterien

Nach der Harmonisierung wird sowohl für OSM als auch für das Basis-DLM eine Topologieprüfung durchgeführt, um überlagerungsfreie Ausgangsdatensätze zu generieren. Dadurch wird eindeutig, anhand welcher der sich räumlich überlagernden Objekte der Qualitätsvergleich erfolgen soll. Zudem werden sehr kleine Objekte (<5 m²) mit demjenigen Nachbarpolygon (unberücksichtigt der jeweiligen Nutzungen) »verschmolzen«, mit dem sie die längste gemeinsame Grenze teilen. Es wird davon ausgegangen, dass Objekte dieser Größe unbeabsichtigt erfasst worden sind. Die Überprüfung der Topologie ergibt 6.716 sich überlagernde Flächen im OSM-Datensatz, die in Abb. 3 rot dargestellt sind.

Die festgestellten Überlagerungen werden schrittweise nach den folgenden Kriterien bereinigt: Die Entscheidung, welches der sich überlagernden Objekte behalten und welches entfernt werden soll, erfolgt aufgrund von Prioritäten, die den einzelnen LULC-Klassen zugeordnet werden. So wird städtisch geprägte Klassen (in Tab. 2 rosa dargestellt) eine höhere Priorität zugewiesen als Klassen mit nicht-städtischen Nutzungen (grün in Tab. 2). Die Priorität ist anwendungsspezifisch begründet und wurde mit Blick auf das Interesse an Stadtstrukturen – in Regiobranding wird der Begriff Kulturlandschaft breit verstanden und schließt sowohl bebaute als auch unbebaute Landschaften mit ein – gewählt.

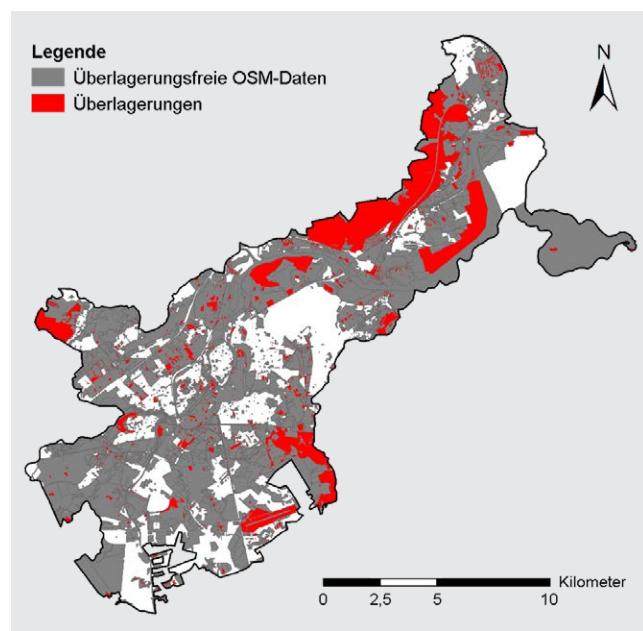


Abb. 3: Flächenhafte Überlagerungen in OSM

Überlagern sich ein städtisch und ein nicht-städtisch geprägtes Objekt, wird die überlagernde Fläche nur vom Objekt mit der nicht-städtischen Nutzung eliminiert. Bei Überlagerungen von zwei städtisch oder zwei nicht-städtischen Objekten wird in einem zweiten Schritt geprüft, welche Detailklasse (städtische Detailklassen sind beispielsweise ‚Wohnen‘ oder ‚Industrie- & Gewerbe‘) diese Objekte repräsentierten. Überlappen sich zwei Polygone derselben Detailklasse, werden die betroffenen Polygone – anders als bei Dorn et al. (2015) – nicht zusammengefasst. Stattdessen wird der sich überlagernde Bereich vom flächenmäßig kleineren Polygon entfernt. Letzteres bleibt bei dieser Herangehensweise als eigenständiges Objekt mit veränderter Geometrie bestehen, während die Geometrie des größeren Polygons gänzlich unverändert bleibt. Ein Zusammenfassen der betroffenen Flächen („merging“) würde dagegen zu einer deutlich geringeren Anzahl von Polygonen gleicher Klassen führen. Für die anschließende Qualitätsprüfung ist dies nicht von Relevanz. Der Vorteil der hier verwendeten Vorgehensweise besteht jedoch darin, dass die ursprüngliche Geometrie des Datensatzes kleinstmöglich verändert wird. Für den konkreten Anwendungsfall bedeutet das, dass der bearbeitete Datensatz für weitere Fachaufgaben – z.B. zur optischen Interpretation von Landnutzungsstrukturen (gemeinsam mit dem Ausgangsdatensatz als Referenz) – genutzt werden kann. Von besonderem Interesse für den Anwendungsfall sind städtische Flächen (insbesondere Wohn- und Gewerbegebiete), Gewässer (insbesondere Fließgewässer), landwirtschaftliche Nutzflächen (Ackerland und Grünland) und Wald, die häufig mit vergleichsweise flächengroßen Polygonen einhergehen. Diese Strukturen sollten nach der Topologiebereinigung möglichst erhalten bleiben, um für fachbezogene Untersuchungen nutzbar zu sein. Deshalb wurden flächengröße Polygone auch bei Überlagerungen von Polygonen mit unterschiedlichen städtischen bzw. mit verschiedenen nicht-städtischen Detailklassen bevorzugt.

Anders als im beschriebenen Fall verfolgen Dorn et al. (2015) den Ansatz, kleinere Polygone zu priorisieren und bevorzugen Landnutzungs- gegenüber Landbedeckungsklassen dort, wo eine Entscheidung allein anhand der Flächengröße nicht möglich ist. Zudem werden Objekte dort bis zu einer Flächengröße von 20 m² (und nicht 5 m²) dem Nachbarpolygon zugeschlagen. Diese unterschiedlichen Vorgehensweisen können sich auf die Ergebnisse der Qualitätsbestimmung auswirken. Dorn et al. (2015) beschreiben in diesem Zusammenhang das Beispiel eines im Wald gelegenen Sees, der bei der hier verfolgten Bereinigungsvariante ggf. aus dem OSM-Datensatz entfernt werden würde: Waldflächen werden häufig aus Luftbildern digitalisiert, wobei kleine Flächen anderer Nutzungen eventuell übersehen werden. Sollte der See zu einem späteren Zeitpunkt noch erfasst werden, würde er das größere Walddatum dann ggf. überlagern und im vorliegenden Fall, aufgrund der Bevorzugung größerer Flächen, entfernt werden.

4 Ergebnisse

4.1 Vollständigkeit

Die Vollständigkeit wurde anhand des von Überlagerungen bereinigten OSM-Datensatzes berechnet. Bei dieser Vorgehensweise ergibt sich gemäß Gl. 1 als Ergebnis eine gut 70-prozentige Bedeckung der Gesamtfläche des Untersuchungsgebiets durch OSM-LULC-Daten (Gl. 6):

$$\text{Vollständigkeit [%]} = \frac{148841869 \text{ m}^2}{211743010 \text{ m}^2} * 100 \approx 70,3 \%. \quad (6)$$

Der erzielte Wert kann als vergleichsweise gut bezeichnet werden. So berechnen Jokar Arsanjani und Vaz (2015) für sieben europäische Stadtregionen Werte zwischen 42 % (Budapest) und 100 % (Bukarest). Berlin als einziger deutscher Vertreter dieser Studie hat demnach eine Vollständigkeit von 61 %. Jokar Arsanjani et al. (2015) ergänzen das Bild, indem sie zusätzlich auch die Stadtregionen von Frankfurt, Hamburg und München untersuchen, die jeweils unter dem Wert für Berlin bleiben. Dorn et al. (2015) ermitteln für die Rhein-Neckar-Region eine Vollständigkeit von 76 %.

Es ist jedoch auf die nur bedingte Vergleichbarkeit der Untersuchungen hinzuweisen. So stellt die Stadt Lübeck ein kleineres und homogeneres Untersuchungsgebiet dar als diejenigen anderer Studien, die in größerem Umfang auch suburbane bis ländlich geprägte Gebiete umfassen. Zudem unterscheiden sich Art und Zahl der ausgewählten Schlüssel in der Lübecker Studie gegenüber früheren Untersuchungen. Jokar Arsanjani und Vaz (2015) blicken z.B. auf die Schlüssel *landuse*, *natural* und *building*, während sich Jokar Arsanjani et al. (2015) auf *landuse* und *natural* beschränken. Bei Dorn et al. (2015) werden dagegen die Schlüssel *landuse*, *leisure*, *natural*, *tourism* und *waterway* analysiert. Den Untersuchungen gemein ist jedoch, dass sie jeweils die Schlüssel *landuse* und *natural* berücksichtigen. Diese haben häufig eine hohe Flächendeckung. Zudem werden Objekte dieser Klassen mitunter großzügig kartiert und überdecken zum Teil flächenkleinere Objekte. Insofern lässt sich erklären, dass Jokar Arsanjani und Vaz (2015), die eine vergleichsweise kleine Zahl an Schlüsseln verwenden, für Bukarest auf eine Vollständigkeit von 100 % kommen.

4.2 Thematische Qualität

Die Konfusionsmatrix, die grundlegend für den Qualitätsvergleich ist, basiert in der vorliegenden Untersuchung auf einer Stichprobe von rund 3.700 Stichprobenpunkten. Diese besteht aus je einem zufällig gesetzten Punkt in jedem Basis-DLM-Polygon, dem sowohl die in OSM als auch die im Basis-DLM eingetragene Klasse zugewiesen wird. Für die Berechnung der Gesamtgenauigkeit werden nur diejenigen zufällig gesetzten Punkte bewertet, an

deren Position sich – neben dem DLM-Objekt – auch ein OSM-Objekt befindet. Dieser Schritt wird damit begründet, dass die in OSM nicht kartierten Gebiete bereits durch den Aspekt der Vollständigkeit negativ in die Qualitätsbeurteilung einfließen. Dieser stichprobenbasierte Vergleich der Klassenzugehörigkeiten in OSM und Basis-DLM führt zu einer Gesamtgenauigkeit von rund 46,6 % (Gl. 7):

$$p_0 [\%] = \frac{1722}{3622} * 100 \approx 46,6 \% . \quad (7)$$

Für den Kappa-Index ergibt sich mit $p_0 \approx 0,466$ und $p_e \approx 0,121$ für die Stadt Lübeck ein Wert von

$$\kappa = \frac{0,466 - 0,121}{1 - 0,121} \approx 0,39 . \quad (8)$$

Landis und Koch (1977) schlagen eine sechsstufige Skala zur Bewertung der Reliabilität von Kappa-Werten vor, die von <0 [= schlechte Übereinstimmung] bis 0,81–1,00 [= (annähernd) vollkommene Übereinstimmung] reicht. Demnach gehört der in Lübeck ermittelte Wert nur zur viertbesten Stufe 0,21–0,40 [= ausreichende Übereinstimmung]. Dabei gilt es aber zu berücksichtigen, dass die Datenaufbereitung und das detailreiche Klassifikations-schema der vorliegenden Untersuchung die Wahrscheinlichkeit von übereinstimmenden Objektklassen beeinflussen können. Dorn et al. (2015) ermitteln mit 0,61 für die Rhein-Neckar-Region einen deutlich höheren Kappa-Index; bei einer allerdings geringeren Anzahl an Klassen. Dort zeigen sich zudem ausgeprägte Unterschiede zwischen den Gemeinden. So finden sich neben Gemeinden mit hoher bzw. sehr hoher Reliabilität auch solche, die vergleichbare oder geringere Kappa-Werte als Lübeck aufweisen. Jokar Arsanjani et al. (2015) berechnen für vier deutsche Stadtregionen Werte von bis 0,41 (Hamburg) bis 0,56 (Frankfurt), allerdings mit Urban Atlas-Daten als Referenz.

Die Tab. 3 zeigt exemplarisch die Produzenten- und Nutzergenauigkeiten für die, bezogen auf den Anwendungsfäll, bedeutsamen LULC-Klassen. Die städtischen Klassen ‚Wohnen‘ und ‚Industrie- & Gewerbe‘ weisen dabei vergleichsweise hohe Produzentengenauigkeiten (PG, und damit eine hohe Vollständigkeit bei der Erfassung) bei jeweils weniger hohen Nutzergenauigkeiten (NG, und damit geringerer Zuverlässigkeit) auf. Im Falle von Wald sind die beiden Genauigkeitsmaße mit gut 60 % ausgeglichen und dabei moderat hoch. Der thematischen Genauigkeit von Wald dürfte zugute kommen, dass bei dieser Klasse keine Detailklassen (Laub-, Nadel-, Mischwald) unterschieden worden sind, die bei einer Erfassung von Luftbildern (zumindest bei deren Aufnahme im Frühjahr bzw. Sommer) die Klassifizierung erschweren könnten. Die thematischen Genauigkeiten der Klassen ‚Ackerland‘, ‚Grünland‘ und ‚Grasland‘ liegen bis auf die NG von Ackerland jeweils – und z.T. deutlich – unter 50 %. Dieser Befund dürfte zum einen in der mitunter schwierigen

Tab. 3: Ergebnisse der Produzenten- und Nutzergenauigkeiten (eine Auswahl)

Klasse	Produzenten-Genauigkeit	Nutzer-Genauigkeit
Ackerland	47,2 %	51,4 %
Fließgewässer	78,9 %	5,8 %
Grasland	7,4 %	35,3 %
Grünland	32,9 %	45,5 %
Industrie & Gewerbe	70,9 %	42,0 %
Wald	62,4 %	61,6 %
Wohnen	88,8 %	59,1 %

Unterscheidbarkeit der Klassen anhand von Luftbildern begründet sein. Zum anderen wurde zuvor bereits erwähnt, dass der Klassifizierungsvorschlag von OSM bezüglich des für landwirtschaftliche Flächen relevanten OSM-Schlüssels *farmland* aufgrund unterschiedlicher Angaben im deutschen und englischen OSM-Wiki nicht eindeutig ist. Eine auffallend geringe NG bei relativ hoher PG weisen die Fließgewässer auf. Dies überrascht umso mehr, als dass Wasser aufgrund seiner leichten Unterscheidbarkeit in früheren Studien vergleichsweise hohe thematische Genauigkeiten erzielt. Die niedrige NG in Lübeck ist jedoch durch ein einziges, überproportional großes Polygon begründet, welches als Fließgewässer klassifiziert worden ist. Dieses Polygon ist als Fehlerfassung anzusehen, da es keine Entsprechung in der Realität aufweist. Ob es ggf. den Stand eines vergangenen Hochwassers wiedergeben soll, aber nicht als solches klassifiziert worden ist, gilt es noch zu prüfen. Dadurch, dass größere Polygone bei der Datenaufbereitung bevorzugt wurden, konnte sich dieses offensichtlich fehlerhafte Polygon bei der Schaffung des eindeutigen Datensatzes durchsetzen und belastet die PG dieser Klasse deutlich.

Die thematischen Genauigkeiten weiterer LULC-Klassen weisen eine Bandbreite von ansprechend hoch (z.B. Kleingarten: PG: 92,6 % und NG: 74,6 %) bis maximal niedrig auf. So haben die nicht-städtischen Klassen ‚Heide‘, ‚Insel‘, ‚Moor‘ und ‚Röhricht‘ sowie die städtischen Klassen ‚Schleuse‘ und ‚Straßenverkehr‘ Genauigkeiten von je 0 %, da für diese Klassen keine Übereinstimmungen bei der Stichprobe gefunden wurden. Dies liegt nicht zuletzt an der geringen Anzahl und Flächengröße mancher dieser Klassen, was das zufällige Auffinden von Übereinstimmungen erschwert. Zudem wurden diese Klassen, die oftmals von anderen – und dabei nach Fläche größeren – Objekten überlagert sind, durch das Priorisieren flächengrößerer Klassen bei der Datenaufbereitung tendenziell benachteiligt.

Auch wenn Unstimmigkeiten i.d.R. auf den OSM-Datensatz zurückzuführen sind, ist zu berücksichtigen, dass das Basis-DLM mit Mindestfassungsgrößen arbeitet. Dagegen obliegt es bei OSM den Mappern, wie genau sie erfassen. So werden in OSM mitunter Objekte kartiert,

die im Basis-DLM nicht erscheinen, da sie kleiner als die Mindest erfassungsgröße sind. Da bei der verfolgten Vorgehensweise das Basis-DLM als Referenz der Wirklichkeit angenommen wurde, werden solche Abweichungen jedoch OSM zur Last gelegt und können sich negativ auf die berechneten Qualitätswerte auswirken.

5 Fazit

LULC-Daten aus OSM weisen in der Hansestadt Lübeck eine Vollständigkeit von rund 70 % auf. Dieser Wert liegt am oberen Ende eines Bereichs, in dem auch die Ergebnisse von OSM-Qualitätsanalysen anderer deutscher Städte bzw. Stadtregionen liegen, die sich thematisch mit Landbedeckungen und Landnutzungen beschäftigen. Dagegen ist die thematische Qualität als eher niedrig einzustufen. So liegt die Gesamtgenauigkeit in Lübeck bei zirka 47 % und der Kappa-Index bei 0,39. Letzterer Wert steht für eine nach Landis und Koch (1977) nur »ausreichende Übereinstimmung« zwischen OSM und dem Basis-DLM. Er liegt zudem unter den Werten, die in anderen deutschen Städten ermittelt worden sind. Beim Vergleich der einzelnen LULC-Klassen sind darüber hinaus große Unterschiede in den Produzenten- und Nutzergenauigkeiten festzustellen.

Dabei gilt es jedoch zu berücksichtigen, dass sich vergleichende Aussagen zu Qualitätsstudien in anderen Regionen nur bedingt treffen lassen, da sich die Methodik der Untersuchungen im Detail unterscheidet. So erfolgte in Lübeck der Vergleich zwischen OSM und dem Basis-DLM zum einen anhand einer vergleichsweise großen Zahl an Klassen, die nicht immer leicht voneinander zu unterscheiden sein dürften, was die relativ niedrige thematische Genauigkeit mit zu erklären vermag. Um OSM mit einem Referenzdatensatz eindeutig vergleichen zu können, werden zum anderen Flächenüberlagerungen – im beschriebenen Fall ebenso wie in anderen Qualitätsstudien, die sich mit LULC in OSM beschäftigen – aufgelöst. Dies erfolgt gemeinhin durch eine an Geometrie und Semantik orientierte Prioritätenfolge, die sich an der Bedeutung für einen gegebenen Anwendungsfall orientiert. Da sich Anwendungsfälle und damit Prioritäten unterscheiden, schränkt dies die Vergleichbarkeit zusätzlich ein. Vor allem aber folgen daraus nicht unwesentliche Veränderungen des OSM-Ausgangsdatensatzes, da im Falle von räumlichen Überlagerungen Objekte geringerer Priorität zugunsten von Objekten höherer Priorität gelöscht werden. OSM-Qualitätseinschätzungen sind insofern stets vor dem Hintergrund solcher Veränderungen zu sehen. Ferner ist zu bedenken, dass zur Beurteilung der thematischen Genauigkeit in vorliegender Untersuchung lediglich Flächen herangezogen wurden, die in OSM kartiert waren. Unvollständigkeit wurde OSM bei Bewertung der thematischen Genauigkeit insofern nicht zur Last gelegt. Das Vorgehen diesbezüglich ist in Veröffentlichun-

gen von potenziellen Vergleichsstudien jedoch nicht in allen Fällen eindeutig beschrieben.

Das Basis-DLM, das in vorliegender Untersuchung als Referenzdatensatz diente, arbeitet mit Mindest erfassungsgrößen. In OSM erfasste Flächen, die der Wirklichkeit entsprechen, aber unter der Erfassungsgröße des Basis-DLMs liegen (und dort deshalb nicht abgebildet sind), werden folglich als Fehler gewertet und belasten die thematische Genauigkeit von OSM. Diese Gegebenheit gilt es bei Qualitätsaussagen zu OSM, die auf das Basis-DLM referenzieren, zusätzlich zu bedenken. Für die Kulturlandschaftsforschung können kleinräumigere Landnutzungen und Landbedeckungen zudem durchaus von Interesse sein, weshalb OSM in diesem Themenfeld eine Datenquelle mit Potenzial darstellen dürfte. Dies gilt insbesondere, wenn man Informationen dieser Art in Zusammenhang mit den Erfassern sieht: Waren es lokale oder externe Mapper, die sich für die Erfassung bestimmter LULC-Objekte verantwortlich zeichnen? Lassen sich aus ihren Interessensgebieten und den erzielten Erfassungsqualitäten Rückschlüsse auf die interne oder externe Wertschätzung einer Kulturlandschaft bzw. bestimmter Kulturlandschaftselemente ziehen? Diese Fragen gilt es in zukünftigen Forschungen zu beantworten. Die LULC-Daten von OSM sind dabei nur ein Aspekt; weitere Datenthemen der vielseitigen Informationsquelle OSM könnten sich hierfür zusätzlich als wertvoll erweisen. Für die grundlegenden und vergleichenden Untersuchungen, die Regiobranding in verschiedenen Bundesländern vornimmt, sind die verlässlichen Geobasisdaten der amtlichen Vermessungsverwaltungen derzeit allerdings nicht zu ersetzen. Auch für die Zukunft erscheint weniger ein »entweder oder«, sondern ein »sowohl als auch« der Nutzung von beiden Datensätzen naheliegend.

Abschließend sei noch auf die Bedeutung des Faktors Zeit für Datenqualitäten von OSM aufmerksam gemacht: Nach Abschluss der Untersuchung und während des Schreibens des Artikels konnten die Autoren augenscheinliche Verbesserungen der Datenqualität im Untersuchungsgebiet feststellen. So wurden zum einen zusätzliche LULC-Objekte durch OSM-Mapper nach Juni 2015 kartiert. Zum anderen erfolgten nachträgliche Wertzuweisungen bei Objekten, für die zur Untersuchungszeit nur die Geometrie, nicht aber die Semantik erfasst wurden war.

Dank

Das Projekt »Regiobranding – Branding von Stadt-Land-Regionen durch Kulturlandschaftscharakteristika« (FKZ 033L121AN) wird durch das BMBF im Förderschwerpunkt Nachhaltiges Landmanagement gefördert. Darüber hinaus bedanken wir uns bei zwei anonymen Gutachtern für ihre wertvollen Hinweise.

Literatur

- Antoniou, V., Schlieder, C. (2014): Participation Patterns, VGI and Gamification. In: Proceedings of the AGILE, Castellon, S. 3–6.
- AdV – Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (2015): Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens (GeoInfoDok), Version 6.0.1. www.adv-online.de/AdV-Produkte/Geotopographie/Digitale-Landschaftsmodelle/Basis-DLM.
- AdV – Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (2016): Produktblatt Digitales Basis-Landschaftsmodell (Basis-DLM). www.adv-online.de/AdV-Produkte/Standards-und-Produktblaetter, Stand 12/2016.
- Arnold, S., Kurstedt, R., Riecken, J., Schlegel, B. (2017): Paradigmenwechsel in der Landschaftsmodellierung – von der Tatsächlichen Nutzung hin zu Landbedeckung und Landnutzung. In: zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement, Heft 1/2017, 142. Jg., S. 30–37. DOI: 10.12902/zfv-0152-2016.
- Blain, C., Levy, S.E., Brent Ritchie, J.R. (2005): Destination Branding: Insights and Practices from Destination Management Organizations. In: Journal of Travel Research: 4/2005, S. 328–338.
- Burley, T.M. (1961): Land use or land utilization? In: Professional Geographer 6, S. 18–20.
- Caprioli, M., Scognamiglio, A., Strisciuglio, G., Tarantino, E. (2003): Rules and standards for spatial data quality in GIS environments. In: Proceedings 21st International Cartographic Conference, Durban, S. 10–16.
- Clawson, M., Steward, C.L. (1965): Land use information. A critical survey of U.S. statistics including possibilities for greater uniformity. Baltimore.
- Cohen, J. (1960): A coefficient of agreement for nominal scales. In: Educational and Psychological Measurement 20, S. 37–46.
- Dorn, H., Törnros, T., Zipf, A. (2015): Quality Evaluation of VGI Using Authoritative Data – A Comparison with Land Use Data in Southern Germany. In: ISPRS International Journal of Geo-Information 3, S. 1657–1671.
- Estima, J., Painho, M. (2013): Exploratory analysis of OpenStreetMap for land use classification. In: Proceedings of the second ACM SIGSPATIAL international workshop on crowdsourced and volunteered geographic information, Orlando, S. 39–46.
- Foody, G.M. (2002): Status of land cover classification accuracy assessment. In: Remote sensing of environment 1, S. 185–201.
- Hansestadt Lübeck (2017): Einwohner/-innen nach Stadtteilen 1990–2016 (Stand 31.12.2016), www.luebeck.de/stadt_politik/statistiken.
- Heipke, C., Mayer, H., Wiedemann C., Jamet, O. (1997): Evaluation of automatic road extraction. In: International Archives of Photogrammetry and Remote Sensing 3, S. 35–46.
- Internationale Organisation für Normung (ISO): ISO 19157:2013 (Geoinformation – Datenqualität), www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32575, letzter Zugriff 9/2016.
- Johnson, B.A., Lizuka, K. (2016): Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping: Case study of the Laguna de Bay area of the Philippines. In: Applied Geography 67, S. 140–149.
- Jokar Arsanjani, J., Vaz, E. (2015): An assessment of a collaborative mapping approach for exploring land use patterns for several European metropolises. In: International Journal of Applied Earth Observation and Geoinformation 35, S. 329–337.
- Jokar Arsanjani, J., Mooney, P., Zipf, A., Schauss, A. (2015): Quality Assessment of the Contributed Land Use Information from OpenStreetMap versus Authoritative Datasets. In: Jokar Arsanjani, J., Mooney, P., Zipf, A., Helbich, M. (Hrsg.): OpenStreetMap in GIScience, Cham, S. 37–58.
- Kempa, D., Herrmann, S. (2014): Regiobranding – Kulturlandschaft als Identitätsträger und Marke einer Region. In: Unimagazin – Forschungsmagazin der Leibniz Universität Hannover 3/4, S. 2–5.
- Kurstedt, R. (2013): ATKIS® Basis-DLM – fachliche Betrachtung amtlicher Geobasisdaten und deren bundesweite Nutzung, In: Meinel, G., Schumacher, U., Behnisch, M. (Hrsg.): Flächennutzungsmonitoring V – Methodik, Analyseergebnisse, Flächenmanagement, S. 141–146.
- Laggner, A. (2016): Das ATKIS-Basis-DLM in der Emissionsberichterstattung. In: zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement, Heft 3/2016, 141. Jg., S. 178–183. DOI: 10.12902/zfv-0109-2016.
- Landesamt für Vermessung und Geoinformation Schleswig-Holstein (LVerMGeo SH 2017): Produktkatalog Ausgabe 2017. Kiel.
- Landis, J.R., Koch, G.G. (1977): The measurement of observer agreement for categorical data. In: Biometrics 33, S. 159–174.
- Meinel, G., Knop, M. (2008): Geobasisdaten in Deutschland – Verfügbarkeit und Qualitätsaspekte des ATKIS® Basis-DLM und der DTK25-(V). In: Schrenk, M., Popovich, V.V., Engelke, D., Elise, P. (Hrsg.): Tagungsband 13. Internationale Konferenz zu Stadtplanung, Regionalentwicklung und Informationsgesellschaft (Real CORP) in Wien, S. 571–582.
- Mondzech, J., Sester, M. (2011): Quality Analysis of OpenStreetMap Data Based on Application Needs. In: Cartographica: The International Journal for Geographic Information and Geovisualization 2, S. 115–125.
- Mooney, P., Corcoran, P., Winstanley, A. (2010): A study of data representation of natural features in OpenStreetMap. In: Proceedings of the 6th GIScience International Conference on Geographic Information Science, S. 150.
- Neis, P., Zielstra, D., Zipf, A. (2013): Comparison of Volunteered Geographic Information Data Contributions and Community Development for Selected World Regions. In: Future Internet, S. 282–300.
- OpenStreetMap (OSM 2016): www.openstreetmap.org, letzter Zugriff 12/2016.
- Schaffert, M., Steensen, T., Wenger, F.C. (2016): Landbedeckungsveränderungen in der Metropolregion Hamburg (1990–2006). Ein räumlich-zeitlicher Vergleich als Beitrag zur Charakterisierung von Kulturlandschaften. In: FuB – Flächenmanagement und Bodenordnung, 4/2016, S. 149–159.
- Schoof, M., Behncke, K., Ehlers, M. (2011): ATKIS-Basis-DLM und OpenStreetMap – Ein Datenvergleich anhand ausgewählter Gebiete in Niedersachsen. In: FOSSGIS-Konferenzbeiträge, S. 118–125.
- Sester, M., Jokar Arsanjani, J., Klammer, R., Burghardt, D., Haunert, J.-H. (2014): Integrating and generalising volunteered geographic information. In: Burghardt, D., Duchene, C., Mackaness, W. (Hrsg.): Abstracting geographic information in a data rich world. Lecture notes in geoinformation and cartography series, Berlin, S. 119–155.
- Story, M., Congalton, R.G. (1986): Accuracy assessment: a user's perspective. In: Photogrammetric Engineering and remote sensing 3, S. 397–399.
- Vaz, E., Jokar Arsanjani, J. (2015): Crowdsourced mapping of land use in urban dense environments: An assesment of Toronto. In: The Canadian Geographer 2, S. 246–255.

Kontakt

M.Sc. Alessa Retat

Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN) – Landesvermessung und Geobasisinformation

Podbielskistraße 331, 30659 Hannover

alessa.retat@lgn.niedersachsen.de

(bis 09/2016 Studium am Geodätischen Institut der Leibniz Universität Hannover)

Dr.-Ing. Markus Schaffert

Leibniz Universität Hannover

Geodätisches Institut

Nienburger Straße 1, 30167 Hannover

schaffert@gih.uni-hannover.de

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