



Master Thesis

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

zum Thema

„Bewertung der Datenqualität der OpenStreetMap als Datengrundlage für einen Verkehrsgraph mit Hilfe offener Daten und Software“

am Anwendungsbeispiel Einzugsbereiche von Park-and-Ride
Anlagen im südlichen Wiener Umland

vorgelegt von

DI Stephanie Kirchmayr-Novak

U1493, UNIGIS MSC Jahrgang 2012

Zur Erlangung des Grades

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

Gutachter:

Ao. Univ. Prof. Dr. Josef Strobl

Wien, 29.Dezebmer 2014

Erklärung

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

Inhalt

1	Kurzfassung	9
2	Abstract	10
3	Einführung	11
3.1	Motivation.....	11
3.2	Aufgabenstellung	11
3.2.1	Raumplanerischer Hintergrund kontextbezogene Forschungsfragen	11
3.2.2	Forschungsfragen der Arbeit und erwartete Ergebnisse	12
3.3	Struktur der Thesis	15
4	Begriffsdefinition	16
4.1	Geographische Daten (Geodaten)	16
4.2	Datenqualität	16
4.3	Offene Daten (Open Data) und Open Data Lizenzen	16
4.4	Open Source Software und freie Software	16
4.5	Einzugsbereich	16
5	Einführung in die OpenStreetMap	18
5.1	Allgemeines.....	18
5.1.1	Datenherkunft und Qualitätskontrolle	18
5.1.2	Lizenzbedingungen	19
5.2	Das Datenmodell.....	19
5.2.1	Gemeinsame Attribute der OSM Objekttypen	19
5.2.2	Objekttypen Nodes, Ways und Relations	20
5.3	Attribuierung der OSM mit Relevanz für Straßengraphen.....	22
5.3.1	Klassifizierung von Straßentypen.....	23
5.3.2	Einbahnen	23
5.3.3	Maximalgeschwindigkeiten	24
5.3.4	Abbiegerelationen	24
5.3.5	Lichtsignalanlagen.....	25
5.4	Datenextraktion aus OSM	25
6	Werkzeuge und Methoden zur Konvertierung von OSM-Rohdaten in einen Verkehrsgraph	27
6.1	Auswahl der Open Source Werkzeuge.....	27
6.2	Open Source Lizenzierung.....	29
6.3	Osmosis.....	30

6.4	osm2po	31
6.4.1	Konvertierung von OSM-Daten in eine routenfähige SQL-Datenbank	31
6.4.2	Umsetzung von Abbiegevorschriften in osm2po	32
6.4.3	Umsetzung von Lichtsignalanlagen in osm2po	32
6.4.4	Ergebnisse der Konvertierung.....	33
6.5	pgRouting	34
6.6	Zusammenfassung der Arbeitsschritte zur Erstellung eines Verkehrsgraphs	35
7	Bestimmung der Qualität von Geodaten	37
7.1	Datenmodell und Anwendungsmodell	37
7.2	Qualitätsmodelle	38
7.3	Qualitätsmerkmale.....	38
7.3.1	Logische Konsistenz	39
7.3.2	Vollständigkeit	41
7.3.3	Lagegenauigkeit	42
7.3.4	Thematische Genauigkeit	42
7.3.5	Zeitliche Genauigkeit	43
7.3.6	Benutzerfreundlichkeit	43
7.3.7	Beziehungen zwischen Datenqualitätselementen.....	43
7.3.8	Evaluierungsprozess zur Bestimmung der Datenqualität	44
8	Evaluierungsansätze für die Analyse der der Datenqualität von OSM-Verkehrsgraphen in der Literatur	46
8.1	Logische Konsistenz.....	47
8.1.1	Topologische Konsistenz	47
8.2	Vollständigkeit.....	47
8.2.1	Vollständigkeit des Netzwerkes	47
8.2.2	Vollständigkeit der Attribute	49
8.3	Lagegenauigkeit	50
8.4	Thematische Genauigkeit.....	50
8.5	Zeitliche Genauigkeit.....	50
8.6	Zusammenfassung der ausgewählten Evaluierungsansätze	51
9	Evaluierung der Datenqualität des Verkehrsgraphs	53
9.1	Spezifikation der Evaluierungsschritte und -stufen	53
9.2	Evaluierung der logischen Konsistenz	56
9.2.1	Formatkonsistenz.....	56
9.2.2	Konzeptuelle Konsistenz	57
9.2.3	Konsistenz der Wertebereiche.....	58
9.2.4	Topologische Konsistenz	60

9.3	Vollständigkeit.....	66
9.3.1	Vollständigkeit der Objekte - Untersuchung möglicher Untererfassung.....	66
9.3.2	Vollständigkeit der Attribute	69
9.4	Thematische Genauigkeit der Kantengeschwindigkeit	73
9.5	Benutzerfreundlichkeit	80
9.6	Zusammenfassung der Evaluierung und Resümee	81
10	Berechnung der Einzugsgebiet von Park-and-Ride Anlagen.....	84
10.1	Aufbereitung der Datengrundlagen	84
10.1.1	Erhebung der Park-and-Ride Anlagen im Untersuchungsgebiet	84
10.1.2	Verortung der Park-and-Ride-Anlagen im Straßengraph.....	85
10.1.3	Berechnung der Fahrzeiten von den Park-and-Ride Anlage zu jedem Netzwerkknoten	85
10.1.4	Extraktion der Park-and-Ride Anlage mit der kürzesten Fahrzeit.....	86
10.2	Fahrzeitbasierte Einzugsbereiche berechnet mit Voronoi-Polygonen.....	87
10.2.1	Voronoi-Polygone	87
10.2.2	Methodik.....	87
10.2.3	Ergebnisse	87
10.3	Fahrzeitisochronen basierend auf Interpolation mit "Regular Spline with tension" (RST)	90
10.3.1	Interpolation mit „Regular Spline with tension“ (RST)	90
10.3.2	Methodik.....	90
10.3.3	Ergebnisse	91
11	Schlussfolgerungen und Ausblick	93
12	Literatur	94

Abbildungsverzeichnis

Abbildung 1:	Untersuchungsregion Wien und südliches Wiener Umland	13
Abbildung 2:	Struktur der Arbeit als Übersicht	15
Abbildung 3:	Vereinfachte Darstellung des OpenstreetMap-Datenmodell.....	20
Abbildung 4:	Darstellung von physischen Kreuzungspunkten und planfreien Kreuzungen in OSM.....	21
Abbildung 5:	Karte der in der OSM verzeichneten Geschwindigkeitslimits (max-speed).....	24
Abbildung 6:	Beispiel für eine Abbiegerelation in der OSM.....	25
Abbildung 7:	Beispiel für die OSM-Modellierung einer einfachen und einer komplexen Kreuzung mit Lichtsignalanlagen	25
Abbildung 8:	Arbeitsschritte zur Erstellung eines Verkehrsgraphs aus der OSM	36
Abbildung 9:	Beschreibung der Datenqualität in Bezug auf den Anwendungsfall	37
Abbildung 10:	Struktur und Zusammenhang der Elemente des Qualitätsmodells für Geodaten für ein ausgewähltes Qualitätsmerkmal	38

Abbildung 11: Qualitätsmerkmale und ihre Qualitätsparameter für Geodaten nach ISO 19157	39
Abbildung 12: Beispiel für Objekte, die nicht den Regeln des konzeptuellen Schemas entsprechen	40
Abbildung 13: Unterscheidung von Evaluierungsmethoden und Evaluierungsumfang	45
Abbildung 14: Reihenfolge der Evaluierungsschritte zur Bestimmung der Datenqualität	45
Abbildung 15: Unterschiede in der Modellierung der OSM (schwarze unterbrochene Linie) und dem GIP (dicke graue Linie)	48
Abbildung 16: Ergebnis des Kürzeste Wege Testroutings mit pgRouting zum Nachweis der Formatkonsistenz	57
Abbildung 17: Potenzielle topologische Fehler der OSM im Straßengraph Wien und südliches Wiener Umland	65
Abbildung 18: Anteil der durch den Straßengraph in einem Umkreis von 50 m erschlossenen Siedlungsfläche	68
Abbildung 19: Gegenüberstellung von Straßen der Straßen einfachsten Ausbauzustandes ("unclassified", strichliert) und Straßen unbekannter Kategorie ("road" in grün) und deren Lage im Siedlungsgebiet (flächig, rot).....	70
Abbildung 20: OSM Straßengraph nach Straßenkategorien	71
Abbildung 21: Karte Thematische Genauigkeit – Klassifikation der Straßen. Autobahnen und Schnellstraßen	74
Abbildung 22: Beispiele für die Fehlklassifikation von Autobahnen und Schnellstraßen im OSM-Graph	75
Abbildung 23: Beispiele für inkorrekte und korrekte Attribuierung von Straßenabschnitten	77
Abbildung 24: Karte der potenziellen Fehler der Geschwindigkeitsattribution	78
Abbildung 25: Testroutings der kürzesten Wege, Ergebnisvergleich OSM/osm2po/pgRouting und ITS Vienna Region	79
Abbildung 26: Karte der Park-and-Ride Anlagen in Wien und im südlichen Wiener Umland	84
Abbildung 27: Voronoi Polygone für die Netzwerkknoten mit Fahrzeitangabe zur zeitlich nächstgelegenen Park-and-Ride Anlage	87
Abbildung 28: Einzugsgebiete der Park-and-Ride Anlagen in Wien und im südlichen Wiener Umland10756	89
Abbildung 29: Anzahl der Einwohner im Einzugsgebiet der Park-and-Ride Anlagen im südlichen Wiener Umland	89
Abbildung 30: Einzugsgebiete der Park-and-Ride Anlagen in Wien und im südlichen Wiener Umland	92

Tabellenverzeichnis

Tabelle 1: Gemeinsame Attribute aller OSM Objekttypen	19
Tabelle 2: Straßenkategorien der OSM	23
Tabelle 3: Werkzeuge und Methoden zur Konvertierung von OSM-Rohdaten in eine Geodatenbank	27
Tabelle 4: Lizenzen der im Rahmen der Arbeit verwendeten Werkzeuge	28
Tabelle 5: Verwendete Open Data Lizenzen	29
Tabelle 6: Auswahl und Übersetzung der OSM-Straßendaten (Tags) in osm2po	32
Tabelle 7: Attribute des Verkehrsgraphs und ihre Datentypen	33
Tabelle 8: Übersicht über die in der Literatur angewandten Evaluierungsansätze für die OSM	51
Tabelle 9: Spezifikation der Datenqualitätsmerkmale, Maßeinheiten und Indikatoren	53

Tabelle 10: Ergebnisse der Bewertung der Formatkonsistenz.....	57
Tabelle 11: Konzeptuelle Anforderungen an einen routenfähigen Graph und Entsprechung des Exportdatensatzes von osm2po	58
Tabelle 12: Ergebnisse der Bewertung der konzeptionellen Konsistenz	58
Tabelle 13: Konzeptuelle Anforderungen an einen routenfähigen Graph und Entsprechung des Exportdatensatzes von osm2po	59
Tabelle 14: Ergebnisse der Bewertung der Konsistenz der Wertebereiche.....	60
Tabelle 15: Gefundene potenzielle topologische Fehler (Toleranzgrenze 0,00001 Grad WGS84)	61
Tabelle 16: Potenzielle und gefundene topologische Fehler	64
Tabelle 17: Ergebnisse der Bewertung der topologischen Konsistenz	64
Tabelle 18: Fläche und Bevölkerung der 271 Gemeinden und 194 Siedlungseinheiten (SE) im Untersuchungsgebiet	67
Tabelle 19: Erschlossene Siedlungseinheiten, Fläche	67
Tabelle 20: Ergebnisse der Bewertung der Vollständigkeit des Netzgraphs.....	68
Tabelle 21: Straßenkategorien im OSM-Straßengraph: Straßenlänge und Anteil am Straßengraph	70
Tabelle 22: Straßen ohne Straßennamen nach Straßenkategorien im OSM-Straßengraph	72
Tabelle 23: Ergebnisse der Bewertung der Vollständigkeit der Attribute	73
Tabelle 24: Länge des Straßennetzes nach Straßenkategorie und Kantengeschwindigkeit in km/h.....	76
Tabelle 25: Ergebnisse der Bewertung der thematischen Genauigkeit der Attribute	80
Tabelle 26: Verfügbarkeit von Informationen für die im Rahmen der Arbeit verwendeten Open Source Produkte	81
Tabelle 27: Ergebnisse der Bewertung der Benutzerfreundlichkeit	81
Tabelle 28: Zusammenfassung der Ergebnisse der Evaluierung des Straßengraphs	82
Tabelle 29: Fläche und Bevölkerung im Einzugsgebiet (EZG) der Park-and-Ride Anlagen (Voronoi Polygone).....	88
Tabelle 30: Durch Park-and-Ride Anlagen erschlossene Fläche und Bevölkerung nach Fahrzeitklassen	92

Verzeichnis der verwendeten Codes

Code 1: Lichtsignalanlage im OSM-Schema (XML-Code)	20
Code 2: Straße im OSM-Schema (XML-Code)	21
Code 3: Relation eine administrative Grenze im OSM-Schema (XML-Code).....	22
Code 4: Osmosis-Befehle zur Extraktion von OSM-Daten für den Verkehrsgraph	30
Code 5: osm2po Befehl zur Konvertierung der OSM-Daten in einen routenfähigen Verkehrsgraph	33
Code 6: SQL Code zur Berechnung des kürzesten Weges zu multiplen Zielen	34
Code 7: SQL-Code zur Berechnung des kürzesten Weges zwischen zwei Knoten zum Nachweis der Formatkonsistenz	56
Code 8: SQL-Befehl zur Prüfung der konzeptuellen Konsistenz.....	57
Code 9: SQL-Code zur Erstellung der Topologie und Grobanalyse des erstellten Graphs	61
Code 10: SQL Code zur Erstellung einer Distanzmatrix vom Knoten 18466 zu allen 68471 Knoten des Straßengraphs	62
Code 11: SQL-Code zur Ermittlung der planfreien Kreuzungen.....	63

Code 12: SQL-Code zur Extraktion der Schleifen des Straßengraph	63
Code 13: SQL-Code zur Identifikation des räumlich nächsten Netzwerkknotens des Straßengraphs um die Park-and-Ride Anlage	85
Code 14: SQL-Code von pgRouting zur Berechnung der Fahrzeit des kürzesten Weges zwischen zwei Punkten (pgr_drivingDistance)	85
Code 15: SQL-Code zur Berechnung der Fahrzeiten von einer Park-and-Ride Anlage zu jedem Netzwerkknoten des Verkehrsgraphs	86
Code 16: SQL-Code zur Extraktion der Park-and-Ride Anlage mit der kürzesten Fahrzeit	86

Abkürzungsverzeichnis

bzw.	beziehungsweise
DBMS	Datenbankmanagement
etc.	et cetera
GIS	Geoinformationssystem
GPLv2 license	GNU General Public License, version 2
GPS	Global Positioning System
OGC	Open Geospatial Consortium
OSM	OpenStreetMap
SQL	Structured Query Language
SRID	Spatial Reference Identifier
VGI	Volunteered geographic information
WGS 84	World Geodetic System 1984
z.B.	zum Beispiel

Glossar

Open Source Software	Software zur freien Nutzung, Vervielfältigung, Weitergabe und Modifikation durch den Nutzer
Freeware	Software zur freien Nutzung, die jedoch durch den Nutzer nicht, oder nur eingeschränkt modifiziert werden kann, da der Quellcode (Source Code) nicht vollständig zur Verfügung gestellt wird.
osm2po	Freeware zur Konvertierung von XML-Daten der OpenStreetMap in einen routenfähigen Verkehrsgraph
pgRouting	Open Source Software für Routings und Netzwerkanalysen
QGIS	Open Source Desktop Geoinformationssystem (ehemals Quantum GIS)
Wiki OpenStreetMap	Online Nachschlagewerk der OpenStreetMap
ws_2po_4pgr	SQL-Datei mit den Kanten des durch osm2po erzeugten ,Straßengraphs

1 Kurzfassung

Die gegenständliche Arbeit führt eine intrinsische Qualitätsevaluierung eines auf Basis der OpenStreetMap erstellten routingfähigen Straßenverkehrsgraphs durch. Für die Evaluierung werden ausschließlich offene Daten und Open Source Werkzeuge verwendet. Die Bewertung bezieht sich auf den konkreten Anwendungsfall „Berechnung von Einzugsbereichen von Park-and-Ride Anlagen im südlichen Wiener Umland für den motorisierten Individualverkehr“. Die intrinsische Bewertung basierend auf den Elementen des zu prüfenden Datensatzes, wird durch extrinsische Stichprobenprüfungen mit frei verfügbaren Referenzdaten ergänzt, wie z. B. der basemap.at, der offiziellen Verwaltungsgrundkarte von Österreich oder des online verfügbaren Routenplaners AnachB von ITS Vienna Region.

Nach Diskussion der Qualitätsmerkmale für Geodaten nach ISO-Standard 19157:2011 und den bislang für die OpenStreetMap durchgeführten intrinsischen und extrinsischen Qualitätsanalysen, wird ein anwendungsspezifischer intrinsischer Evaluierungsprozess aufgesetzt und Qualitätsanforderungen für jedes Qualitätsmerkmal definiert.

Insgesamt konnten vier Qualitätsmerkmale intrinsisch untersucht und bewertet werden. Geprüft wurden die logische Konsistenz, die Vollständigkeit der Objekte und der Attribute, die thematische Genauigkeit sowie die Benutzerfreundlichkeit.

Die Ergebnisse zeigen, dass die Extraktion eines routenfähigen Straßengraphs auf Basis der OpenStreetMap mit Hilfe von Open Source und Freeware Werkzeugen in einer Form möglich ist, die Routing- und Netzwerkanalysen erlaubt. Der aus der OpenStreetMap erstellte Verkehrsgraph ist logisch konsistent, die Benutzerfreundlichkeit ist durch ausreichende Dokumentation zu den Daten der OpenStreetMap sowie der Funktionalitäten der Open Source Software und Freeware gegeben. Die Vollständigkeit der Objekte und Attribute entspricht den Anforderungen an einen makroskopischen Verkehrsgraph: Im hochrangigen Straßennetz sind alle Straßen verfügbar und korrekt attribuiert und das niederrangige Straßennetz ist insofern vollständig, als alle Siedlungseinheiten für die Berechnung der Einzugsbereiche erschlossen sind. Insgesamt lässt die stichprobenartige visuelle Kontrolle der Vollständigkeit der Objekte darauf schließen, dass das Straßennetz in der Untersuchungsregion umfassend verfügbar ist. Durch die intrinsische Prüfung der thematischen Genauigkeit kann aufgezeigt werden, welche Einstellungen bei der Konvertierung der OpenStreetMap-Daten benutzerdefiniert adaptiert werden müssen, um den Straßenverkehrsgraph für die Anwendung zu optimieren. Dabei handelt es sich um die standardmäßige Geschwindigkeitsattribuierung von Autobahnauffahrten sowie den Ausschluss von „Straßen einfachsten Ausbauzustandes“ (zumeist asphaltierte Wirtschaftswege), welche kaum verkehrliche Bedeutung haben und daher nicht in den Verkehrsgraphen aufgenommen werden sollten.

Die Umsetzung des Anwendungsfalls zeigt, dass der OSM-Straßengraph für die Berechnung von Einzugsbereichen mit Voronoi-Polygonen sehr gut geeignet ist. Vor der Verwendung des Straßengraphs für detailliertere Netzwerkanalysen, wie z.B. der Berechnung von Isochronen um die Park-and-Ride Anlagen, sind jedoch zusätzliche Adaptionen bzw. Kalibrierungen des Verkehrsgraphen erforderlich, um mit den Ergebnissen keine Scheingenauigkeiten zu kommunizieren. Zu diesen Adaptionen gehört insbesondere die Berücksichtigung von Wartezeiten an Ampelanlagen, welche in der OpenStreetMap erfasst sind, vom Konvertierungsprogramm standardmäßig jedoch nicht berücksichtigt werden.

Generell kann bei der intrinsischen Bewertung des Verkehrsgraphs bemängelt werden, dass wichtige Parameter des Straßenverkehrsgraphs wie die Korrektheit der Abbiegerelationen und Einbahnregelungen ohne einen routenfähigen Referenzstraßengraph nicht evaluiert werden können. Dieser Mangel ist jedoch hinsichtlich der Anforderungen an einen Straßenverkehrsgraph zur Erzeugung von Einzugsbereichen für Park-and-Ride Anlagen vernachlässigbar. Für kleinräumigere Routings und Netzwerkanalysen muss der erstellte Straßengraph jedoch durch den Nutzer weiter adaptiert werden.

2 Abstract

This work assesses the quality of a street traffic graph elaborated based on data of the OpenStreetMap (OSM) applying direct intrinsic methods. The use case defining the quality requirements of the task is calculating catchment areas of park-and-ride facilities in the Southern region of Vienna. The evaluation and the data used for the task are restricted to open data and open source software, respectively freeware. In order to gain more insight into the strengths and weaknesses of the data, the main intrinsic approach is complemented by an extrinsic evaluation for randomly selected sample areas, again using only freely available open data resources.

After the discussion of quality parameters defined by ISO-Standard 19157:2011, the work presents the intrinsic and extrinsic evaluation methods heretofore applied to assess the data quality of the OpenStreetMap. Based on this knowledge, a user specific evaluation process is designed determining the quality requirements for each quality character intrinsically assessable.

In total four quality characteristics are evaluated. These are the logical consistency, the completeness of objects and attributes, the thematic accuracy as well as the usability. No statements are made for positional accuracy and temporal quality.

The results show, that the extraction of routable street traffic graphs based on OSM data is possible with Open Source and Freeware tools. The street traffic graph is logically consistent. The usability is ensured by a broad documentation of the OSM data as well as the software tools applied. The completeness of the objects conforms to the quality requirements set for a macroscopic street traffic graph: the main street network is represented and all settlement areas - i.e. inhabitants of the region - are connected to the graph. A visual control of the completeness of objects against the official and freely available transport infrastructure maps implies that the street network of the region is correctly represented.

The assessment of the thematic accuracy shows that the street traffic graph holds plausible speed limits. It further reveals which settings should be adapted by the user during conversion of the OSM-data, in order to optimize the graph for the use-case. These adaptations include the adjustment of the default speed attribution for slip roads as well as the exclusion of "streets with least traffic relevance".

Finally the implementation of the use-case illustrates, that the street traffic graph is adequate for calculating catchment areas with Voronoi-Polygons. Before using the graph for more detailed analysis, as for example the designation of distinct driving time isochrones, it is advised to further adapt the traffic graph. Particularly, delays for stops at traffic signals should be included into the calculation of driving times in order to display more reliable results, especially in urban areas. Traffic signals are included in the OSM data model, but the conversion program used does not consider them by default.

The intrinsic evaluation of the street traffic graph is well applicable for the case of macroscopic traffic modelling and the selected use case. It does however not include the assessment of parameters like the accuracy of turn restrictions, one-way regulations or locally specific speed restrictions, which require the comparison to a reference traffic graph. These weaknesses are not of significant relevance for the use case, but have to be considered for small scale network analysis and routing applications.

3 Einführung

3.1 Motivation

Für viele Bereiche der Stadt- und Regionalplanung, zum Beispiel der Standortbewertung, ist sehr hilfreich Erreichbarkeiten und Erreichbarkeitspotenziale für bestimmte Standorte bzw. Einrichtungen berechnen zu können. Eine dafür wesentliche Grundlage ist ein Verkehrsgraph. Die Anschaffung von kommerziellen Verkehrsgraphen (z. B. Produkte von Navteq) und Verkehrsmodellsoftware (z. B. von PTV¹) ist mit sehr hohen Kosten verbunden und die derzeit verfügbaren offiziellen Datengrundlagen (z. B. die Graphenintegrations-Plattform GIP) stehen nicht allen öffentlichen Einrichtungen zur Verfügung bzw. sind für private Planungsbüros nicht zugänglich.

Demgegenüber steht mit der OpenStreetMap eine umfangreiche Geodatenbasis zur freien Nutzung zur Verfügung, in der auch das Verkehrswegenetz mit vielen Attributen enthalten ist. Ferner werden viele GIS-Funktionalitäten, darunter Netzwerkanalysen, bereits als Open Source Software angeboten.

Ziel der vorliegenden Arbeit ist es daher, die Einsatzfähigkeit dieser offenen Daten mit Hilfe von Open Source Software zu prüfen und anhand eines konkreten Anwendungsbeispiels zu testen.

3.2 Aufgabenstellung

Für das südliche Wiener Umland soll die Qualität eines aus der OpenStreetMap erzeugten Verkehrsgraphs bewertet werden. Da die finanziellen Mittel für die Studie beschränkt sind und weder ein Verkehrsmodell der Region noch eine Lizenz für GIS-Software zur Verfügung stehen, soll die Aufgabe mit Hilfe von "Open Source Software" und „Open Data Ressourcen“ gelöst werden. Diese Restriktion schränkt sowohl die verwendbaren Funktionalitäten, als auch die Verfügbarkeit von Referenzdatensätzen zur Bewertung des Verkehrsgraphs ein.

Dazu wird zunächst die Qualität und Eignung der Geodaten der OpenStreetMap zum Aufbau eines Verkehrsgraphs analysiert und der daraus ermittelte Verkehrsgraph bewertet. Im Anschluss daran werden Einzugsbereiche für die Park-and-Ride Anlagen im südlichen Wiener Umland berechnet und basierend auf den Erkenntnissen der Datenqualitätsanalyse interpretiert.

3.2.1 Raumplanerischer Hintergrund kontextbezogene Forschungsfragen

Die Stadt Wien hat es sich zum Ziel gesetzt, die von Wien beeinflussbaren Treibhausgasemissionen pro Kopf bis zum Jahr 2030 um 35% und bis zum Jahr 2050 um 80% zu senken, basierend auf den Werten des Jahres 1990. Darunter fallen neben den Emissionen für Gebäudeheizung und Energiebereitstellung auch die Emissionen im Wiener Straßennetz (Magistrat der Stadt Wien, 2014a, S. 32). Durch unterschiedliche Maßnahmen gelang es bereits, die Wege der WienerInnen immer stärker auf die Nutzung des Umweltverbundes (Fußwege, Radwege, öffentlicher Verkehr) zu verlagern. Im Jahr 2012 lag der Modal Split-Anteil des Umweltverbundes bei 73% der Wege, d.h. für nur 27% der Wege wurde der Pkw genutzt (Magistrat der Stadt Wien, 2014b, S. 106). Demgegenüber liegt der Anteil des motorisierten Individualverkehrs im grenzüberschreitenden Stadtverkehr von und nach Wien im Jahr 2010 mit 68% der Fahrten noch sehr hoch (Rittler, 2011) - das Ziel der Stadt Wien für das Jahr 2030 liegt auf einer Absenkung auf 55% aller Wege (Magistrat der Stadt Wien, 2013, S. 10). Für das Erreichen dieser Zielsetzung reichen Maßnahmen auf Wiener Stadtgebiet nicht aus. Um eine nachhaltige Verhaltensänderung einläuten zu können, ist eine Kooperation mit der Umlandregion erforderlich.

Eine der im Rahmen der Diskussion aufgeworfenen Planungsfragen ist jene nach der energieeffizienten Neuorganisation der Verkehrswege in der Stadtregion Wien, u.a. im Rahmen des Ausbaus der Park-and-Ride Anlagen in Niederösterreich und dem Burgenland. Folgende Fragestellungen ergeben sich aus der Problemstellung:

¹ PTV Planung Transport Verkehr AG

- Wie groß ist der Einzugsbereich der bestehenden Park-and-Ride Anlagen?
- Wie sind die Einzugsbereiche der einzelnen Park-and-Ride Anlagen in Bezug auf die Fahrzeit gegeneinander abgegrenzt?
- Wieviele Personen haben in den Einzugsbereichen ihren Hauptwohnsitz?
- Wie ist das Verhältnis der Personen im Einzugsbereich zum Angebot an Stellplätzen der Park-and-Ride Anlagen?

Diese grundsätzlichen Fragen sind die Basis für die Beantwortung wichtiger verkehrsplanerischer Fragestellungen wie

- Welches ist die attraktivste Park-and-Ride Anlage für WienpendlerInnen in Bezug auf das gegenwärtige Verkehrsangebot und die verfügbare Stellplatzanzahl?
- Reichen die Kapazitäten der Park-and-Ride Anlagen aus, sollten 10% der Pkw-Pendlerfahrten auf den öffentlichen Verkehr verlagert werden?
- Wo könnten zukünftig die Kapazitätsgrenzen erreicht werden?
- An welchen Standorten sollten neue Park-and-Ride Anlagen errichtet werden?

Die Beantwortung dieser Fragestellung erfordert neben einem Straßenverkehrsgraph auch die Hinterlegung des öffentlichen Verkehrsangebots am Bahnhof, der Anzahl der PendlerInnen nach Wien sowie der Berechnung von gewichteten Erreichbarkeitspotenzialen. Die Lösung dieser Aufgabenstellungen ist aufgrund des damit verbundenen Arbeitsaufwandes nicht mehr Teil der gegenständlichen Arbeit.

3.2.2 Forschungsfragen der Arbeit und erwartete Ergebnisse

Lange Zeit lag die Feststellung der Datenqualität von Geodaten in der Verantwortung der Datenbereitsteller, welche die Konformität der Daten mit den offiziellen Datenqualitätsstandards prüfen und gewährleisten mussten. Diese Konformitätstests waren auf die Beseitigung von Fehlern im Rahmen des Erstellens der Geodatenbanken fokussiert und auf eine bestimmte Art der Anwendung ausgerichtet (Veregin, 1999).

Die Eignung der Geodaten für eine andere technische und oder semantische Anwendung wird dabei in letzter Instanz zum Nutzer der Daten verlagert, der anhand seiner Anforderungen die Eignung der Daten („fitness to use“) selbst bewerten muss.

Von Oort (2006, S.6) gliedert diese Qualitätsprüfung durch den Nutzer in drei Arbeitsschritte:

1. Der Suche nach einem Geodatensatz, welcher die benötigten Informationen (in der erforderlichen Form) für die vorgesehene Anwendung enthält.
2. Der Prüfung der legalen und finanziellen Beschränkungen die dem Zugriff auf diese Geodaten bzw. Ihrer Nutzung entgegenstehen könnten.
3. Der Feststellung, ob die Datenqualität der Geodaten für die Anwendung ausreichend ist, bzw. die Qualitätsmängel im Sinne der Anwendung akzeptiert werden können.

In der vorliegenden Arbeit steht der dritte Arbeitsschritt im Vordergrund, da insbesondere der erste der beiden Punkte durch die Fragestellung bereits weitgehend determiniert ist. Es sollen in Hinblick auf die inhaltliche Fragestellung der Ermittlung von Einzugsbereichen von Park-and-Ride Anlagen die folgenden Fragen beantwortet werden:

1. Welche Informationen stehen aus der OpenStreetmap für die Erstellung eines Verkehrsgraphs zur Verfügung?
2. Mit welchen Open Source Anwendungen können Daten aus der OpenStreetMap zur Entwicklung eines Verkehrsgraphs für das südliche Wiener Umland extrahiert bzw. genutzt werden? Welche legalen und finanziellen Beschränkungen sind zu berücksichtigen?
3. Wie kann die Datenqualität von Geodaten bestimmt werden?
4. Welche Evaluierungsansätze wurden bislang auf die OpenStreetMap angewandt?
5. Welche Qualität hat ein aus der OpenStreetMap erzeugter Verkehrsgraph?

6. Ist mit dem erstellten Verkehrsgraph und weiterer Open-Source Software die Beantwortung der oben genannten kontextbezogenen Fragestellungen möglich?

Erwartete Ergebnisse

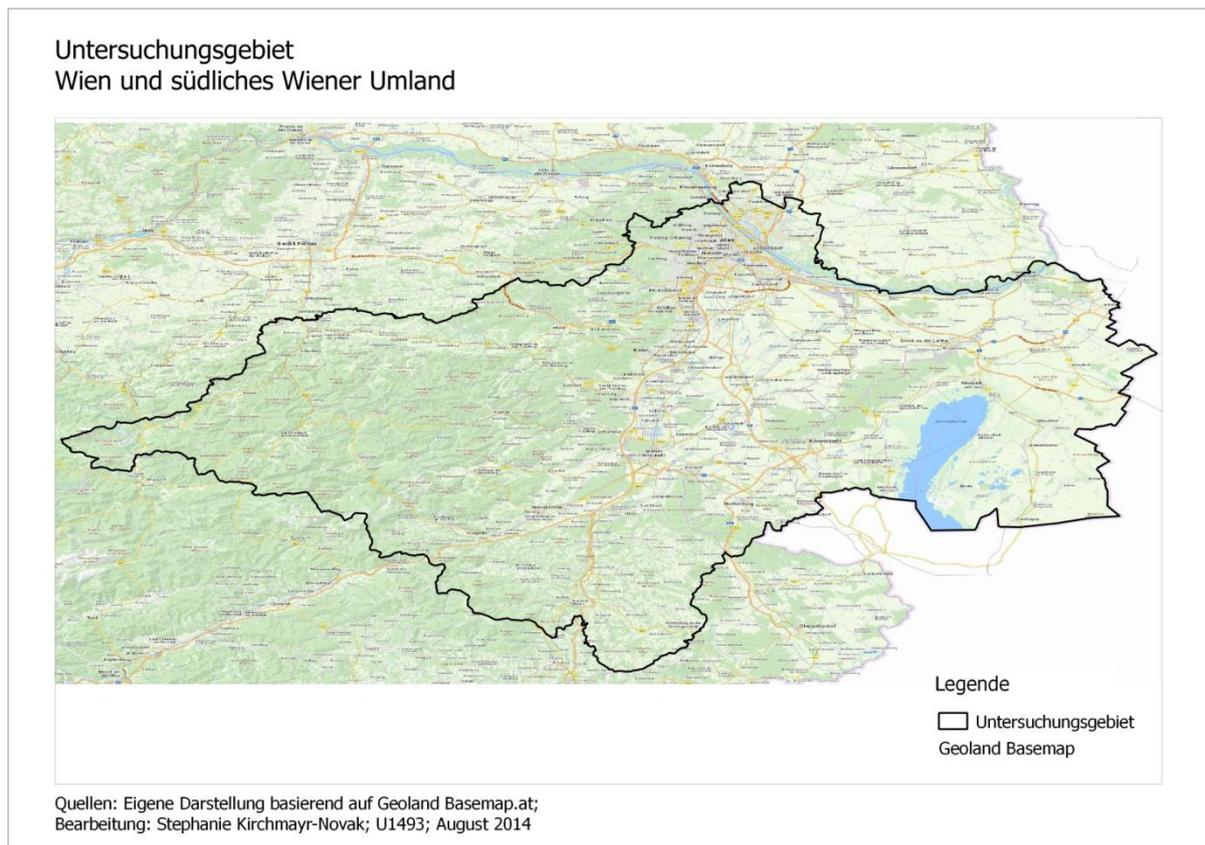
Als Ergebnis stehen ein Straßenverkehrsgraph für die südliche Wiener Umlandregion zur Verfügung sowie ein Bewertungsmodell für Straßenverkehrsgraphen. Darüber hinaus werden die Resultate der Qualitätsbewertung hinsichtlich des konkreten Anwendungsfalls interpretiert und die Ergebnisse der Einzugsbereichsberechnung zur Beantwortung der kontextbezogenen Forschungsfragen herangezogen.

Untersuchungsgebiet

Die Analyse wird für das südliche Wiener Umland durchgeführt. und inkludiert die Gemeinden der politischen Bezirke Eisenstadt, Eisenstadt Umgebung, Rust, Neusiedl am See Mattersburg, Baden, Mödling, Wiener Neustadt, Wiener Neustadt Land, Bruck an der Leitha, Lilienfeld, Neunkirchen, sowie die südlichen Gemeinden von Wien Umgebung.

Die Basis für die Analyse bildet der Datensatz der OpenStreetmap für das südliche Wiener Umland mit Stand vom 12. Oktober 2013.

Abbildung 1: Untersuchungsregion Wien und südliches Wiener Umland



Quelle: Eigene Darstellung basierend auf der Geoland basemap.at per August 2014

Werkzeuge, Datengrundlagen und Nicht-Ziele der Arbeit

Die Arbeit greift ausschließlich auf Open Source Software zurück.

Auch hinsichtlich der verwendeten Referenzdatensätze zur Bewertung der Datenqualität der OpenStreetmap ist die vorliegende Arbeit auf die frei verfügbaren Datenquellen bzw. offiziellen Open Data Kataloge für Österreich beschränkt. Diese Restriktionen erschweren bzw. verunmöglichen vor allem die extrinsische Bewertung des Verkehrsgraphs durch den Vergleich mit einem offiziellen Referenzverkehrsgraph und führen dazu, dass nicht alle Qualitätsparameter geprüft werden können. Die einzige offizielle Open Source

Datengrundlage für das Untersuchungsgebiet ist die Verwaltungsgrundkarte von Österreich „basemap.at“, welche im alle zwei Monate aktualisiert wird, jedoch nur als Rasterkarte zur Verfügung steht.² Damit sind die Möglichkeiten der Analysen auf Grundlage dieses Datensatzes auf visuelle Überprüfungen beschränkt.

Der Fokus der Arbeit liegt auf der Bewertung eines Straßenverkehrsgraphs für die Region. Eine Kalibrierung des Straßenverkehrsgraphs zur Anpassung der Fahrzeiten an reale Verhältnisse ist nicht Teil der Arbeit. Der Anwendungsfall dient in erster Linie dazu, die erforderlichen Qualitätsparameter für die Bewertung des Verkehrsgraphs zu determinieren. Damit verbundene verkehrsplanerische Fragestellungen wie die Analyse von multimodalen Verkehrsgraphen die auch das Angebot der öffentlichen Verkehrsmittel berücksichtigen, werden nicht behandelt.

Beabsichtigtes Publikum

Die Arbeit richtet sich an Personen, die mit Hilfe der OpenStreetMap einen Verkehrsgraph erzeugen möchten und sich über die Machbarkeit, die Datenqualität sowie verfügbare Open Source Werkzeuge informieren möchten.

² Stadt Wien und Österreichische Länder bzw. Ämter der Landesregierung (2014): <http://www.basemap.at/> (Zugriff: 01.09.2014)

3.3 Struktur der Thesis

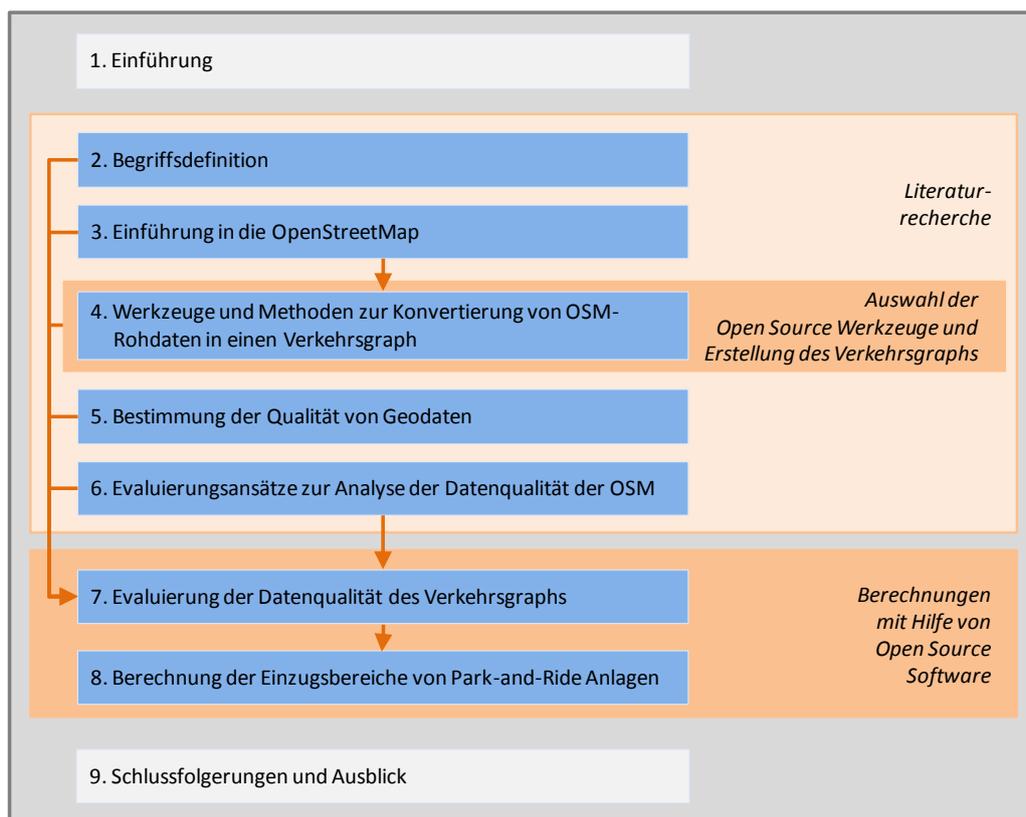
Die Arbeit ist in acht Abschnitte gegliedert, die in Abbildung 2 einer graphischen Übersicht zusammengefasst sind. Nach der Definition der wichtigsten Begriffe in Abschnitt 2 wird im dritten Teil der Arbeit die OpenStreetMap und die wichtigsten Parameter der in Hinblick auf die Erstellung eines Verkehrsgraphs dargestellt. In Abschnitt 4 wird schließlich beschreiben mit welchen Open Source Anwendungen Daten aus der OpenStreetMap zur Entwicklung eines Verkehrsgraphs für das südliche Wiener Umland herangezogen werden können und welche legalen und finanziellen Beschränkungen zu berücksichtigen sind. Dabei wird auch erläutert, wie mit Hilfe dieser Werkzeuge der Verkehrsgraph erstellt werden kann.

In Abschnitt 5 wird definiert, durch welche Merkmale und Parameter die Qualität von Geodaten definiert werden kann. Dieser Ansatz wird im nächsten Abschnitt 6 konkretisiert, indem die bisher verwendeten Evaluierungsansätze zur Bewertung der OpenStreetMap genauer beleuchtet werden. Diese Abschnitte stützen sich auf die Erkenntnisse aus Literaturrecherchen.

Der erzeugte Verkehrsgraph wird in Abschnitt 7 entsprechend der für den Anwendungsfall formulierten Evaluierungsschritte hinsichtlich seiner Qualität und Eignung bewertet. Basierend auf diesem Wissen, können in Abschnitt 8 die Einzugsbereiche der Park-and-Ride Anlagen berechnet und ihre Ergebnisse interpretiert werden.

Im letzten Abschnitt werden die Ergebnisse zusammengefasst und Schlussfolgerungen diskutiert.

Abbildung 2: Struktur der Arbeit als Übersicht



Quelle: Eigene Darstellung

4 Begriffsdefinition

Im folgenden Abschnitt werden die für die Arbeit wichtigsten Begriffe definiert, bzw. ihr Verständnis im Rahmen der Arbeit erläutert.

4.1 Geographische Daten (Geodaten)

Daten mit implizitem oder explizitem Bezug zu einem Punkt relativ zum Planet Erde werden als geographische Daten bezeichnet (ISO 19109:2005, Definition 4.12). Im Rahmen der vorliegenden Arbeit werden geographische Daten als Geodaten bezeichnet.

4.2 Datenqualität

Der Begriff "qualitas" stammt aus der lateinischen Sprache und bedeutet "Beschaffenheit, Eigenschaft" (Scholze-Stubenrecht 1997). In der ISO Norm 19113 „Qualitätsgrundsätze für Geoinformation“ wird Qualität beschrieben als „die Gesamtheit aller Eigenschaften eines Produktes die es befähigen, konkretisierte und implizierte Anforderungen zu erfüllen“.³ Die Qualität von Daten ist damit immer im Zusammenhang mit ihrer Anwendung zu bewerten.

4.3 Offene Daten (Open Data) und Open Data Lizenzen

Offene Daten sind Daten die von jedem frei genutzt, vervielfältigt und weiterverteilt werden können, auch zu kommerziellen Zwecken, solange auf ihre Quelle verwiesen wird. Oft wird hinzugefügt, dass Veränderungen des Datenstandes und allenfalls daraus entstandene Werke zu gleichen Bedingungen weitergegeben werden müssen.⁴ Nicht als offene Daten gelten Daten deren Nutzung beschränkt wird, sei es durch das Untersagen von Veränderungen des Datenstandes oder der kommerziellen Nutzung der Daten.⁵

4.4 Open Source Software und freie Software

Nach Definition der Free Software Foundation (FSF, 2008) wird Software dann als „frei“ verstanden, wenn die Nutzer der Software die Freiheit haben,

- die Software für jeden Zweck zu nutzen
- das Produkt entsprechend ihrer Anforderungen zu adaptieren
- Kopien des Produkts weiterzugeben und
- modifizierte Versionen des Programmes zu verbreiten.

Im Unterschied zu Freeware wird den Nutzern bei freier Software auch der Quellcode (engl. source code) zugänglich gemacht. Bei Freeware, die ebenfalls kostenfrei bezogen werden kann, ist dies nicht der Fall. Die Nutzer können die Software also nicht oder nur über konkret verfügbar gemachte Parameter adaptieren (Rey, 2008).

4.5 Einzugsbereich

Unter Einzugsbereichen versteht man die Abgrenzung und Zuweisung einer Fläche bzw. einer Menge von Punkten zu einem vorgegebenen Anziehungspunkt, z.B. einer Haltestelle. In der Verkehrsplanung werden

³ Frei übersetzt durch die Autorin aus DIN EN ISO 19113 2005 Deutsches Institut für Normung: Geoinformation - Qualitätsgrundsätze. Berlin, Beuth, 2005. Originaltext: „Quality: totality of characteristics of a product that bear on its ability to satisfy stated and implied needs.“

⁴ siehe OpenDefinition.org <http://okfn.org/opendata/> (15.02.2014)

⁵ siehe Max Plank Gesellschaft (2003): Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen vom 22. Oktober 2003

Einzugsbereiche zumeist genutzt um Haltestellen des öffentlichen Verkehrs Einwohner, Arbeitsplätze oder andere Marktgruppen im Einzugsbereich der Haltestelle, z.B. als potenzielle Nutzer, zuzuweisen. Zur Ermittlung der Einzugsbereichen werden in der Verkehrsplanung zwei unterschiedliche methodische Ansätze verwendet: der euklidische Abstand über die Luftdistanz oder die Netzwerkberechnung über Verkehrswege (Jermann, 2004).

Im Rahmen der gegenständlichen Arbeit werden die Einzugsbereiche von Park-and-Ride Anlagen, d.h. Stellplatzanlagen für Fahrgäste der Bahn, die zum Bahnhof mit dem privaten Pkw fahren, ihre Fahrt von dort aus jedoch mit dem Zug fortsetzen, behandelt.

5 Einführung in die OpenStreetMap

Wie in Kapitel 6 Qualität von Geodaten noch dargelegt wird, ist für die Beschreibung der Datenqualität das Datenmodell der Geodaten von zentraler Bedeutung. Aus diesem Grund wird nun zunächst das Datenmodell der OpenStreetMap (OSM) vorgestellt, bevor die Qualität von Geodaten allgemein sowie die bislang angewandten, konkreten Evaluierungsmethoden zur Bewertung der Datenqualität der OSM erörtert werden.

5.1 Allgemeines

5.1.1 Datenherkunft und Qualitätskontrolle

Das Projekt Openstreetmap (OSM) ist eine freie geographische Datenbank (www.openstreetmap.org) mit dem Ziel, möglichst alle geographischen Merkmale der Erde zu erfassen. Die Arbeiten an der OSM wurden 2004 von Steve Coast in Großbritannien initiiert. Mit der OSM wurde eine Kartengrundlage entwickelt, in welche freie Geodaten mit Open Source Software eingearbeitet und auch bezogen werden können, ohne die strengen Lizenzbedingungen zu brechen oder hohe Lizenzgebühren zu bezahlen (Bennett 2010, S.8ff). Die OpenStreet-Map ist unter der Open Database License (ODbL) verfügbar (siehe Kapitel 5.1.2.)

Im Jahr 2013, ist die OpenStreetMap eine geographische Datenbank mit Einträgen aus aller Welt. Per November 2014 waren über 1,87 Million Benutzer registriert⁶, wobei nur ein kleiner Teil der Benutzer auch Daten in die Datenbank lädt – per Juli 2014 haben rund 400.000 Nutzer zumindest eine Veränderung in der OSM eingetragen (Neis, 2014).

Die Daten der OpenStreetMap werden in der Projektion WGS 84 (SRID 4326)⁷ gespeichert und in der Regel im Laufe einer Begehung und mit Hilfe von handelsüblichen GPS-Empfängern aufgezeichnet und danach mit zusammen mit den Beobachtungen während der Begehung (z.B. Straßennamen, Abbiegerelationen) in die OSM eingepflegt. Teilweise wurden jedoch auch offizielle freie Datenbestände in die OSM eingebracht, z.B. der TIGER⁸-Datensatz des US Census Bureaus (Neis, Zielstra, Zipf 2012).

Die Sammlung und Aufbereitung der Geodaten erfolgt durch die freiwillige Kollaboration vieler Individuen (crowdsourcing) und gehört damit neben Wikimapia oder Flickr den vielen aktuellen Beispielen für volunteered geographic information (Goodchild, 2007). Volunteered geographic information (VGI) ist damit eine auf geographische Information spezifizierte Form des allgemeineren Begriffes „user generated content“ (UGC) der von der OECD beschrieben wurde als „über das Internet öffentlich zugänglich gemachter Inhalt, der ein gewisses Maß an kreativem Aufwand erfordert und der außerhalb von professionellen Arbeitsabläufen erstellt wird“ (OECD, 2007). Der starke Trend zu VGI profitiert von der - zumindest in westlichen Ländern - starken Durchdringung von tragbaren, kostengünstigen GPS-Empfängern, Web 2.0 Technologien und leistungsfähigen Breitbandverbindungen (O'Reilly, 2005).

Es gibt keine zentral geregelte Qualitätskontrolle der OpenStreetMap wie bei kommerziellen Produkten, allerdings profitiert OSM wie andere „crowdsourced data“ von dem räumlichen Wissen der lokalen Beiträge sowie der Tatsache, dass crowdsourcing von vielen Individuen durchgeführt wird, die sich gegenseitig kontrollieren (Linus's law, Raimond, 2002).⁹ Zusätzlich gibt es zahlreiche Tools zur Qualitätsüberprüfung, wie Openstreetbugs (für Österreich <http://openstreetbugs.schokokeks.org/?lon=15.115&lat=48.068&zoom=9>, 13.02.2013). Ein Indiz für die Qualität wäre damit die Anzahl der registrierten aktiven lokalen Nutzer.

⁶ Neis, P: (2014): OSMstats. <http://osmstats.neis-one.org/> (Zugriff: 14.11.2014)

⁷ <http://spatialreference.org/ref/epsg/4326/> (15.10.2013)

⁸ „Topologically Integrated Encoding and Referencing“ (TIGER), vgl. Wiki OpenStreetMap (2014): TIGER <http://wiki.openstreetmap.org/wiki/TIGER> (Zugriff: 14.11.2014).

⁹ Linus's law, benannt nach Linus Towards, den Initiator von Linux Kernel besagt: „Given enough eyeballs, all bugs are shallow“. Raymond (2002).

Ein Ziel dieser Arbeit ist es festzustellen, ob die Qualität der OSM-Daten in der Ostregion Österreichs (Bundesländer Niederösterreich, Burgenland, Wien) ausreichend sind, um einen Verkehrsgraph daraus zu erzeugen, mit dem Einzugsbereichsberechnung durchgeführt werden können.

5.1.2 Lizenzbedingungen

Die OpenStreetMap (OSM) kann derzeit unter der Open Database License 1.0 ("ODbL") genutzt werden, die es erlaubt die Datenbank und ihren Inhalt zu nutzen, (kommerziell) zu vervielfältigen und öffentlich zugänglich zu machen. Werke dürfen basierend auf der Datenbank erstellt werden, aber auch die Datenbank selbst darf modifiziert, transformiert und als Datengrundlage verwendet werden. Diese freie Nutzung gilt unter der Einschränkung, dass bei allen öffentlichen Nutzungen der Datenbank auf die OSM und die ODbL-Lizenz verwiesen werden muss, alle Veröffentlichungen der adaptierten Datenbank wiederum unter ODbL zur Verfügung gestellt werden müssen und die Nutzung der dieser adaptierten Datenbank offen nach ODbL möglich sein muss. Das hat zur Folge, dass auch kommerzielle Anwender, die ihre Datenbank mit OSM-Daten ergänzen, die so entstandene "derivative Datenbank" wiederum unter einer ODbL-Lizenz zur Verfügung stellen müssen.

Demgegenüber müssen Produkte die auf der OSM-Datenbank aufbauen diese aber nicht verändern, jene Teile ihres Produkts die nicht auf der Datenbank aufbauen (z.B. überlagernde Datenlayer), nicht zur Verfügung stellen¹⁰.

5.2 Das Datenmodell

Aus mathematischer Sicht ist die OSM ein gemischter Graph, der aus Knoten und Kanten besteht. Je nach Attribuierung sind diese Objekttypen miteinander verbunden oder isoliert (Bennett, 2010).

Das Datenmodell der OSM besteht aus drei grundlegenden Objekttypen (data primitives): Nodes (Punkte oder Knoten), Ways (Linien) und Relations (Beziehungen). Flächen werden durch geschlossene Ways dargestellt. Die Attribuierung der Objekte erfolgt über sogenannte „Tags“, die aus einem Schlüssel (key) und einem Wert (Value) bestehen und in der Form `key=value` (z.B. `highway=motorway`) angegeben werden (Ramm, Topf 2010). Das Datenmodell lässt grundsätzlich das Vergeben beliebiger Tags zu, trotzdem wurde im Rahmen der Wiki-Datendokumentation der OSM ein informeller Standard für die Attribuierung der Objekttypen vereinbart. (http://wiki.openstreetmap.org/wiki/Map_features, Zugriff: 14.02.2013).

5.2.1 Gemeinsame Attribute der OSM Objekttypen

Jedes Objekt erhält bei seiner Erstellung eine eindeutige numerische Kennung (ID), die auch nach etwaigem Löschen des Objekts nicht wieder vergeben wird. Jeder Objekttyp hat einen eigenen Nummernraum, ein Node und ein Way demnach gleiche ID-Nummern haben, sind aber trotzdem eindeutig definiert (Ramm, Topf 2010).

Das Standardformat in dem das Datenmodell dargestellt wird ist XML. Derzeit können Daten vom OpenStreetMap Server auch nur in diesem Format geladen werden (Bennett, 2010). Ergänzend stellen unterschiedliche Anbieter Auszüge der OSM über ihre Homepage zur Verfügung.

Für jedes erzeugte Element werden die Attribute ID, User, User-Id (uid), timestamp, Datenstand (visible), Version und Changeset vergeben (Tabelle 1). Jedem Element kann zusätzlich eine beliebige Menge an Tags zugewiesen werden.

Tabelle 1: Gemeinsame Attribute aller OSM Objekttypen

Name	Datenformat	Beschreibung
ID	integer	Identifizierung des Elements. Jeder Objekttyp hat eigenen Nummernraum.
user	string	Name des Bearbeiters, der das Objekt zuletzt modifiziert hat. (Bearbeiter können ihre Namen ändern, aber nicht die uid)

¹⁰ Openstreetmap Wiki: http://wiki.openstreetmap.org/wiki/Open_Data_License/Legal_Structure (Zugriff: 19.07.2014).

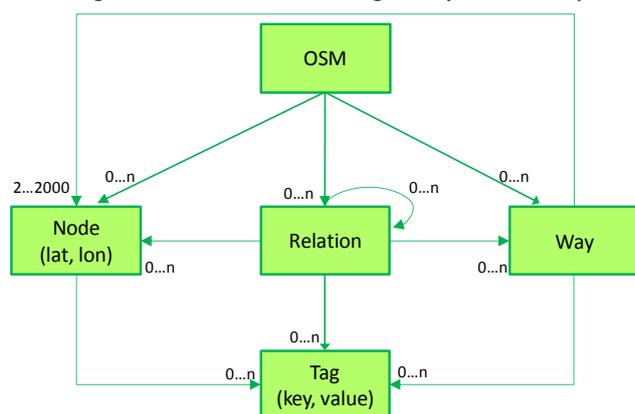
Name	Datenformat	Beschreibung
uid	integer	Die numerische Nutzerkennung, wird einmal vergeben und kann nicht mehr geändert werden.
timestamp	W3C Date und Zeit Formate	Zeitpunkt der letzten Modifikation
visible	"true"	Zeigt, ob das Objekt aus der Datenbank gelöscht wurde (=false) oder nicht (=true). Gelöschte Objekte können nur mehr als "History" abgerufen werden.
	"false"	
version	integer	Die Bearbeitungsversion des Objektes. Neue Objekte beginnen mit Version "1". Der Server erhöht die Versionsnummer, wenn eine neue Version des Objektes durch einen Benutzer geladen wird.
changeset	integer	Zeigt das Changeset in dem das Objekt erstellt oder aktualisiert wurde.

Quelle: Openstreetmap Wiki: <http://wiki.openstreetmap.org/wiki/Elements> (Zugriff: 14.02.2013)

5.2.2 Objekttypen Nodes, Ways und Relations

Die folgenden Absätze beschreiben die drei grundlegenden Objekttypen der OSM. Die Informationen dafür wurden dem Wiki der Openstreetmap (wiki.openstreetmap.org), Bennett 2010 oder Ramm, Topf 2010 entnommen. Abbildung 3 zeigt eine vereinfachte Darstellung des OSM Datenmodells aus Nodes, Ways und Relations. Geographische Koordinaten sind nur in den Nodes enthalten. Ein Way kann aus 2 bis 2000 Nodes bestehen. Nodes, Ways und Relations können keine oder beliebig viele Relations haben. Zusätzlich ist es möglich Nodes, Ways und Relations beliebig viele Tags (key-value Kombinationen) zuzuweisen.

Abbildung 3: Vereinfachte Darstellung des OpenstreetMap-Datenmodell



Quelle: Eigene Darstellung nach Schlesinger 2011

Nodes

Nodes sind die einzigen Elemente in der OSM für die geographische Koordinaten (Länge und Breite) gespeichert werden, optional kann auch eine dritte Dimension verzeichnet werden (z.B. `layer=*` oder `level=*`). Die Koordinaten beziehen sich auf die Projektion „World Geodetic System 1984“ kurz WGS 84 (EPSG:4326) und werden in vollen Grad mit sieben Nachkommastellen erfasst. Nodes können eine alleinstehende Bedeutung haben, wenn sie z.B. eine Kreuzung markieren (`highway=crossing`) oder als Eck- bzw. Stützpunkte für Ways dienen. Es ist nicht auszuschließen, dass ein und dasselbe Objekt in der Natur, z. B. eine Kirche, einmal als Punkt und einmal als Fläche in der OSM verortet ist. Beispielhaft kann eine Lichtsignalanlage im OSM-Schema als XML-Code wie folgt dargestellt werden.

Code 1: Lichtsignalanlage im OSM-Schema (XML-Code)

```
<node id="25496583" lat="51.5173639" lon="-0.140043" version="1" changeset="203496"
user="80n" uid="1238" visible="true" timestamp="2007-01-28T11:40:26Z">
<tag k="highway" v="traffic_signals"/>
</node>
```

Quelle: <http://wiki.openstreetmap.org/wiki/Node>

Ways

Ein Way ist eine geordnete Liste von mindestens zwei bis maximal 2.000 Nodes. Ein Node kann zu keinem, einen oder mehreren Ways gehören. Ways können als offene Graphen lineare Polylinien, Straßen oder Wasserzüge etc. repräsentieren oder auch geschlossene Graphen sein, wenn der letzte Node in der Liste dem ersten Node entspricht. Geschlossene Graphen werden dann als flächenhafte Objekte interpretiert, wenn sie ein Tag mit der Bezeichnung `area=yes` besitzen.

Ways sind gerichtete Linien, die vom ersten bis zum letzten verknüpften Node einer Richtung folgen, auch wenn diese semantisch bedeutungslos sein kann. Ein physischer Kreuzpunkt von zwei Ways wird nur dann erfasst, wenn sich zwei Ways einen Node teilen. So muss eine einfache Straßenkreuzung, die ein Wechseln von einem Way zum anderen erlaubt (z.B. das Rechtsabbiegen), durch einen gemeinsamen Node modelliert werden. Überlagern sich die Straßen nur ohne sich einen Punkt zu teilen, wird angenommen, dass sie sich nicht in selben Niveau befinden wie es z. B. bei einer Brücke über eine Autobahn der Fall ist. Zwei Beispiele für eine physische Kreuzung und eine planfreie Kreuzung sind in Abbildung 4 dargestellt.

Abbildung 4: Darstellung von physischen Kreuzungspunkten und planfreien Kreuzungen in OSM

Kreuzung mit physischen Kreuzungspunkten



Unterführung (planfreie Kreuzung)



Quelle: Openstreetmap

Aufgrund dieser Gliederung in physische und planfreie Kreuzungen definiert durch Kanten und Knoten sind die Geodaten aus der OSM ohne weitere Bearbeitung nicht routenfähig. Außerdem müssen die Abbiegerelationen, die als Relationen definiert werden (siehe Kapitel 5.3.4), erst für einen Routingalgorithmus übersetzt werden.

Beispielhaft kann eine Straße im OSM-Schema als XML-Code wie folgt dargestellt werden.

Code 2: Straße im OSM-Schema (XML-Code)

```
<way id="5090250" visible="true" timestamp="2009-01-19T19:07:25Z" version="8"
changeset="816806" user="Blumpsy" uid="64226">
<nd ref="822403"/>
<nd ref="21533912"/>
<nd ref="821601"/>
<nd ref="21533910"/>
<nd ref="135791608"/>
<nd ref="333725784"/>
<nd ref="333725781"/>
<nd ref="333725774"/>
<nd ref="333725776"/>
<nd ref="823771"/>
<tag k="highway" v="residential"/>
<tag k="name" v="Clipstone Street"/>
<tag k="oneway" v="yes"/>
</way>
```

Quelle: <http://wiki.openstreetmap.org/wiki/Way>

Relations

Relationen repräsentieren geographische oder logische Beziehungen zwischen anderen Elementen. Bei Relationen handelt es sich um sortierte Listen von beliebig vielen Objekten oder Teilnehmern (Members), die Nodes, Ways und bzw. oder andere Relationen sein können. Einem Member einer Relation kann optional eine Rolle zugewiesen werden, welche die Beziehung neben den Tags (key-value Kombinationen) genauer definieren. Eine Rolle kann z. B. spezifizieren, ob eine Grenze den inneren oder den äußeren Rand eines Gebietes umschreibt.

Einige Relationstypen sind so weit verbreitet, dass sie zu einem informellen Standard geworden sind. Dazu gehören Abbiegerestriktionen (`type=restriction`, z. B. `restriction=no_left_turn`), Routen (`type=route`, z. B. `route=bicycle`), Multipolygone (`type=multipolygon`) oder Grenzen (`type=boundary`, z. B. `boundary=administrative`).

Relationen sind mit einer Liste ihrer Tags und einer Liste ihrer Teilnehmer (Members) verbunden. Jeder Teilnehmer hat folgende Attribute: seinen Objekttyp, die ID-Kennung sowie seine Rolle (role). Die Teilnehmer werden dabei nach dem Datum ihrer ersten Speicherung in der OSM gelistet, die Reihenfolge ist daher nicht notwendigerweise signifikant. Wie bei allen anderen Objekttypen erfolgt bei der Erstellung von Relations keine Prüfung ihrer Formate oder Standards, was eine Validierung der Relations durch den Nutzer der Daten erfordert. Beispielhaft kann eine administrative Grenze im OSM-Schema als XML-Code wie folgt dargestellt werden.

Code 3: Relation eine administrative Grenze im OSM-Schema (XML-Code)

```
<relation id="1">
<tag k="type" v="boundary" />
<tag k="boundary" v="administrative" />
<tag k="land_area" v="administrative" />
<tag k="admin_level" v="2" />
<tag k="name" v="light green country" />
<member type="way" id="AB" role="outer" />
<member type="way" id="AC" role="inner" />
</relation>
```

<http://wiki.openstreetmap.org/wiki/Relation> und <http://wiki.openstreetmap.org/wiki/Relation:boundary>

Bei den Relationen erfolgt die Verknüpfung der Relation mit ihren Mitgliedern über die Tabelle `relation_members`. Darin wird auch der `member_type` festgelegt, d.h. der Objekttyp Node, Way oder Relation. Dies ist erforderlich, da über die ID-Kennung allein keine eindeutige Identifizierung der Objekte möglich wäre. Zusätzlich wird in dieser Tabelle die Rolle der Relation vermerkt. Das Datenschema enthält darüber hinaus noch Einträge zu Changesets, GPS-Tracks und den Benutzerdaten, die hier nicht gesondert aufgeführt werden.

5.3 Attribuierung der OSM mit Relevanz für Straßengraphen

Im Folgenden werden die wichtigsten Attribute der OSM für den später erzeugten Straßengraph erläutert. Wie die jeweilige Information bei der Berechnung des Straßengraphs durch das Konvertierungsprogramm übersetzt wird, wird in Kapitel 6.4 dargelegt.

Das Datenmodell der OSM sowie die aufzunehmenden Objekte und Objektklassen für Straßen sind im "OpenStreetMap Wiki" dokumentiert. Für Deutschland steht eine sehr detaillierte Ausführung über aufzunehmende Objekte und Objektklassen zur Verfügung¹¹, welche für Österreich weitgehend übernommen werden kann auch wenn die Attribuierung, z.B. der Geschwindigkeitsbeschränkungen, auf die Österreichische Straßenverkehrsordnung angepasst werden muss.

¹¹ OpenStreetMap Wiki (2013): http://wiki.openstreetmap.org/wiki/Attribuierung_von_Str%C3%9Fen_in_Deutschland; zuletzt geändert 20 April 2014 (30.07.2014).

5.3.1 Klassifizierung von Straßentypen

Das Attribut **highway** ist das Hauptattribut für Straßen in der OSM, unabhängig ob diese vom motorisierten Straßenverkehr genutzt werden (z.B.: Autobahnen als **highway=motorway**) oder nur dem Fußgängerverkehr zur Verfügung stehen (z.B. Fußwege als **highway=pedestrian**). In Tabelle 2 werden die wichtigsten Straßenkategorien aufgelistet und deren Entsprechung im Österreichischen Straßennetz erläutert.

Die einzelnen Tags sind weitgehend intuitiv benannt, es ist jedoch zu beachten, dass zwischen nicht klassifizierten Straßen, die in der OSM mit dem Tag-key **highway=road** gekennzeichnet sind und Straßen mit einfachem Ausbauzustand (ohne Mittellinie) und geringer Verkehrsbedeutung unterschieden wird, die unter die Kategorie **highway-unclassified** fallen.

Tabelle 2: Straßenkategorien der OSM

OSM-Tag	Straßenkategorie	Beschreibung
highway.motorway	Autobahn	Straße mit baulich getrennten Richtungsfahrbahnen; in Österreich alle Autobahnen
highway.trunk	Schnellstraße	Autobahnähnliche Straße, Autostraße, Schnellstraße; in Österreich: alle Schnellstraßen und Kraftfahrstraßen
highway.primary	Haupt-/Bundesstraßen B	Hauptverbindungsstraßen mit besonderer Kennzeichnung in Österreich: Bundesstraßen
highway.secondary	Landesstraßen L	in Österreich: Gut ausgebaute Landesstraße
highway.tertiary	Nebenstraßen	in Österreich: weniger gut ausgebaute Landesstraße für Regionalverkehr, Verbindungsstraße für kleinere Orte
highway.residential	Erschließungsstraßen	Erschließungsstraße mit Wohnbebauung
highway.unclassified	Straße einfachsten Ausbauzustandes	Öffentliche Nebenstraße mit einfachstem Ausbauzustand, oft ohne Mittellinie
highway.road	Straße unbekannter Klassifikation	Platzhalter für Straßen, für die noch keine Zuordnung gewählt wurde.
highway.track	Forststraße / Waldweg	Wege mit überwiegend Forst- und landwirtschaftlichem Verkehr
highway.living-street	Wohnstraße	Verkehrsberuhigte Straße
highway.pedestrian	Fußgängerzone, Fußweg	Weg, Platz oder Straße der nur Fußgängern zugänglich ist.
highway.cycleway	Radweg	
highway.footway	Fußweg	

Quelle: *wiki.openstreetmap* (2014): Key: Highway. <http://wiki.openstreetmap.org/wiki/DE:Key:highway> (Zuletzt modifiziert am 08.08.2014)

Die Straßenattribuierung ist für die Auswertung der Datenqualität (Vollständigkeit und Genauigkeit der Attribute) ausschlaggebend. Darüber hinaus wird bei fehlender Angabe zur Maximalgeschwindigkeit die Standardgeschwindigkeit der Objektklasse vergeben (siehe Kapitel 6.4). Die Straßenattribuierung ist also ein entscheidender Einflussfaktor für die Routenwahl.

5.3.2 Einbahnen

Einbahnen werden in der OSM mit dem Tag **oneway=*** gekennzeichnet und zwar in der Richtung, in der ein Weg genutzt werden darf. So bedeutet **oneway=yes**, dass die Straße nur in die so gekennzeichnete Fahrtrichtung befahren werden darf, bei **oneway=-1** ist das Befahren nur in die Gegenrichtung möglich. Keine Beschränkung liegt vor, wenn die Straße mit **oneway=no** gekennzeichnet ist.¹²

¹² Wiki OpenstreetMap (2014): De:Key:oneway. <http://wiki.openstreetmap.org/wiki/DE:Key:oneway> (Zugriff: 14.11.2014).

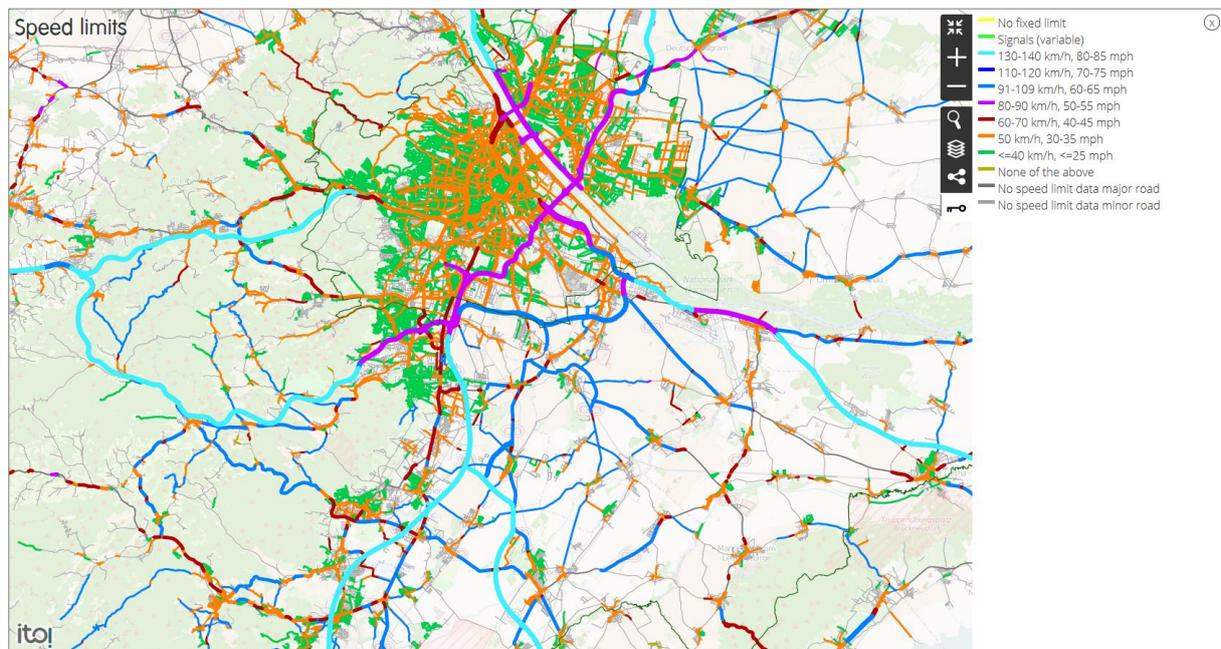
5.3.3 Maximalgeschwindigkeiten

Die maximal zulässige Geschwindigkeit wird über den Schlüssel „**maxspeed**“ beschrieben. Standardmäßig ist die Eingabe der Werte in Kilometer pro Stunde vorgesehen. Weicht die Werteinheit der Angabe von diesem Standardwert ab, soll diese durch ein Leerzeichen getrennt am Ende des Wertes angegeben werden. Ergänzend kann die Informationsquelle für die Angabe der zulässigen Höchstgeschwindigkeit über den Schlüssel **source:maxspeed=*** definiert werden.¹³

Grundsätzlich besteht auch die Möglichkeit unterschiedliche Geschwindigkeitsvarianten für einen Straßenabschnitt anzugeben, z. B. in Abhängigkeit von Tageszeit, Wetterbedingungen aber auch nach Fahrzeugart, Fahrtrichtung und Fahrspur. Wie die folgende Karte (Abbildung 5) zeigt, sind Maximalgeschwindigkeiten für das Hauptstraßennetz in und um Wien weitgehend verfügbar, während für das niederrangige Straßennetz auf die Standardwerte nach Straßenklasse zurückgegriffen werden muss.

Dabei ist anzumerken, dass bei Verkehrsgraphen nicht primär die Maximalgeschwindigkeit für die Kantengewichtung von Bedeutung ist, sondern die Durchschnittsgeschwindigkeit – welche in den Verkehrsmodellen zumeist für unterschiedliche Zeitintervalle (Morgenspitzenstunde, Tagesdurchschnittswert) unterschieden wird. Inwieweit solche Spezifikationen auch für den Verkehrsgraph aus der OSM herangezogen werden können, wird in Kapitel 6.4 beschrieben.

Abbildung 5: Karte der in der OSM verzeichneten Geschwindigkeitslimits (max-speed)



Quelle: ITO Map, <http://www.itoworld.com/> (Zugriff: 14.11.2014).

5.3.4 Abbiegerelationen

In der OSM werden die an einer Kreuzung oder Einmündung möglichen Abbiegemanöver über Relations definiert. Auf den Kanten und an den Knotenpunkten sind generell alle länderspezifisch standardmäßigen Manöver erlaubt, wenn nicht explizit andere Vorschriften oder Abbiegeverbote oder -gebote angegeben werden. So können beispielsweise Straßen für Pkws gesperrt werden (**vehicle=no**) sowie Abbiegeverbote (**no-right-turn**) oder Gebote definiert werden (**only straight on**). Soll eine Verbindung untersagt werden, muss dies in Form einer Restriktion geschehen. Eine Restriction-Relation besteht aus drei Elementen:

¹³ Wiki OpenstreetMap (2014): DE:Key:maxspeed. <http://wiki.openstreetmap.org/wiki/DE:Key:maxspeed> (Zugriff: 14.11.2014).

- dem way von dem aus die Beschränkung gilt (from),
- dem way zu dem die Beschränkung gilt (to) und
- einem verbindenden Element, dass entweder durch einen Knoten (node) repräsentiert sein kann oder über einen oder mehreren ways mit der Rolle „via“.¹⁴

Die Restriktion kann als Verbot formuliert sein (z. B. `no_left_turn`) oder als Gebot (`only_left_turn`).

Abbildung 6: Beispiel für eine Abbiegerelation in der OSM



Relationsbeschreibung in OSM:
`type=restriction`
`restriction=only_right_turn`

Von Weg „a“ aus sind nur Rechtsabbiegemanöver nach Weg „b“ erlaubt, womit in der Folge das Abbiegen nach links verboten wird.

Quelle: Wiki OpenstreetMap (2014): <http://wiki.openstreetmap.org/wiki/DE:Relation:restriction> (Zugriff 16.11.2014)

5.3.5 Lichtsignalanlagen

Ampeln bzw. anderen Lichtsignale zur Steuerung des Verkehrsflusses werden in der OSM über das Tag `highway=traffic_signals` auf Knoten (nicht auf Kanten) dargestellt. Bei einfachen Kreuzungen mit vier Kanten wird der Kreuzungsknoten mit dem Tag markiert.

Bei komplexeren Kreuzungen wird im Wiki der OpenStreetMap empfohlen, die Tags für Lichtsignalanlagen als Knoten an der Haltelinie der Straße bzw. Fahrbahn zu setzen. Zusätzlich kann über das Tag `traffic_signals:direction=forward/backward` festgelegt werden, dass die Lichtsignalanlage nur in eine Richtung wirksam ist.¹⁵

Abbildung 7: Beispiel für die OSM-Modellierung einer einfachen und einer komplexen Kreuzung mit Lichtsignalanlagen



Quelle: Wiki Openstreetmap (2014): http://wiki.openstreetmap.org/wiki/DE:Tag:highway%3Dtraffic_signals (Zugriff 16.11.2014)

5.4 Datenextraktion aus OSM

Kopien der gesamten OpenStreetMap werden regelmäßig als „Planet Files“¹⁶ zwischengespeichert und zum Download zur Verfügung gestellt. Zumeist werden jedoch nicht alle weltweiten Einträge benötigt, sondern nur ein räumlich begrenzter Ausschnitt. Vordefinierte, erstbearbeitete OSM-Auszüge werden von mehreren OSM-Dienstleistern angeboten, wie z. B. vom Unternehmen Geofabrik¹⁷. Das Unternehmen bietet auf seiner Homepage Planet Files für das Land Österreich an, sowohl im XML-Format, als auch als ESRI Shapefile (Bennett, 2010 S. 173ff). Die Shapefiles sind jedoch aufgrund ihrer Struktur nur für die Darstellung der Inhalte, nicht aber für

¹⁴ Wiki OpenstreetMap (2014): <http://wiki.openstreetmap.org/wiki/DE:Relation:restriction> (Zugriff 16.11.2014)

¹⁵ Wiki OpenstreetMap (2014): http://wiki.openstreetmap.org/wiki/DE:Tag:highway%3Dtraffic_signals (Zugriff 16.11.2014)

¹⁶ Vgl. <http://planet.openstreetmap.org/> (Zugriff: 1.10.2013)

¹⁷ Geofabrik (2014): <http://www.geofabrik.de/de/> (Zugriff: 1.11.2014)

die Konvertierung in einen Verkehrsgraph geeignet, da wichtige Informationen für ein Routing fehlen, wie Zugangsbeschränkungen für bestimmte Straßenkategorien, die Unterscheidung zwischen bestehenden, geplanten und im Bau befindlichen Straßen sowie Informationen zu Abbiegevorschriften und Barrieren (Zimmermann, 2010).

6 Werkzeuge und Methoden zur Konvertierung von OSM-Rohdaten in einen Verkehrsgraph

Aus der Fülle der verfügbaren Werkzeuge und Methoden zur Konvertierung OSM-Rohdaten in einen Verkehrsgraphen werden im Folgenden jene Werkzeuge ausgewählt und vorgestellt, die als Open Source bzw. Freeware Produkte zur Verfügung stehen. Neben der Funktionalität der Werkzeuge, wird auch die jeweils zur Anwendung kommende Nutzungslizenz vorgestellt.

6.1 Auswahl der Open Source Werkzeuge

Für die Extraktion größerer Datenmengen aus der OSM sind mehrere Werkzeuge und Methoden anwendbar, wie Tabelle 3 zeigt. Diese sind jedoch nur bedingt für die Extraktion der Daten geeignet, weil sie entweder keine benutzerdefinierte Selektion bestimmte Attribute bzw. Inhalte zulassen oder die Ergebnisse nicht zur Entwicklung eines Verkehrsgraphs geeignet sind. Kostenpflichtige Anwendungen wurden aufgrund der Fragestellung, die auf Open Source Software fokussiert ist, grundsätzlich von der Selektion ausgeschlossen.

Nach eingehender Prüfung der Werkzeuge wurden für die Erstellung des Verkehrsgraphs die Werkzeuge Osmosis und Osm2po ausgewählt, da diese in Kombination sowohl eine benutzerdefinierte Selektion relevanter Features ermöglichen, als auch die Erstellung eines routenfähigen Verkehrsgraphs. Die diesbezüglich zu beachtenden Open Source Lizenzen sowie die wichtigsten Open Source Werkzeuge werden im nächsten Kapitel beschrieben.

Tabelle 3: Werkzeuge und Methoden zur Konvertierung von OSM-Rohdaten in eine Geodatenbank

Werkzeuge und Methoden	Ausgabeformat	Benutzerdefinierte Selektion möglich?	Anmerkungen	Für Verkehrsgraphen nutzbar
QGIS ¹⁸	ESRI Shapefile	Nein	Erweiterung für QGIS. Es werden alle Tag Informationen in je eine Spalte kopiert.	Nein
Osm2pgsql + ¹⁹	PostGIS Datenbank	Ja (mit style file)	Konvertiert OSM-Daten in PostgreSQL Datenbank, vor allem für Mapnik-Anwendungen Punkte (Nodes) und Wege (Ways) werden zu Line-strings und Polygonen konvertiert, die Topologische Information geht verloren.	Nein
Osmosis ²⁰	OSM OSM.bz2	Ja	Konkrete räumliche und thematische Auswahl von Daten möglich.	Nein
Osm2po ²¹	PostgreSQL-File	Ja	Für die Anwendung von pgRouting adaptiert. Freeware, Source Code nicht verfügbar, benutzerdefinierte Adaptionen über Konfigurationsdatei möglich.	Ja
FME ²²	ESRI Shapefile, ESRI Geodatabase	Ja	Kostenpflichtig	Nein
GeoConverter ²³	ESRI Shapefile, ESRI Geodatabase	Nein	Nur für Files bis zu 100 MB	Nein

¹⁸ Wiki OpenStreetMap (2014): http://wiki.openstreetmap.org/wiki/Importing_Planet.osm_into_QUANTUMGIS (Zugriff: 13.04.2014)

¹⁹ <http://wiki.openstreetmap.org/wiki/Osm2pgsql> (Zugriff: 13.04.2014)

²⁰ <http://wiki.openstreetmap.org/wiki/Osmosis> (Zugriff: 13.04.2014)

²¹ <http://osm2po.de/> (Zugriff: 13.04.2014)

²² <http://www.safe.com/solutions/for-applications/openstreetmap/> (Zugriff: 13.04.2014)

²³ <http://geoconverter.hsr.ch/> (Zugriff: 13.04.2014)

Werkzeuge und Methoden	Ausgabeformat	Benutzerdefinierte Selektion möglich?	Anmerkungen	Für Verkehrsgraphen nutzbar
OSM2GIS ²⁴	ESRI Shapefile, KML, MapInfo, GML, SQLite	Nein	Datenmenge beschränkt, nur räumliche, keine thematische Auswahl	Nein
Osm2nds ²⁵	ArcGIS Netzwerk Datenset, GDB	Ja	Nur nutzbar mit ArcGIS 9.3 und damit indirekt auch kostenpflichtig.	Ja

Quelle: Eigene Erhebung

Für die Durchführung der Arbeit wurde insgesamt auf 9 verschiedene Werkzeuge zurückgegriffen, die unter verschiedenen Lizenzen zur Verfügung gestellt werden. Tabelle 4 listet diese Werkzeuge, ihre Funktionalität und die relevanten Lizenzen auf. Neben der OSM an sich, wurde das Desktop-Gis QGIS genutzt, das zuletzt in Version 2.4 vorlag. Ergänzend wurde PostgreSQL als objektrelationales Datenbanksystem mit der Erweiterung PostGIS für räumliche und geographische Objekte genutzt.

Die Datenextraktion aus den XML-Daten der OSM wurde mit den Werkzeugen Osmosis und osm2po durchgeführt, zur Durchführung der Netzwerkberechnungen diente pgRouting.

Tabelle 4: Lizenzen der im Rahmen der Arbeit verwendeten Werkzeuge

#	Werkzeug	Funktionalität	Lizenzierung
1	OpenStreetMap	Datenbank mit Vektordateninformationen, z. B. von Straßennetzen	Open Data Commons Open Database License (ODbL)
2	QGIS (QUANTUMGIS) 2.4 Chugiak	Desktop-GIS	Creative Commons Attribution-ShareAlike 3.0 Lizenz (CC BY-SA)
3	Osmosis 0.44	Kommandozeilen-Java-Applikation zur Verarbeitung von OSM-Daten	Public Domain Certification (zurückgezogenes Rechtswerkzeug von creative commons)
4	osm2po 4.7.7	Konvertierungssoftware und Routing Engine für XML-Daten der OSM	Osm2po License
5	pgRouting 2.0	Erweiterung für PostGIS und PostgreSQL mit Routing- und Netzwerkfunktionalitäten	The PostgreSQL License GNU General Public License, version 2 (GPLv2) Boost Software License -Version 1.0 MIT-X License CC BY-SA 3.0
6	pgAdmin III	Administration und Management Werkzeug für PostgreSQL	PostgreSQL License
7	PostgreSQL 9.2 (PS)	Open Source, objektrelationales Datenbanksystem	PostgreSQL License
8	PostGIS 2.0 (PG)	Erweiterung für räumliche Datenbanken (PostgreSQL) für Lageabfragen mit SQL	GNU General Public License (GPLv2)
9	Java Runtime Environment 6.0+ (7.0) (J);	Laufzeitumgebung für Java, Softwarplattform, virtuelle Maschine (JVM)	Oracle Binary Code License Agreement for the Java SE Platform Products and JavaFX

Quelle: Eigene Darstellung

²⁴ Arnaud Vandecasteele (geotribu), <http://www.osm974.re/osm2gis/> (Zugriff: 13.04.2014)

²⁵ <http://52north.org/downloads/geoprocessing/osm2nds> (Zugriff: 13.04.2014)

6.2 Open Source Lizenzierung

Die folgende Kurzbeschreibung von Open Data und Open Source Software Lizenzen beruht auf Angaben der Open Knowledge Foundation (2014) und Klampferer (2014).

Grundsätzlich muss bei der Nutzung aller Open Data Lizenzen angemessene Urheber- und Rechteangaben machen. Diese Art der Bearbeitung wird mit „BY“ (requires attribution) gekennzeichnet. Zudem muss zumeist ein Link zur Lizenz beifügt und angegeben werden, ob Änderungen an den Daten vorgenommen wurden. Dieses Lizenzattribut tragen alle verwendeten Lizenzen.

Oft dürfen anderen Nutzern bei der Weiterverwendung der lizenzierten Daten oder Software keine über die Originallizenzierung hinausgehenden, zusätzlichen Klauseln Nutzungen rechtlich untersagen, welche die Originallizenz erlaubt.²⁶ Die dezidierte Vorgabe, veränderten Datenbeständen (z.B. Source Code) unter gleichen Lizenzbedingungen weitergeben zu müssen, kommt bei vielen Open Data Lizenzen vor. Diese Weitergabe unter gleichen Bedingungen wird mit „SA“ (requires share alike) gekennzeichnet. Dies trifft für die Lizenzen ODbL, CC-BY-SA 3.0 und GPLv2 license zu (siehe Tabelle 5).

Die Nutzung insbesondere von Open Source Software kann auch bei freier Zurverfügungstellung beschränkt werden, in dem z.B. keine Bearbeitung des Source Codes erlaubt wird („ND“ - no derivatives) oder die kommerzielle Nutzung untersagt wird („NC“ - non-commercial). Die Nutzung von Osm2po fällt unter die Bearbeitungsbeschränkung „ND“, die Software darf darüber hinaus auch nicht verkauft oder anderweitig vertrieben werden.

Tabelle 5 listet die im Rahmen der Arbeit relevanten Lizenzen auf. Wichtig sind die Lizenzen Creative Commons Attribution-ShareAlike 3.0 License (CC BY-SA) unter die sowohl das Werkzeug QGIS, als auch die im Rahmen der Arbeit verwendeten Open Source Daten (mit Ausnahme der OSM) fallen sowie die GNU General Public License, version 2 (GPLv2 license).

Tabelle 5: Verwendete Open Data Lizenzen

Lizenz	Lizenzattribute				Informationen zur Lizenzbeschreibung	Open Data Software und Open Data unter dieser Lizenz
	BY	SA	ND	NC		
Open Database License (ODbL)	X	X			http://opendatacommons.org/licenses/odbl/	OpenStreetMap
Creative Commons Attribution-ShareAlike 3.0 License (CC BY-SA)	X	X			http://creativecommons.org/licenses/by-sa/3.0/	QGIS, pgRouting
Open Government Data Österreich Lizenz CC-BY 3.0 AT	X				http://www.basemap.at/	basemap.at, data.gv.at - offene Daten Österreichs
The PostgreSQL License	X				http://www.postgresql.org/about/License/	PostgreSQL, pgAdmin III
GNU General Public License, version 2 (GPLv2 license)	X	X			http://www.gnu.org/licenses/gpl-2.0.html	PostGIS, pgRouting
Boost Software License	X				http://poost.org/LICENSE_1_0.txt	pgRouting
MIT-X License	X				http://opensource.org/licenses/MIT	Teile des Codes von pgRouting
Osm2po License	X		X	(x)	http://osm2po.de/ Software oder Teile daraus dürfen nicht verkauft, vermietet etc. werden. Source Code nicht vollständig verfügbar.	Osm2po

²⁶ Siehe z.B. Creative Commons (2014): Namensnennung - Weitergabe unter gleichen Bedingungen 3.0 Österreich (CC BY-SA 3.0 AT) <https://creativecommons.org/licenses/by-sa/3.0/at/> (Zugriff: 14.11.2014).

Lizenz	Lizenzattribute				Informationen zur Lizenzbeschreibung	Open Data Software und Open Data unter dieser Lizenz
	BY	SA	ND	NC		
Public Domain Certification ²⁷	Frei von Copyright Restriktionen				http://creativecommons.org/licenses/publicdomain/	Osmosis
<i>BY.. Namensnennung (requires attribution), SA.. Weitergabe unter gleichen Bedingungen (share alike), ND.. keine Bearbeitung des Source Codes (no derivatives), NC.. keine kommerzielle Nutzung (non-commercial).</i>						

Quelle: Eigene Darstellung basierend auf den Angaben der Open Source Software Bereitsteller und <http://opendefinition.org/licenses/> (14.11.2014).

6.3 Osmosis

Osmosis ist Java-basiertes Programm zur universellen Filterung und Konvertierung von OSM-Daten, welches über eine DOS-Kommandozeile exekutiert wird. Osmosis liest die zu bearbeitenden Daten nicht am Stück, sondern schrittweise ein und ist dadurch in der Lage mit großen Datenmengen – wie sie bei einem Verkehrsgraph anfallen – zu operieren. (Ramm, Topf 2010, S. 279).

Im Zuge der Arbeit wird Osmosis dazu genutzt die für gesamt Österreich vorliegenden Rohdaten der OSM, welche von Geofabrik frei zur Verfügung gestellt werden, auf das Bearbeitungsgebiet zu beschränken und in weiterer Folge nur jene Daten zu extrahieren, die für die Erstellung des Verkehrsgraphs von Relevanz sind. Der Filter „bounding-polygon“ ermöglicht es, ein Polygon für die räumliche Selektion einzubinden.²⁸ Darüber hinaus wird im letzten Arbeitsschritt die komprimierte OSM-Datenbank des Typs „bz2“ in eine nicht komprimierte Datenbank umzuwandeln, welche in weiterer Folge von osm2po eingelesen werden kann.

Code 4: Osmosis-Befehle zur Extraktion von OSM-Daten für den Verkehrsgraph

Schritt 1: Beschneidung des Österreichdatensatzes auf das Untersuchungsgebiet mit Hilfe des Polygons „WSUMS.poly“ inklusive der über das Untersuchungsgebiet hinausführenden Straßen (completeWays="yes").

```
. | osmosis --read-xml file="austria.osm.bz2"
--bounding-polygon file="WSUMS.poly" completeWays="yes"
--write xml file="Oregplus.osm.bz2"
```

Schritt 2: Selektion der Straßeninformationen aus dem OSM-Datensatz (accept-ways highway=)*

```
. | osmosis --read-xml file="Oregplus.osm.bz2"
--tag filter accept-ways highway=*
--write xml file="Oplushighway.osm.bz2"
```

Schritt 3: Extraktion der komprimierten Datenbank für die weitere Bearbeitung

```
. | osmosis --read-xml file="Oplushighway.osm.bz2"
--write xml file="Oplushighway.osm"
```

Quelle: Eigene Darstellung nach Ramm, Topf (2012)

Weiterführende Informationen zur Osmosis bieten Ramm, Topf (2010).

²⁷ Im Jahr 2010 zurückgezogenes Rechtswerkzeug von creative commons, ersetzt „CC0 Public Domain Dedication“ und die „Public Domain Mark“, vgl. <http://creativecommons.org/retiredlicenses> (Zugriff: 22.11.2014).

²⁸ Zur Formatierung der Bounding-Polygon-Datei siehe beschrieben unter http://wiki.openstreetmap.org/wiki/Osmosis/Polygon_Filter_File_Format (02.02.2013).

6.4 osm2po

Osm2po (Möller, 2012²⁹) ist ein Konvertierungsprogramm unter CC BY-SA 3.0 Lizenz, welches das Straßen- und Wegenetz aus den Daten der OpenStreetMap extrahiert, für Routings aufbereitet und als SQL-Datei zur Verfügung stellt. Dadurch können die Ergebnisdaten in eine PostGIS-Datenbank geladen und durch Routingprogramme, wie z. B. pgRouting (Pasche et. al., 2013)³⁰ weiterverarbeitet werden.

Osm2po arbeitet inkrementell. Im ersten Schritt, dem Tiling, werden alle Knoten der OSM-Daten geographisch in Quadranten sortiert und die darin befindlichen Wege und Relationen vorgefiltert extrahiert. Danach werden mehrere Quadranten gegen die Wege aufgelöst (Joining) und im dritten Schritt, der Segmentierung, die Wege eingelesen, auf Kreuzungen untersucht und gegebenenfalls segmentiert (Möller, 2012). Die Standardkonvertierungsparameter können in der Konfigurationsdatei `osm2po.config` nachgelesen und adaptiert werden. Standardmäßig werden von osm2po nur die Straßeninformationen, die von Pkws befahren werden dürfen (`wtr.finalMask=car`) aus den OSM-Daten ausgelesen.

6.4.1 Konvertierung von OSM-Daten in eine routenfähige SQL-Datenbank

Bei der Konvertierung durch osm2po, wird die Struktur des OSM-Datensatzes von XML in eine SQL-Datenbank mit den Kanten des Verkehrsgraphs und je 19 Attribute mit definierten Datentypen übersetzt. Tabelle 7 zeigt die Attribuierung des von osm2po erzeugten Verkehrsgraphs und erläutert sie.

Wesentliche Veränderungen des OSM-Datensatzes durch osm2po sind:

- die Umformung der Vektordaten in einen routenfähigen Verkehrsgraph. So werden jeder Kante ein Quell- und ein Zielknoten (source, target) zugewiesen;
- die Berechnung der Kantenlänge aus den Koordinaten der Quell- und Zielpunkte (km);
- die Konvertierung der OSM-Tags für Straßenkategorien von key-value Kombinationen in einen Zahlencode, z.B. der "clazz" 11 für `highway=motorway`;
- die Einführung einer Standardgeschwindigkeit "kmh" für jede Kante auf Basis der Straßenkategorie (z. B. 120 km/h für Autobahnen). Stehen Informationen zur Maximalgeschwindigkeit (`maxspeed`) aus dem OSM-Datensatz zur Verfügung, ersetzen diese die Standardgeschwindigkeit;
- die Berechnung von Kosten (in Zeiteinheiten) für das Überfahren der Kante in Fahrtrichtung sowie die Gegenrichtung aus der Länge der Kante (km) und der Standardgeschwindigkeit (kmh). OSM-Tags für Einbahnen (`oneway=yes`) werden in der nicht befahrbaren Richtung mit dem Wert von 1.000.000 Stunden "gesperrt";
- die Übersetzung von Abbiegevorschriften in routingfähige Attribute-

Bei der Übersetzung der OSM-Daten in einen routenfähigen Verkehrsgraph werden im Konvertierungsalgorithmus vier Parameter definiert, welche die Kanten des Graphen klassifizieren (siehe auch Tabelle 6):

- die laufende Nummer zur Identifikation des erlaubten (Haupt-)Verkehrsträgers (1 für highway, 2 für route, 3 für railway),
- die Straßenklasse, welche den Straßentyp wiedergibt,
- die standardmäßige Kantengeschwindigkeit in km/h,
- die Definition der für die Straßenkategorie erlaubten Verkehrsmittel.

Sofern die OSM-Daten Informationen über die höchstzulässige Geschwindigkeit auf der Straße bzw. der Kante enthalten (OSM-Tag `maxspeed`), werden die Standardwerte der Kantengeschwindigkeit mit der höchstzulässigen Maximalgeschwindigkeit überschrieben (Moeller 2012, `osm2po.config`).

Tabelle 6 zeigt die in der Konfigurationsdatei festgelegten Standardwerte für die Übersetzung der OSM-Daten. Wie bereits erwähnt wird die Kantengeschwindigkeit, wenn keine OSM-Daten zur Verfügung stehen, den

²⁹ Möller, C. (2012): osm2po. <http://www.osm2po.de/> (15.10.2013)

³⁰ Pasche et. al. (2013): pgRouting 2.0.0. <http://pgrouting.org/> (15.10.2013)

Kanten auf Basis der Straßenklasse zugewiesen. Es erfolgt keine weitere Unterteilung nach der Lage der Straßen, z.B. Innerort- und Überlandstraßen. Diese können vom Benutzer adaptiert werden. Im Rahmen der vorliegenden Arbeit wurde mit den untenstehenden Standardwerten gearbeitet und nur die Segmente ausgewählt, auf welchen das Befahren mit Pkws erlaubt ist (car).

Tabelle 6: Auswahl und Übersetzung der OSM-Straßendaten (Tags) in osm2po

OSM-Tags für Straßen	Verkehrsträger	Straßenklasse	Standardwert der Kantengeschwindigkeit in km/h	erlaubte(s) Verkehrsmittel
highway.motorway	1	11	120	car
highway.motorway_link	1	12	30	car
highway.trunk	1	13	90	car
highway.trunk_link	1	14	30	car
highway.primary	1	15	70	car
highway.primary_link	1	16	30	car
highway.secondary	1	21	60	car
highway.secondary_link	1	22	30	car
highway.tertiary	1	31	40	car bike
highway.residential	1	32	50	car bike
highway.road	1	41	30	car bike
highway.unclassified	1	42	30	car bike
highway.service	1	51	5	car bike
highway.living_street	1	63	7	car bike foot
highway.pedestrian	1	62	5	bike foot
highway.track	1	71	10	bike foot
highway.path	1	72	10	bike foot
highway.cycleway	1	81	5	bike
highway.footway	1	91	5	foot
highway.steps	1	92	5	foot
route.ferry	2	1	10	ferry
route.shuttle_train	2	2	50	rail car
railway.rail	3	3	50	rail

Quelle: Moeller, 2012, *osm2po.config*

6.4.2 Umsetzung von Abbiegevorschriften in osm2po

Das Datenmodell der OSM sieht vor, dass Straßen nur dann über einen physischen Kreuzungspunkt verfügen, wenn sie einen gemeinsamen Knoten haben. Das Konvertierungsprogramm osm2po interpretiert Kreuzungen in diesem Sinne. Überall, wo Straßen einen modellierten Kreuzungsknoten haben, wird auch ein physischer Kreuzungspunkt angenommen. Straßen, die sich ohne gemeinsamen Knoten kreuzen, werden als planfreie Kreuzungen (Unterführung, Brücke, etc.) in den Verkehrsgraph aufgenommen. Abbiegevorschriften werden, wenn vorhanden, aus den Relationen ausgelesen. osm2po liest ausschließlich Relationen ein, wenn die folgenden vier Parameter verfügbar sind: die Abbiegevorschrift, die ID der Kante die überfahren wird, die ID des als Kreuzung dienenden Knotens sowie die ID der Kante auf die abgebogen werden soll (Moeller, 2012, S. 12).

6.4.3 Umsetzung von Lichtsignalanlagen in osm2po

Standardmäßig werden für OSM-Tags für Lichtsignalanlagen von osm2po keine zusätzlichen Wartezeiten berücksichtigt. Mit Implementierung der mit osm2po mitgelieferten Erweiterung „TrafficSignalsNodeTagResolver“ kann eine standardisierte, durchschnittliche Wartezeit an Ampeln festgelegt werden. Dabei werden wird die

Wartezeit jedoch nur „virtuell“ im Web-Router bei Überfahrt des jeweiligen Knoten aufgeschlagen, die Kantengeschwindigkeit jedoch nicht verändert.

Um die Wartezeit den Kosten der Kantenüberquerung (Fahrzeit) aufzuschlagen, müssen in der Datei `osm2po.config` der „`pgVertexWriter`“ sowie der „`TrafficSignalsNodeTagResolver`“ aktiviert werden, indem die Kommentarzeichen davor („`#`“) entfernt werden. Dadurch wird eine Vertex-Tabelle erstellt, welche die jeweils angegebene Durchschnittswartezeit den Knoten mit Lichtsignalanlage zuweist. In einem zweiten Schritt muss über eine SQL-Abfrage dieser Wartezeitauflage den betroffenen Kanten in der Kantentabelle entweder über den Quell- oder den Zielknoten zugeschlagen werden.³¹

Im Rahmen der Arbeit werden zunächst die Standardeinstellungen verwendet und bewertet und danach Empfehlungen für erforderliche benutzerdefinierte Adaptionen gemacht.

6.4.4 Ergebnisse der Konvertierung

Werden keine benutzerdefinierten Veränderungen vorgenommen, kann die Konvertierung des lokal gespeicherten OSM-Files "WSUMs-osm.bz2" über einen einzigen Befehl erfolgen (Code 5).

Code 5: osm2po Befehl zur Konvertierung der OSM-Daten in einen routenfähigen Verkehrsgraph

```
java -jar osm2po-core-4.7.7-signed.jar prefix=ws WSUMs.osm.bz2
```

Quelle: Eigene Berechnung nach Möller (2012)

Als Ergebnis erstellt `osm2po` neben dem Log-File mit der Dokumentation der Arbeitsschritte eine SQL-Datei „`ws_2po_4pgr.sql`“, die den routenfähigen Verkehrsgraph enthält und mit Hilfe von PostGIS bzw. PostgreSQL weiterbearbeitet werden kann.

Die Attribute und ihre Datentypen werden in Tabelle 7 dargestellt und erläutert. Neben einem Identifier für den erzeugten Verkehrsgraph werden auch Informationen aus der OSM in den Datensatz mit aufgenommen, wie die Identifier der Kante und der Quell- und Zielknoten aus der OSM, der Straßenname und verfügbare Metadaten. Darüber hinaus wird der Kante eine Straßenklasse zugewiesen, ihre Länge ermittelt mit Hilfe der Kantengeschwindigkeit die Kosten in beide Fahrrichtungen ermittelt.

Tabelle 7: Attribute des Verkehrsgraphs und ihre Datentypen

Attribut	Datentyp	Erläuterung
id	integer NOT NULL, PRIMARY KEY (id)	Identifier
osm_id	bigint	Identifier der Kante gemäß OSM
osm_name	character varying	Name der Straße gemäß OSM
osm_meta	character varying	Metadateninformationen gemäß OSM
osm_source_id	bigint	Identifier des "Quellknotens" der Kante gemäß OSM
osm_target_id	bigint	Identifier des "Zielknotens" der Kante gemäß OSM
clazz	integer	Klasse der Straße als Zahlencode
flags	integer	Klasse des Verkehrsmittels (z. B. Pkw, Bahn, Fußgänger)
source	integer	Identifier des Quellknotens
target	integer	Identifier des Zielknotens
km	double precision	Länge der Kante in km
kmh	integer	Standardgeschwindigkeit
cost	double precision	Kosten in Fahrtrichtung in Stunden (km / kmh)
reverse_cost	double precision	Kosten in Gegenfahrtrichtung in Stunden (km/kmh)

³¹ Möller, C. (2013): Add vertices costs to postgresql routable network. <http://gis.stackexchange.com/questions/80144/add-vertices-costs-to-postgresql-routable-network> (Zugriff: 12.11.2014)

Attribut	Datentyp	Erläuterung
x1	double precision	X-Koordinate des Quellknotens
y1	double precision	Y-Koordinate des Quellknotens
x2	double precision	X-Koordinate des Zielknotens
y2	double precision	Y-Koordinate des Zielknotens
geom_way	geometry(LineString4326)	Geometrie der Kante

Quelle: *osm2po*

Weiterführende Informationen zu *osm2po* bietet Moeller, C. 2012 sowie die Online-Plattform „github“ <https://github.com/> (Zugriff 22.11.2014).

6.5 pgRouting

PgRouting ist eine Erweiterung für die geographischen Datenbanken PostGIS und PostgreSQL, welche diese um Routing- und Netzwerkanalysefunktionen ergänzt. Ein Vorläufer der Software wurde unter dem Namen *pgDijkstra* vom französisch-schweizer Unternehmen *Camptocamp*³² erstellt, um kürzeste Wege zu berechnen. Später wurden weitere Algorithmen vom japanischen Unternehmen *Orkney*³³ ergänzt und die Software in *pgRouting* umbenannt. Derzeit wird die Software von den Unternehmen *Georepublic*³⁴ und *iMaptools*³⁵ sowie einer wachsenden Nutzergemeinde gepflegt und verwaltet (Takubo et al. 2013, S. 3). Der Großteil der Funktionalitäten von *pgRouting* ist unter der GPLv2 License verfügbar (siehe auch Kapitel 6.2 Open Source Lizenzierung).

PgRouting wurde in der Version 2.0.0 verwendet, die eine Vielzahl von Algorithmen zur Verfügung stellt. Wichtige Funktionen umfassen die Konversion von GIS-Daten in eine routenfähige topologische Datenbank, eine Grobanalyse von routenfähigen Graphen um Inkonsistenzen und auffällige Anomalien im Graph aufzudecken sowie ein breit gefächertes Repertoire an Routingalgorithmen. Letztere umfassen Kürzeste-Wege-Berechnungen nach Johnson's, Floyd-Warshall oder Dijkstra Algorithmus (*pgr_apsJohnson*, *pgr_apsWarshall* bzw. *pgr_dijkstra*). Neben einfachen Quelle-Ziel Berechnungen, können auch komplexere Analyse zu multiplen Zielen durchgeführt werden, die jedoch deutlich mehr Rechnerleistung erfordern.

Code 6: SQL Code zur Berechnung des kürzesten Weges zu multiplen Zielen

PgRouting Funktion zur Berechnung des kürzesten Weges zu multiplen Zielen

```
varchar pgr_kdijkstraCost(text sql, integer source, integer [] targets, boolean directed,
boolean has_rcost);
```

und die SQL query wie folgt definiert ist:

```
SELECT id, source, target, cost [,reverse_cost] FROM edge_table]
```

wobei

<i>id</i>	... Identifizier der Kante,	<i>reverse_cost</i>	... Kantenfahrzeit in Gegenrichtung
<i>source</i>	... Identifizier des Quellknotens,	<i>directed</i>	... wahr, wenn der Graph gerichtet ist,
<i>target []</i>	... Array von Identifiern der Zielknoten,	<i>has_rcost</i>	... wahr wenn die Spalte <i>reverse_cost</i> für die
<i>cost</i>	... Fahrzeit auf der Kante,		Kantenfahrzeit in Gegenrichtung benutzt wird.

Und als Ergebnis der Berechnung folgende Parameter zurückgegeben werden:

<i>seq</i>	... Zeilensequenz	<i>cost</i>	... Kosten um den Weg von <i>id1</i> zu <i>id2</i> zurück zu
<i>id1</i>	... Identifizier des Quellpunkts		legen. Es wird der Wert -1.0 zurück gegeben, wenn
<i>id2</i>	... Identifizier des Zielpunkts		kein Weg zwischen den Punkten besteht

Quelle: eigene Darstellung nach Takubo et al., 2013.

³² <http://www.camptocamp.com/> (Zugriff: 12.11.2014)

³³ <http://www.orkney.co.jp/> (Zugriff: 12.11.2014)

³⁴ <http://georepublic.de/en/> (Zugriff: 12.11.2014)

³⁵ <http://imaptools.com/> (Zugriff: 12.11.2014)

Wichtig für die gegenständliche Arbeit ist die Möglichkeit auch Fahrzeiten von einem Startknoten zu anderen Knoten im Netz auszuwerten (`pgr_drivingDistance` bzw. `pgr_kdijkstraCost`), siehe Beispiel in Code 6.

Im Allgemeinen benötigen die Routingalgorithmen eine SQL-Abfrage die ein oder mehrere der folgenden Attribute enthalten: `id` (integer, Identifier einer Kante), `source` (integer, Identifier des Quellknotens), `target` (integer, Identifier des Zielknotens), `cost` (double precision, Fahrzeit auf der Kante), `reverse_cost` (double precision, Fahrzeit auf der Kante in Gegenrichtung), `the_geom` (Geometrie).

Als Ergebnis werden standardmäßig die folgenden Attribute zurückgegeben: `seq` (ID zur Identifikation der Wegrichtung), `id1` (der Identifier des Knotens), `id2` (Identifier der Kante), `cost` (Fahrzeit auf der Kante). Die Vergabe der Parameter `id1` und `id2` variiert je Funktion. So kann `id2` auch für den Identifier des Zielknotens verwendet werden, wenn es die Logik der Funktion erfordert (siehe Beispiel in Code 6).

Alle im Rahmen der Arbeit verwendenden `pgRouting`-Algorithmen werden beim jeweiligen Arbeitsschritt erläutert. Weiterführende Informationen zu `pgRouting` bieten Takubo et al. 2013 sowie die Online-Plattform „github“ <https://github.com/> (Zugriff 22.11.2014).

6.6 Zusammenfassung der Arbeitsschritte zur Erstellung eines Verkehrsgraphs

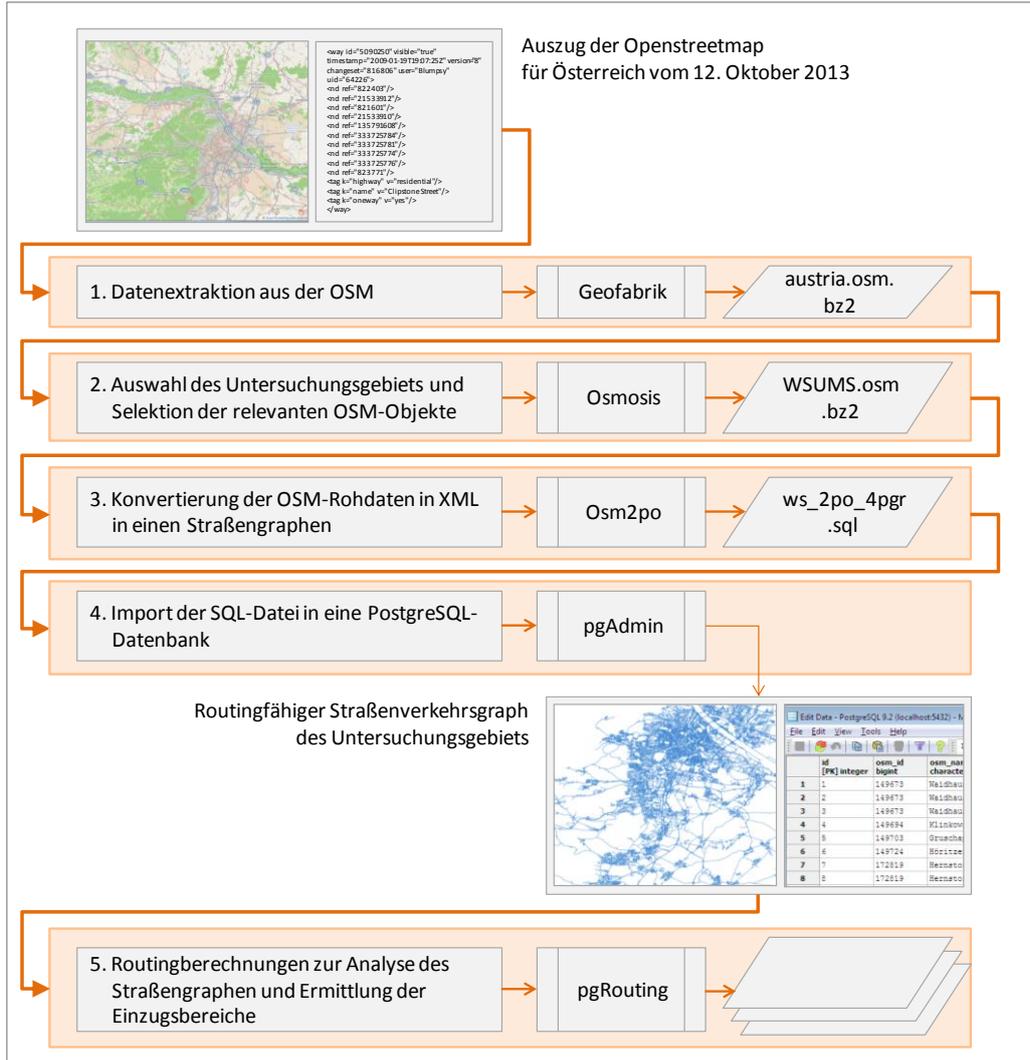
Die Arbeitsschritte zur Erstellung eines routenfähigen Straßenverkehrsgraphs aus OSM-Rohdaten werden in Abbildung 8 zusammengefasst.

Als Basis für den Verkehrsgraph dient ein Auszug der OpenStreetmap für Österreich vom 12. Oktober 2013. Dieser Datensatz wird mit Osmosis auf das Untersuchungsgebiet zugeschnitten und danach auf die relevanten straßen- und routingbezogenen Information reduziert.

Die so erstellten OSM-Rohdaten wurden im nächsten Schritt mit Hilfe von `osm2po` in einen routenfähigen Verkehrsgraph konvertiert. Die daraus erstellte SQL-Datei der Straßenkanten kann danach für die weitere Bearbeitung in eine PostgreSQL Datenbank geladen werden.

Für die topologische Analyse des Straßengraphs sowie die weiteren Routingberechnungen werden mit `pgRouting` durchgeführt.

Abbildung 8: Arbeitsschritte zur Erstellung eines Verkehrsgraphs aus der OSM



Quelle: Eigene Darstellung

7 Bestimmung der Qualität von Geodaten

Geodaten werden stets mit Hilfe von Datenmodellen erstellt, welche implizit oder explizit das erforderliche Maß an Abstraktion und Generalisierung in Relation zum realen Phänomen spezifizieren. Daher müssen Geodaten im Rahmen der Qualitätsbewertung nicht in Hinblick auf „die“ objektive Realität bewertet werden, sondern haben in Bezug auf ihre Datenmodellspezifikation (Brassel et al 1995).

Diese Definition erfordert nicht, dass eine „Wahrheit“ existiert, die beobachtet werden kann. Dies würde in vielen Fällen zu Problemen in der Bewertung führen, wenn z.B. die reale Welt nicht (mehr) beobachtbar ist, wie im Falle von historischen Daten oder wenn aufgrund des hohen Aufwandes und der hohen Kosten keine Beobachtung des realen Phänomens möglich ist. Die objektive Realität wird also nicht adressiert, wenn es um die Qualitätsbewertung von Geodaten geht, sondern die Realisierung des Datenmodells in Geodaten und deren Nutzbarkeit für einen konkreten Anwendungsfall (Vereign 1999).

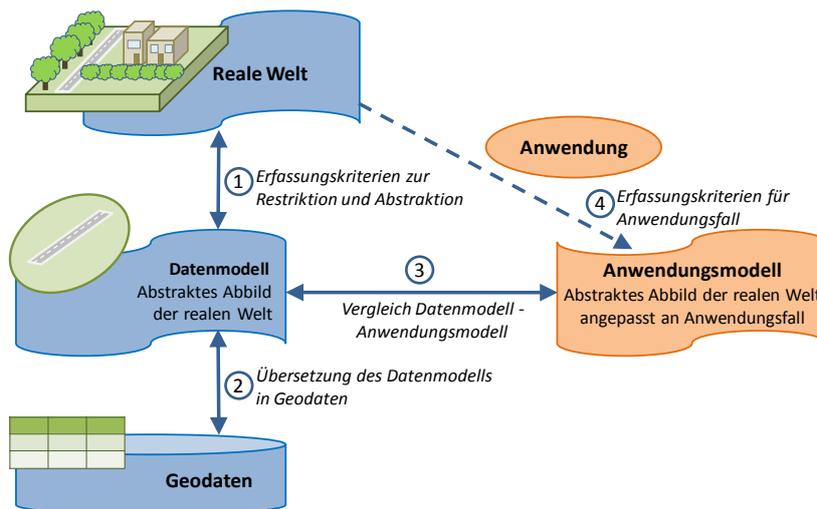
Ausgehend vom eingangs dargelegten Begriff der Qualität von Daten im Allgemeinen und Geodaten im Speziellen, wird im Folgenden erörtert wie die Qualität von Geodaten anhand von Datenmodellen, Qualitätsmodellen und Qualitätsmerkmalen definiert werden kann.

7.1 Datenmodell und Anwendungsmodell

Die Übersetzung der realen Welt in Geodaten folgt konkreten Erfassungskriterien, die in konzeptionellen Modellen definiert und an die konkrete Aufgabenstellung angepasst werden. Ein solches konzeptionelles Modell wird als Datenmodell bezeichnet und ist ein vereinfachtes, widerspruchsfreies, abstraktes Abbild der "Realität" (vgl. Joos 2000, S.35f).

Die Übersetzungsschritte von der realen Welt zu Geodaten sind in Abbildung 9 dargestellt. Im ersten Schritt (1) wird durch Restriktionen und Abstraktion ein Datenmodell erstellt. Die im Datenmodell festgeschriebenen Erfassungskriterien beschreiben die Art und Weise in welcher die Objekte und Attribute der realen Phänomene in den Geodaten abgebildet werden sollen (z.B. Straßenzüge als Liniengeometrie, flächenhafte Gewässer als Polygone aber erst ab einer Fläche von 50 m²).

Abbildung 9: Beschreibung der Datenqualität in Bezug auf den Anwendungsfall



Quelle: eigene Darstellung nach Joos (2000)

Die Übersetzung in Geodaten erfolgt, wenn die "reale Welt" entsprechend dem Datenmodell in - in Schritt 2 - in Geodaten übersetzt wird. Die "reale Welt" wird so in räumliche Objekte (z.B. Straßen, Grundstücksflächen, Gebäude) unterteilt, deren Eigenschaften durch Attribute beschrieben werden und deren Ausprägung durch quantitative oder qualitative Werte festgelegt wird. Das Ausmaß in welchem die Geodaten dem widerspruchsfreien Datenmodell entsprechen wird als Datenqualität bezeichnet.

Während die Modellqualität die Eigenschaften und die Konsistenz des Datenmodells in Relation zur realen Welt beschreibt (1), bezieht sich die Datenqualität auf die Eigenschaften der Daten in Relation zu den Festlegungen des Datenmodells (2), vgl. Joos 2000.

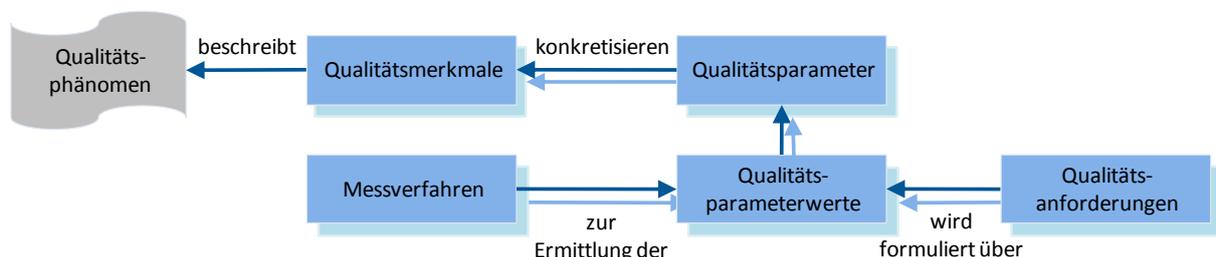
Geodaten werden meist nicht für einen einzigen Zweck erhoben, sondern sollen - im Sinne effizienter Ressourcenverwendung - für möglichst viele Anwendungen zur Verfügung stehen. Um zu entscheiden, ob die Daten und ihre Qualität für eine bestimmte Anwendung geeignet sind, muss zunächst festgelegt werden wie die Daten für den Anwendungsfall beschaffen sein müssen (4). Erst dann kann das so entworfene Anwendungsmodell mit dem Datenmodell verglichen und festgestellt werden, ob dieses prinzipiell den Erfordernissen entspricht (3).

7.2 Qualitätsmodelle

In Qualitätsmodellen werden jene Kriterien und Methoden definiert, welche den Grad der Entsprechung von (Geo-)Daten hinsichtlich der von ihnen abgebildeten (Auszüge) der realen Welt beschreiben lassen (Joos 2000). Sie erfassen die Datenqualität durch mehrere Qualitätsmerkmale, welche jeweils ein bestimmtes Kriterium der Übereinstimmung mit der realen Welt beschreiben und ihrerseits durch qualitative und quantitative Qualitätsparameter konkretisiert werden (Wiltschko; Kaufmann 2005). Unterschiedliche Messverfahren können zur Ermittlung der Qualitätsparameterwerte herangezogen werden. Welche Qualitätsparameterwerte innerhalb der Toleranzgrenzen liegen und damit den Erwartungen an die Datenqualität entsprechen, wird über die Qualitätsanforderungen formuliert. Abbildung 10 stellt diese Zusammenhänge graphisch dar.

Wie auch beim Datenmodell, werden die Anforderungen an die Qualitätsmodelle in Hinblick auf die gewünschte Anwendung der Daten formuliert und alle weiteren Schritte danach definiert.

Abbildung 10: Struktur und Zusammenhang der Elemente des Qualitätsmodells für Geodaten für ein ausgewähltes Qualitätsmerkmal



Quelle: Eigene Darstellung nach Wiltschko, Kaufmann (2005, S. 793)

Zur objektiven, stringenten Beschreibung der Datenqualität von Geodaten in Allgemeinen sind mehrere Qualitätsmodelle erarbeitet worden. Zu den wichtigsten Beispielen zählen die ISO-Standards der Normenfamilie ISO19XXX. Darüber hinaus wird das von den ISO-Standards abgeleitete DDGI-Qualitätsmodell (DIN PAS 1071) häufig genannt, aber auch das Qualitätsmodell der International Cartographic Association (ICA) (siehe Guptill; Morrison 1995). Eine Gegenüberstellung unterschiedlicher Qualitätsmodelle findet sich in Chen (2011), eine ausführliche Beschreibung von Qualitätsmodellen zeigt Joos (2000).

7.3 Qualitätsmerkmale

Im Folgenden werden die theoretischen Eigenschaften der von ISO empfohlenen Qualitätsmerkmale erläutert, um daraus in der Folge für den Anwendungsfall der "Berechnung von Einzugsbereichen von Park-and-Ride Anlagen im südlichen Wiener Umland" zur Verfügung zu stehen.

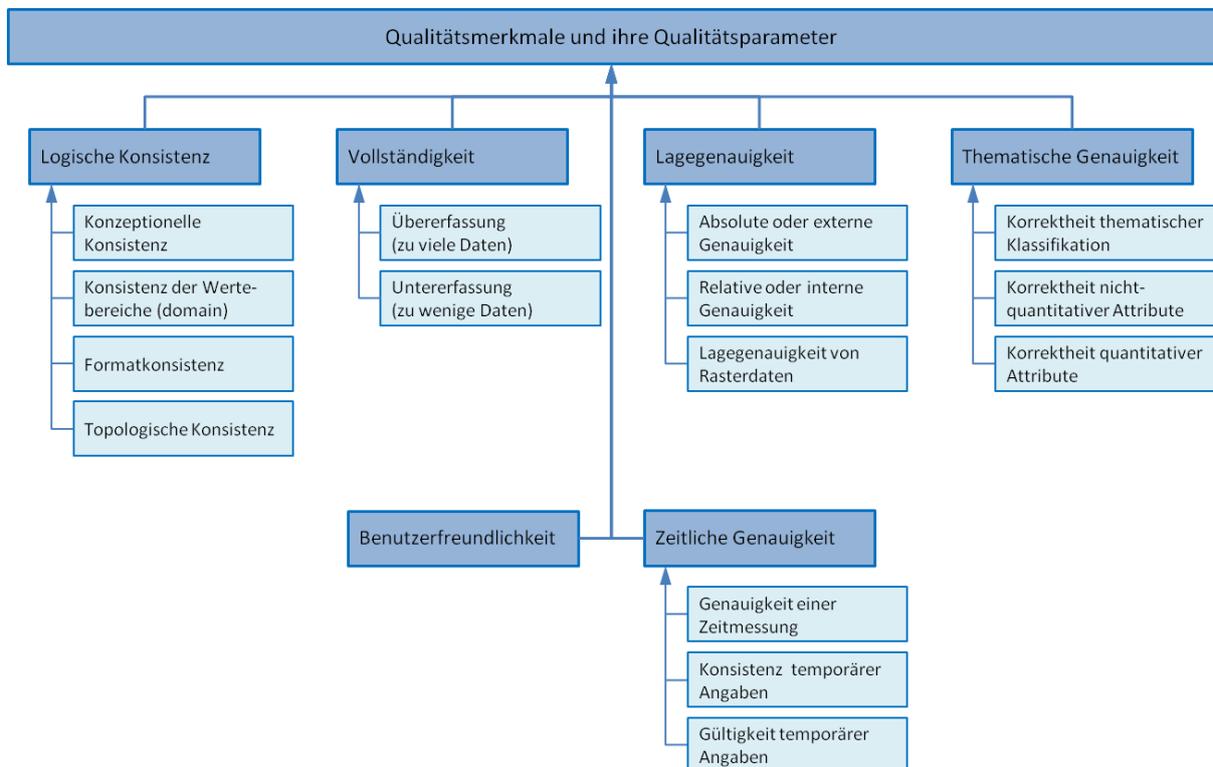
Das Qualitätsmodell nach ISO 19157 enthält die sechs Qualitätsmerkmale (data quality elements) logische Konsistenz, Vollständigkeit, thematische Genauigkeit, zeitliche Genauigkeit, Lagegenauigkeit und Benutzerfreundlichkeit. Diese sind zusammen mit den dazugehörigen Qualitätsparametern in Abbildung 11 dargestellt.

Unter dem Qualitätsmerkmal Benutzerfreundlichkeit wird die Bewertung all jener Nutzeranforderungen im konkreten Anwendungsfall zusammengefasst, die durch die übrigen Qualitätsmerkmale nicht abgedeckt werden können und nicht standardisierbar sind (siehe Kapitel 7.3.6).

Die logische Konsistenz kann gänzlich ohne den Vergleich zu Referenzmessungen bewertet werden (siehe Kapitel 7.3.1). Demgegenüber bezieht sich die Evaluierung der Vollständigkeit, der thematischen Genauigkeit und der Lagegenauigkeit auf den Vergleich der Geodaten mit der "realen Welt" oder mit als "wahr" definierten Referenzmessungen. Die Bestimmung der zeitlichen Genauigkeit folgt sowohl logischen Kriterien, als auch dem Vergleich mit Referenzmessungen bzw. der realen Welt (siehe Kapitel 7.3.5).

Die in Abbildung 11 dargestellten sechs Qualitätsmerkmale und die dazugehörigen Qualitätsparameter werden in der Folge detaillierter ausgeführt und bilden den Rahmen für die spätere Qualitätsbewertung in Abschnitt 9.

Abbildung 11: Qualitätsmerkmale und ihre Qualitätsparameter für Geodaten nach ISO 19157



Quelle: Eigene Darstellung und Übersetzung nach ISO 19157:2011, Draft text for DIS final ISO/CD 19157 und Lange (2013)

7.3.1 Logische Konsistenz

Logische Konsistenz ist erreicht, wenn alle digitalen Objekte des Datensatzes die logischen Regeln für Datenstruktur, Attribuierung und Beziehung des Datenmodells einhalten (Joos 2000) und sich die Informationen im Datensatz nicht widersprechen. Im Standard ISO 19157:2011 werden vier Qualitätsparameter unterschieden: konzeptuelle Konsistenz, begriffliche Konsistenz, Formatkonsistenz und topologische Konsistenz (ISO 2011). Im Gegensatz zu den Qualitätsmerkmalen Vollständigkeit und Lagegenauigkeit kann die Einhaltung der Konsistenzregeln in GIS in der Regel durch eine intrinsische Datenprüfung operationalisiert überprüft werden. Die Feststellung der Konsistenz der Daten ist für sich allein kein Garant für ihre Genauigkeit (Veregin, 1999).

Konzeptuelle Konsistenz

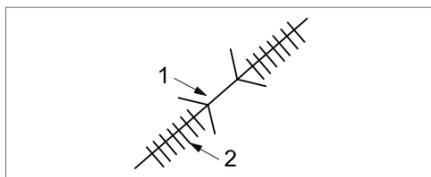
Die konzeptuelle Konsistenz bezeichnet die Einhaltung der Regeln des konzeptuellen Datenmodellschemas in dem die Gliederung und die Inhalte der aus der realen Welt aufzunehmenden Objekte und Objektklassen definiert sind. Dazu gehört die Festlegung des Erfassungsgebietes, der aufzunehmenden Attribute sowie der Struktur des Datensatzes in Form von Dimension, Repräsentationsform der Geometrie (Vektor- oder Raster-

modellierung), des geodätischen Bezugssystems, des Datentyps der Attribute sowie der Regeln zur Objektbildung der Wertebereiche für Attribute. Eine umfassende Beschreibung der konzeptuellen Modellierung bietet Joos, 2000 (S. 7-25).

"Im konzeptionellen Datenmodell sind implizit oder explizit Regeln vereinbart, die von allen Objekten der entsprechenden Objektklassen eingehalten werden müssen" (Joos 2000 S. 72). Im Gegensatz zur logischen Konsistenz, die sehr allgemein formuliert werden kann, muss die konzeptionelle Konsistenz auf die Bedeutung der Daten abgestimmt sein, wodurch für jede Objektklasse eigene Regeln definiert werden müssen. Diese können sich sowohl auf die Geometrie, die Sachdaten, ihre Attributwerte oder auf die Relationen zwischen Attributwerten beziehen. (Joos 2000 S. 72)

So kann innerhalb eines Straßengraphs ein Widerspruch gegen das konzeptionelle System entstehen, wenn z. B. ein Straßenzug, der über eine Brücke führt als Brücke (1) und nicht als Straßenzug (2) über eine Brücke definiert ist (siehe Abbildung 12).

Abbildung 12: Beispiel für Objekte, die nicht den Regeln des konzeptuellen Schemas entsprechen



Quelle: ISO 19157:2011, S. 56.

Konsistenz der Wertebereiche

Die Konsistenz der Wertebereiche beschreibt welche Wertebereiche für Attribute gemäß des konzeptionellen Datenmodellschemas einzuhalten sind. Dazu gehören zwingend erforderliche Attributwerte, Begriffsdefinitionen, Datentypen, Bezugseinheiten sowie Minimal- und Maximalwerte. Ein Beispiel für eine Begriffsdefinition ist die Vorschreibung der Eindeutigkeit von Identifikatoren, wie sie bei Objektschlüsseln üblich ist. Konsistenzfehler treten z. B. bei der Nutzung nicht spezifizierter Zeichen, unkorrekter Feldlänge auf. (Joos 2000, S. 69f)

Formatkonsistenz

Die Formatkonsistenz beschreibt die Übereinstimmung der Datenspeicherung der Attribute mit der physikalischen Struktur des Datensatzes. Die physikalische Konsistenz wird durch die Regeln von (objektrelationalen) Datenbankmanagementsystemen (DBMS), z.B. PostgreSQL, gewährleistet und ist bereits in Geoinformationssystemen implementiert (vgl. Joos 2000, S. 69). In den OpenGIS Standards der OGC³⁶ werden die Standards für Objekttypen von Geoinformationen, die erforderlichen Funktionen um sie zu verändern sowie zwei standardmäßige Metadatenansichten definiert (SPATIAL_REF_SYS und GEOMETRY_COLUMNS), siehe OGC 2010 (S.V bzw. S. 14). In der erstgenannten Sicht wird die Referenzidentifikationsnummern der räumlichen Koordinatensysteme als „SRID“ (Spatial Reference Identifier) dargestellt. Fehlt beispielsweise im internen Speicherformat der räumlichen Objekte des Geodatensatzes die Angabe zur SRID, treten Validitätsprobleme auf (PostGIS Project Steering Committee o. D., Kapitel 4).

Topologische Konsistenz

Die topologische Konsistenz umfasst die Korrektheit der topologischen Kodierung des Datensatzes in Knoten, Kanten und Maschen sowie räumlichen Beziehungen (vgl. Joos 2000, S. 13).

Die Forschungsgesellschaft Mobilität (2011) beschreibt die Anforderungen an konsistente Verkehrsgraphen wie folgt: Die topologische Konsistenz von routenfähigen Straßengraphen ist gegeben, wenn alle Straßen im

³⁶ OGC (2010): OpenGIS Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option. Version: 1.2.1. Status: Corregendum http://portal.opengeospatial.org/files/?artifact_id=25354 (14.11.2014)

Graphen miteinander topologisch verbunden und mögliche Ausnahmen, wie ein Fährenbetrieb über ein Gewässer, klar definiert sind. Darüber hinaus muss bei sich überschneidenden Kanten korrekt abgebildet werden, ob es sich bei dieser Überschneidung um einen physischen Kreuzungspunkt handelt, oder eine planfreie Kreuzung wie z.B. bei Brücken oder Unterführungen vorliegt.

Im konzeptuellen Datenmodell der OSM ist z.B. definiert, dass Straßen nur dann einen physischen Kreuzungspunkt haben, wenn diese über einen gemeinsamen Knoten verfügen (siehe Kapitel 5.2).
Vollständigkeit

7.3.2 Vollständigkeit

Auch für die Ermittlung der Vollständigkeit eines Geodatensatzes bildet das Datenmodell eine wesentliche Grundlage. Ein Geodatensatz ist dann vollständig erfasst, wenn alle im Datenmodell als für die Erhebung relevant spezifizierten räumlichen Objekte, ihre Attribute und Beziehungen enthalten sind (PAS1071 2007).

Objekte sind vollständig, "wenn jedem abstrakten Objekt des Datenmodells genau ein digitales Objekt des Datensatzes mit allen Attributwerten und Beziehungen" zugeordnet ist und umgekehrt. Attribute sind vollständig erfasst, wenn den digitalen Objekten je ein Attributwert für jede Eigenschaft der abstrakten Objekte zugewiesen wird "und dieser Attributwert auch einer Eigenschaft des zugehörigen abstrakten Objekts entspricht". (Joos 2000, S. 44)

Vollständigkeit wird nach ISO 19157:2011 durch zwei Qualitätsparameter gemessen, die das Abweichen des Datensatzes von den Spezifikationen des Datenmodells wiedergeben:

- Der **Übererfassung**, d.h. dem Vorhandensein nicht erforderlicher (überzähliger) Objekte, Eigenschaften oder Beziehungen bzw.
- der **Untererfassung**, d.h. dem Fehlen von digitalen Objekten, ihrer Eigenschaften oder Beziehungen in einem Datensatz.

Die **Vollständigkeit der Objekte** wird zumeist in der Differenzbetrachtung mit einem als wahr akzeptierten Referenzdatensatz geprüft oder durch ein Stichprobenverfahren im Vergleich mit der realen Welt.

Die Prüfung der **Vollständigkeit der Attribute** kann operationalisiert werden, ebenso die Vollständigkeit der Attributwerte (bei erforderlichen Attributwerten wie z.B. bei Objektschlüsseln). Erstere lässt sich durch das Prüfen der Existenz von erforderlichen Attributspalten ermitteln, letztere über die Suche nach Objekten im Datensatz, deren zwingend erforderliche Attribute den Attributwert NULL³⁷ enthalten (vgl. Joos 2000, S. 73).

Fehlende Attributwerte können durch drei semantisch unterschiedliche Wertangaben ausgedrückt werden (ISO 2011, S. 150):

- ein leeres Feld ohne Wertangabe für das Attribut;
- eine Darstellungsform für "nicht anwendbar", wenn kein entsprechender Wert existiert z.B. weil eine Straße offiziell keinen Namen hat;
- eine Darstellungsform für unbekannte, fehlende Werte wenn das Attribut zwar anwendbar, der Wert der Ausprägung jedoch unbekannt ist.

Der Standard ISO 19157:2011 empfiehlt zwingend erforderliche Attribute ohne Wert als logische Inkonsistenz zu bewerten. Nicht anwendbare Werte für zwingend erforderliche Attribute sollten nicht in die Bewertung der Vollständigkeit der Attribute einfließen, während unbekannte, fehlende Werte für erforderliche Attribute als Unvollständigkeit zu werten sind.

³⁷ NULL dient in SQL als Eintrag für fehlende oder nicht anwendbare Informationen.
Vgl. https://en.wikipedia.org/wiki/Null_%28SQL%29 (Zugriff: 01.12.2013)

7.3.3 Lagegenauigkeit

Die Lagegenauigkeit bezeichnet die Genauigkeit der Koordinatenangaben von räumlichen Objekten in einem räumlichen Referenzsystem. Meist wird unterschieden zwischen

- **absoluter oder externer Genauigkeit** der angegebenen Koordinatenwerte im Vergleich zu akzeptierten oder als wahr bekannten Werten im gleichen Referenzsystem und
- **relativer oder interner Genauigkeit** der relativen Position von räumlichen Objekten in einem Datensatz im Vergleich zur anderen Datensätzen des gleichen Gebiets (Van Oort 2005, S. 15).
- Die **Lagegenauigkeit von Rasterdaten** wird als Nähe der Koordinatenangaben zu als wahr akzeptieren Werten definiert (ISO/DIS 19157, S. 11).

Fehler in der Lagegenauigkeit können in allen drei räumlichen Dimensionen gemessen werden. Üblich ist die Darstellung des horizontalen Fehlers (horizontal error) gemessen in der Distanz zwischen dem wahren Wert und dem Datenbankwert. Systematische Fehler können durch Verzerrungen entstehen, z. B. unpräzisen Datengrundlagen oder Messgeräten. Sie werden als Durchschnittsfehler angegeben. Darüber hinaus gibt es "zufällig" auftretende Fehler, die auf Irrtümer etc. zurückzuführen sind (Veregin, 1999).

Die Genauigkeit der Koordinatenangaben wird von der Erfassungstechnik determiniert. Während geodätische Vermessungen hochpräzise Ergebnisse liefern können, bieten Lageerfassungen aus Karten geringere Auflösungen (Lange 2013, S. 261). Die Datengrundlagen der OSM werden in der Regel mit herkömmlichen GPS-Endgeräten und Luftbildern erfasst, die eine durchschnittliche Lagegenauigkeit von 15 m aufweisen (Krampe, 2009).

Die Lagegenauigkeit von Linien, z.B. Straßen, wird zumeist durch Pufferzonen um die Referenzlinie definiert, innerhalb derer die erfasste Linie des Datensatzes mit einer bestimmten Wahrscheinlichkeit liegt. Dieser Bewertungsansatz wurde u.a. von Hakley (2010) oder Graser et al. (2013) angewandt, um die Lagegenauigkeit des OSM Straßengraphs zu eruieren.

Relativ zum amtlichen österreichischen Verkehrsgraph der Graphenintegrationsplattform "GIP" wird der OSM in der Großregion Wien sehr gute Lagegenauigkeit für das hochrangige Straßennetz bescheinigt - mit Ausnahme der Autobahnzufahrten die in den Datensätzen unterschiedlich klassifiziert sind und daher nicht automatisiert verglichen werden können (Graser et al. 2013).

7.3.4 Thematische Genauigkeit

Die thematische Genauigkeit umfasst die korrekte Auswahl und Zuweisung der Eigenschaften von räumlichen Objekten im Rahmen der Klassifizierung der Daten im Vergleich zu einer zur Verfügung stehenden Grundbeobachtung (z. B. Bodenverifikation³⁸ oder Referenzdaten). Fehler können durch inkorrekte Klassifizierung auftreten (z.B. Eintrag einer Autobahn als Nebenstraße) oder durch unscharfe Abgrenzung der verwendeten Datenklassen für das gleiche räumliche Objekt (Caprioli et al., 2003).

Die korrekte Erfassung nicht quantitativer Attribute wie z. B. Flächennutzung wird meist durch die Vorgabe von Zuweisungstabellen operationalisiert. Die Onlinedokumentation der OpenStreetMap bietet eine ausführliche Liste an Attributen und Beispielen für die korrekte Klassifizierung des Straßennetzes³⁹.

Die Genauigkeit quantitativer Attribute wird im Vergleich zu einem als wahr bekannten oder akzeptierten Wert determiniert. Darüber hinaus dürfen sich redundante quantitative Attribute nicht widersprechen, so muss z. B. die Konformität der Werte der Bevölkerungszahl eines Gebiets, dessen Fläche und seiner Bevölkerungsdichte gewährleistet sein (Caprioli et al. 2003).

³⁸ Bodenverifikation (ground truth) bezeichnet ergänzende, bodennahe Beobachtungen zur Interpretation von Fernerkundungsdaten. Vgl. Short, N. (2005): Ground truth, the 'Multi'-Concept. Imaging spectroscopy. In: The remote sensing tutorial. Section 13. Last update 1.11.2005. https://www.fas.org/irp/imint/docs/rst/Sect13/Sect13_1.html. (Zugriff: 15.10.2013).

³⁹ OSM Wiki, http://wiki.openstreetmap.org/wiki/DE:Map_Features#Wege (Zugriff : 19.07.2014).

7.3.5 Zeitliche Genauigkeit

Unter zeitlicher Genauigkeit werden mehrere Charakteristika zusammengefasst die in der Folge erläutert werden. Zunächst wird darunter die Genauigkeit der in den Datensätzen angegebenen temporären Attribute verstanden, die zumeist den Zeitpunkt (oder das Zeitintervall) der Datenerfassung repräsentieren und oft für die gesamte Datenbank Gültigkeit haben. Dargestellt werden kann z. B. die Korrektheit eines Attributs im Zeitablauf, wie etwa dem „Einwohnerstand per 1.1.2014“ (ISO, 2011; S. 154).

Demgegenüber wird unter temporärer Konsistenz die korrekte chronologische Erfassung von Ereignissen verstanden. Ein Beispiel ist die Unterscheidung zwischen einer Straße in Planung, einer Straße in Errichtung und einer Straße im Bestand, wobei das Planungsdatum vor dem Errichtungsdatum liegen muss um zeitliche Konformität zu gewährleisten. Um die zeitliche Konsistenz eines routenfähigen Straßengraphs sicherzustellen, müssen alle Daten des Straßengraphs für den gleichen definierten Zeitraum gültig sein. Temporäre Konsistenz kann auch über die begriffliche Konsistenz (siehe 7.3.1) definiert werden, wenn es z. B. um die begriffliche Abgrenzung des Status „Planung“, „Bau“ und „Errichtung“ geht (ISO 2011; S. 154).

Temporäre Validität bezieht sich schließlich auf die Ausprägung der Angabe an sich, z. B. die gewählte Form der Datumsangabe (z. B. YY-MM-DD) und kann ebenfalls Teil der Formatkonsistenz (siehe 7.3.1) sein (ISO 2011; S. 155).

Konkret abgegrenzt wird die zeitliche Genauigkeit von der Aktualität des Datensatzes, die in Hinblick auf die Nutzeranforderungen definiert werden muss und nicht objektiv festgelegt werden kann. Ein historischer Datensatz kann beispielsweise hohe zeitliche Genauigkeit für die Nutzeranforderungen aufweisen, repräsentiert aber nicht den Datenstand in Bezug auf die aktuelle Situation. Neben den ISO-Standards sehen andere Standards, z.B. CEN/TC287⁴⁰, die Klassifikation von "veraltet" (out_of_date), "gültig" (valid) und "noch nicht gültig" (not_yet_valid) vor. (Caprioli et al. 2003)

7.3.6 Benutzerfreundlichkeit

Die Bewertung der Benutzerfreundlichkeit basiert auf den konkreten Nutzeranforderungen im Anwendungsfall. Alle Qualitätsmerkmale können angewendet werden, um die Benutzerfreundlichkeit zu bewerten. Wenn spezifische Nutzeranforderungen nicht beschrieben werden können soll dieses Kriterium dazu genutzt werden, um nutzerrelevante Qualitätsinformationen hinsichtlich der Eignung eines Datensatzes für eine bestimmte Anwendung zu beschreiben oder um die Konformität der Daten in Hinblick auf ein Set von Anforderungen darzustellen.

7.3.7 Beziehungen zwischen Datenqualitätselementen

Geographische Phänomene erstrecken sich über Inhalt (theme), Raum (space) und Zeit (time) und stehen in Beziehung zueinander. Die Datenqualität darf also nicht isoliert für einzelne Qualitätsmerkmale gemessen werden, da die einzelnen Merkmale in Abhängigkeit zueinander stehen. Eine Änderung des Zeithorizonts kann eine Änderung der thematischen Genauigkeit zur Folge haben, wenn z.B. eine Straße von 2 Spuren zum Zeitpunkt X auf 4 Spuren zum Zeitpunkt Y ausgebaut wurde (vgl. Vereg, 1999).

Nach ISO 19157:2011 (ISO 2011, S. 150f) gibt es Beziehungen zwischen Datenqualitätselementen die dazu führen, dass bestimmte Qualitätsaspekte auf unterschiedliche Art ausgedrückt werden können. Die Zuordnung ist immer von der semantischen Auslegung im Anwendungsfall abhängig:

- Bei unbekannter räumlicher Verteilung kann eine Abweichung entweder der (fehlenden) Genauigkeit der Attributwerte oder der Lagegenauigkeit zugeordnet werden, z.B. eine falsche Höhenangabe entweder dem inkorrekten Attribut (Höhe) oder einer inkorrekten Koordinate (z-Wert).

⁴⁰ Technisches Komitee 287 des "Comité Européen de Normalisation" (CEN).

- Bei Phänomenen, die sich im zeitlichen Ablauf verändern, z. B. Flussläufen, kann ein Fehler einer inkorrekten Lageangabe (Lagegenauigkeit) oder einer inkorrekten Zeitangabe (Flusslauf vor 20 Jahre) zugeordnet werden.
- Bei falschen Zeitbezügen, kann ein Fehler entweder als Genauigkeit von Attributwerten oder als falsche Zeitangabe Berücksichtigung finden (z.B. Einwohnerstand per 1.1.2014).
- Auch die Abgrenzung zwischen Vollständigkeit und Lagegenauigkeit kann bei der Bewertung von Datensätzen gegenüber Referenzmessungen unterschiedlich interpretiert werden. Haben die zu vergleichenden Objekte das gleiche Thema, werden Lageabweichungen als Lagegenauigkeit interpretiert, handelt es sich jedoch Objekte unterschiedlicher Themen, wird die Vollständigkeit des Datensatzes bewertet und angenommen, dass entweder das Objekt der Referenzmessung mit falschem Thema aufgenommen wurde bzw. im Datensatz ein Objekt aus der Referenzmessung fehlt und gleichzeitig ein Objekt zu viel aufgenommen wurde. (ISO 2001, S. 150f)

Zur Klarstellung wird daher meist festgelegt in welcher Reihenfolge die Qualitätsmerkmale geprüft werden, wie im folgenden Kapitel dargelegt.

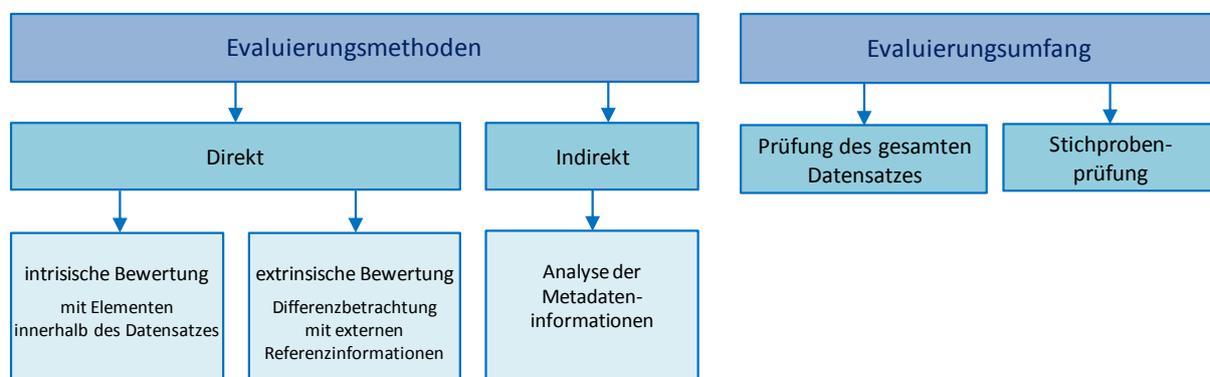
7.3.8 Evaluierungsprozess zur Bestimmung der Datenqualität

Die Qualitätsevaluation der Daten umfasst mehrere Arbeitsschritte und Evaluierungsmethoden. Nach ISO 19157:2011 kann zwischen direkten und indirekten Methoden unterschieden werden.

- Direkte Evaluationsmethoden bewerten die Datenqualität eines Datensatzes anhand der Elemente innerhalb dieses Datensatzes (intrinsisch) oder in der Differenzbetrachtung mit externen Referenzinformationen (extrinsisch).
- Indirekte Evaluierungsmethoden schließen auf die Datenqualität oder schätzen diese mit Metadaten und anderen Informationen über die Daten, z.B. über deren Herkunft.

In der ISO Norm wird empfohlen, direkte vor indirekten Evaluierungsmethoden zu nutzen. Direkte Evaluierungsmethoden werden weiter danach unterteilt, ob sie interne Informationsquelle für die Evaluierung heranziehen oder externe.

Des Weiteren wird unterschieden, ob im Zuge der Evaluierung der gesamte Datensatz oder nur Stichproben untersucht werden. Generell kann von der Qualität einer Teilmenge des Datensatzes, z.B. eines einzelnen Attributes, nicht auf die Qualität des gesamten Datensatzes geschlossen werden. Daher wird empfohlen die Formatkonsistenz, die logische Konsistenz und die Vollständigkeit für möglichst den gesamten Datensatz zu untersuchen (ISO, 2011). Die Evaluierung der Genauigkeit, kann auch auf Teile des Datensatzes beschränkt sein. Abbildung 13 stellt die Unterscheidung von Evaluierungsmethoden und –umfang graphisch dar.



Quelle: Eigene Darstellung nach ISO 19157:2011

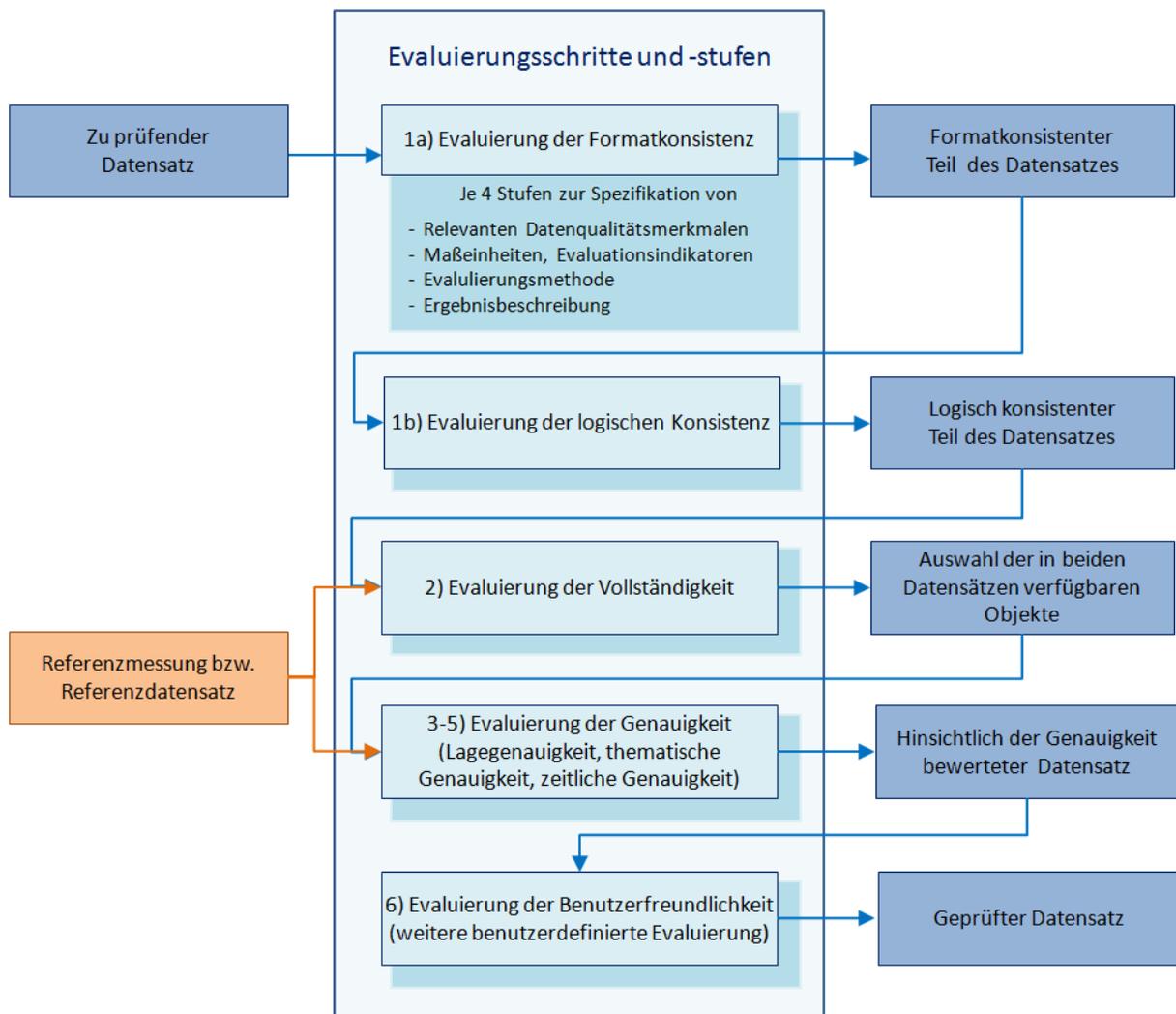
Die ISO Norm ISO 19157:2011 empfiehlt die Evaluierungsschritte in folgender Reihenfolge vorzunehmen. Zuerst soll die Formatkonsistenz geprüft werden. Daten, die nicht gelesen oder interpretiert werden können, werden identifiziert, dokumentiert und von der weiteren Bearbeitung ausgeschlossen. Danach wird die

logische Konsistenz des Datensatzes überprüft und abermals nicht konforme Daten ausgeschlossen. Erst dann wird die Vollständigkeit des Datensatzes überprüft. Hier wird der vorliegende Datensatz zumeist mit einer Referenzmessung bzw. einem Referenzdatensatz verglichen. Schließlich wird die Genauigkeit (Lagegenauigkeit, thematische Genauigkeit und zeitliche Genauigkeit) überprüft. Dieser Analyseschritt bezieht sich wiederum auf einen Referenzdatensatz und kann auf Auswahlgebiete räumlich begrenzt sein. (ISO 19157:2011, S. 149f). Schließlich wird der Datensatz auf Basis hinsichtlich weiterer anwendungsspezifischer Anforderungen geprüft.

Abbildung 13: Unterscheidung von Evaluierungsmethoden und Evaluierungsumfang

Jeder einzelne Evaluierungsschritt wird in vier Stufen geteilt, die Spezifikation der relevanten Datenqualitätsmerkmale, die Spezifikation der Maßeinheiten bzw. Evaluationsindikatoren, die Spezifikation der Evaluierungsmethoden oder -prozeduren und die Ergebnisbeschreibung. Abbildung 14 illustriert die Reihenfolge der Evaluierungsschritte.

Abbildung 14: Reihenfolge der Evaluierungsschritte zur Bestimmung der Datenqualität



Quelle: Eigene Darstellung nach ISO 19157:2011 (ISO 2011)

8 Evaluierungsansätze für die Analyse der der Datenqualität von OSM-Verkehrsgraphen in der Literatur

Seit der Etablierung der OSM im Jahr 2004 wurde die Datenqualität der Straßeninformationen in mehreren Untersuchungen geprüft. Die erste systematische Evaluierung der OSM-Daten wurde für England durchgeführt, wo historisch gesehen der Bearbeitungszeitraum am längsten ist. Aber auch im deutschsprachigen Raum wo das Engagement der OSM-Nutzer besonders hoch ist, wurde bald danach begonnen die Qualität der OSM zu bewerten.

Für England führte Hakley (2010) bereits im Jahr 2008 die erste systematische Qualitätsanalyse für die OSM durch, in der er mit einer extrinsischen Evaluierung den OSM-Graph mit einem Referenzdatensatz höherer Genauigkeit (Straßengraph "Ordnance Survey Meridian 2") verglich. Hakley folgte dabei methodisch den Ansätzen von Goodchild und Hunter (1997) sowie Hunter (1999) und ermittelte die Lagegenauigkeit, die Vollständigkeit des Netzwerks sowie die Vollständigkeit ausgewählter Attribute. Neben einer großflächigen Vergleichsanalyse führte Hakley zusätzlich eine visuelle Kontrolle von fünf zufällig gewählten Rasterflächen von 113 km² Größe durch, um die Vollständigkeit der OSM-Daten genauer zu überprüfen.

Die Methoden von Hakley zur großflächigen Bewertung des OSM-Straßengraphs wurden in der Folge von mehreren Autoren aufgenommen und erweitert, z.B. für Großbritannien (Ather 2009), für Deutschland (Zielstra, Zipf 2010, Neis et al. 2011) und auch für Österreich (Thaller 2009, Graser et al. 2013).

Neis et al. (2011) zeigen am Beispiel Deutschland, dass die OSM Datenbank bereits im Jahr 2011 um 27% mehr Informationen im Straßennetz und Routinginformationen für Fußgänger enthält, als das kommerzielle Produkt TomTom (Tom Tom Multinet Datenset 2011, ehemals TeleAtlas). Wenn auch nicht fehlerlos, insbesondere hinsichtlich der für das Routing essentiellen Abbiegevorschriften, zeigt sich aus ihren Evaluierungen, dass die absolute und relative Zahl der Fehler in der OSM Datenbank zurückgeht.

Thaller (2009) untersucht für Österreich den OSM-Datensatz von 2008 hinsichtlich Lagegenauigkeit, des Umfangs inhaltlicher Elemente (topographische und thematische Inhalte), der Abdeckung des Straßennetzes für städtische und ländliche Räume (Gesamtstraßenlänge in km) sowie der Vollständigkeit und Richtigkeit des Namensguts. Die Ergebnisse der Analysen zeigen, dass bereits im Jahr 2008 die Abdeckung des Straßennetzes durch die OSM in Wien (73%) und Wien Umgebung (80%) relativ hoch ist, während periphere Gebiete durch die OSM weitaus weniger erfasst sind.

Für die Großregion Wien wurden von Graser et al. (2013) die Differenzen zwischen dem Straßengraph der OSM und dem offiziellen Österreichischen Referenzgraph der Graphenintegrations-Plattform (GIP)⁴¹ analysiert. Die Ergebnisse zeigen nach Isolation methodischer Differenzen eine hohe Übereinstimmung der Lagegenauigkeit und vergleichbare Netzwerklängen. Hinsichtlich der Vollständigkeit der Attribuierung von Straßennamen und Kantengeschwindigkeiten, verfügt die OSM über eine geringere Erfassungstiefe, als die GIP. Über den Ansatz von Hakley hinaus, wurde von Graser et al. auch die Qualität der routingrelevanten Informationen zu Abbiegevorschriften und Einbahnführung untersucht. Hier zeigten sich Unterschiede und unvollständige bzw. fehlerhafte Erfassungen der OSM vor allem im niederrangigen Straßennetz und in periphereren Gebieten, während die OSM in innerstädtischen Gebieten genauer war. Graser et al. führen diese Differenzen auf die in zentralen Lagen höhere Anzahl an OSM-Usern mit Ortskenntnis zurück.

Im Gegensatz zu den oben genannten Studien, beschreiben Rehl K. et. al. (2012) eine an die ISO Norm 19157:2011 angelehnte Methodik zur intrinsischen Qualitätsevaluation von Verkehrsgraphen. Dabei werden Vollständigkeit, logische Konsistenz und thematische Genauigkeit der OSM intrinsisch untersucht. Da für die

⁴¹ Im Rahmen der Graphenintegrations-Plattform, werden die digitalen Geodaten für das gesamte Verkehrsnetz Österreich, auf welche sich alle Behörden beziehen vernetzt. Das Ergebnis wird als gemeinsamer österreichweiter Verkehrsgraph etabliert, siehe auch <http://www.gip.gv.at/> (17.07.2014).

gegenständliche Arbeit kein offizieller Referenzdatensatz zur Verfügung steht, wird beim Aufsetzen der Evaluierungsmethodik dieser Ansatz in großen Teilen weiterverfolgt.

In der Folge werden die in der Literatur angewandten Evaluierungsansätze für die Qualitätsmerkmale logische Konsistenz, Vollständigkeit, Lagegenauigkeit und thematische sowie zeitliche Genauigkeit vorgestellt.

8.1 Logische Konsistenz

Dargestellt werden Evaluierungsansätze für die Ermittlung der topologischen Konsistenz. Für die zeitliche Konsistenz, Formatkonsistenz und die Konsistenz der Wertebereiche wurden keine speziellen Evaluierungsansätze für die OSM gefunden.

8.1.1 Topologische Konsistenz

Ein Prüfalgorithmus zur Identifikation von Topologiefehlern wurde von Rehl et al. (2011) entwickelt. Dabei werden zwei Arten von topologischen Fehlern untersucht. Kreuzende Segmente ohne gemeinsamen Knoten und Segmente in unmittelbarer Nähe zueinander, aber ohne Verbindung werden dabei als potenzielle Problemfälle erfasst.

Kreuzende Segmente ohne Kreuzungspunkt werden dann als problematisch definiert, wenn kein Segment davon als Brücke oder Tunnel attribuiert wurde und der in OSM definierte "Layer", d.h. Straßenebene, identisch ist. Semantisch gesehen, sind diese Segmente planfreie Kreuzungen wie Unterführungen oder Brücken, werden in der OSM aber auch zum Auflösen komplexer Kreuzungen genutzt. Neis et al. (2011) untersuchen diese Art von topologischen Fehler ebenfalls.

Parallel verlaufende Kanten oder aufeinander zuführende Kanten mit einem Abstand von maximal fünf Metern werden dann als potenzielle Fehler definiert, wenn sie über keinen gemeinsamen Knoten verfügen. Damit können fehlende Kreuzungsverbindungen, sogenannte "undershoots" identifiziert werden, aber auch Doppelerhebungen von Straßen bzw. Straßenabschnitten. Auch dieser Fehlertyp wird von Neis et al. (2011) untersucht, wenn auch die Fehlertoleranz durch einem Abstand von nur einem Meter definiert ist.

Um die Verbundenheit des Graphs festzustellen untersuchten Rehl et al. (2011) darüber hinaus die Anzahl der isolierten Kanten mit dem Ziel die Existenz von Inselgraphen ohne Verbindung zum Hauptgraph zu identifizieren. Die Berechnung wurde mit Hilfe von Tarjan's Algorithmus (Tarjan, 1971) durchgeführt, der auch die Zahl der befahrbaren Kanten, die Anzahl der Kanten die nur ein eine Richtung befahrbar sind (Einbahnen), die Zahl stark verbundenen Komponenten und der Straßen die nur in eine Richtung mit der größten verbundenen Komponente verbunden sind ermittelt. Wesentlich für einen routenfähigen Straßengraph ist aber in erster Linie, ob jeder Knoten von jedem anderen Knoten aus erreichbar ist - also die Abwesenheit von isolierten Kanten.

8.2 Vollständigkeit

Hinsichtlich der Bewertung der Vollständigkeit wird zwischen der Vollständigkeit des Netzwerkes und jener der Attribute unterschieden. Evaluierungsansätze werden für beide Varianten in der Folge dargestellt.

8.2.1 Vollständigkeit des Netzwerkes

Extrinsische Bewertung der Vollständigkeit des Netzgraphs

Verfügbare Studien wählten für die Bewertung der Vollständigkeit(Q_v) des OSM-Straßengraphs zumeist die Differenzbetrachtung mit einem als vollständigen definierten Referenzdatensatz, wobei die Netzwerklängen der Datensätze verglichen werden (Haklay 2010, Neis et al. 2012, Thaller 2009, Graser et al. 2013).

$$Q_v = \sum_{i=1}^n \text{Netzwerklänge OSM} - \sum_{i=1}^n \text{Netzwerklänge Referenzdatensatz}$$

wobei n = Anzahl der Kanten je Untersuchungsgebiet bzw. je räumlicher Einheit im Untersuchungsgebiet

Während Hakaly (2010) die Netzwerklängen für das in 1 km große Raster geteilte Untersuchungsgebiet vergleicht und dabei auch Aussagen über die Zusammenhänge zwischen dem Einkommensniveau je Raster und Vollständigkeit des OSM-Straßennetze macht, differenzieren Neis et al. (2011) nach der Bevölkerungsdichte administrativer Einheiten und leiten Ergebnisse für stark und gering besiedelte Gemeinden ab sowie Ergebnisse differenziert nach Wegetypen. Graser et al. (2013) klassifizieren die Ergebnisse ebenfalls nach Straßenkategorie.

Im Stadtgebiet von London überprüfte Haklay (2010) die Vollständigkeit zusätzlich durch eine visuelle Kontrolle von fünf zufällig gewählten Rasterflächen von 113 km² Größe.

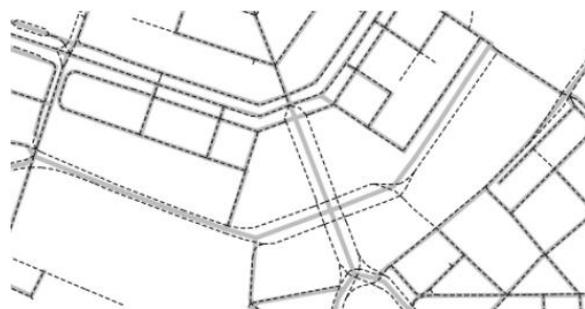
Haklay (2010) stellt in seiner Untersuchung für England räumliche Differenzen in der Vollständigkeit des OSM-Graphs fest. Während die Zentren der großen Städte sehr hohe Vollständigkeit aufweisen, sinkt die Erfassungsqualität zu den Rändern der Stadt ab. Insbesondere an den Stadtgrenzen und in peripheren Räumen sind die OSM-Daten unvollständig. Darüber hinaus attestiert Haklay auch soziale Unterschiede zwischen den mit 76% OSM-Abdeckung vollständiger erfassten reicheren und unvollständiger erfassten ärmeren Regionen mit nur 46% OSM-Abdeckung.

Auch für Deutschland stellen Neis et al (2011) fest, dass im Jahr 2011 Gebiete mit unvollständig erfasster Netzwerklänge vor allem in dünn besiedelten Gebieten, d.h. in Kleinstädten mit 10.000 - 50.000 Einwohnern und in ländlichen Gebieten mit weniger als 10.000 Einwohnern liegen. Sie prognostizieren jedoch - bei weiter fortschreitender Erhebungsarbeit - eine weitgehend vollständige Erfassung bis Mitte des Jahres 2012.

In der aktuellsten Untersuchung von Graser et al. (2013) zeigen die Ergebnisse nach Straßenkategorien unterschiedlich hohe Übereinstimmung. Während die Differenz der Netzwerklänge bei Bundesstraßen mit 2% nur geringfügig abweicht, werden für andere Straßenkategorien deutlich höhere Abweichungen von 15% für Erschließungsstraßen bis zu 28% für Autobahnen ausgewiesen. Graser et al. Führen dies auf die Differenzen im Datenmodell des GIP-Verkehrsgraphs und der OSM-Datenbank zurück. So sind im herangezogenen Referenzgraph der GIP im Gegensatz zur OSM keine Forst- und Feldwege enthalten, Autobahnrampen nicht den Autobahnen zugeordnet und bei den Erschließungsstraßen nicht immer für jede Richtungsfahrbahn eine Kante modelliert, wenn diese z.B. durch eine Fußgängerinsel getrennt sind (

).

Abbildung 15: Unterschiede in der Modellierung der OSM (schwarze unterbrochene Linie) und dem GIP (dicke graue Linie)



Quelle: Graser et. al, 2013 p. 13

Die Ergebnisse von Haklay und Neis et al. zeigen, dass für die Interpretation der OSM-Daten eine Differenzierung nach Bevölkerungsdichte und zentralen sowie peripheren Lagen insofern sinnvoll ist, als die Zahl der OSM-Nutzer und damit die Vollständigkeit der OSM-Daten mit der Bevölkerungsdichte steigt. Darüber hinaus zeigt sich eine Klassifizierung der Ergebnisse nach Straßenkategorien wie bei Graser et al. als hilfreich für die Interpretation von unterschiedlichen Erhebungsstandards zwischen den Datensätzen. Diese Unterscheidung ist auch für die Bewertung der Relevanz von Unter- bzw. Übererfassungen relevant, da Lücken im hochrangigen Straßennetz Routingergebnisse deutlich stärker verzerren als unvollständig erfasste Feldwege.

Intrinsische Bewertung der Vollständigkeit des Netzgraphs

Rehrl. et al. (2011) wählen für die Bewertung der Vollständigkeit einen intrinsischen Ansatz und werten den Grad der Erschließung von CORINE-Landcover Siedlungsgebieten aus. Dabei werden Puffer von 50m bis 400m um die Straßenkanten gelegt und berechnet, welcher Prozentsatz der Siedlungsflächen mit Straßeninfrastruktur erschlossen ist.

Eine darüber hinaus gehende, räumliche Differenzierung des Erschließungsgrades nach der Bevölkerungsdichte erscheint den Ergebnissen der extrinsischen Betrachtung folgend für die Interpretation jedoch hilfreich.

8.2.2 Vollständigkeit der Attribute

Straßennamen, Straßenklasse und Höchstgeschwindigkeit

Wichtige Qualitätsparameter für routenfähige Graphen sind die Vollständigkeit der Straßennamen sowie Angaben zur höchstzulässigen Geschwindigkeit.

An der Verfügbarkeit von Straßennamen wird von Haklay (2010) die tatsächliche physische Ortskenntnis des Erhebers abgeleitet, da bei der OSM-Digitalisierung ausdrücklich darauf hingewiesen wird, Straßennamen nicht aus bestehenden Karten zu kopieren, um das Copyright nicht zu verletzen.

Neis et al. (2011) argumentieren ähnlich, wenn sie die Gründe für fehlende Straßennamen in der OSM interpretieren. Für Routing-Anwendungen ist die Verfügbarkeit von Straßennamen darüber hinaus wichtig, wenn es um Quell- oder Zielangaben der Wegeführung oder um die Orientierung der Nutzer geht. Außerdem können aus dem Straßennamen bereits bei der Digitalisierung der Straßenzüge wichtige Informationen zum Straßenverlauf, der Kategorisierung von Straßen oder der Abbiegebeziehungen abgeleitet werden. (Neis et al. 2011)

Für die Auswahl des kürzesten Weges ist die Geschwindigkeit auf der Kante eine wesentliche Größe. Im Datenmodell der OSM wird die standardmäßige Geschwindigkeitsbeschränkung der Straßenklasse unterstellt, wenn keine explizite Angabe gemacht wird.⁴² Diesem Ansatz folgt das Konvertierungstool `osm2po`, wenn es aus den OSM-Daten einen routenfähigen Straßengraph erstellt. Eine Online-Ansicht der verfügbaren OSM-Tags für Geschwindigkeitsbeschränkungen (`maxspeed`) bietet das Unternehmen ITOmap auf seiner Website⁴³ zumindest für das hochrangige Straßennetz, siehe auch Abbildung 5.

Haklay (2010) untersucht die Attribuierung der Straßennamen, wie auch Neis et al. (2011) und Graser et al. (2013). Letztere erweitern die Untersuchung um die Vollständigkeit der Geschwindigkeitsangaben. In der Großregion Wien verfügen 78% des OSM-Straßennetzes über Straßennamen, während nur 43% des OSM-Straßennetzes eine Höchstgeschwindigkeit zugewiesen ist (Graser et al. 2013). In Deutschland sind die "namenlosen" Straßen vor allem in und um Wohngebiete zu finden (Neis et al. 2011), in England vor allem in den ärmeren Regionen, was Haklay auf die geringere Anzahl von OSM-Teilnehmern zurückführt.

Auch Rehrl et al. (2012) untersuchen die Vollständigkeit der Attribuierung der Straßennamen und der Geschwindigkeitsbeschränkungen der OSM. Sie kommen in ihrer Untersuchung zu dem Schluss, dass das OSM-Tag "`maxspeed`" aufgrund der unvollständigen Erfassung von nur 9,1% der Netzwegelänge in ihrem Untersuchungsgebiet für die Berechnung realistischer Reisezeiten nicht verwendet werden kann. Sie führen daher andere Attribute wie die Straßenklasse sowie die Lage der Straße (z.B. innerhalb von Siedlungsgebieten) als wesentlich für die Berechnung realistischer Kantengeschwindigkeiten an.

Abbiegevorschriften

Neis et al. (2011) werten die Anzahl der Abbiegevorschriften, seien es Verbote oder Anweisungen, nach Straßenkategorie und Bevölkerungsdichte im OSM-Straßengraph aus und setzen sie in Relation zu der Anzahl der Abbiegevorschriften im Referenzgraph. Insgesamt ist für Deutschland im Jahr 2011 die Anzahl der in der

⁴² OSM Wiki, verfügbar unter: http://wiki.openstreetmap.org/wiki/OSM_tags_for_routing/Maxspeed#Austria (19.08.2014)

⁴³ Ito World, verfügbar unter <http://www.itoworld.com/map/124?lon=16.38678&lat=48.12537&zoom=10> (19.08.2014)

OSM-Datenbank verfügbaren Abbiegevorschriften mit einem Verhältnis von 1:9 um ein vielfaches geringer als im Referenzgraph von TomTom. Auch Graser et al. (2013) vergleichen die Anzahl der Abbiegevorschriften in beiden Straßengraphen. Da im Referenzgraph aber alle Abbiegevorschriften implizit definiert sind, im OSM-Graph aber alle Abbiegerelationen erlaubt sind, solange sie nicht explizit verboten werden, kann dieselbe Kreuzungssituation in der OSM mit deutlich weniger Abbiegevorschriften modelliert werden, als im Referenzgraph. Graser et al. (2013) kommen daher zu dem Schluss, dass ein Vergleich der Anzahl an Abbiegevorschriften allein nicht repräsentativ für die Vollständigkeit des OSM-Graphs ist. Sie wählen daher eine systematische Untersuchung mit einem routenbasierten Ansatz zur Kalkulation der Abbiegerelationen. Dabei wird die Wegelänge gleicher Abbiegemanöver für die OSM und den Referenzgraph ermittelt und die thematische Genauigkeit der Abbiegerelationen untersucht (siehe Kapitel 8.4).

8.3 Lagegenauigkeit

Extrinsische Bewertung der Lagegenauigkeit

Die Lagegenauigkeit wird erstmals von Hakley (2010) basierend auf den Methoden von Goodchild und Hunter (1997) und Hunter (1999) ermittelt und in der Folge von anderen Autoren immer wieder aufgegriffen (siehe Thaller 2009, Graser et al. 2013). Dabei werden Pufferanalysen verwendet, um jenen Anteil der Straßenlinien zu ermitteln, der innerhalb einer bestimmten Distanz des gleichen Objekts im Referenzdatensatz mit höherer Genauigkeit liegt (Hakley 2010). In England ermittelt Hakley damit für Autobahnen und einem Puffer von 20m eine durchschnittlich 80% Übereinstimmung zwischen den Datensets. In einer weiteren visuellen Untersuchung werden in fünf zufällig ausgewählten Gebieten je 100 Straßenabschnitte beider Datensets miteinander verglichen. Dabei kommt er zu regional unterschiedlichen Ergebnissen der Lagegenauigkeit der OSM-Daten die von Gebiet zu Gebiet zwischen 3 und 8 m vom Referenzdatensatz abweichen.

8.4 Thematische Genauigkeit

Extrinsische Bewertung der thematischen Genauigkeit von Abbiegevorschriften und Einbahnen

Graser et al. (2013) untersuchen die thematische Genauigkeit von Abbiegevorschriften und Einbahnstraßen im Hilfe eines routingbasierten Ansatzes. Dabei werden bei den Abbiegevorschriften die Liniengeometrien für alle Abbiegeverbote im OSM-Graph mit der Liniengeometrie des kürzesten Weges im Referenzgraph verglichen, der vom Start- zum Endpunkt des OSM-Abbiegeverbots führt. Ist die Fahrt von A nach B im OSM-Graph verboten, wird die Länge der Liniengeometrie von A nach B in der OSM mit der Liniengeometrie im Referenzgraph für dieselbe Relation verglichen. Gibt es im Referenzgraph eine Liniengeometrie, die dem Abbiegeverbot im OSM-Graph entspricht, ist das OSM-Abbiegeverbot falsch implementiert, da es eine Relation verhindert, die laut Referenzgraph erlaubt sein müsste.

Die thematische Genauigkeit der Einbahnstraßen wird ähnlich ermittelt. Es wird für alle Einbahnen der OSM die Wegelänge vom Startpunkt A zum Endpunkt B mit der Wegelänge des Referenzgraphs in die entgegengesetzte Relation (von B nach A) verglichen. Die Einbahn ist dann korrekt in OSM implementiert, wenn der Referenzgraph nur einen deutlich längeren Weg für die Strecke von B nach A findet, als OSM für A nach B ausweist.

8.5 Zeitliche Genauigkeit

Neis et al. (2011) werteten zur Evaluierung der zeitlichen Genauigkeit auch die Zeitangaben, die sogenannten "timestamps", der einzelnen Objekte in der OSM aus. Jede Änderung eines Objekts führt zu einer Aktualisierung dieses Zeitstempels. Mitte des Jahres 2011 stammte je ein Drittel der Straßenobjekte der OSM für Deutschland aus dem Jahr 2011, dem Jahr 2010 oder den Jahren 2008 bis 2009. Die Aktualität der OSM kann für Deutschland also mit Geodaten offizieller Provider verglichen werden, wobei die systematische Aktualisierung der OSM-Daten jedoch nicht vorausgesetzt werden kann. Tatsächlich werden erst die nächsten Jahre zeigen, ob eine regelmäßige Aktualisierung der OSM-Daten stattfinden wird, da bislang die größten Anstrengungen darauf konzentriert waren, "weiße Flecken" auf der OSM-Karte zu entfernen (Neis et al., 2011).

Die Konsistenz der temporären Angaben wird in den vorgestellten Evaluierungsansätzen jeweils als Teil der Vollständigkeitsbewertung untersucht und wurde nicht explizit ausgewertet. Sie kann ohne Vergleich mit einem Referenzdatensatz nicht dargestellt werden.

8.6 Zusammenfassung der ausgewählten Evaluierungsansätze

Zumeist wird die Qualität der Daten der OpenStreetMap mit Hilfe eines Referenzgraphs extrinsisch ermittelt. Ausnahme ist die Feststellung der topologischen Konsistenz die intrinsisch, d.h. basierend auf den Daten des zu untersuchenden Datensatzes, ermittelt wird, die Vollständigkeit der Namensattribuierung innerhalb des Datensatzes und die zeitliche Genauigkeitsprüfung anhand der OSM-Timestamp Angabe. Für die übrigen Qualitätsmerkmale stellen nur Rehl et al. (2012) intrinsische Untersuchungen an. Die Benutzerfreundlichkeit wird als solche in der Literatur nicht eigens behandelt.

Für die Interpretation der Bewertungsergebnisse scheint eine Differenzierung nach Straßenkategorien und Siedlungsdichten bzw. zwischen städtischen und peripheren Regionen oftmals sinnvoll.

Tabelle 8 fasst die in der Literatur angewandten Evaluierungsansätze in einer Übersicht zusammen.

Tabelle 8: Übersicht über die in der Literatur angewandten Evaluierungsansätze für die OSM

Qualitätsmerkmal	Autor	Typ	Evaluierungsansatz
Topologische Konsistenz	Rehl et al.	Intr.	Verbundenheitsprüfung mit Tarjan's Algorithmus (Tarjan, 1971) zur Feststellung von isolierten Kanten, Einbahnen, stark verbundenen Komponenten und Straßen die nicht in beide Fahrtrichtungen mit der größten verbundenen Komponente verbunden sind.
	Rehl et al. Neis et al.	Intr.	Anzahl der topologischen Fehler durch nicht verbundene, überschneidende Segmente die keine Brücken oder Tunnel sind. Anzahl der angrenzenden, nicht verbundene Segmente die sich bis auf 5m nähern, aber keinen gemeinsamen Knoten haben.
Lagegenauigkeit	Haklay, Thaller, Graser et al.	Extr.	Prozentanteil der Netzwerklänge des Graphs innerhalb einer definierten Entfernung (Puffer) zum Referenzgraph (nach Goodchild & Hunter 1997)
Vollständigkeit der Objekte	Hakaly Graser et al. Neis et al.	Extr.	Vergleich der Netzwerklängen für das in kleinere räumliche Einheiten aufgeteilte Untersuchungsgebiet. Diese können Rasterzellen, administrative Einheiten oder andere räumliche Abgrenzungen z. B. administrative Einheiten nach Bevölkerungsdichte sein, z. B. Hakaly: ... 1km Rasterzellen Neis et al.: ... administrative Grenzen, klassifiziert nach Bevölkerungsdichte Graser et al.: ... benutzerdefinierte Regionen, Ergebnisse klassifiziert nach Straßenkategorie
	Hakaly	Extr.	Visuelle Überprüfung der Vollständigkeit des Netzgraphs für fünf zufällig gewählte Rasterzellen von 113km ² Größe
	Rehl et al.	Intr.	Anteil des durch den Verkehrsgraph erschlossenen Siedlungsgebiets an der Fläche des gesamten Siedlungsgebiets gemäß CORINE Landcover
Vollständigkeit der Attribute	Hakaly	Extr.	Prozentanteil des Straßennetzes mit Straßennamen
	Neis et al.	Intr.	Prozentanteil des Straßennetzes mit und ohne Straßennamen im Untersuchungsgebiet
	Graser et al.	Extr.	Prozentanteil des Straßennetzes mit Straßennamen und Informationen zur Geschwindigkeitsbeschränkung
	Rehl et al.	Intr.	Prozentanteil des Straßennetzes mit Straßennamen und Informationen zur Geschwindigkeitsbeschränkung
	Neis et al.	Extr.	Anzahl der Abbiegevorschriften nach Straßenkategorie und Stadtklassifikation

Qualitätsmerkmal	Autor	Typ	Evaluierungsansatz
Thematische Genauigkeit - Abbiegevorschriften	Graser et al.	Extr.	Routingbasierter Ansatz Vergleich der Liniengeometrie aller Abbiegeverbote im Graph mit der Liniengeometrie des kürzesten Weges im Referenzgraph der vom Start- und Endpunkt des Abbiegeverbots führt. (Ähneln sich die verbotene Liniengeometrie und der kürzeste Weg im Referenzgraph, ist das Abbiegeverbot falsch implementiert.)
Thematische Genauigkeit - Einbahnstraßen	Graser et al.	Extr.	Routingbasierter Ansatz Vergleich der Wegelänge vom Startpunkt zum Endpunkt aller Einbahnen in Fahrtrichtung mit der Wegelänge im Referenzgraph für die entgegengesetzte Relation (von Endpunkt zu Startpunkt).
Zeitliche Genauigkeit	Neis et al.	Intr.	Jahr der letzten Änderung der Straßengraphobjekte (timestamp)

Extr... Extrinsische Evaluierung in Differenzbetrachtung mit einem Referenzgraph;

Intr... Intrinsische Evaluierung ohne Referenzgraph.

Quelle: Eigene Darstellung

9 Evaluierung der Datenqualität des Verkehrsgraphs

Der erstellte Verkehrsgraph wird im Folgenden mit direkten Evaluierungsmethoden bewertet. Dabei werden in erster Linie Bewertungen basierend auf internen Informationen vorgenommen, die bei entsprechender Datenverfügbarkeit um eine Differenzbetrachtung zu einem Referenzdatensatz ergänzt werden.

Analog zum Standard ISO 19157:2011 folgt der Evaluierungsprozess vier Bearbeitungsschritten, der Spezifikation der relevanten Datenqualitätsmerkmale, der Spezifikation der Maßeinheiten bzw. Evaluationsindikatoren, der Spezifikation der Evaluierungsmethoden oder -prozeduren sowie der Ergebnisbeschreibung.

9.1 Spezifikation der Evaluierungsschritte und -stufen

Der Verkehrsgraph wird hinsichtlich der vier Qualitätsmerkmale logische Konsistenz, Vollständigkeit, thematische Genauigkeit und Benutzerfreundlichkeit untersucht und bewertet. Das Qualitätsmerkmal Lagegenauigkeit kann aufgrund des fehlenden Referenzgraphs nicht geprüft werden, dies gilt ebenso für das Qualitätsmerkmal zeitliche Konsistenz, welches aber im Rahmen der Vollständigkeit der Objekte weitgehend abgedeckt werden kann.

Je Qualitätsmerkmal wird zumindest ein Indikator festgelegt, über welchen die Qualität des Vektorgraphs in Bezug auf die spätere Einzugsbereichsberechnung festgelegt wird.

Da für das Untersuchungsgebiet kein Referenzstraßengraph für die extrinsische Bewertung des Verkehrsgraphs zur Verfügung steht, wird die intrinsische Qualität des Datensatzes in erster Linie durch definierte Qualitätsmerkmale ohne externe Referenzdaten evaluiert. Ergänzend wird nach Möglichkeit mit frei verfügbaren Referenzdaten versucht, die Qualitätsbewertung zu vervollständigen. Die Evaluierung soll weitgehend für den gesamten Datensatz erfolgen.

Aus den in den Kapiteln 7.3 Qualitätsmerkmale und 8 Evaluierungsansätze für die Analyse der der Datenqualität von OSM-Verkehrsgraphen in der Literatur vorgestellten Qualitätsmerkmalen und -parametern werden für die Fragestellung von Einzugsbereichsberechnung und Erreichbarkeitsanalysen die in Tabelle 9 dargestellten Evaluierungsansätze und Indikatoren abgeleitet und die Qualitätsanforderungen in Hinblick auf den Anwendungsfall definiert.

Tabelle 9: Spezifikation der Datenqualitätsmerkmale, Maßeinheiten und Indikatoren

Qualitätsmerkmal / -parameter	Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung
1. Logische Konsistenz			
1.1 Formatkonsistenz	Können die OSM-Daten nach den durchgeführten Bearbeitungs- und Konvertierungsschritten in einer PostgreSQL/ PostGIS-Datenbank als Verkehrsgraph weiter verarbeitet werden?	Binär (ja / nein)	Ja
1.2 Konzeptuelle Konsistenz	Ist das Datenmodell in sich konsistent und entspricht es dem für ein Datenrouting erforderlichen Datenmodell?	Binär (ja / nein)	Ja
1.3 Konsistenz der Wertebereiche	Sind alle für das Routing erforderlichen Daten im generierten Datensatz enthalten?	Binär (ja / nein)	Ja
	Sind die Wertebereiche der Attribute konsistent mit diesem Datenmodell?	Binär (ja / nein)	Ja
1.4 Topologische Konsistenz	Anzahl der topologischen Fehler und %-Anteil der fehlerhaften Kanten bzw. Knoten an allen Kanten bzw. Knoten in Anlehnung an Rehr et al. (2012) und die Forschungsgesellschaft Mobilität (2011).		Anteil fehlerhafter Objekte < 5% von allen Objekten

Qualitätsmerkmal / -parameter	Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung
	Isolierte Einzelsegmente (Einzelkanten) ohne Verbindung zum "Hauptstraßengraph"	Isolierte Straßen-segmente: Anteil der isolierten Straßen-segmente bestehend aus einer oder mehreren Kanten ohne Verbindung zum Hauptstraßengraph	
	Isolierte Segmente aus mehr als einer Kante ohne Verbindung zum Hauptstraßengraph		
	Knoten, die Lücken im Netz darstellen. Diese Knoten haben, wie Sackgassen, nur zu einer Kante eine Verbindung, sind aber semantisch keine Sackgassen, sondern unvollständig digitalisierte Straßenzüge.	Lücken im Straßennetz: Anteil der Knoten, an welche nur eine Kante anschließt, die aber keine Sackgassen sind	
	Inkorrekte planfreie Kreuzungen die weder Brücken oder Unterführungen noch für die Auflösung komplizierter Fahrtrationen herangezogen werden.	Fehlende Schnittpunkte kreuzender Kanten: Anteil der sich schneidenden Kanten, die keine planfreien Kreuzungen sind.	
	Kanten, deren Quellknoten gleichzeitig auch Zielknoten ist, aber keine Schleifen darstellen.	Schleifen: Anteil der Kanten, deren Anfangs- und Endpunkte fälschlicher Weise auf denselben Knoten verweisen	
2. Vollständigkeit			
2.1 Vollständigkeit der Objekte	Erschließungsquote und Erschließungsgrad der Siedlungsgebiete Anteil der Siedlungsgebiete und der Siedlungsfläche, die durch Straßeninfrastruktur erschlossen wird. Die Siedlungsfläche wird dem frei verfügbaren Geodaten-satz der Siedlungseinheiten, Gebietsstand 2010 von Statistik Austria entnommen. Der Erschließungsgrad wird in vier Varianten gerechnet, indem ein Puffer von 50m - 400m um die Straßenkanten gelegt wird.	Erschließungsquote: Anteil der Siedlungsgebiete (Siedlungseinheiten) in %, der durch Straßen in einem Umkreis von 50, 100, 150, 200 und 400 m erschlossen wird.	Erschließungsquote von 100%
		Erschließungsgrad: Anteil der Siedlungsfläche in %, der durch Straßen in einem Umkreis von 50, 100, 150, 200 und 400 m erschlossen wird.	Erschließungsgrad von 95%
2.2 Vollständigkeit der Attribute	Anzahl, Anteil und Länge der Kanten räumlich differenziert und nach Straßenkategorie, für die Angaben zur Straßenklasse verfügbar sind.		
		Anteil der Straßenlänge mit Angabe zur Straßenkategorie an der Gesamtstraßenlänge in %	100 %
		Anteil der Straßenlänge der Kategorie „einfachster Ausbauzustand“ an der Gesamtstraßenlänge in %	< 10 %

Qualitätsmerkmal / -parameter	Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung
	für die die Namensangaben verfügbar sind.	Anteil der Straßenlänge nach Straßenkategorie an der Gesamtstraßenlänge in %	
		Autobahnen, Schnellstraßen	< 3 %
		Haupt-, Landesstraßen	< 3 %
		Nebenstraßen, Erschließungsstraßen	< 20 %
		Straßen unbekannter Kategorie und einfachsten Ausbaustandes	keine
3. Thematische Genauigkeit	Sind die Straßenkategorien korrekt attribuiert? Wurde die korrekte Kantengeschwindigkeit zugewiesen?	Anzahl der falsch attribuierten Autobahnabschnitte	< 5
	Gibt es unplausible Geschwindigkeitsattribuierungen die durch visuelle Kontrolle identifiziert werden können?	Anteil der Straßenlänge je Straßenkategorie mit unplausiblen Geschwindigkeitsangaben	< 3 %
4 Benutzerfreundlichkeit	Ist die Dokumentation der verwendeten Softwareprodukte ausreichend, um den Inhalt des Straßengraphs und die Ergebnisse der durchgeführten Berechnungen korrekt zu interpretieren?	Binär (ja / nein)	Ja

Quelle: Eigene, erweiterte Darstellung auf Basis von Hakley (2010), Rehl et al. (2012) und Graser (2013)

9.2 Evaluierung der logischen Konsistenz

Die logische Konsistenz der Datenbank wird auf vier Ebenen gemessen: der formatbezogenen, der konzeptuellen, der Konsistenz der Wertebereiche und der topologischen Konsistenz. Im konkreten Anwendungsfall ist die vollständige, uneingeschränkte Einhaltung der Konsistenzvorschriften eines routingfähigen Verkehrsgraphs erforderlich.

9.2.1 Formatkonsistenz

Evaluierungsmethode

Nur formatkonsistente Dateien können mit den Werkzeugen osm2po, QGIS und pgRouting weiter verarbeitet werden. Wesentliche Konsistenzprüfungen werden automatisiert bei der Nutzung des objektrelationaler Datenbankmanagements (DBMS) von PostgreSQL sowie der Erweiterung PostGIS durchgeführt.

Grundsätzlich muss beim Import der Daten auf die Definition der SRID geachtet werden. OSM-Daten werden in der Projektion WGS 84 (SRID 4326)⁴⁴ erhoben, entsprechend wurde die SRID 4326 beim Import der von osm2po erstellten SQL-Tabelle in die Datenbank definiert.

Die Formatkonsistenz wird geprüft, in dem der von osm2po generierte SQL-Datendatz mit dem routenfähigen Verkehrsgraph in eine PostgreSQL-Datenbank geladen und in QGIS visualisiert wird. Die Formatkonsistenz des Verkehrsgraphs ist gegeben, wenn der Import in eine PostgreSQL-Datenbank fehlerfrei möglich war, die Daten in QGIS visualisiert werden können und ein Testrouting erfolgreich war.

Code 7: SQL-Code zur Berechnung des kürzesten Weges zwischen zwei Knoten zum Nachweis der Formatkonsistenz

PgRouting Funktion zur Berechnung des kürzesten Weges

```
Varchar pgr_dijkstra(text sql, integer source, integer target, Boolean directed, Boolean has_rcost);
```

und die SQL query wie folgt definiert ist:

```
SELECT id, source, target, cost [,reverse_cost] FROM edge_table]
```

wobei

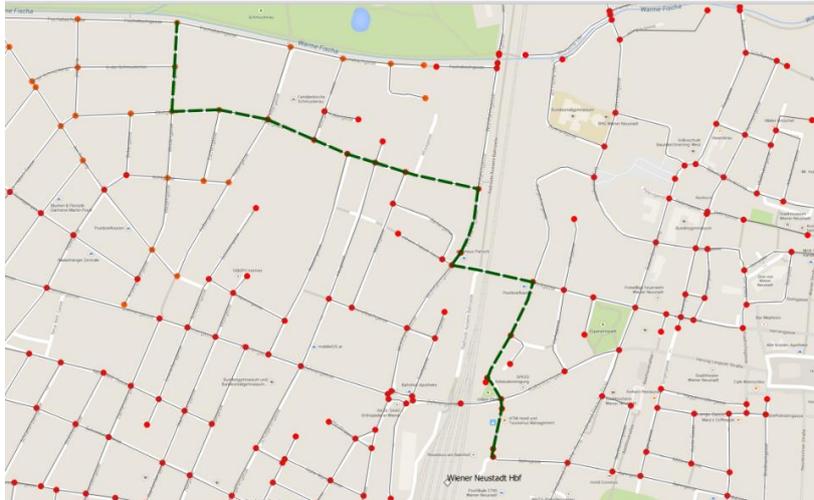
<i>id</i>	<i>... Identifier der Kante,</i>	<i>reverse_cost</i>	<i>... Kantenfahrzeit in Gegenrichtung</i>
<i>source</i>	<i>... Identifier des Quellknotens,</i>	<i>directed</i>	<i>... wahr, wenn der Graph gerichtet ist,</i>
<i>target</i>	<i>... Identifier des Zielknotens,</i>	<i>has_rcost</i>	<i>... wahr wenn die Spalte reverse_cost für die</i>
<i>cost</i>	<i>... Kantenfahrzeit,</i>		<i>Kantenfahrzeit in Gegenrichtung benutzt wird.</i>

Berechnung des kürzesten Weges aus der Kantentabelle ws_2po_4pgr

```
SELECT seq, id1 as node, id2 as edge, geom_way
FROM pgr_dijkstra(
  'SELECT id, source, target, cost, FROM ws_2po_4pgr', 51567, 36993, false, false
) AS di
JOIN ws_2po_4pgr ws
ON di.id2=ws.id;
```

Quelle: Eigene Darstellung nach Tabuko, 2013

⁴⁴ <http://spatialreference.org/ref/epsg/4326/> (15.10.2013)

Abbildung 16: Ergebnis des Kürzeste Wege Testroutings mit pgRouting zum Nachweis der Formatkonsistenz

Quelle: Eigen Darstellung basierend auf OSM, QGIS und pgRouting

Ergebnis

Die OSM-Daten können nach der Bearbeitung und Konvertierung in einer PostgreSQL-/PostGIS-Datenbank als routenfähiger Verkehrsgraph eingesetzt werden (Tabelle 10). Die Formatkonsistenz ist damit gegeben.

Tabelle 10: Ergebnisse der Bewertung der Formatkonsistenz

Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis der Bewertung
Können die OSM-Daten nach den durchgeführten Bearbeitungs- und Konvertierungsschritten in einer PostgreSQL/ PostGIS-Datenbank als Verkehrsgraph weiter verarbeitet werden?	Binär (ja / nein)	Ja	Ja

Quelle: Eigene Darstellung

9.2.2 Konzeptuelle Konsistenz

Evaluierungsmethode

In PostGIS kann die Validität von Geodaten auf Konformität mit den OpenGIS Spezifikationen des Open Geospatial Consortiums (OGC) mit der booleschen Funktion `ST_IsValid()` : Integer geprüft werden. Ist die Geometrie wohlgeformt, wird "true" zurückgegeben (Brinkhoff, 2005, S. 91). Mit der Funktion `ST_IsValidDetail` werden im Fall von Konformitätsfehlern auch die Gründe für unzureichende Validität ausgegeben. Zumeist werden mit dieser Funktion Polygone, nicht aber lineare Objekte überprüft. PostGIS erweitert die Definition der Funktion `ST_IsValid` insofern, als "false" zurückgegeben wird, wenn ein Streckenzug (line string) weniger als zwei eindeutige Eckpunkte hat.⁴⁵

Code 8: SQL-Befehl zur Prüfung der konzeptuellen Konsistenz

```
SELECT ST_IsValid(geom_way)
FROM ws_2po_4pgr;
```

Quelle: Eigene Darstellung

In pgRouting stehen für die topologische Überprüfung des Graphs weitere Funktionen zur Verfügung (siehe Kapitel 8.1.1). Für die spätere Nutzung durch pgRouting sind die in Tabelle 11 dargestellten Parameter und

⁴⁵ http://postgis.net/docs/manual-2.0/using_postgis_dbmanagement.html#OGC_Velocity (15.10.2013)

Wertebereiche erforderlich, welche der Verkehrsgraph für die weitere Berechnung beinhalten bzw. welchen er entsprechen muss.

Ergebnis

Wie Tabelle 11 zeigt, gibt es für jede relevante Spalte von pgRouting eine korrekte Entsprechung im erzeugten Vektorgraph. Die einzige Ausnahme ist das Attribut "rcost_len", die reverse Länge in Gegenfahrtrichtung. Von osm2po werden nur die Zeitkosten in die Gegenfahrtrichtung (reverse_cost bzw. rcost_time) ermittelt. Im Grunde sind damit alle zwingend erforderlichen Attribute mit korrekten Datentypen enthalten.

Tabelle 11: Konzeptuelle Anforderungen an einen routenfähigen Graph und Entsprechung des Exportdatensatzes von osm2po

Für pgRouting erforderliche Attribute			Äquivalente Attribute des Exportvektorgraphs von osm2po	
Attribut	Datentyp	Erläuterung	Attribut	Datentyp
id	integer, NOT NULL	Identifizier	id	integer, NOT NULL
source	integer	Identifizier des Quellknotens	source	integer
target	integer	Identifizier des Zielknotens	target	integer
cost_len	double precision	Länge der Kante	km	double precision
cost_time	double precision	Zeitkosten in Fahrtrichtung	cost	double precision
rcost_len	double precision	Länge der Kante in Gegenfahrtrichtung	nicht verfügbar	
rcost_time	double precision	Zeitkosten der Kante in Gegenfahrtrichtung	reverse_cost	double precision
x1	double precision	X-Koordinate des Quellknotens	x1	double precision
x2	double precision	Y-Koordinate des Quellknotens	x2	double precision
y1	double precision	X-Koordinate des Zielknotens	y1	double precision
y2	double precision	Y-Koordinate des Zielknotens	y2	double precision

Quelle: Takubo et al. (2013), SQL-Datensatz aus osm2po (Moeller, 2012)

Das Datenmodell ist in sich konsistent und entspricht dem für ein Datenrouting mit pgRouting gestellten Anforderungen (Tabelle 12).

Tabelle 12: Ergebnisse der Bewertung der konzeptionellen Konsistenz

Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis der Bewertung
Ist das Datenmodell in sich konsistent und entspricht es dem für ein Datenrouting erforderlichen Datenmodell?	Binär (ja / nein)	Ja	ja

Quelle: Eigene Darstellung

9.2.3 Konsistenz der Wertebereiche

Evaluierungsmethode

Überprüft wird die Eindeutigkeit von Identifikatoren, die vollständige Verfügbarkeit zwingend erforderlicher Attribute sowie die Einhaltung Wertebereiche gemäß der Begriffsdefinition.

Die Durchführung der Datenanalyse erfolgt indem

- das Attribut „id“ auf Doppelnennungen überprüft wird,
- die Anzahl der Kanten ermittelt wird, für die zwingend erforderliche Attribute nicht verfügbar d.h. leer sind,

- die Anzahl der Kanten, für die die Attributwerte zwingend erforderlicher und anderer wichtiger Attribute außerhalb der Wertebereiche liegen (durch Sortierung und Ermittlung der Minimal- und Maximalwerte).

Ergebnis

Das Attribut „id“ ist als Identifikator zwingend erforderlich und muss eindeutig identifizierbar sein, beide Anforderungen werden von den Attributwerten des Datensatzes erfüllt.

Tabelle 13: Konzeptuelle Anforderungen an einen routenfähigen Graph und Entsprechung des Exportdatensatzes von osm2po

Attribut	Datentyp	Erläuterung	Kanten ohne Informationen		Wertebereiche	
			Anzahl	Anteil in %	min	max
Für pgRouting zwingend erforderliche Attribute						
id	integer, NOT NULL	Identifizier	0	0	1	87.295
source	integer	Identifizier des Quellknotens	0	0	1	68.471
target	integer	Identifizier des Zielknotens	0	0	1	68.473
clazz	Integer	Code der Straßenklasse	0	0	11	42
geom_way	geometry linestring	Geometrie der Kante	0	0	.	.
cost_len (km)	double precision	Länge der Kante	0	0	0,00	14,56
cost_time (cost)	double precision	Zeitkosten in Fahrtrichtung	0	0	0,00	0,24
rcost_time (reverse_cost)	double precision	Zeitkosten der Kante in Gegenfahrtrichtung	0	0	0	1.000.000
Weitere für die Überprüfung des Verkehrsgraphs erforderliche Attribute						
osm_id	bigint	Identifizier der Kante gemäß OSM	0	0	199.732	2490022962
osm_source_id	bigint	Identifizier des Quellknoten in OSM	0	0	199.732	2490022960
osm_target_id	bigint	Identifizier des Zielknoten in OSM				
osm_name	character var.	Straßenname aus OSM-Datensatz	13208	15,1	0	0
osm_meta	character var.	Metadateninformation	87295	100,0	0	0
kmh	Integer	Geschwindigkeit auf der Kante	0	0	5	130
rcost_len	double precision	Länge der Kante in Gegenfahrtrichtung	nicht verfügbar			
x1	double precision	X-Koordinate des Quellknotens	0	0	15	17
x2	double precision	Y-Koordinate des Quellknotens	0	0	47	48
y1	double precision	X-Koordinate des Zielknotens	0	0	15	17
y2	double precision	Y-Koordinate des Zielknotens	0	0	47	48

Quelle: Eigene Darstellung

Die Attribute source und target sind Objektschlüssel der Straßenknoten sind ebenfalls zwingend erforderlich, genauso wie das Geometrieattribut der Kanten (geom_way) und die Straßenklasse (clazz). Mithilfe dieser Attribute können die übrigen für das Routing erforderlichen Attribute ermittelt werden. Für alle diese Attribute sind Attributwerte im Datensatz vollständig verfügbar.

Die Spalten mit den Identifikatoren der OSM (osm_id, osm_name, osm_meta, osm_source_id, osm_target_id), sind für das Routing direkt nicht relevant sind, sondern vielmehr für die Rückkopplung der Daten mit dem Originaldatensatz der OSM. Die Attribute osm_id, osm_source_id und osm_target_id sind vollständig, die Rückkopplung mit dem OSM-Originaldaten also möglich. Das Attribute osm_name ist zu 85% besetzt, während

dem Attribut `osm_meta` für keine der Kanten Werte zugewiesen sind, wodurch die Qualität hinsichtlich des Routings aber nicht beeinträchtigt wird.

Die Wertebereiche liegen innerhalb der erforderlichen Wertebereiche. Der Wert von 1.000.000 für „`rcost_time`“ weist auf eine Einbahn hin und die Kante in Gegenrichtung nicht befahren werden darf.

Tabelle 14 fasst die Ergebnisse in einer Übersicht zusammen, die Qualitätsanforderungen sind vollständig erfüllt.

Tabelle 14: Ergebnisse der Bewertung der Konsistenz der Wertebereiche

Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis der Bewertung
Sind alle für das Routing erforderlichen Daten im generierten Datensatz enthalten?	Binär (ja / nein)	Ja	Ja
Sind die Wertebereiche der Attribute konsistent mit diesem Datenmodell?	Binär (ja / nein)	Ja	Ja

Quelle: Eigene Darstellung

9.2.4 Topologische Konsistenz

Spezifikation der relevanten Datenqualitätsmerkmale, Maßeinheiten und Indikatoren

Für einen routenfähigen Verkehrsgraph ist es erforderlich, dass alle Straßen logisch miteinander verbunden sind, d.h. innerhalb des Straßennetzes, jeder Punkt von jedem Punkt innerhalb des Netzes erreicht werden kann (Forschungsgesellschaft Mobilität 2011, Rehl et al. 2012). Dies ist die wesentlichste Anforderung an einen routenfähigen Verkehrsgraph. Darüber hinaus muss bei Kreuzungen in der OSM richtig definiert sein, ob es sich um eine planfreie Kreuzung (mit Brücke, Unterführung etc.) oder eine Kreuzung mit einem physischen Knotenpunkt handelt, damit die Kreuzung und die möglichen Abbiegerelationen korrekt in die routenfähige Datenbank übernommen werden können (Möller, 2012).

Basierend auf diesen Anforderungen werden folgende vier Qualitätsparameter definiert:

- Qualitätsparameter T1: Isolierte Straßensegmente
Anteil der isolierten Straßensegmente bestehend aus einer oder mehreren Kanten, die keine Verbindung zum Hauptstraßengraph haben
- Qualitätsparameter T2: Lücken im Straßennetz
Anteil der Knoten, an welche nur eine Kante anschließt, die aber keine Sackgassen sind
- Qualitätsparameter T3: Fehlende Schnittpunkte kreuzender Kanten
Anteil der sich schneidenden Kanten, die keine planfreien Kreuzungen sind.
- Qualitätsparameter T4: Schleifen
Anteil der Kanten, deren Anfangs- und Endpunkte auf denselben Knoten verweisen.

Datengrundlagen

Die Bewertung der topologischen Konsistenz kann durch intrinsische Analyse erfolgen, erfordert also keinen externen Datensatz. Trotzdem wird für die Interpretation der ermittelten potenziellen Fehler der OSM-Straßengraph visuell mit öffentlich verfügbaren Referenzstraßengraphen verglichen (Basemap.at, Bing-Maps, Google-Maps).

Evaluierungsmethode 1: Grobanalyse

Zunächst erfolgt eine automatisierte Überprüfung des Straßengraphs mit Hilfe vordefinierter Algorithmen des pgRouting-Tools.

Algorithmen des PgRouting-Tools

PgRouting bietet für eine erste Analyse von Straßengraphen die Befehle "pgr_createTopology" und "pgr_analyzeGraph" (Code 9). Der erste Befehl erstellt die Topologie und eine Tabelle aller Quell- und Zielpunkte der Kanten. Der zweite Befehl durchsucht den Straßengraph auf Lücken und isolierte Einzelkanten und gibt darüber hinaus die Anzahl der im Graph gefundenen Sackgassen, Schleifen und planfreien Kreuzungen aus. In der Tabelle der Quell- und Zielknoten werden zudem jene Punkte attribuiert, die Endpunkte von Sackgassen sind sowie jene Punkte, welche potenzielle Lücken darstellen.

Code 9: SQL-Code zur Erstellung der Topologie und Grobanalyse des erstellten Graphs

PgRouting Funktion zur Erstellung der Topologie

```
Varchar pgr_createTopology(text edge_table, double precision tolerance, text
the_geom:='the_geom', text id:='id', text source:='source', text target:='target', text
rows_where:='true')
```

Erstellung der Topologie für den osm2po-Datensatz ws_2po_4pgr mit dem geometrischen Attribut geom_way'

```
SELECT pgr_createTopology('ws_2po_4pgr',0.00001,'geom_way');
```

PgRouting Funktion zur Prüfung der Topologie

```
Varchar pgr_analyzeGraph(text edge_table, double precision tolerance, text
the_geom:='the_geom', text id:='id', text source:='source', text target:='target', text
rows_where:='true')
```

Prüfung der Topologie für den osm2po-Datensatz ws_2po_4pgr mit dem geometrischen Attribut geom_way'

```
SELECT pgr_analyzeGraph('ws_2po_4pgr',0.00001,'geom_way');
```

Quelle: pgRouting Manual 2.0.0, eigene Berechnung.

Die in Tabelle 15 dargestellte Ausgabe ist Ausgangspunkt für die weitere Analyse. Bei einer Toleranz von 0.00001 Grad ("snapping distance") wurden insgesamt fünf potenzielle Lücken gefunden, d.h. Knoten die nicht ausreichend topologisch an den Straßengraph angebunden sind. Darüber hinaus gibt es 91 isolierte Kanten-segmente, das heißt Kanten deren Quell- und Zielknoten Sackgassen sind, die also keine Verbindung zum übrigen Straßengraph haben. Für die Fehleranalyse sind jedoch auch isolierte Straßenzüge relevant, die aus mehreren Kantensegmenten bestehen. Diese werden in der Folge ermittelt.

Ergebnisse der Grobanalyse

Insgesamt wurden 10.936 Sackgassen, 1.314 planfreie Kreuzungen und 120 Schleifen gefunden. Die potenziellen Lücken und die isolierten Segmente deuten eindeutig auf topologische Fehler hin. Die Zahl der Sackgassen erstaunt hingegen nicht, da es im Bearbeitungsgebiet viele Siedlungsgebiete mit niederrangigen Anrainerstraßen gibt welche häufig Sackgassen aufweisen. Aus gleichem Grund muss die Zahl der Schleifen nicht auf Fehler in der Topologie hinweisen. Die hohe Zahl an planfreien Kreuzungen kann durch die hohe Zahl an Autobahnen und Eisenbahnverbindungen erklärt werden.

Tabelle 15: Gefundene potenzielle topologische Fehler (Toleranzgrenze 0,00001 Grad WGS84)

Potenzieller topologischer Fehler	Anzahl der Fälle
Isolierte Segmente bestehend aus nur einer Kante (isolated segments)	91
Potenzielle Lücken in der Nähe von Sackgassen (Potential gaps found near dead ends)	5
Sackgassen (dead ends)	10.936
Planfreie Kreuzungen (Intersections detected)	1.314
Schleifen (Ring geometries)	120

Quelle: eigene Berechnung mit pgRouting

Evaluierungsmethode 2: Detailanalyse

Im zweiten Schritt werden die automatisch ermittelten Ergebnisse im Rahmen einer Detailanalyse durch eigene Berechnungsverfahren überprüft, um ein besseres Bild von der Qualität des Datensatzes zu erhalten.

Straßenzüge aus einem oder mehreren Segmenten

Für die Qualität des Straßengraphs sind alle isolierten Straßenzüge aus einem oder mehreren Segmenten relevant, die keine Verbindung zum Straßengraph haben. Mehrkantige isolierte Subnetzwerke werden von der Formel "pgr_analyzeGraph" allerdings nicht erkannt. Daher wurde in einem weiteren Schritt eine Distanzmatrix von einem zentralen Sackgassenknoten am Wiener Stephansplatz zu allen Knoten des Straßengraphs berechnet. Dazu wurde die pgRouting-Funktion „pgr_kdijkstra“ verwendet. Die Ergebnistabelle weist für alle Knoten des Netzes die Wegekosten ausgehend vom Startknoten aus, Knoten ohne Verbindung zum Startknoten werden mit "-1" attribuiert. Um die kartographische Darstellung der Daten zu ermöglichen, wird in der Ergebnistabelle die Geometrie der Punkte hinzugefügt (Code 10)⁴⁶. Darüber hinaus können über die eindeutige ID des Zielknotens (Attribut target), die in der Distanzfunktion ermittelten Kosten auf die Kanten übertragen werden. Insgesamt werden anstelle der 91 isolierten Segmente 403 isolierte Kanten mit 536 isolierten Knoten identifiziert. Das entspricht bei 68.471 Knoten und 87.295 Kanten insgesamt einer sehrn geringe Fehlerquote von 0,78% der Knoten bzw. 0,46% aller Kanten.

Wie die Karte in Abbildung 17 zeigt liegt rund die Hälfte der isolierten Straßensegmente am Rande des Bearbeitungsgebiets. Innerhalb eines Puffers von 1 km um die Grenze des Bearbeitungsgebiets liegen 180 von 403 isolierten Kanten, welche 42 km von insgesamt 90 km isolierter Kantenlänge entsprechen. Diese Fehler sind in erster Linie auf das Ausschneiden des Straßengraphs aus dem Gesamtnetz für Österreich zurück zu führen. Trotz des Osmosis-Befehls complete-ways beim Beschneiden der OSM auf das Bearbeitungsgebiet sind offensichtlich einige rudimentäre Straßensegmente importiert worden.

Code 10: SQL Code zur Erstellung einer Distanzmatrix vom Knoten 18466 zu allen 68471 Knoten des Straßengraphs

Berechnung der kürzesten Wege und Anfügung des Ergebnisses an die den Kantenstraßengraph ws_2po_4pgr

```
CREATE TABLE kdijkstraCost_all AS
  SELECT seq, id1 as source, id2 as target, cost
  FROM pgr_kdijkstraCost(
    'SELECT id, source, target, cost FROM ws_2po_4pgr',
    18466, array[1,2,...68471],false, false);

ALTER TABLE kdijkstraCost_all ADD column geom geometry(Point,4326);

UPDATE kdijkstraCost_all k
  SET geom=ws.the_geom
  FROM ws_2po_4pgr_vertices_pgr ws --ws_2po_4pgr_vertices_pgr als Liste aller Knoten des Verkehrsgraphs
  WHERE k.target=ws.id;

ALTER TABLE ws_2po_4pgr ADD COLUMN kdij_cost double precision;

UPDATE ws_2po_4pgr SET kdij_cost=(
  SELECT cost FROM kdijkstraCost_all k
  WHERE ws_2po_4pgr.target=k.target)
  WHERE EXISTS (
    SELECT cost FROM kdijkstraCost_all k
    WHERE ws_2po_4pgr.target=k.target);
```

Quelle: pgRouting Manual 2.0.0, eigene Berechnung.

Potenzielle Lücken im Straßennetz

Die fünf potenziellen Lücken im Straßennetz können aufgrund der geringen Anzahl visuell mit einem Referenzdatensatz (Orthophotos von BingMaps und GoogleMaps) geprüft werden. Jeder der identifizierten Lücken ist auf einen topologischen Fehler zurückzuführen. Ein Fehler beruht auf einem "Overshot" während der

⁴⁶ Der zugrunde liegende PgRouting Algorithmus wird in Kapitel 6.5 pgRouting, Code 6 erläutert.

Digitalisierung: die betroffene Anrainerstraße ist nicht an die Zufahrtsstraße angebunden, weil der falsche Quellknoten keinen Kreuzungspunkt mit dieser Zufahrtsstraße hat. Die übrigen vier Knoten sind topologisch falsch gesetzt. Es handelt sich dabei nicht um Quell- oder Zielknoten, sondern um "Konstruktionspunkte" im Straßenverlauf bzw. "vergessene" Konstruktionspunkte. Die von ihnen ausgehenden Kanten sind stets Teilstücke von Kanten, welche bereits durch ein anderes dargestellt werden. Bei einer Korrektur könnten „überzähligen“ Punkte und die zu ihnen verlaufende Kante gelöscht werden.

Untersuchung der planfreien Kreuzungen

Die Zahl der planfreien Kreuzungen wurde neben der automatischen Auswertung nochmals mit dem PostGIS-Befehlen "ST_Intersection" und "ST_Crosses" überprüft, um jene Fälle zu identifizieren in welchen sich zwei unterschiedliche Kanten ohne Knotenpunkt kreuzen (Code 11). Die Berechnung ermittelt 1.327 planfreie Kreuzungen, d.h. 13 Kreuzungen mehr als die automatische Analyse der pgRouting-Funktion „pgr_analyzeGraph“. Wie erwartet befinden sich die planfreien Kreuzungen zum überwiegenden Teil (87,8%) im Bereich von Autobahnen und Schnellstraßen in Wien und im Wiener Umland (siehe auch Abbildung 17). Die übrigen 162 planfreien Kreuzungen zur korrekten Darstellung von Über- und Unterführungen, aber auch für das Auflösen komplizierter Kreuzungsrelationen benutzt.

Innerhalb der 162 planfreien Knoten abseits der Autobahnen- und Schnellstraßen konnten bei einer visuellen Kontrolle mit einem Referenzdatensatz (Orthophotos von BingMaps und GoogleMaps) nur 4 tatsächliche topologische Fehler gefunden werden. Geht man davon aus, dass die planfreien Kreuzungen auf Autobahnen und Schnellstraßen alle korrekt sind, beträgt die Fehlerquote 0,4%.

Code 11: SQL-Code zur Ermittlung der planfreien Kreuzungen

```
CREATE TABLE intersections as
SELECT
    ST_Intersection(a.geom_way, b.geom_way)::geometry(point,4326)AS geom,
    Count(Distinct a.id)
FROM
    ws_2po_4pgr as a,
    ws_2po_4pgr as b
WHERE
    ST_Crosses(a.geom_way, b.geom_way)
    AND a.id != b.id
GROUP BY
    ST_Intersection(a.geom_way, b.geom_way);
ALTER TABLE intersections ADD COLUMN id serial PRIMARY KEY;
```

Quelle: Adaptiert nach <http://gis.stackexchange.com/questions/20835/identifying-road-intersections-using-postgis>.

Untersuchung der Schleifen

Um die Schleifen im Straßengraph darstellen zu können, wurde eine eigene Tabelle mit allen Kanten berechnet, welche den gleichen Anfangs- und Endknoten haben (Code 12). Im Gegensatz zum pgRouting Algorithmus "pgr_analyzeGraph" mit 120 identifizierten Schleifen werden 130 Schleifen ermittelt. Alle Schleifen wurden visuell mit Hilfe eines Referenzdatensatzes (Orthophotos von BingMaps und GoogleMaps) überprüft. Innerhalb der Schleifen konnten keine topologischen Fehler gefunden werden.

Code 12: SQL-Code zur Extraktion der Schleifen des Straßengraph

```
create TABLE schleifen as
select id, source, target, km, geom_way FROM ws_2po_4pgr WHERE source=target;
```

Quelle: pgRouting Manual 2.0.0, eigene Berechnung.

Ergebnisse der Detailanalyse

In Tabelle 16 werden die Ergebnisse der automatischen Fehlerberechnung mit "pgr_analyze_graph" mit der vorgenommenen Detailanalyse verglichen. Definitiv topologisch falsch sind alle gefundenen Lücken und

isolierten Segmente. Innerhalb der planfreien Kreuzungen konnten nur vier Digitalisierungsfehler identifiziert werden, bei allen gefundenen Schleifen handelt es sich um korrekt dargestellte Straßensegmente.

Basierend auf den durchgeführten Prüfungen ist die topologische Qualität des Straßengraphs bei einem Fehleranteil von unter 1% sehr hoch. In Bezug auf die Anforderung eines routenfähigen Verkehrsnetzes, alle Straßen logisch miteinander zu verbinden und Inselnetzwerke auszuschließen, besteht der Verkehrsgraph mit einer Fehlerquote von 0,59% den Qualitätstest.

Tabelle 16: Potenzielle und gefundene topologische Fehler

Art des potenziellen topologischen Fehlers	Einheit	Pre-Test*	Detailanalyse		Anteil fehlerhafter Objekte an allen Objekten in %**
		Anzahl der Fälle	Anzahl der Fälle	davon Fehler in %	
Isolierte Segmente bestehend aus nur einer Kante (isolated segments)	Kanten	91	91	100%	0,13%
Isolierte Segmente aus einer oder mehreren Kanten	Kanten	n.a.	403	100%	0,59%
Potenzielle Lücken in der Nähe von Sackgassen (Potential gaps found near dead ends)	Knoten	5	5	100%	0,01%
Sackgassen (dead ends)	Knoten	10.936	10.936	0%	0,00%
Planfreie Kreuzungen (Intersections detected)	Knoten	1.314	1.327	0,4%	0,00%
Schleifen (Ring geometries)	Kanten	120	130	0%	0,00%

* Pretest mit der pgRouting-Funktion "pgr_analyze_Graph"

** Anteil der gefundenen fehlerhaften Objekte an den 68471 Knoten bzw. 87295 Kanten des Straßengraphs

Quelle: Auswertung mit pgRouting und eigene Berechnungen.

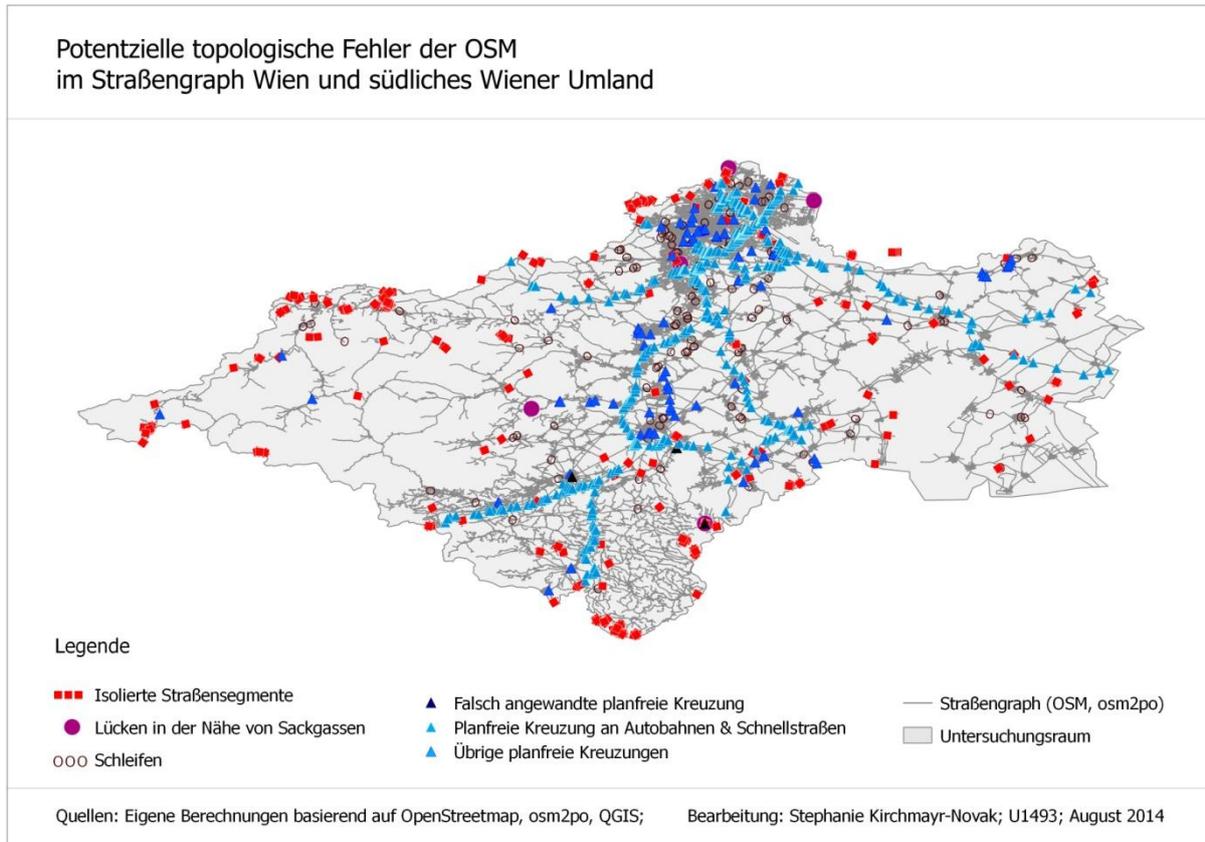
Die Karte in Abbildung 17 zeigt die automatisch gefunden potenziellen und die analysierten Fehler. Die vom Netzwerk isolierten Knoten sind zum Großteil an den Grenzen des Bearbeitungsgebiets zu finden und dürften daher weniger auf Fehler in der OSM, als auf die Beschneidung des Straßengraphs auf das Bearbeitungsgebiet zurückzuführen sein. Darüber hinaus gibt es isolierte Straßensegmente vor allem in peripheren Gebieten, die vermutlich nicht regelmäßig von der OSM-Community gewartet werden. Tabelle 17 fasst die Ergebnisse der Bewertung der topologischen Konsistenz zusammen. Der Datensatz entspricht den Qualitätsanforderungen.

Tabelle 17: Ergebnisse der Bewertung der topologischen Konsistenz

Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis der Bewertung
Isolierte Einzelsegmente (Einzelkanten) ohne Verbindung zum "Hauptstraßengraph"	Isolierte Straßensegmente	Anteil fehlerhafter Objekte < 5% von allen Objekten	0,13%
Isolierte Segmente aus mehr als einer Kante ohne Verbindung zum Hauptstraßengraph			0,46%
Knoten, die Lücken im Netz darstellen. Diese Knoten haben, wie Sackgassen, nur zu einer Kante eine Verbindung, sind aber semantisch keine Sackgassen, sondern unvollständig digitalisierte Straßenzüge.	Lücken im Straßennetz		0,01%
Inkorrekte planfreie Kreuzungen die weder Brücken oder Unterführungen noch für die Auflösung komplizierter Fahrtrelationen herangezogen werden.	Fehlende Schnittpunkte kreuzender Kanten		0,00%
Kanten, deren Quellknoten gleichzeitig auch Zielknoten ist, aber keine Schleifen darstellen.	Fehlerhafte Schleifen		0,00%

Quelle: Eigene Darstellung

Abbildung 17: Potenzielle topologische Fehler der OSM im Straßengraph Wien und südliches Wiener Umland



Quelle: Eigene Auswertung basierend auf Openstreetmap, osm2po, pgRouting, QGIS

9.3 Vollständigkeit

Der Definition aus Kapitel 7.3.2 Vollständigkeit folgend, wird zwischen der Vollständigkeit der Objekte und der Vollständigkeit der Attribute unterschieden.

9.3.1 Vollständigkeit der Objekte - Untersuchung möglicher Untererfassung

Spezifikation der relevanten Datenqualitätsmerkmale, Maßeinheiten und Indikatoren

Da für die Evaluierung der Vollständigkeit neben der OpenStreetMap kein Referenzvektorgraph des Untersuchungsgebiets für die extrinsische Evaluierung zur Verfügung steht, wird unter Berücksichtigung der Bedürfnisse der Einzugsbereichsberechnung ein anderer Ansatz zur Qualitätsbewertung herangezogen. Dabei ist darauf hinzuweisen, dass die Vollständigkeit der OSM Straßendatenbank ohne einen Referenzdatensatz nicht exakt bestimmt werden kann (Barron, Neis, Zipf, 2013).

Im Zusammenhang mit der intrinsischen Qualitätsbewertung stellen Rehr et. al. (2012) fest, dass bei der Berechnung von Einzugsbereichen die Vollständigkeit des Verkehrsgraphs über den Grad der Erschließung von Siedlungsgebieten definieren lässt. Nur sofern jedes der Siedlungsgebiete an den Verkehrsgraph angebunden ist, kann eine Erreichbarkeitsanalyse durchgeführt werden. Daraus werden zwei Qualitätsparameter wie folgt definiert:

Qualitätsparameter V1: Erschließungsquote
Anteil der Siedlungsgebiete (Siedlungseinheiten), der durch Straßen in einem
Umkreis von 50, 100, 150, 200 und 400 Metern erschlossen wird

Qualitätsparameter V2: Erschließungsgrad der Siedlungsgebiete (Siedlungseinheiten)
Anteil der Siedlungsfläche, der durch Straßen in einem Umkreis von 50, 100, 150, 200
und 400 Metern erschlossen wird

Datengrundlagen

Als Geodatenquelle für die Siedlungsgebiete in Österreich steht ein frei verfügbarer Datensatz von Statistik Austria zu Verfügung, der die Siedlungseinheiten Österreichs zum Gebietsstand 2010 enthält.⁴⁷

Für das Untersuchungsgebiet weist der Geodatensatz 194 Siedlungseinheiten mit einer Gesamtfläche von 724 km² aus, dies entspricht rund 10% der Gesamtgemeindefläche im Untersuchungsgebiet (Tabelle 18). Der Großteil der übrigen 90% der Fläche entfällt auf forst- oder landwirtschaftlich genutzte Flächen, Verkehrsflächen und ähnliche. Von den 194 Siedlungseinheiten werden 2,35 Mio. Einwohner erfasst, dies entspricht 97,4 % der Gemeindebevölkerung im Untersuchungsgebiet. Aufgrund dieses hohen Einwohnererfassungsgrades, sind die Siedlungseinheiten als Berechnungsparameter für die Vollständigkeitsberechnung geeignet.

⁴⁷ Statistik Austria (2010): Siedlungseinheiten Österreichs 1: 2.000.000, Gebietsstand 2010. Verfügbar unter http://www.statistik.at/web_de/klassifikationen/regionale_gliederungen/siedlungseinheiten/index.html (Zugriff 18.08.2014)

In Anlehnung an die Statistische Kommission der Vereinten Nationen werden Siedlungseinheiten als kleinste vergleichbare Einheit als "zusammenhängend verbautes Gebiet von Wohnhäusern, industriellen, gewerblichen, sonstigen wirtschaftlichen und kulturellen Einrichtungen" ermittelt. Die Definition von Statistik Austria grenzt Siedlungseinheiten als zusammenhängend verbaute Gebiete mit mindestens 500 Einwohnern unabhängig von administrativen Gliederungen ab. Insgesamt weist Statistik Austria für das Jahr 2010 1.629 Siedlungseinheiten aus. Diese Siedlungseinheiten enthalten 1.653.456 Gebäude, dies sind 71,1% aller Gebäude in Österreich. Die Summe aller Einwohner in den definierten Siedlungsgebieten beläuft sich auf 6.682.076, das entspricht 79,8% der Gesamtbevölkerung Österreichs.

Tabelle 18: Fläche und Bevölkerung der 271 Gemeinden und 194 Siedlungseinheiten (SE) im Untersuchungsgebiet

Raum	Fläche in km ²			Bevölkerung 2010		
	Gemeinden*	Siedlungseinheiten	Anteil der SE in %	Gemeinden	Siedlungseinheiten	Anteil der SE in %
Untersuchungsgebiet insgesamt	7.067	724	10,3	2.408.000	2.346.000	97,4
davon auf Wiener Stadtgebiet	416	313	75,4	1.690.000	1.690.000	100,0
davon im Wiener Umland	6.652	411	6,2	718.000	656.000	91,4

* Berechnung auf Basis frei verfügbarer, aber generalisierter Geodaten von Statistik Austria und OSM (offiziell beträgt die Gemeindefläche für Wien 414,65 km²; ⁴⁸ der generalisiert Datensatz gibt 415,72 km² aus).

Anmerkung: Der Datensatz der Siedlungseinheiten wurde bzgl. der Siedlungseinheit Wien um jenen Niederösterreichischen Gebietsteil bereinigt, der im Norden über das Stadtgebiet Wiens und damit über das Untersuchungsgebiet herausragt (47 km² mit 67.600 Einwohnern).

Evaluierungsmethode und Ergebnisse

Die Erschließungsquote und der Erschließungsgrad werden für vier Varianten ermittelt, indem vier Puffer (50 m, 100 m, 200 m und 400 m) um die Straßenkanten gelegt werden. Die Ergebnisse sind in Tabelle 19 zusammengefasst.

Die Erschließungsquote entspricht bereits im Umkreis von 50 m um den Straßengraph 100 %, d.h. jede der 194 Siedlungseinheiten ist durch den OSM-Straßengraph erschlossen.

Der Erschließungsgrad, d.h. der Anteil der Fläche der erschlossenen Siedlungsgebiete, beträgt bei einem 50 m-Puffer bereits 68 %, wobei 98 % der Siedlungsflächen zumindest zur Hälfte erschlossen sind. Wird der Puffer auf 400 m erhöht, steigt der Anteil der erschlossenen Siedlungsfläche auf bis zu 99,5 % an.

Tabelle 19: Erschlossene Siedlungseinheiten, Fläche

Puffer um Straßenkanten	Erschlossene Siedlungseinheiten		Erschlossene Siedlungsfläche		Siedlungseinheiten mit Erschließungsgrad von 50% oder mehr	
	Anzahl	Erschließungsquote in %	in ha	Erschließungsgrad in %	Anzahl	in %
50-m Puffer	194	100,0	49.428	68,3	190	97,9
100-m Puffer	194	100,0	63.910	88,4	193	99,5
200-m Puffer	194	100,0	70.446	97,4	193	99,5
400-m Puffer	194	100,0	71.958	99,5	194	100,0
Untersuchungsgebiet	194	100,0	72.332	100,0	194	100,0

Quelle: Eigene Berechnungen auf Basis Statistik Austria und OSM

Es ist zu beachten, dass in dieser Auswertung nicht nur die Qualität des OSM-Graphs eine Rolle spielt, sondern auch die Größe der Baublöcke bzw. der Grundstücke relevant ist.

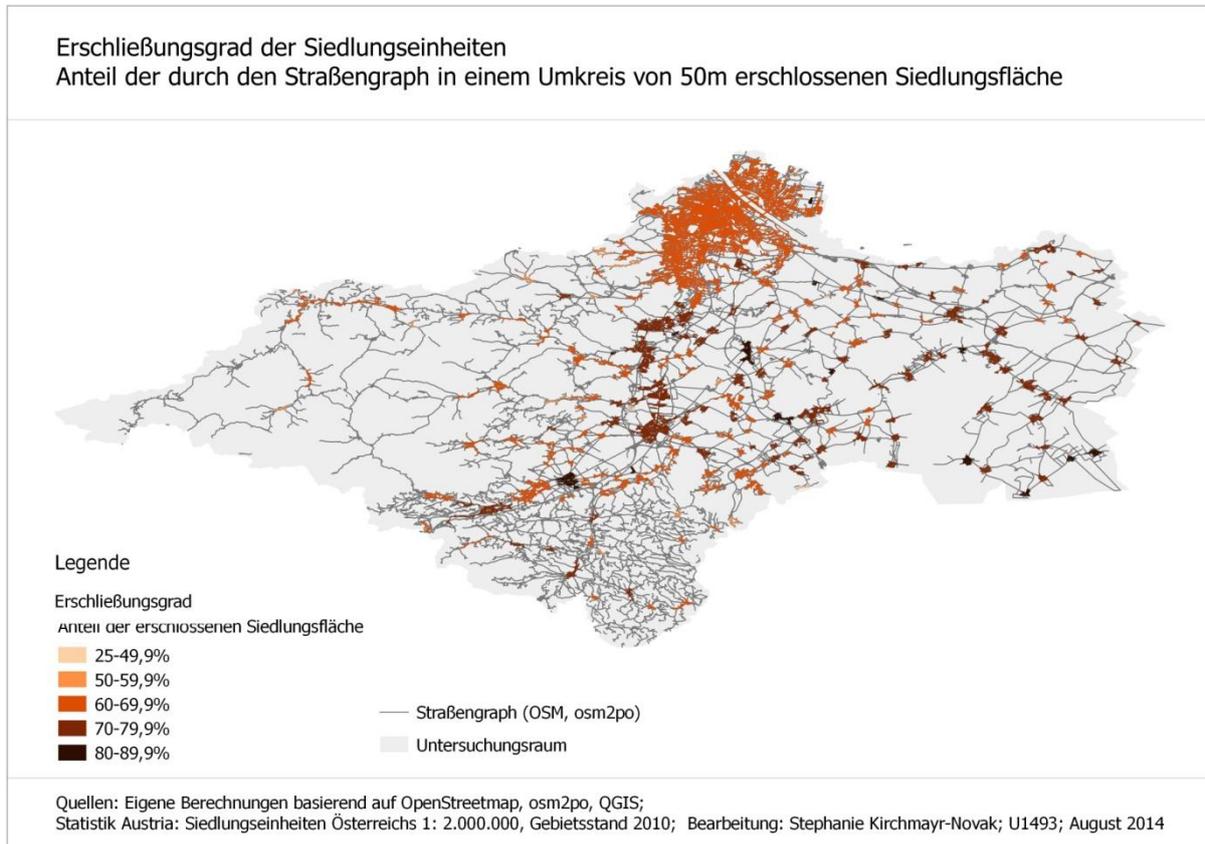
Würde der Verkehrsgraph bereits bei einem 50 m-Puffer rund um die Straßenkanten die Siedlungsflächen zu 100 % erfassen, dürfte die Breite der Baublöcke – also der Abstand zwischen zwei Straßenkanten – maximal 100 m betragen. Eine entsprechende Erschließung würde in vielen, insbesondere ländlichen, Gemeinden nicht der Vorstellung der Bewohner von Lebensqualität entsprechen. Abbildung 18 verdeutlicht die Unterschiede zwischen den Siedlungseinheiten mit geringer Erschließung im peripheren Raum im Westen und Süden gegen-

⁴⁸ Statistik Austria (2014): Übersicht der Bundesländer, verfügbar unter http://www.statistik.at/web_de/klassifikationen/regionale_gliederungen/bundeslaender/ (18.08.2014)

über dem gut erschlossenen städtischen Siedlungsgebiet entlang der Süd-Autobahn (Baden, Mödling, Wiener Neustadt), aber auch der Straßendörfer im Seewinkel.

Es muss auch darauf hingewiesen werden, dass der hier verfolgte Evaluierungsansatz der Feststellung einer möglichen Untererfassung im Straßengraph dient. Er ermöglicht aber keine Aussage darüber, ob der Straßengraph Übererfassungen enthält, also Objekte, die für die Berechnung der Erreichbarkeit nicht erforderlich bzw. nicht Teil des spezifizierten Datenmodells sind. Diese Fragestellung wird im nächsten Kapitel behandelt.

Abbildung 18: Anteil der durch den Straßengraph in einem Umkreis von 50 m erschlossenen Siedlungsfläche



Quelle: Eigene Berechnung basierend auf OSM, osm2po, QGIS; Statistik Austria: Siedlungseinheiten Österreichs 1:2.000.000, Gebietsstand 2010

Tabelle 20 fasst die Ergebnisse der Bewertung der topologischen Konsistenz zusammen. Der Datensatz entspricht den Qualitätsanforderungen.

Tabelle 20: Ergebnisse der Bewertung der Vollständigkeit des Netzgraphs

Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis der Bewertung
Erschließungsquote und Erschließungsgrad der Siedlungsgebiete Anteil der Siedlungsgebiete und der Siedlungsfläche, die durch Straßeninfrastruktur erschlossen wird. Die Siedlungsfläche wird dem frei verfügbaren Geodatenatz der Siedlungseinheiten, Gebietsstand 2010 von Statistik Austria entnommen. Der Erschließungsgrad wird in vier Varianten gerechnet, indem ein Puffer von 50m - 400m um die Straßenkanten gelegt wird.	Erschließungsquote: Anteil der Siedlungsgebiete (Siedlungseinheiten) in %, der durch Straßen in einem Umkreis von 50, 100, 150, 200 und 400 m erschlossen wird.	100 %	100 %
	Erschließungsgrad: Anteil der Siedlungsfläche in %, der durch Straßen in einem Umkreis von 50, 100, 150, 200 und 400 m erschlossen wird.	95 %	97,9 – 100 % (50 m – 400 m Puffer)

Quelle: Eigene Darstellung

9.3.2 Vollständigkeit der Attribute

Spezifikation der relevanten Datenqualitätsmerkmale, Maßeinheiten und Indikatoren

Wesentlichen Einfluss auf die Performanz der Routingberechnungen hat die Anzahl der Kanten im Straßengraph. Enthält der Straßengraph Straßenkanten die für die Berechnung der Erreichbarkeit im Sinne des Datenmodells nicht erforderlich, spricht man von Übererfassung. Im Zusammenhang mit der Erreichbarkeitsanalyse treten Übererfassungen auf wenn im Straßengraph Straßenzüge enthalten sind, die nicht oder nur sehr eingeschränkt mit dem Pkw befahren werden dürfen (z. B. Radwege, private Forststraßen, Straßen ohne verkehrliche Erschließungsbedeutung wie etwa Güterwege).

Ohne Referenzdatensatz können im Bereich der Objekteerfassung keine Aussagen zu einer möglichen Übererfassung gemacht werden. Es ist jedoch möglich, über die Vollständigkeit der Attribute Aussagen über eine möglich Übererfassung der Objekte zu machen. Straßenzüge, für welche die Angabe der Straßenkategorie fehlt oder unbekannt ist, können Teile des nicht bzw. eingeschränkt befahrbaren Straßennetzes darstellen.

Hinsichtlich der Vollständigkeit der Attribute werden folgende Qualitätsparameter definiert:

Definition der Qualitätsparameter

- Qualitätsparameter A1: Fehlende oder unbekannte Attribuierung der Straßenkategorie
Länge der Straßenkanten und Anteil im Straßengraph, für die keine spezifischen Angaben zur Straßenklasse verfügbar sind (Straßen unbekannter Kategorie: Merkmal "clazz" = 41 "road")
- Qualitätsparameter A2: Fehlende Attribuierung der Straßennamen
Länge der Straßenkanten und Anteil im Straßengraph, für die kein Straßename angegeben ist.

Datengrundlagen

Die Bewertung der topologischen Konsistenz kann durch intrinsische Analyse erfolgen, erfordert also keinen externen Datensatz. Trotzdem wird für die Interpretation der ermittelten potenziellen Fehler der OSM-Straßengraph visuell mit öffentlich verfügbaren Referenzstraßengraphen (Basemap.at, Orthophotos von Bing-Maps und Google-Maps) verglichen.

Evaluierungsmethode und Ergebnisse

Fehlende Attribuierung der Straßenkategorie

Zunächst wird für den Straßengraph die Straßenlängen nach Straßenkategorien ausgewertet.

Allen Kanten im Untersuchungsgebiet ist im Verkehrsgraph eine Straßenkategorie zugewiesen. Insgesamt entfallen im Untersuchungsgebiet 17% der Straßen auf das hochrangige Straßennetz, 12% auf das sekundäre Straßennetz (Landesstraßen) und 71% auf Gemeindestraßen (Tabelle 21). Die Ergebnisse verdeutlichen den hohen Anteil an Gemeindestraßen (Neben- und Erschließungsstraßen), der in Österreich bei durchschnittlich 69%⁴⁹ liegt. Der Anteil an Autobahnen und Schnellstraßen ist im Bearbeitungsgebiet mit 5,7% fast dreimal höher als der Österreichische Durchschnittswert von 2%⁴⁹. Dieser Umstand ist der hohen Autobahnen- und Schnellstraßendichte rund um Wien geschuldet.

Der Anteil der Straßen ohne Straßenkategorie ist mit 5 km oder 0,04 % äußerst gering. Wie die Karte in Abbildung 20 zeigt, liegen diese Straßenabschnitte in peripheren Gebieten innerhalb oder außerhalb von Siedlungsgebieten. Insgesamt handelt es sich um sehr kurze Straßenabschnitte, bei welchen entweder vergessen wurde die Straßenkategorie auszuweisen oder deren Straßenkategorie zum Zeitpunkt der Erhebung noch unklar war.

⁴⁹ BMVIT (2011): Verkehr in Zahlen – Ausgabe 2011. Kapitel 3 Infrastruktur, S. 37.

Tabelle 21: Straßenkategorien im OSM-Straßengraph: Straßenlänge und Anteil am Straßengraph

Clazz	OSM Straßentyp	Straßenkategorie	km	Anteil an allen Straßen in %
11	highway.motorway	Autobahn	646	4,8
12	highway.motorway_link	Autobahnauffahrt	249	1,9
13	highway.trunk	Schnellstraße	86	0,6
14	highway.trunk_link	Schnellstraßenauffahrt	20	0,1
15	highway.primary	Haupt-/Landesstraßen B	1.275	9,5
16	highway.primary_link	Auffahrt Haupt-/Landesstraßen B	21	0,2
21	highway.secondary	Landesstraßen L	1.554	11,6
22	highway.secondary_link	Auffahrt Landesstraßen L	2	0,0
31	highway.tertiary	Nebenstraßen	1.082	8,1
32	highway.residential	Erschließungsstraßen von Siedlungen	6.677	49,8
42	highway.unclassified	Straßen einfachsten Ausbaustandes	1.798	13,4
41	highway.road	Straßen unbekannter Kategorie	5	0,04
Straßenkategorien nach Gruppen				
11-15	Hochrangiges Straßennetz (Autobahnen, Schnellstraßen und Bundesstraßen)		2.297	17,1
21-22	Sekundäres Straßennetz (Landesstraßen)		1.556	11,6
31-32, 42	Gemeindestraßen (Nebenstraßen, Erschließungsstraßen, übrige Straßen)		9.557	71,2
41	Straßen unbekannter Kategorie		5	0,0
Insgesamt			13.415	100,0

Quelle: Eigene Berechnungen basierend auf Straßengraph der OSM und osm2po

Der Unterschied zwischen Straßen unbekannter Kategorie und Straßen mit einfachstem Ausbaustand wird in Abbildung 19 deutlich. Es handelt sich dabei um längere Straßenabschnitte, zum Teil ehemalige Güterwege, die zwar asphaltiert aber von sehr geringer Verkehrsbedeutung sind. Obwohl es sich bei dieser Straßenkategorie nicht um eine fehlerhafte Erhebung der OSM handelt, ist fraglich ob sie für die Berechnung der Erreichbarkeiten von Relevanz ist. Grundsätzlich ist davon auszugehen, dass sich der Pendlerverkehr auf das gut ausgebaute Straßennetz konzentriert und diese Straßen – wenn überhaupt – nur punktuell als Ausweichstrecken verwendet werden. Aufgrund ihrer marginalen Verkehrsbedeutung ist es daher vorzuziehen, sie aus dem Datenmodell bzw. dem Datensatz zu entfernen und so die Performanz des Straßengraphs zu erhöhen.

Abbildung 19: Gegenüberstellung von Straßen der Straßen einfachsten Ausbaustandes ("unclassified", strichliert) und Straßen unbekannter Kategorie ("road" in grün) und deren Lage im Siedlungsgebiet (flächig, rot)

Beispiel 1: Straße unbekannter Kategorie in Antau



Beispiel 2: Straße unbekannter Kategorie in Hainfeld



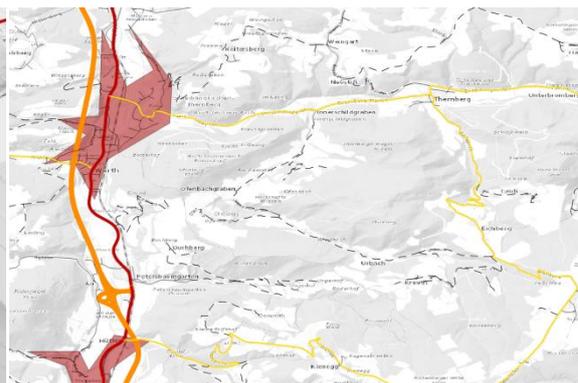
Beispiel 3: Straße einfachsten Ausbaustandes durch

Beispiel 4: Nebenstraßen geringer verkehrlicher

Weinberge bei Purbach

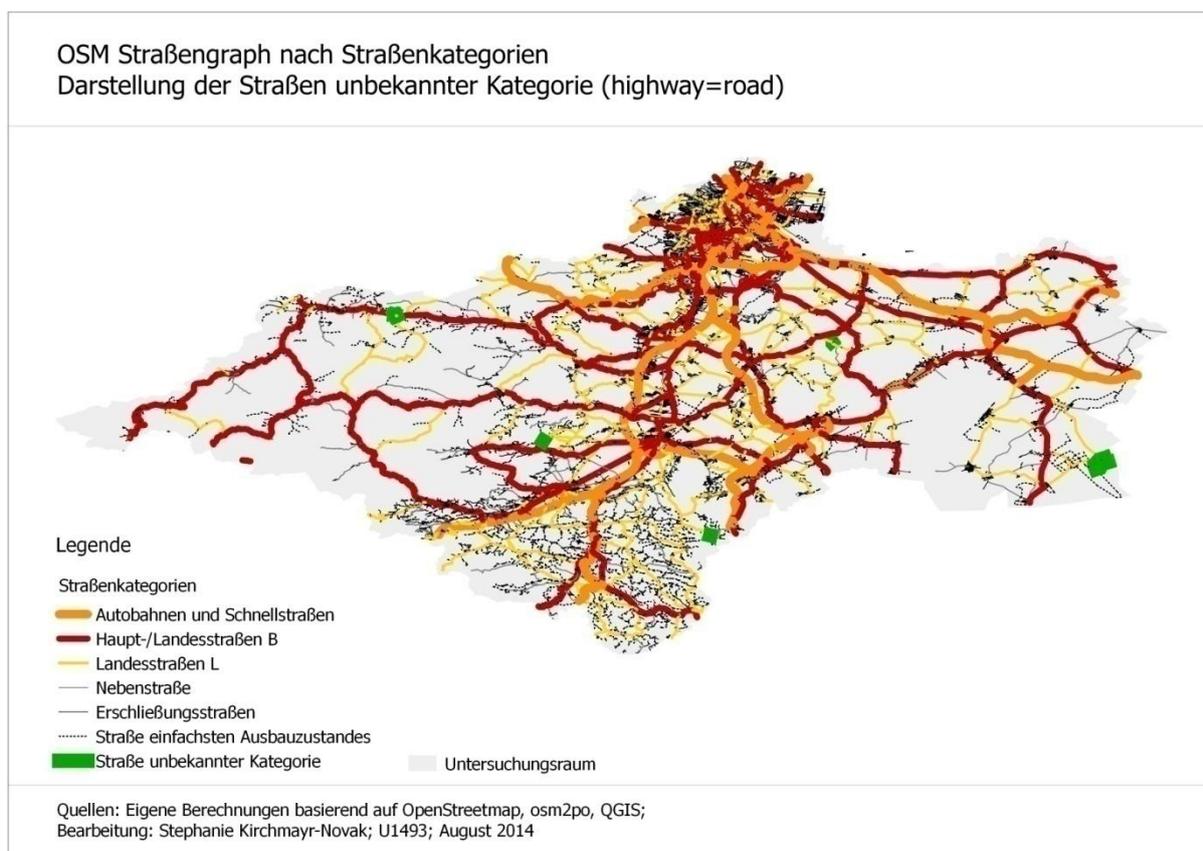


Bedeutung bei Grimmenstein



Quelle: Straßengraph nach OSM und osm2po, Geoland Basemap.at

Abbildung 20: OSM Straßengraph nach Straßenkategorien



Quelle: Eigene Berechnung basierend auf OSM, osm2po, QGIS

Fehlende Attribuierung der Straßennamen

Für den Straßengraph wird die Straßenlänge aller Straßen ausgewertet, die über keinen Straßennamen verfügen.

Insgesamt beträgt die Länge der Straßenkanten für die kein Straßennamen angegeben ist 2.500 km, das sind rund 19 % des Straßengraphs. Rund die Hälfte davon entfällt auf Straßen einfachsten Ausbauszustandes (1.200 km), die zu 68 % keinen Straßennamen aufweisen. Straßen unbekannter Kategorie fehlen erwartungsgemäß – mit Ausnahme von zwei Straßenabschnitten – die Angabe der Straßennamen.

Auffällig ist, dass beim Großteil der Auffahrten zu Autobahnen sowie Haupt- und Landesstraßen in der OSM kein Straßenname angeführt wird. Dem Datenmodell der OSM entsprechend sollten diese Straßenabschnitte jeweils den Namen des Straßenzugs erhalten, zu dem sie gehören⁵⁰, d.h. die Auffahrt zur Autobahn A1 sollte der Namen „A1“ zugewiesen werden. Berechnet man den Anteil der Straßen ohne Namen ohne die Auffahrten, fehlt das Straßennamenattribut nur bei 1,7 % der Autobahnen und Schnellstraßen.

Auch bei den Haupt- und Landesstraßen fallen die Auffahrten negativ auf. Insgesamt wird die Vorgabe der OSM zur Benennung von Auffahrten nur kaum erfüllt. Großzügig ausgelegt dürfte es sich hier um ein allgemeines Informationsdefizit der OSM-Erheber handeln und weniger um die unzureichende Ortskenntnis der Erheber. Dieser Abweichung wird daher im Zuge der Qualitätsbeurteilung weniger Bedeutung zugemessen.

Bei den Neben- und Erschließungsstraßen liegt der Anteil der Straßenlänge ohne Straßennamen bei 10 % bzw. 13 %. Hier ist eher davon auszugehen, dass der Name bei der Erhebung vergessen wurde, oder der Erheber für die Aufnahme der Daten nicht vor Ort war. Der Wert liegt jedoch unter den 20 %, die als Qualitätsvorgabe genannt wurden und kann daher noch akzeptiert werden.

Der relativ hohe Anteil an Straßen ohne Namen bei den Straßen einfachsten Ausbauzustandes ist durch ihre geringe Verkehrsbedeutung zu erklären und die Tatsache, dass diese Straßen meist nicht mit Straßenschildern versehen sind und eine Namensaufnahme im Gelände damit entsprechend verunmöglicht wird. Da sich die Frage stellt, ob diese Straßen nicht ohnedies aus dem Straßengraph entfernt werden sollten, sind sie für diese Qualitätsbewertung kaum von Bedeutung.

Tabelle 22: Straßen ohne Straßennamen nach Straßenkategorien im OSM-Straßengraph

Clazz	OSM Straßentyp	Straßenkategorie	km insg.	km ohne Straßennamen	Anteil der Straßen ohne Namen in %
11	highway.motorway	Autobahn	646	0	0,0
12	highway.motorway_link	Autobahnauffahrt	249	230	92,4
13	highway.trunk	Schnellstraße	86	13	14,6
14	highway.trunk_link	Schnellstraßenauffahrt	20	7	35,4
15	highway.primary	Haupt-/Landesstraßen B	1 275	5	0,4
16	highway.primary_link	Auffahrt Haupt-/Landesstraßen B	21	15	71,9
21	highway.secondary	Landesstraßen L	1 554	19	1,2
22	highway.secondary_link	Auffahrt Landesstraßen L	2	1	59,0
31	highway.tertiary	Nebenstraßen	1 082	111	10,3
32	highway.residential	Erschließungsstraßen von Siedlungen	6 677	885	13,3
41	highway.road	Straßen unbekannter Kategorie	5	4	93,3
42	highway.unclassified	Straßen einfachsten Ausbauzustandes	1 798	1 217	67,7
Straßenkategorien nach Gruppen					
11-15	Hochrangiges Straßennetz (Autobahnen, Schnellstraßen und Bundesstraßen)		2 297	270	11,8
21-22	Sekundäres Straßennetz (Landesstraßen)		1 556	20	1,3
31-32, 42	Gemeindestraßen (Nebenstraßen, Erschließungsstraßen, Straßen einfachsten Ausbauzustand)		9 557	2 213	23,2
41	Straßen unbekannter Kategorie		5	4	93,3

⁵⁰ OSM Wiki (2014):

De:Tag:highway=motorway_link,http://wiki.openstreetmap.org/wiki/DE:Tag:highway%3Dmotorway_link (08.08.2014)

Insgesamt	13 415	2 507	18,7
-----------	--------	-------	------

Quelle: Eigene Berechnungen basierend auf Straßengraph der OSM und osm2po

Tabelle 23 fasst die Ergebnisse der Bewertung zusammen. Hinsichtlich der Angaben zur Straßenkategorie ist liegt die Qualität des Datensatzes innerhalb der hohen Anforderungen, lediglich die Aufnahme der Straßen „einfachsten Ausbauzustandes“ in den Verkehrsgraph ist zu hinterfragen.

Die Attribuierung der Straßennamen wird als gut bewertet. Nimmt man die Auffahrten zu Autobahnen- und Schnellstraßen aus der Bewertung heraus, entspricht der Verkehrsgraph – mit Ausnahme der Neben- und Erschließungsstraßen – den gestellten Qualitätsanforderungen. Letztere liegen mit einer Fehlerquote von 23,2 % namenloser Straßen leicht über der Qualitätsanforderung von unter 20 %.

Tabelle 23: Ergebnisse der Bewertung der Vollständigkeit der Attribute

Evaluiierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis der Bewertung
Anzahl, Anteil und Länge der Kanten räumlich differenziert und nach Straßenkategorie für die Angaben zur Straßenklasse verfügbar sind.	Anteil der Straßenlänge mit Angabe zur Straßenkategorie an der Gesamtstraßenlänge in %	100 %	99,96 %
	Anteil der Straßenlänge der Kategorie „einfachster Ausbauzustand“ an der Gesamtstraßenlänge in %	< 10 %	13,4 %
Anzahl, Anteil und Länge der Kanten räumlich differenziert und nach Straßenkategorie für die Namensangaben verfügbar sind.	Anteil der Straßenlänge nach Straßenkategorie an der Gesamtstraßenlänge in %		
	Autobahnen, Schnellstraßen (mit Auffahrten / ohne Auffahrten)	< 3 %	11,8 % / 1,7%
	Haupt-, Landesstraßen	< 3 %	1,3 %
	Nebenstraßen, Erschließungsstraßen	< 20 %	23,2 %
	Straßen unbekannter Kategorie und einfachsten Ausbauzustandes	keine	93,3 %

Quelle: Eigene Darstellung

9.4 Thematische Genauigkeit der Kantengeschwindigkeit

Spezifikation der relevanten Datenqualitätsmerkmale, Maßeinheiten und Indikatoren

Die Überprüfung der thematischen Genauigkeit ist aufgrund des Fehlens eines Referenzdatensatzes nur beschränkt durchführbar. Relevant für die Qualitätsbewertung ist grundsätzlich die Feststellung, ob die Kantengeschwindigkeit korrekt eingegeben wurde, die Abbiegerelationen fehlerfrei modelliert wurden oder die modellierten Einbahnregelungen den aktuellen Bestimmungen entsprechen.

Um dennoch eine Aussage zur thematischen Genauigkeit treffen zu können, soll die Straßenkategorisierung des Verkehrsgraphs visuell und mit extrinsischen Mitteln auf seine Richtigkeit überprüft werden. Es wird die Hypothese aufgestellt, dass eine falsche Straßenkategorisierung, eine fehlerhafte Geschwindigkeitsangabe auf der Kante zur Folge haben kann, welche wiederum das Routingergebnis verfälscht.

Das intrinsisch angewandte Verfahren untersucht die Länge des Straßennetzes nach Straßenkategorien und Kantengeschwindigkeit um festzustellen, ob bestimmten Straßenkategorien unplausible Geschwindigkeiten zugeordnet wurden.

Darüber hinaus wird für das Beispiel der Park-and-Ride Anlage Wiener Neustadt exemplarisch eine extrinsische Bewertung durchgeführt indem über Test routings der kürzeste Weg berechnet und mit den Online-Routing Ergebnissen des Projekts ITS Vienna Region verglichen wird.

Datengrundlagen

Für die visuelle Überprüfung werden die Rasterkarte der freien Österreichischen Verwaltungsgrundlage basemap.at sowie das Routing der des Verkehrstelematik-Projekts ITS Vienna Region herangezogen, welches auf dem Referenzgraph der Graphenintegrationsplattform GIP beruht und unter <http://www.anachb.at> verfügbar ist.

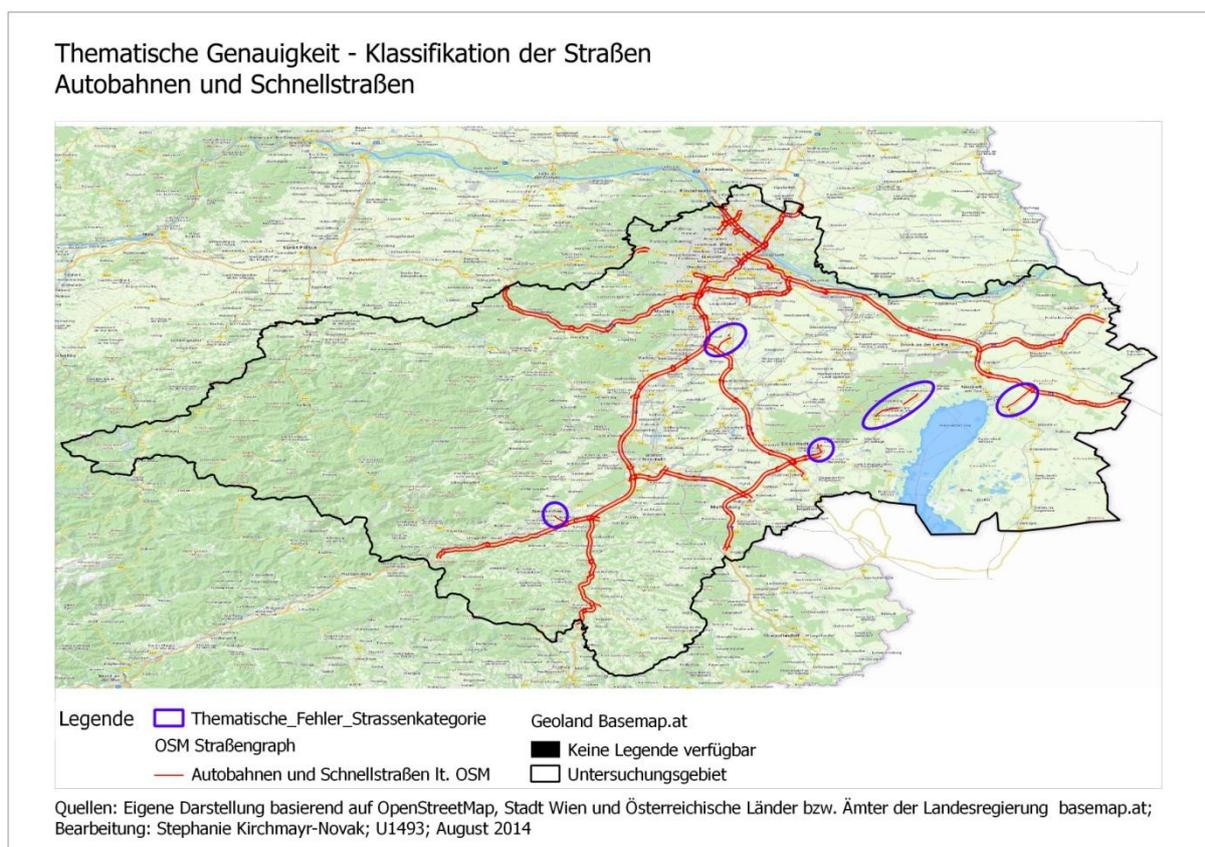
Evaluierungsmethode und Ergebnisse

Überprüfung der Attribuierung von Autobahnen- und Schnellstraßen und ihren Auffahrten

Der aus der OSM erzeugte Verkehrsgraph wird als Layer über die basemap.at gelegt. Da bei Autobahnen und Schnellstraßen in der OSM jede Fahrtrichtung erhoben wird, kann durch Versatz der Straßensignatur eine „Umrandung“ der Autobahnen- und Schnellstraßen der basemap.at erzeugt werden. Danach wird die Übereinstimmung der Datensätze visuell kontrolliert, die gefundenen fünf Fehler sind in Abbildung 21 dargestellt.

Mit einer Ausnahme wurden jeweils die Zubringerstraßen zu Autobahnauffahrten bereits aus Autobahnauffahrten modelliert. Dies wirkt sich im erzeugten Verkehrsgraph insoweit aus, als für Autobahnauffahrten standardmäßig eine Geschwindigkeit von 30 km/h vergeben wird, sofern die OSM keine Angabe zur höchstzulässigen Maximalgeschwindigkeit enthält. Insgesamt haben diese Abschnitte im Verkehrsgraph eine zu geringe Geschwindigkeit. Ein weiterer Fehler ist die falsche Klassifikation der Landesstraße B 50 Burgenland Straße zwischen Donnerskirchen und Breitenbrunn als Schnellstraße. Die Geschwindigkeiten sind jedoch vom OSM-Benutzer als Maximalgeschwindigkeit korrekt eingetragen worden, die fehlerhafte Klassifikation hat also keine Auswirkungen auf die Kantengeschwindigkeit. Zwei Beispiele für eine falsche Klassifikation der Autobahnen und Schnellstraßen sind in Abbildung 22 dargestellt.

Abbildung 21: Karte Thematische Genauigkeit – Klassifikation der Straßen. Autobahnen und Schnellstraßen

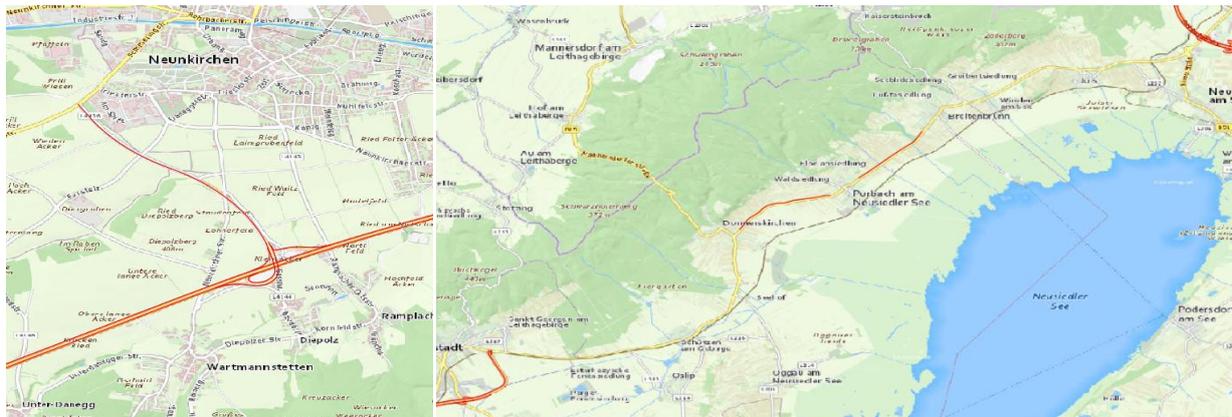


Quelle: Eigene Darstellung basierend auf Openstreetmap und basemap.at

Abbildung 22: Beispiele für die Fehlklassifikation von Autobahnen und Schnellstraßen im OSM-Graph

Beispiel 1: Landesstraße L4136 zu weit als Autobahnauffahrt klassifiziert

Beispiel 2: Landesstraße B 50 Burgenland Straße teilweise als Schnellstraße klassifiziert (rote Markierung)



Quelle: Eigene Darstellung auf Basis von OSM und basemap.at

Überprüfung der Angaben zur Kantengeschwindigkeit für das gesamte Straßennetz

Die visuelle Überprüfung für das übrige Straßennetz ist aufgrund des umfangreicheren Straßennetzes aufwändiger und fehleranfälliger in der visuellen Kontrolle. Daher wird im zweiten Schritt der Überprüfung der thematischen Genauigkeit hinsichtlich der Kantengeschwindigkeit im Netz ein anderer Evaluierungsansatz gewählt. Um unplausible Kantenattribuierungen zu identifizieren, wird die Länge des Straßennetzes nach Straßenkategorie und der jeweiligen Kantengeschwindigkeit ausgewertet.

Dabei sollten folgende Geschwindigkeitsangaben nicht unter- bzw. überschritten werden:

- Autobahnen und Schnellstraßen: $60 \leq x \leq 130$ km/h
- Auffahrten zu Autobahn, Schnellstraßen und Haupt- bzw. Landesstraßen: $30 \leq x \leq 100$ km/h
- Haupt- und Landesstraßen B: $50 \leq x \leq 100$ km/h
- Landesstraßen L: $50 \leq x \leq 100$ km/h
- Nebenstraßen: $50 \leq x \leq 100$ km/h
- Erschließungsstraßen von Siedlungen: $10 \leq x \leq 50$ km/h
- Straßen unbekannter Kategorie: $10 \leq x \leq 50$ km/h
- Straßen einfachsten Ausbaustandes: $10 \leq x \leq 30$ km/h

Tabelle 24 zeigt die Auswertung nach Straßenkategorie und Kantengeschwindigkeit, wobei die Länge des Straßennetzes je Kategorie dargestellt wird. Die jeweils nicht erlaubten Geschwindigkeiten je Straßenkategorie sind in der Tabelle grau hinterlegt.

Insgesamt liegt der Großteil der Straßen innerhalb der festgelegten Geschwindigkeitsgrenzen. Folgende Ausnahmen und potenzielle Fehler wurden festgestellt und visuell überprüft:

- Auffahrten zu Autobahnen, Haupt- und Landesstraßen:
Ein großer Teil der Auffahrten – egal zu welcher Straßenkategorie – ist mit 30 km/h attribuiert, dies ist auf die Standardattribuierung von osm2po zurückzuführen, die dann übernommen wird, wenn aus dem OSM-Datensatz keine Geschwindigkeitsangabe abgeleitet werden kann. Diese Standardgeschwindigkeit ist generell als zu gering anzusehen, kann aber aufgrund der unterschiedlichen baulichen Situationen schwer verallgemeinert werden.
- Autobahnauffahrten:
19 km (7,7 %) der Autobahnauffahrten haben eine Geschwindigkeit von 130 km/h; da es sich hier um Autobahnkleblattauffahrten oder andere kurvenreiche Abschnitte handelt, ist diese Attribuierung allein aufgrund der baulichen Situation als zu hoch anzusehen.

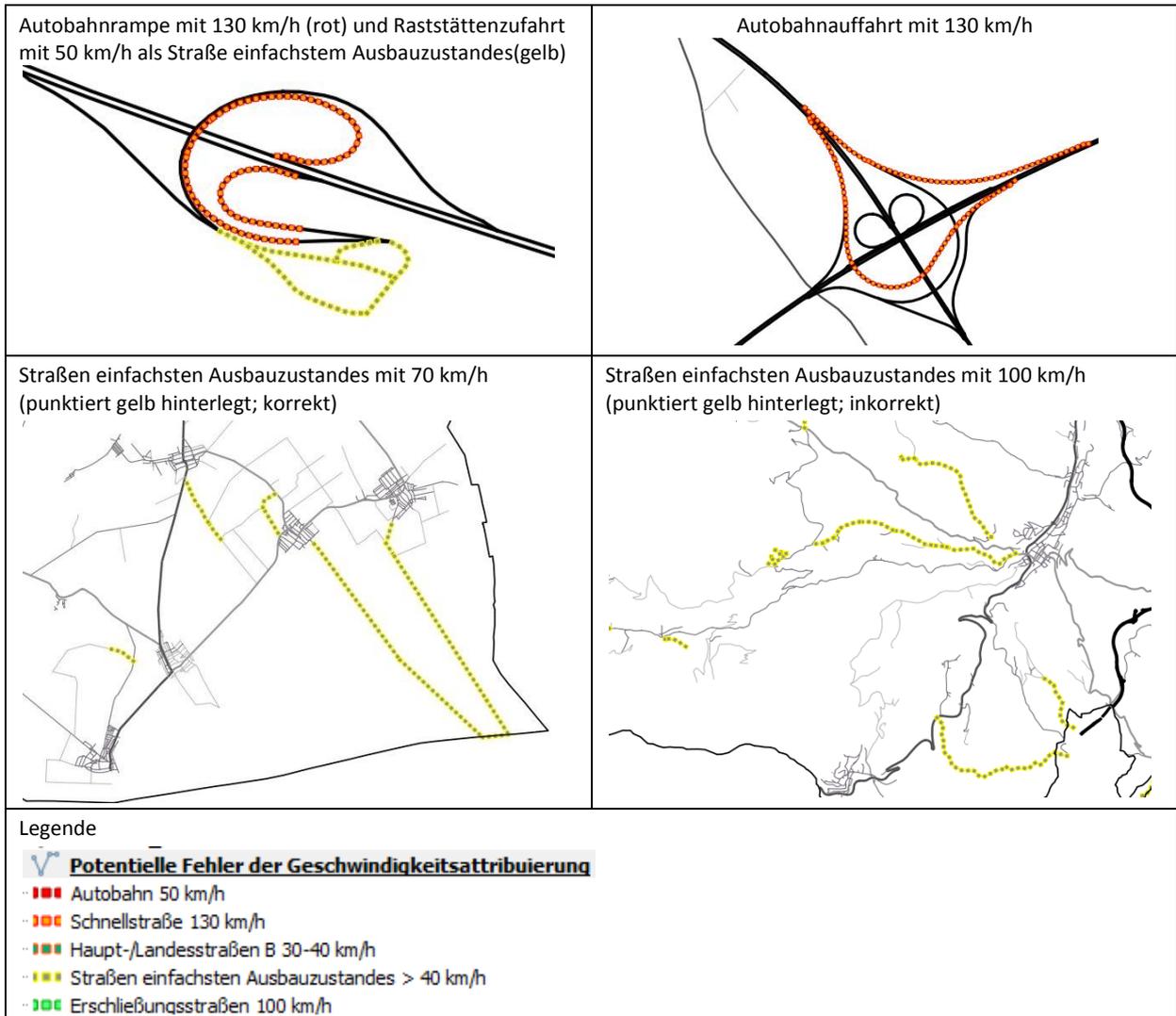
- **Autobahnen:**
700m (0,1 %) der Autobahnen sind mit nur 50 km/h attribuiert, es handelt sich um korrekte Geschwindigkeitsangaben der Einmündung der A4 bzw. A22 in die Stadtstraßen Wiens;
- **Haupt- und Landesstraßen B:**
2 km (0,2 %) der Haupt- und Landesstraßen B sind mit 30 km/h; 7,5 km (0,6%) mit 40 km/h attribuiert; die Geschwindigkeitsangaben sind in allen 9 Abschnitten korrekt vermerkt, zumeist handelt es sich um Ortsdurchfahrten oder Bahnübergänge
- **Erschließungsstraßen von Siedlungen:**
7 km der Erschließungsstraßen sind mit 100 km/h attribuiert, dies betrifft zwölf Straßenabschnitte welche in fünf Fällen korrekt attribuiert sind, in 8 Fällen von kurvenreichen Erschließungsstraßen ist die Geschwindigkeit jedoch zu hoch angeführt.
- **Straßen einfachsten Ausbaustandes:**
139 km (7,8 %) dieser Straßen sind mit 40 km/h oder mehr attribuiert, 72 km (3,9 %) davon sogar mit 70 bis 100 km/h. Die hohen Geschwindigkeitsangaben von 70 bis 100 km/h sind in der Mehrzahl der Fälle inkorrekt, da sie auf den teilweise kurvenreichen Strecken im Voralpengebiet kaum realistisch erscheinen. Jedenfalls dürften die Angaben von 100 km/h stets die erlaubte höchstzulässige Geschwindigkeit überschreiten.

Tabelle 24: Länge des Straßennetzes nach Straßenkategorie und Kantengeschwindigkeit in km/h

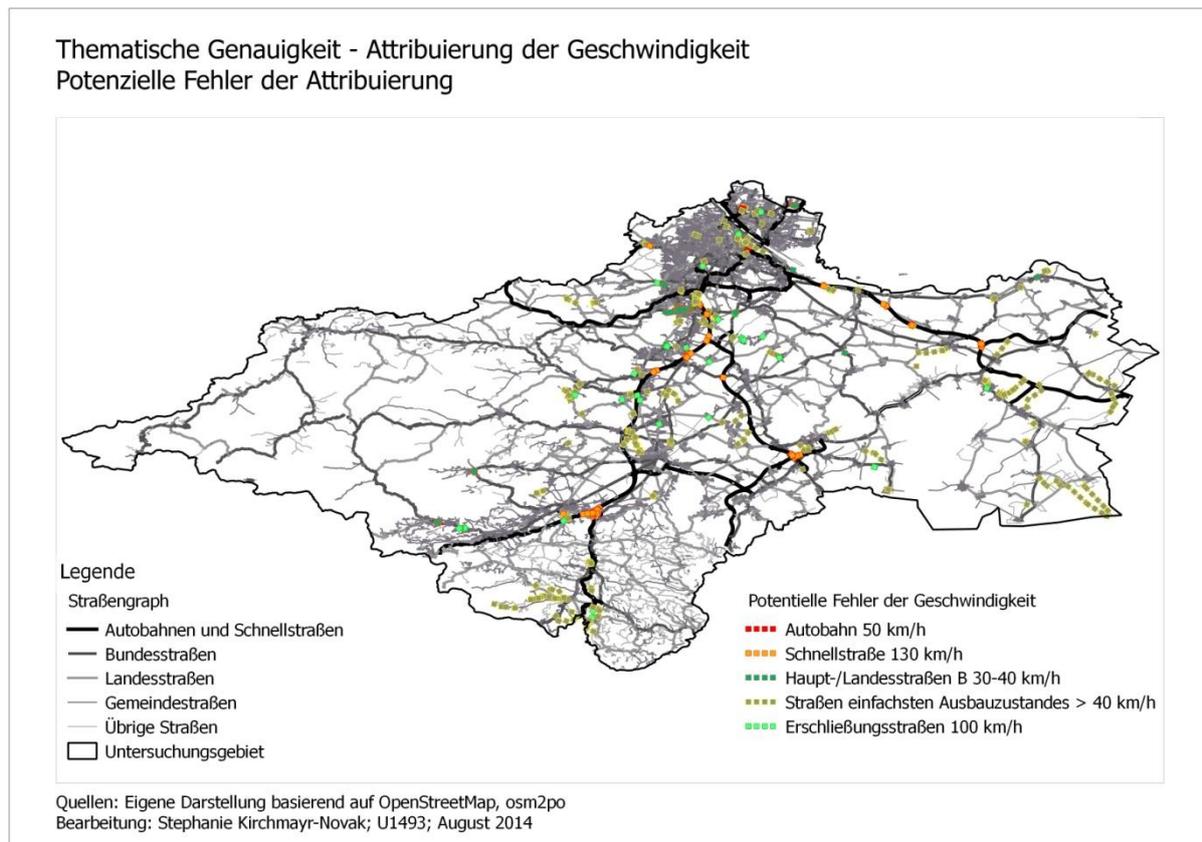
Straßenkategorie	Länge in km	Länge des Straßennetzes nach Kantengeschwindigkeit [km/h]													
		10	15	20	30	40	50	60	70	80	90	100	120	130	Insg.
11 Autobahn	646						0,1*		0,2	15,8		11,9	4,5	67,4	100,0
12 Autobahnauffahrt	249				61,4	0,3	6,5	0,8	2,9	17,1		3,3		7,7	100,0
13 Schnellstraße	86						0,4		8,3	2,5	44,1	44,6			100,0
14 Schnellstraßen-auffahrt	20				83,4		5,6		6,9	3,8		0,4			100,0
15 Haupt-/Landesstraßen B	1.275				0,2*	0,6*	37,0	0,4	46,7	0,3		14,9			100,0
16 Auffahrt Haupt-/Landesstr. B	21				56,0		38,7		0,5			4,8			100,0
21 Landesstraßen L	1.554				1,2	1,1	32,9	52,7	2,4			9,6			100,0
22 Auffahrt Landesstraßen L	2				18,1		74,9		3,3			3,8			100,0
31 Nebenstraßen	1.082				6,3	56,2	31,5	0,3	1,4			4,4			100,0
32 Erschließungsstraßen von Siedlungen	6.677	0,1			23,3	2,1	74,3					0,1			100,0
41 Straßen unbekannter Kategorie	5			0,3	22,4		77,3								100,0
42 Straßen einfachsten Ausbaustandes	1.800				92,2	0,8*	2,5*	0,4*	1,9*	0,2		1,8			100,0
Insgesamt	13.417	0,0	0,0	0,0	26,0	5,9	47,4	6,2	5,2	1,2	0,3	4,1	0,2	3,4	100,0

Anmerkung: die grau markierten Felder stellen unplausible Geschwindigkeitsangaben dar, die mit * gekennzeichneten Kategorien konnten nach visueller Kontrolle als vollständig korrekt identifiziert werden.

Quelle: eigene Darstellung basierend auf OSM und osm2po

Abbildung 23: Beispiele für inkorrekte und korrekte Attribuierung von Straßenabschnitten

Quelle: eigene Darstellung basierend auf OSM, QGIS und osm2po

Abbildung 24: Karte der potenziellen Fehler der Geschwindigkeitsattribution

Quelle: Eigene Darstellung basierend auf OSM und osm2po

Test routings der kürzesten Wege

Mit dem Dijkstra-Algorithmus aufbauenden pgRouting-Befehl `pgr_dijkstra` (siehe Code 14 in Kapitel 7.3.1) wurde die Fahrzeit für den kürzesten Weg zwischen zwei Punkten berechnet und mit dem Routingergebnis für den kürzesten Weg im Straßenverkehr von ITS Vienna Region (für einen Mittwoch im Dezember zwischen 23.00 und 24.00 Uhr) verglichen.

Die Ergebnisse sind für drei Beispiele in Abbildung 25 dargestellt.

Grundsätzlich fließt in die Routenwahl und Fahrzeitangabe von ITS Vienna Region auch die durchschnittliche Verkehrsbelastung der Tageszeit mit ein, laufend über Echtzeiterhebung von Taxiunternehmen eingespielt wird. Daher wurde für das Vergleichsrouting bewusst ein Zeitraum geringer Verkehrsbelastung gewählt, der ein möglichst unbelastetes Netz widerspiegeln soll.

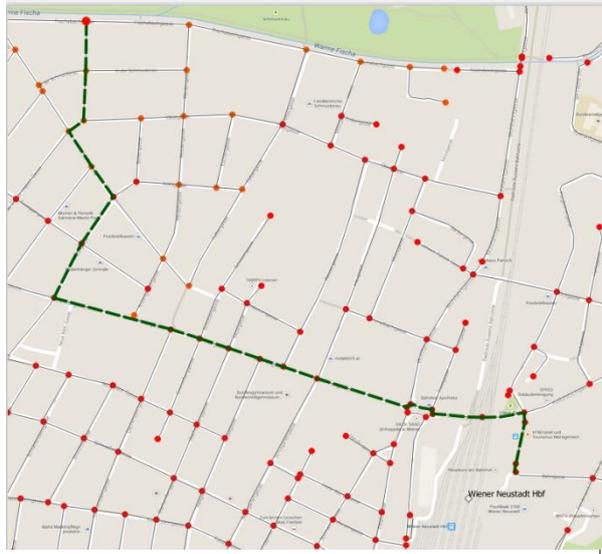
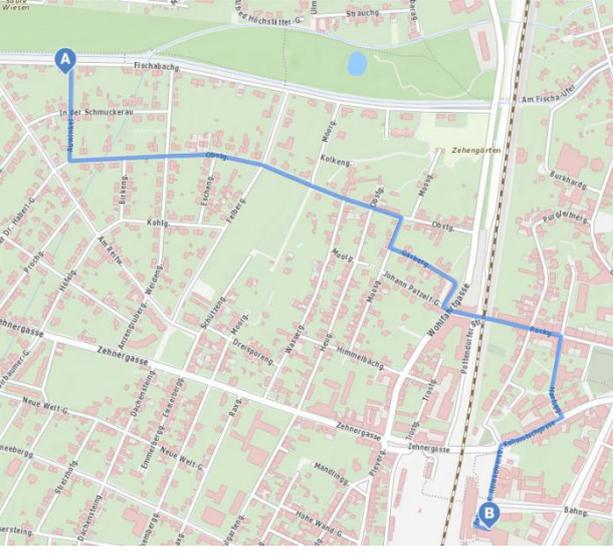
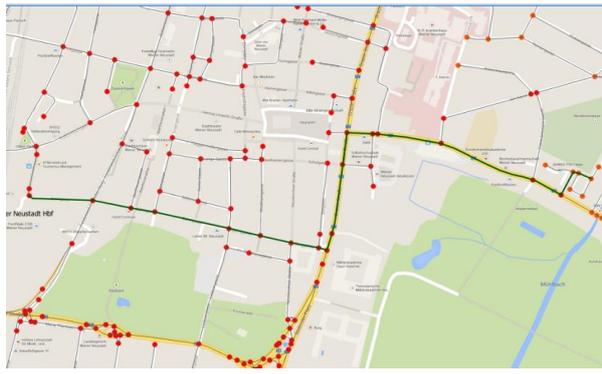
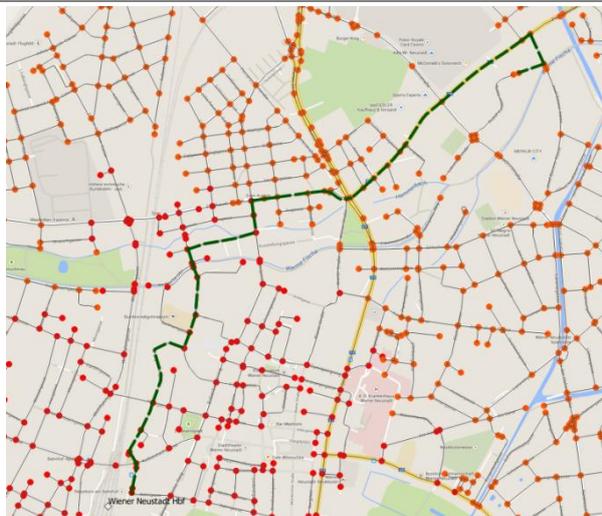
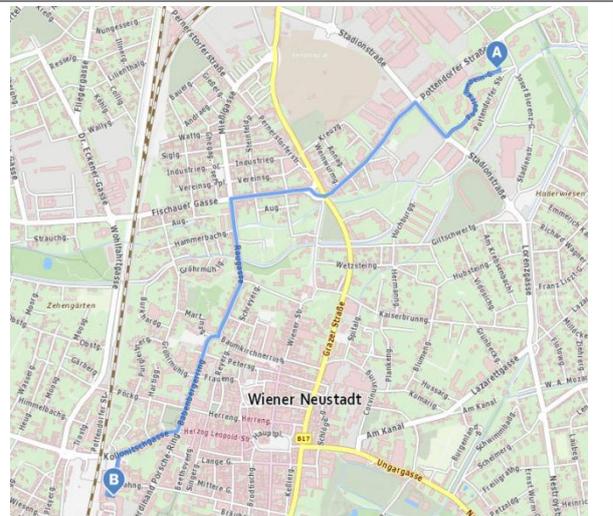
Dennoch weichen in jedem der Beispiele sowohl die Routenwahl, als auch die Fahrzeitberechnung voneinander ab, wobei die Routenwahl eine geringere Diskrepanz zeigt, als die Fahrzeitberechnung.

Es wird vermutet, dass der Grund für die unterschiedlichen Ergebnisse in der Berücksichtigung durchschnittlicher Wartezeiten an Lichtsignalanlagen liegt, die von ITS Vienna Region berücksichtigt werden, in den Konvertierungsgrundeinstellungen von osm2po aber nicht berücksichtigt wurden. Geht man von rund einer Minute durchschnittlicher Wartezeit an Lichtsignalanlage aus, sind die Fahrzeitergebnisse bereits vergleichbar.

Für diese Begründung spricht, dass die Geschwindigkeitsangabe als weiterer, wichtiger Einflussparameter mit 50 km/h innerhalb des Stadtgebiets von Wiener Neustadt im Straßengraph von osm2po als korrekt eingeschätzt werden kann.

Grundsätzlich unterschätzt der von osm2po erstellte Straßenverkehrsgraph also die Fahrzeit bereits auf kurzen Strecken von 2 bis 3 km, was bei der Interpretation der später zu erstellenden Einzugsbereiche der Park-and-Ride Anlagen zu berücksichtigen ist.

Abbildung 25: Test routings der kürzesten Wege, Ergebnisvergleich OSM/osm2po/pgRouting und ITS Vienna Region

Beispiel 1: Wiener Neustadt: Fischabachstraße 75 - Bahnhofplatz [Knoten 37892 -> 36993]	
OSM, osm2po, pgRouting: 2,67 Minuten	ITS Vienna Region: 5 Minuten (inkl. 2 Ampeln)
	
Beispiel 2: Wiener Neustadt: Burgenlandgasse 8 - Bahnhofplatz [Knoten 37891 -> 36993]	
OSM, osm2po, pgRouting: 2,07 Minuten	ITS Vienna Region: 3 Minuten (inkl. 1 Ampel)
	
Beispiel 3: Wiener Neustadt: Rudolf Kumbain-Gasse 15 - Bahnhofplatz [Knoten 37706 -> 36993]	
OSM, osm2po, pgRouting: 4,0 Minuten	ITS Vienna Region: 6 Minuten (inkl. 3 Ampeln)
	

Quelle: Eigene Darstellung basierend auf OSM, osm2po, pgRouting, QGIS und ITS Vienna Region

Zusammenfassung und Schlussfolgerungen zur Bewertung der thematischen Genauigkeit

Die visuelle Kontrolle der thematischen Genauigkeit ist aufwändig und unvollständig. Sie dient in erster Linie dazu, sich mit den Stärken und Schwächen des Verkehrsgraphs vertraut zu machen und ist nicht mit einer systematischen, flächendeckenden Analyse zu vergleichen, die in der Differenzbetrachtung mit einem routenfähigen Referenzstraßengraph gemacht werden kann.

Bei der visuellen Kontrolle wird nur die Plausibilität der Ergebnisse überprüft und vernachlässigt, dass auch Straßenzüge mit plausiblen Geschwindigkeitsangaben aufgrund der konkreten rechtlichen oder baulichen Situation falsch attribuiert sein können. Darüber hinaus ist anzumerken, dass der Straßengraph keine Informationen hinsichtlich der Straßenkapazität wie Kurvigkeit und Steigung enthält, welche die tatsächlich mögliche Geschwindigkeit auf der Kante stark beeinflussen können.

Zusammenfassend kann gesagt werden, dass nur wenige schwerwiegende Fehler gefunden wurden. Diese umfassen die mit 130 km/h attribuierten Autobahnauffahrten, die Straßen einfachsten Ausbaustandes mit einer Geschwindigkeitsangabe von 100 km/h und die in osm2po definierte Standardattribuierung der Autobahnauffahrten mit 30 km/h, welche grundsätzlich zu hinterfragen ist. Die selektive und qualitative Auswertung der Testroutings wird aufgrund der geringen Anzahl der Testfälle nicht als vollwertiger Evaluierungsansatz in die Bewertung aufgenommen. Tabelle 25 fasst diese Ergebnisse zusammen.

Tabelle 25: Ergebnisse der Bewertung der thematischen Genauigkeit der Attribute

Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis der Bewertung
Straßenkategorie: Sind die Straßenkategorien korrekt attribuiert? Wurde die korrekte Kantengeschwindigkeit zugewiesen?	Anzahl der falsch attribuierten Autobahnabschnitte	< 5	5 falsch kategorisierte Autobahnabschnitte
Geschwindigkeitsattribuierung: Gibt es unplausible Geschwindigkeitsattribuierungen die durch visuelle Kontrolle identifiziert werden können?	Anteil der Straßenlänge je Straßenkategorie mit unplausiblen Geschwindigkeitsangaben	< 3 %	Autobahnauffahrten: mit 30 km/h: 61 %; mit 130 km/h: 7,7 % Straßen einfachsten Ausbaustandes: mit > 70 km/h: 2%

Quelle: Eigene Darstellung

9.5 Benutzerfreundlichkeit

Spezifikation der relevanten Datenqualitätsmerkmale, Maßeinheiten und Indikatoren

Wesentlich für die Benutzerfreundlichkeit im Anwendungsfall ist die Verfügbarkeit ausreichend detaillierte und verständliche Dokumentation der verwendeten Open Source Software.

Evaluierungsmethode und Ergebnisse

Es wird für jedes im Rahmen der Arbeit verwendete Open Source Produkt überprüft, ob eine Dokumentation inklusive Anwendungsbeispielen zur Verfügung steht bzw. darüber hinaus eine Möglichkeit besteht direkte Fragen an die Ersteller der Software bzw. andere Anwender der Software zu stellen die gute Aussicht auf Beantwortung hat. Die Ergebnisse sind in Tabelle 26 dargestellt.

Jedes der gewählten Open Source Produkte verfügt über eine eigene Homepage auf welcher eine Dokumentation zum Produkt und Informationen über Anwendungsbeispiele wie Handbücher oder Berichte von externen Nutzern abgerufen werden können. Darüber hinaus stehen für jedes Produkt eine Mailingliste oder ein Online Forum zur Verfügung.

Tabelle 26: Verfügbarkeit von Informationen für die im Rahmen der Arbeit verwendeten Open Source Produkte

Open Source Produkt	Aktuelle Dokumentation verfügbar (Zugriff per 3.11.2014)	Dokumentation enthält Anwendungsbeispiele	Anwenderforum mit aktuellen Fragen und Antworten verfügbar
QGIS	QGIS Benutzerhandbuch 2.2 http://docs.qgis.org/2.6/de/docs/user_manual/index.html	Schulungs-handbuch mit Beispielen verfügbar	Maillingliste http://www.qgis.org/de/site/getinvolved/maillinglists.html#qgis-maillinglists Online Forum http://gis.stackexchange.com
Osmosis	Umfassende Dokumentation in Ramm, Topf (2010, 279ff)	ja	Entwickler Mailingliste: http://lists.openstreetmap.org/listinfo/osmosis-dev http://trac.openstreetmap.org/query?component=osmosis
Osm2po	Dokumentation wird im Download zusammen mit Programm zur Verfügung gestellt.	Anwendungsbeispiele einer externen Nutzerin verfügbar http://anitagraser.com/2011/12/15/an-osm2po-quickstart/ http://anitagraser.com/2011/12/18/osm2po-part-2-pgrouting-on-osm-the-easy-way/	Online Forum http://gis.stackexchange.com
pgRouting	Dokumentation in mehreren Sprachen verfügbar http://docs.pgrouting.org/	Workshop Manuals verfügbar http://workshop.pgrouting.org/	User Mailinglist3 und GIS StackExchange: http://lists.osgeo.org/mailman/listinfo/pgrouting-users Kommerzieller Support in Deutschland, Japan, US, Schweiz und Frankreich

Quelle: Eigene Erhebung

Die Qualitätsanforderungen hinsichtlich der Benutzerfreundlichkeit werden durch die verfügbare Software Dokumentation erfüllt, wie auch in Tabelle 27 dargestellt wird.

Tabelle 27: Ergebnisse der Bewertung der Benutzerfreundlichkeit

Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis der Bewertung
Ist die Dokumentation der verwendeten Softwareprodukte ausreichend, um den Inhalt des Straßengraphs und die Ergebnisse der durchgeführten Berechnungen korrekt zu interpretieren?	Jedes Open Source Produkt verfügt über eine verständliche, ausreichend detaillierte Dokumentation.	Binär (ja / nein)	Ja

Quelle: Eigene Darstellung

9.6 Zusammenfassung der Evaluierung und Resümee

Insgesamt konnten ohne Zugriff auf einen routingfähigen Referenzstraßenverkehrsgraph vier der sechs Qualitätsmerkmale nach dem Qualitätsmodell von ISO 19157:2011 untersucht und bewertet werden. Unbewertet blieben nur die Lagegenauigkeit und die zeitliche Genauigkeit, sofern sie nicht durch die Prüfung der Vollständigkeit der Objekte abgedeckt werden konnte.

Die Ergebnisse werden in Tabelle 28 nochmals zusammenfassend aufgeführt. Die anfangs gestellten Qualitätsanforderungen konnten weitgehend vollständig erreicht oder übertroffen werden.

Nicht vollinhaltlich erfüllt werden konnten die Qualitätsanforderungen an die Vollständigkeit der Attribute und die thematische Genauigkeit. Geringe Abweichungen von den Qualitätsanforderungen zeigte der Anteil der Straßen mit Angabe einer Straßenkategorie bzw. der Anteil der Straßen „einfachsten Ausbauzustandes“ innerhalb des Straßengraphs.

Größere Abweichungen von den Qualitätsanforderungen konnten hinsichtlich des Anteils an Straßen festgestellt werden, deren Geschwindigkeitsattribuierung aufgrund ihrer standardmäßigen baulichen Ausgestaltung in Österreich zu hinterfragen ist. Problematisch sind hier insbesondere die zum überwiegenden Teil (61 %) mit einer Geschwindigkeitsangabe von 30 km/h zu langsam attribuierten Autobahnauffahrten. Da diese Attribuierung auf die Standardeinstellung der Software osm2po zurückgeht (siehe Kapitel 6.4 osm2po), wird empfohlen diesen Wert in der Konfigurationsdatei von osm2po auf 50 km/h anzuheben.

Generell kann bei der intrinsischen Bewertung des Verkehrsgraphen bemängelt werden, dass wichtige Parameter des Verkehrsgraphen wie die Korrektheit der Abbiegerelationen und Einbahnregelungen ohne einen routenfähigen Referenzstraßengraph nicht bewertet werden konnten.

Dieser Mangel ist jedoch hinsichtlich der Aufgabenstellung einen makroskopischen Verkehrsgraphen für die Erzeugung von Einzugsbereichen für Park-and-Ride Anlagen zu erstellen vernachlässigbar, da hier gröbere Aussagen über die Fahrzeit ausreichen. Für konkrete Routingverfahren ist dieser Mangel signifikanter. Die gegenständliche Arbeit kann hinsichtlich der Genauigkeit der Abbiegerelationen und Einbahnregelungen nur auf die Auswertungen von Graser et al. (2014) verweisen, die dem Straßengraphen in der Wiener Großregion – welche räumlich enger abgegrenzt ist als das gegenständliche Untersuchungsgebiet – eine gute Qualität hinsichtlich der oben genannten Parameter Abbiegerelationen und Einbahnregelungen aussprechen.

Tabelle 28: Zusammenfassung der Ergebnisse der Evaluierung des Straßengraphen

Qualitätsmerkmal / -parameter Evaluierungsansatz	Maßeinheit und Indikator	Qualitätsanforderung	Ergebnis
1.1 Formatkonsistenz			
Können die OSM-Daten nach den durchgeführten Bearbeitungs- und Konvertierungsschritten in einer PostgreSQL/ PostGIS-Datenbank als Verkehrsgraph weiter verarbeitet werden?	Binär (ja / nein)	Ja	Ja
1.2 Konzeptuelle Konsistenz			
Ist das Datenmodell in sich konsistent und entspricht es dem für ein Datenrouting erforderlichen Datenmodell?	Binär (ja / nein)	Ja	Ja
1.3 Konsistenz der Wertebereiche			
Sind alle für das Routing erforderlichen Daten im generierten Datensatz enthalten?	Binär (ja / nein)	Ja	Ja
Sind die Wertebereiche der Attribute konsistent mit diesem Datenmodell?	Binär (ja / nein)	Ja	Ja
1.4 Topologische Konsistenz			
Anzahl der topologischen Fehler und %-Anteil der fehlerhaften Kanten bzw. Knoten			
Isolierte Einzelsegmente (Einzelkanten) ohne Verbindung zum "Hauptstraßengraph"	Isolierte Straßensegmente	Anteil fehlerhafter Objekte < 5% von allen Objekten	0,13%
Isolierte Segmente aus mehr als einer Kante ohne Verbindung zum Hauptstraßengraph			0,46%
Knoten, die Lücken im Netz darstellen. Diese Knoten haben, wie Sackgassen, nur zu einer Kante eine Verbindung, sind aber semantisch keine Sackgassen, sondern unvollständig digitalisierte Straßenzüge.	Lücken im Straßennetz		0,01%
Inkorrekte planfreie Kreuzungen die weder Brücken oder Unterführungen noch für die	Fehlende Schnittpunkte kreuzender Kanten		0,00%

Qualitätsmerkmal / -parameter Evaluierungsansatz	Maßeinheit und Indikator	Qualitäts- anforderung	Ergebnis
Auflösung komplizierter Fahrtrouten herangezogen werden.			
Kanten, deren Quellknoten gleichzeitig auch Zielknoten ist, aber keine Schleifen darstellen.	Fehlerhafte Schleifen		0,00%
2.1 Vollständigkeit des Netzgraphs			
Erschließungsquote und Erschließungsgrad der Siedlungsgebiete Anteil der Siedlungsgebiete und der Siedlungsfläche, die durch Straßeninfrastruktur erschlossen wird. Die Siedlungsfläche wird dem frei verfügbaren Geodatenatz der Sied- lungseinheiten, Gebietsstand 2010 von Statistik Austria entnommen. Der Erschließungsgrad wird in vier Varianten gerechnet, indem ein Puffer von 50m - 400m um die Straßenkanten gelegt wird.	Erschließungsquote: Anteil der Siedlungsgebiete (Siedlungseinheiten) in %, der durch Straßen im Umkreis von 50, 100, 150, 200 und 400 m erschlossen wird.	Erschließungs- quote von 100%	100 %
	Erschließungsgrad: Anteil der Siedlungsfläche in %, der durch Straßen im Um- kreis von 50, 100, 150, 200 und 400 m erschlossen wird.	Erschließungs- grad von 95%	97,9 – 100 % (50 m – 400 m Puffer)
2.2 Vollständigkeit der Attribute			
Anzahl, Anteil und Länge der Kanten räumlich differenziert und nach Straßenkategorie,			
für die Angaben zur Straßenklasse verfügbar sind.	Anteil der Straßenlänge mit Angabe zur Straßenkategorie an der Gesamtstraßenlänge in %	100 %	99,96 %
	Anteil der Straßenlänge der Kategorie „einfachster Ausbauzustand“ an der Gesamtstraßenlänge in %	< 10 %	13,4 %
für die die Namensangaben verfügbar sind.	Anteil der Straßenlänge nach Straßenkategorie an der Gesamtstraßenlänge in %		
	Autobahnen, Schnellstraßen	< 3 %	11,8 % / 1,7%
	Haupt-, Landesstraßen	< 3 %	1,3 %
	Nebenstraßen, Erschließungsstraßen	< 20 %	23,2 %
	Straßen unbekannter Kategorie und einfachsten Ausbauzustandes	keine	93,3 %
3. Thematische Genauigkeit			
Sind die Straßenkategorien korrekt attribuiert? Wurde die korrekte Kantengeschwindigkeit zugewiesen?	Anzahl der falsch attribuierten Autobahnabschnitte	< 5	5 falsch kategorisierte Autobahnabschnitte
Gibt es unplausible Geschwindigkeitsattribuierungen die durch visuelle Kontrolle identifiziert werden können?	Anteil der Straßenlänge je Straßenkategorie mit unplau- siblen Geschwindig- keitsangaben	< 3 %	Autobahnauffahrten: mit 30 km/h: 61 %; mit 130 km/h: 7,7 % Straßen einfachsten Ausbauzustandes: mit > 70 km/h: 2%
4 Benutzerfreundlichkeit			
Ist die Dokumentation der verwendeten Software- produkte ausreichend, um den Inhalt des Straßen- graphs und die Ergebnisse der durchgeführten Berechnungen korrekt zu interpretieren?	Binär (ja / nein)	Ja	Ja

Quelle: Eigene Darstellung

10 Berechnung der Einzugsgebiet von Park-and-Ride Anlagen

Im letzten Evaluierungsschritt soll die Qualität des erzeugten Straßenverkehrsgraphs anhand einer Anwendung überprüft werden. Dazu werden zwei Ansätze zur Berechnung der Einzugsgebiete der Park-and-Ride Anlagen auf Basis des Straßenverkehrsgraphs und der Fahrzeit zu den Park-and-Ride Anlagen herangezogen.

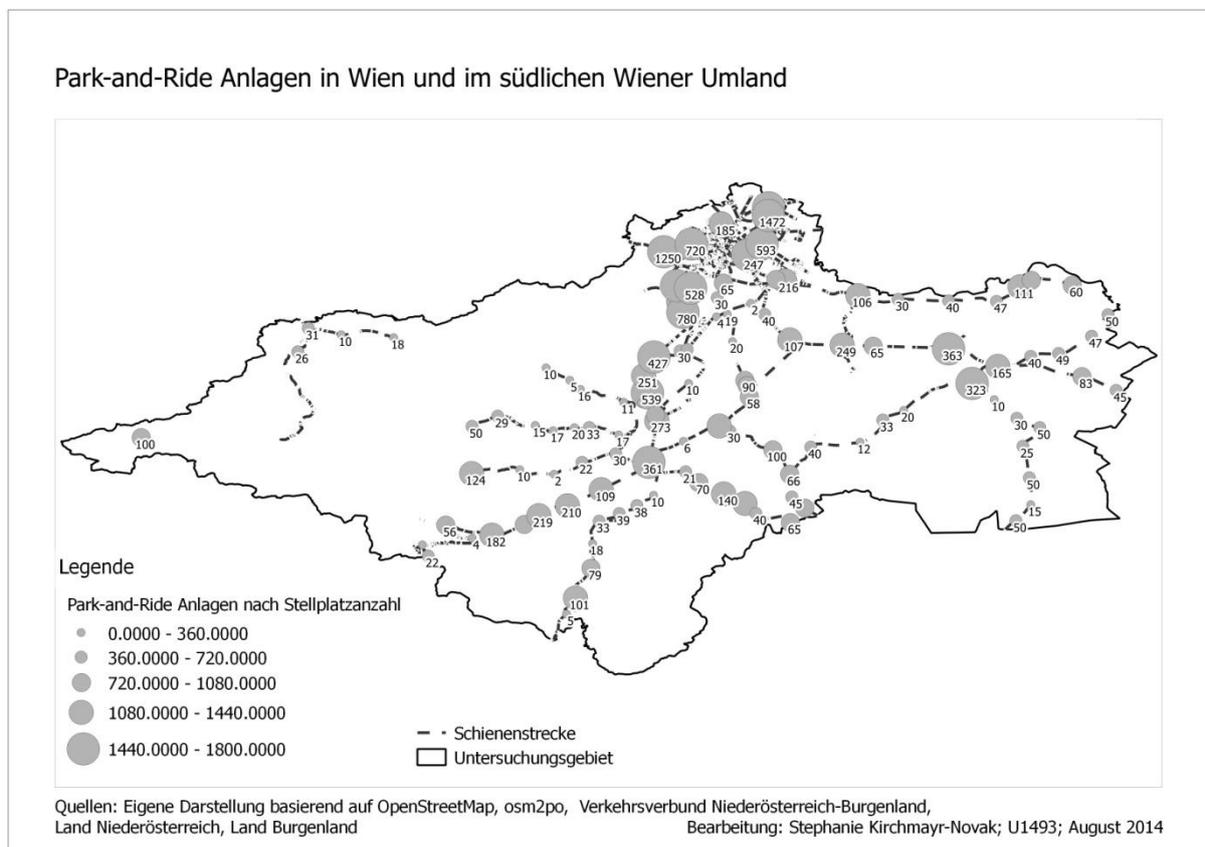
Die Fahrzeit zu Park-and-Ride Anlagen ist grundsätzlich ein kontinuierliches Phänomen, welches mit der Distanz stetig ansteigt, allerdings räumlich auf den Verlauf der Verkehrswege beschränkt ist. Im ersten Ansatz werden die Einzugsgebiete mit Hilfe von Voronoi-Polygonen ermittelt und hinsichtlich ihrer Fläche, der Anzahl der Einwohner sowie der Anzahl der Park-and-Ride Stellplätze pro Einwohner im Einzugsgebiet ausgewertet. Im zweiten Ansatz erfolgt die Ermittlung der Einzugsgebiete durch Interpolation der Fahrzeit zwischen den Straßengraphknotenpunkten im Regular Spline with Tension-Verfahren (RST).

10.1 Aufbereitung der Datengrundlagen

10.1.1 Erhebung der Park-and-Ride Anlagen im Untersuchungsgebiet

Nach Recherchen beim Verkehrsverbund Niederösterreich-Burgenland (VVNB), dem Land Niederösterreich und dem Land Burgenland gibt es in Wien und im südlichen Wiener Umland insgesamt 117 Park-and-Ride Anlagen bzw. Bahnstationen mit Parkplätzen, die in Abbildung 26 dargestellt sind und zwischen zwei und 1.800 Stellplätzen bieten. Das Industrieviertel ist durch die Südbahnstrecke und ihre Nebenstrecken sehr gut erschlossen, auch im Nordburgenland stehen im Verlauf der Pressburger Bahn, der Ostbahn und der Nebenstrecken ausreichend Stellplätze zur Verfügung. Im Niederösterreichischen Bezirk Lilienfeld im Westen des Untersuchungsgebiets gibt es hingegen nur fünf Park-and-Ride Anlagen, deren Bahnstrecken jeweils nicht direkt, sondern nur über St. Pölten mit Wien verbunden sind.

Abbildung 26: Karte der Park-and-Ride Anlagen in Wien und im südlichen Wiener Umland



Quelle: Eigene Darstellung basierend auf OSM und osm2po

10.1.2 Verortung der Park-and-Ride-Anlagen im Straßengraph

Voraussetzung für die weiteren Berechnungen ist es, jeder digitalisierten Park-and-Ride Anlage einen Knotenpunkt im Straßengraph zuzuordnen, durch welchen er für die weitere Berechnung repräsentiert wird. Code 13 zeigt den SQL-Code, mit welchem der nächste Netzwerkknoten jeder Park-and-Ride Anlage identifiziert und der Tabelle der Park-and-Ride Anlagen (`wsums_p_r`) hinzugefügt wird.

Code 13: SQL-Code zur Identifikation des räumlich nächsten Netzwerkknotens des Straßengraphs um die Park-and-Ride Anlage

```
ALTER TABLE wsums_p_r ADD COLUMN nearest_node integer;
CREATE TABLE temp AS
  SELECT a.gid, b.id, min(a.dist)
  FROM
    (SELECT hst.gid, min(st_distance(hst.geom, nodes.the_geom)) AS dist
     FROM wsums_p_r hst, ws_2po_4pgr_vertices_pgr nodes
     GROUP BY hst.gid
    ) AS a,
    (SELECT hst.gid, nodes.id, st_distance(hst.geom, nodes.the_geom) AS dist
     FROM wsums_p_r hst, ws_2po_4pgr_vertices_pgr nodes
    ) AS b
  WHERE a.dist=b.dist AND a.gid=b.gid
  GROUP BY a.gid, b.id;
UPDATE wsums_p_r SET nearest_node =
  (SELECT id FROM temp WHERE temp.gid=wsums_p_r.gid);
```

Quelle: eigene Berechnung adaptiert nach Graser, A. (2011)

10.1.3 Berechnung der Fahrzeiten von den Park-and-Ride Anlage zu jedem Netzwerkknoten

Im nächsten Schritt wird für jede Park-and-Ride Anlage die Fahrzeit zwischen dem Standort der Anlage und jedem Netzwerkknoten berechnet. Der in Code 15 dargestellte vereinfachte Ansatz wurde von Graser, 2011 übernommen und basiert auf dem pgRouting-Algorithmus „`pgr_drivingDistance`“. Dieser Dijkstra-Algorithmus ermittelt den kürzesten Weg zwischen den Punkten berechnet daraus die Fahrzeit (`cost`), wobei der kürzeste Weg nicht notwendiger Weise auch der schnellste Weg sein muss. Die Parameter des Algorithmus und seine Schreibweise sind in Code 14 dargestellt. Eine detailliertere Beschreibung des Algorithmus und seiner Parameter bietet TAKUBO et al. (2013, S. 66 f).

Code 14: SQL-Code von pgRouting zur Berechnung der Fahrzeit des kürzesten Weges zwischen zwei Punkten (`pgr_drivingDistance`)

PgRouting Funktion zur Berechnung des kürzesten Weges

```
Pgr_costResult [] pgr_drivingDistance(text sql, integer source, double precision distance,
boolean directed, boolean has_rcost);
```

und die SQL query wie folgt definiert ist:

```
SELECT int4 id, int4 source, int4 target, float8 cost [,float8 reverse_cost] FROM edge_table
```

wobei

<code>id</code>	... Identifizier der Kante,	<code>distance</code>	... Wert in der Einheit der Kantenkosten,
<code>source</code>	... Identifizier des Quellknotens,	<code>directed</code>	... wahr, wenn der Graph gerichtet ist,
<code>target</code>	... Identifizier des Zielknotens,	<code>has_rcost</code>	... wahr wenn die Spalte <code>reverse_cost</code> für die
<code>cost</code>	... Kantenfahrzeit,		Kantenfahrzeit in Gegenrichtung benutzt wird.

Quelle: Takubo (2013)

Insgesamt werden die in Code 3 dargestellten vier Bearbeitungsschritte durchgeführt, die Erstellung der Fahrzeitentabelle, die Integration der Fahrzeiten und Attribuierung der Park-and-Ride Anlage sowie der Integration eines eindeutigen Schlüssels als Kombination der Netzwerkknoten-ID und der Fahrzeit.

Code 15: SQL-Code zur Berechnung der Fahrzeiten von einer Park-and-Ride Anlage zu jedem Netzwerkknoten des Verkehrsgraphs*Erstellung der Berechnungstabelle*

```
CREATE TABLE drivetimes (
  node integer,          -- Netzwerkknoten-ID
  cost double precision, -- Fahrzeit in Stunden
  the_geom geometry(Point,4326), -- Geometrie des Netzwerkknoten
  hst integer);         -- ID des die Park-and-Ride Anlage repräsentierenden Knotens
```

Berechnung der Entfernung der Park-and-Ride Anlage zu allen Netzpunkten

```
INSERT INTO drivetimes
  (SELECT di.id1 AS node,di.cost,nd.the_geom
   FROM pgr_drivingDistance(
     'SELECT id, source, target, cost FROM ws_2po_4pgr', 29371, 10, false, false
   ) AS di
  JOIN ws_2po_4pgr_vertices_pgr nd
    ON di.id1 = nd.id);
```

Attribuierung der Ergebnisse mit dem Knotencode der Park-and-Ride-Anlage

```
UPDATE drivetimes SET hst = 29371 WHERE hst IS NULL;
```

Hinzufügen eines eindeutigen Schlüssels „key“ als Kombination der Netzwerkknoten-ID und der Fahrzeit

```
ALTER TABLE drivetimes ADD column key;
UPDATE drivetimes SET key = concat(node,cost);0
```

Quelle: eigene Berechnung adaptiert nach Takubo (2013) und Graser, A. (2011)

10.1.4 Extraktion der Park-and-Ride Anlage mit der kürzesten Fahrzeit

Der in Code 15 dargestellte Algorithmus listet die Fahrzeiten zwischen allen 117 Park-and-Ride Anlagen und allen Netzwerkknoten auf. Aus dieser umfangreichen Liste wird im nächsten Schritt für jeden Netzwerkknoten die Park-and-Ride-Anlage mit der kürzesten Fahrzeit extrahiert und in eine neue Tabelle geschrieben. Auch dieser Liste wird ein eindeutiger Schlüssel aus Knoten-ID und Fahrzeit zur Park-and-Ride Anlage hinzugefügt. Dieser Schlüssel wird im letzten Schritt dazu genutzt, die ID des die Park-and-Ride Anlage repräsentierenden Knotens (hst) hinzuzufügen.

Code 16: SQL-Code zur Extraktion der Park-and-Ride Anlage mit der kürzesten Fahrzeit*Erstellung einer Tabelle durch Extraktion der Park-and-Ride Anlage mit der kürzesten Fahrzeit*

```
CREATE TABLE catchment_node AS
SELECT node, min(cost) AS cost_min, the_geom
FROM drivetimes
GROUP BY node, the_geom;
```

Hinzufügen eines eindeutigen Schlüssels „key“ als Kombination der Netzwerkknoten-ID und der Fahrzeit

```
ALTER TABLE catchment_node ADD column key text;
UPDATE catchment_node SET key = concat(node,cost_min);
```

Hinzufügen der Knoten-ID der Park-and-Ride Anlage

```
ALTER TABLE catchment_node ADD column hst integer;
UPDATE catchment_node SET hst =
  (SELECT hst
   FROM drivetimes
   WHERE catchment_node.key = drivetimes.key);
```

Quelle: eigene Berechnung adaptiert nach Graser, A. (2011)

10.2 Fahrzeitbasierte Einzugsbereiche berechnet mit Voronoi-Polygonen

Für die erste Umsetzung der Fahrzeiten von jedem Netzwerkknoten zur jeweils am schnellsten erreichbaren Park-and-Ride Anlage in Einzugsbereiche der Park-and-Ride Anlagen wurden Voronoi-Polygone ermittelt.

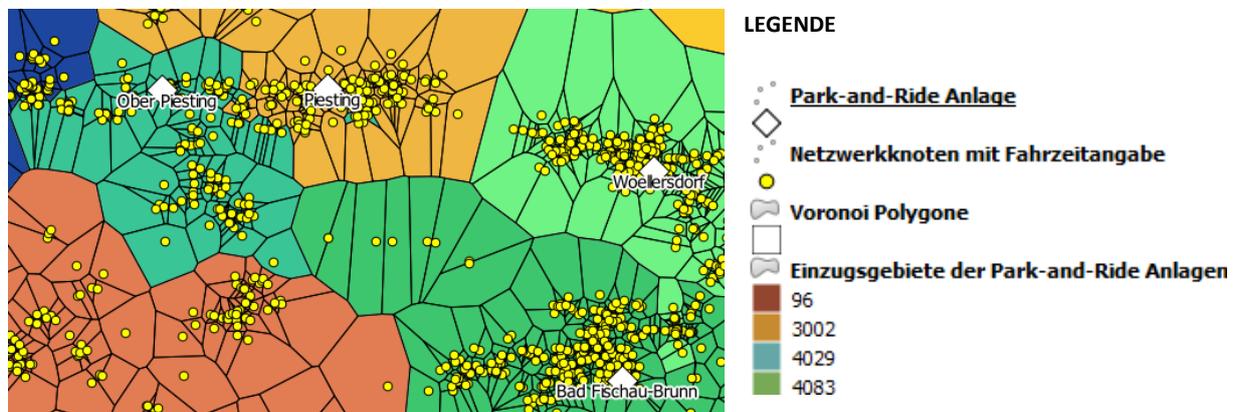
10.2.1 Voronoi-Polygone

Bei dieser nicht-statistischen Methode zur Schätzung räumlicher Zusammenhänge wird das gesamte Untersuchungsgebiet in diskrete, homogene Polygone aufgeteilt. Es handelt sich um eine Methode zur Darstellung diskreter Phänomene. Jedes Voronoi-Polygon besteht jeweils aus allen Punkten, die einem bekannten Messpunkt räumlich näher liegen als allen anderen bekannten Messpunkten und das gesamte Polygon erhält den Messwert dieses Messpunktes (vgl. Aurenhammer, 1991). Aufgrund der scharfen Abgrenzung der interpolierten Werte an der Polygongrenze, sind die Ergebnisse oft etwas zu grob, da sprunghafte Wertänderungen entstehen können, welche in der Natur nicht zu finden sind (Longley et. al. 2005, S. 333f). Je dichter die bekannten Messpunkte daher räumlich aneinander liegen, umso feingliedriger und genauer sind die Voronoi Polygone. Aufgrund der Engmaschigkeit des vorliegenden Straßengraphs bzw. seiner Netzwerkknotenpunkte bietet die Berechnung der Voronoi Polygone ausreichend detaillierte Ergebnisse.

10.2.2 Methodik

Die Voronoi Polygone werden auf Basis der Fahrzeit jedes Netzwerkknotens zur zeitlich nächstgelegenen Park-and-Ride Anlage berechnet. Die Umsetzung ist mit Hilfe der Vektorgeometriewerkzeuge von QGIS leicht und mit geringer Rechnerleistung möglich. Für jeden Knotenpunkt mit Fahrzeitangabe wird ein Voronoi Polygon berechnet, dessen Kanten durch Halbierung der euklidischen Distanz zu den nächsten Knotenpunkten mit Fahrzeitangabe ermittelt werden (Abbildung 27). Anschließend werden die einzelnen Voronoi Polygone über die zeitlich nächstgelegene Park-and-Ride Anlage zu einem Polygon zusammengefasst, welches das Einzugsgebiet der Park-and-Ride Anlage bildet.

Abbildung 27: Voronoi Polygone für die Netzwerkknoten mit Fahrzeitangabe zur zeitlich nächstgelegenen Park-and-Ride Anlage



Quelle: Eigene Darstellung basierend auf OpenStreetMap, osm2po, QGIS

10.2.3 Ergebnisse

Von den insgesamt 117 Park-and-Ride Anlagen verbleiben ohne die Einzugsbereiche auf Wiener Stadtgebiet 105 Anlagen mit Relevanz für das südliche Wiener Umland. Der Straßengraph ist für die Berechnung von Einzugsbereichen mit Voronoi-Polygonen sehr gut geeignet. Die Methode liefert plausible Ergebnisse, da sich systematische Fehler (z.B. durch nicht Berücksichtigung von Fahrzeitaufschlägen für Ampelwartezeiten, Wiedergabe der Fahrzeiten im unbelasteten Straßennetz), über das gesamte Untersuchungsgebiet weitgehend ausgleichen.

Insgesamt lässt die Berechnung der Voronoi-Polygone gute Schlüsse über die Verteilung der Park-and-Ride Anlagen, deren Bevölkerungsabdeckung sowie der Stellplatzverfügbarkeit zu.

Abbildung 28 zeigt die bereits nach Park-and-Ride Anlagen zusammengefassten Einzugsgebiete, Tabelle 29 listet Minimum-, Maximum und Mittelwert der Fläche, Einwohner und Stellplätze je Park-and-Ride Anlage auf. Durchschnittlich ist ein Einzugsgebiet 64 km² groß und umfasst 6.215 Einwohner. Die Größe der Einzugsgebiete variiert stark zwischen 5 und 341 km², abhängig von der Dichte der Park-and-Ride Anlagen. Je weiter man in die peripheren Lagen des Untersuchungsgebiets kommt, umso weiter sind die Park-and-Ride Anlagen voneinander entfernt und umso größer die Fläche der Einzugsgebiete und damit auch die Anzahl der Einwohner im Einzugsgebiet. Auch das Verhältnis Einwohner zu Stellplätzen ist in peripheren Lagen tendenziell höher und damit der Versorgungsgrad mit Stellplätzen schlechter, als im Zentralraum des Niederösterreichischen Industrieviertels oder im nördlichen Burgenland. Dies wird auch in Abbildung 29 deutlich, welche die Anzahl der Einwohner je Stellplatz illustriert.

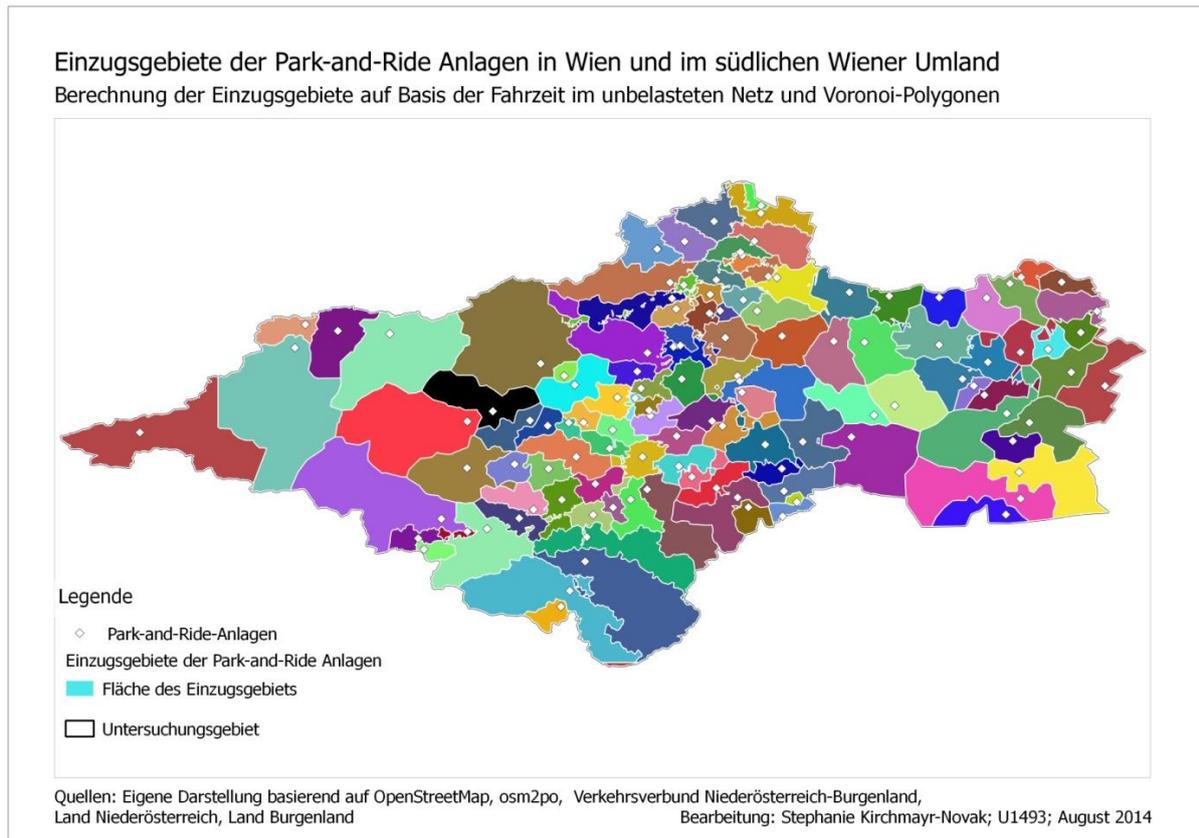
Tabelle 29: Fläche und Bevölkerung im Einzugsgebiet (EZG) der Park-and-Ride Anlagen (Voronoi Polygone)

Rang	Name der Park-and-Ride Anlage	Fläche des Voronoi Poygons in km ²	Einwohner im EZG (SE)*	Stellplätze je Park-and-Ride Anlage	Einwohner im EZG pro Stellplatz
.	Minimum	5	190	2	5
.	Maximum	341	31.998	780	3.740
.	Mittelwert	64	6.215	92	233
Park-and-Ride Anlagen mit der größten und kleinsten Fläche der Voronoi-Polygone					
1	Lilienfeld	341,0	5.802	26	223
2	Wienerbruck-Josefsberg	332,8	2.060	100	21
3	Weissenbach-Neuhaus	281,0	4.838	10	484
4	Payerbach-Reichenau	252,8	3.360	56	60
5	Edlitz-Grimmenstein	231,3	5.049	79	64
101	Ober Piesting	10,4	1.536	20	77
102	Achau	7,9	1.210	19	64
103	Eichberg	6,8	244	4	61
104	Wittmannsdorf NÖ	5,4	2.873	20	144
105	Baumgarten	4,8	307	57	5
Park-and-Ride Anlagen mit den meisten und wenigsten Einwohnern im Einzugsbereich (Einwohner auf Basis der Siedlungseinheiten)					
1	Baden	109,5	31.998	427	75
2	Brunn-Maria Enzersdorf	62,3	29.566	187	158
3	Wiener Neustadt Hbf	37,0	27.880	361	77
4	Wien Liesing	102,3	21.664	320	68
5	Moedling	23,5	21.242	780	27
101	Waldegg	23,5	850	17	50
102	Semmering	16,5	697	22	32
103	Baumgarten	4,8	307	57	5
104	Eichberg	6,7	244	4	61
105	Breitenstein	24,6	190	15	13

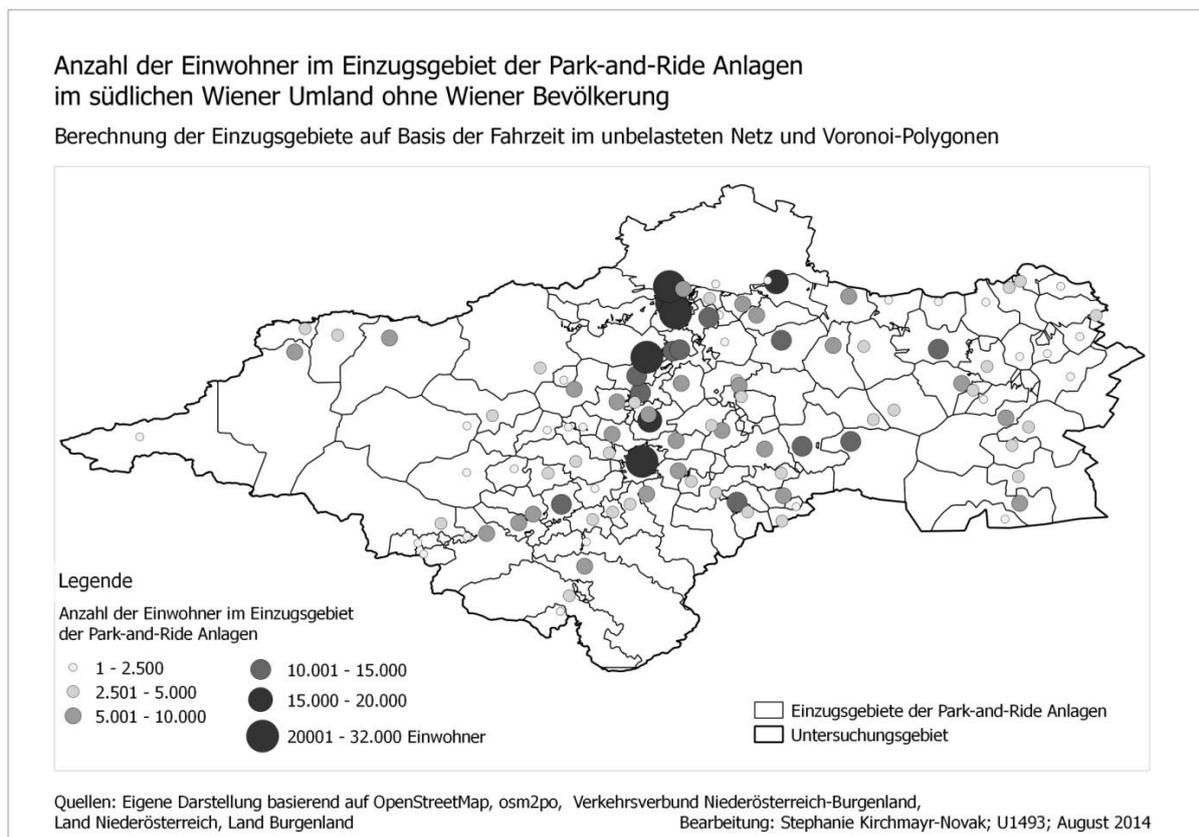
EZG.. Einzugsgebiet, SE.. Siedlungseinheiten

* Vereinfachte Berechnung der erschlossenen Bevölkerung anteilige Umlegung der Siedlungsfläche je Fahrzeitklasse

Quelle: eigene Berechnungen

Abbildung 28: Einzugsgebiete der Park-and-Ride Anlagen in Wien und im südlichen Wiener Umland¹⁰⁷⁵⁶

Quelle: Eigene Darstellung basierend auf OSM, osm2po, pgRouting und QGIS

Abbildung 29: Anzahl der Einwohner im Einzugsgebiet der Park-and-Ride Anlagen im südlichen Wiener Umland

Quelle: Eigene Darstellung basierend auf OSM, osm2po, pgRouting und QGIS

10.3 Fahrzeitisochronen basierend auf Interpolation mit "Regular Spline with tension" (RST)

Um einen besseren Eindruck über die Qualität des Straßengraphs und der daraus abgeleiteten Fahrzeiten zu den Park-and-Ride Anlagen zu bekommen, wird im nächsten Schritt die Fahrzeit zwischen den Straßenknoten flächig interpoliert und daraus Isochronen der Fahrzeit abgeleitet. Dazu wird das „Regular Spline with Tension“-Verfahren (RST) angewandt.

10.3.1 Interpolation mit „Regular Spline with tension“ (RST)

Bei den Spline-Verfahren handelt es sich um exakte geostatistische Methoden, die zur Gruppe der radialen Basisfunktionen gehören und von zwei Anforderungen ausgehen: die Interpolationsfunktion verläuft durch die Datenpunkte oder sehr nah an diesen, d.h. die Summe der Abweichungen von den Datenpunkten ist minimal und die Funktion so glatt wie möglich (Mitas & Mitasova, 1999). Die Verfahren ähneln dem Versuch, eine Platte oder Gummimatte durch alle Messpunkte zu biegen und dabei die Spannung der Platte (tension) zu variieren (Johnston et al. 2001, S. 126f).

Spline Verfahren werden verwendet, um glatte Oberflächen aus einer großen Anzahl von Datenpunkten zu berechnen. Sie sind jedoch ungeeignet, wenn die Messwerte auf kurzer horizontaler Distanz stark schwanken, da die Oberflächen aus gewichteten Durchschnittswerten der benachbarten Messpunkte ermittelt werden (Mitas & Mitasova, 1993, Johnston et al. 2001, S.126f).

Bei der RST-Interpolation sind zwei Parameter wesentlich für das Ergebnis. Der Parameter „Tension“ bestimmt die Spannung der Oberfläche. Je höher die Tension, desto geringer der Einfluss der einzelnen Messpunkte. Die Oberfläche ähnelt dann einer Membran, die bei Hoch- und Tiefpunkten schnell wieder zu einer Trendoberfläche zurückkehrt. Der Parameter „Smoothing“ kontrolliert die Glättung der Oberfläche. Mit diesem Parameter kann die Interpolation in eine Schätzfunktion umgewandelt werden, um Datensätze zu generalisieren oder lokale Störungen zu eliminieren. Darüber hinaus sind auch die Richtungsabhängigkeit („Anisotropy“) und die maximale und minimale Distanz zwischen den Messpunkten wichtige Einflussparameter (Hofierka et al. 2013).

10.3.2 Methodik

Interpolation der Fahrtzeit

Die Interpolation der Fahrtzeit wird mit dem GRASS GIS Modul „v.surf.rst“ durchgeführt. Der verwendete Algorithmus beruht auf regionaler quadtree Segmentierung (Mitasova 1992a, Mitasova 1992b, Mitasova et al., 1995) und kann große Datenmengen verarbeiten. Während der Segmentierung wird das gesamte Gebiet in mehrere überlappende Segmente aufgeteilt, danach werden die Oberflächenwerte je Segment interpoliert, bevor die endgültige Oberfläche berechnet wird.

Ein wesentlicher Vorteil der GRASS GIS Funktion ist, dass das zu interpolierende Gelände mit Hilfe einer Maske zugeschnitten werden kann. Dies ist aufgrund der unregelmäßigen Form des Untersuchungsgebiets ein Vorteil gegenüber anderen Methoden, die an den Rändern des Gebiets zu stark verzerrten Ergebnissen führen.

Die Funktion wird GRASS GIS durch vier Parameter kontrolliert (Neteler & Mitasova, 2008):

- dmin – der erlaubten Minimaldistanz zwischen den Datenmesspunkten, um fast identische Messpunkte zu entfernen (Standardwert 0,001);
- dmax – der erlaubten Maximaldistanz zwischen den Datenmesspunkten auf einer geraden Linie, um zusätzliche Schätzpunkte einzufügen (Standardwert 2,5);
- npmin der minimalen Anzahl von Punkten für die Schätzung in einem Segment (Standardwert 300);
- segmax – der maximalen Anzahl von Messpunkten in einem Interpolationssegment (Standardwert 40).

Für die Berechnung wurden die Standardeinstellungen und der voreingestellte Parameterwert von 40 für die Tension verwendet.

Berechnung der Fahrzeitisochronen

Aus der generierten Rasteroberfläche der Fahrzeit können im nächsten Schritt Isochronen als Konturlinien der Oberfläche extrahiert werden. Dafür werden Isochronen für Fahrzeiten in 2,5 Minuten- und 5 Minuten-Schritten gebildet. Abbildung 30 stellt die Ergebnisse für das Untersuchungsgebiet dar.

Berechnung der erschlossenen Siedlungsfläche und Bevölkerungszahl nach Fahrzeitisochrone

Zur Ermittlung der erschlossenen Siedlungsfläche und Bevölkerungszahl nach Fahrzeitisochronen wurden die Siedlungsgebiete mit den Fahrzeitisochronen verschnitten und die Bevölkerung in einer vereinfachten Berechnung anteilig zur Fläche den Fahrzeitisochronen zugewiesen. Tabelle 30 listet die erschlossene Fläche und Bevölkerung nach Fahrzeitklassen auf, die im Folgenden beschrieben werden.

10.3.3 Ergebnisse

Die gute Erreichbarkeit der Park-and-Ride Anlagen im Zentralraum des Niederösterreichischen Industrieviertels und des nördlichen Burgenlandes wird in Abbildung 30, welche die Einzugsgebiete der Park-and-Ride Anlagen mit interpolierten Fahrtzeiten und den daraus abgeleiteten Isochronen zeigt, abermals deutlich. Demgegenüber sind die peripheren und weniger besiedelten Gebiete im Bezirk Lilienfeld deutlich schlechter versorgt.

Insgesamt liegen rund 50% der Fläche und 54% der Bevölkerung im südlichen Wiener Umland eine Park-and-Ride Anlage innerhalb der 2,5 Minuten Fahrzeitisochrone (Tabelle 30). Weitere 23-24 % der Fläche bzw. der Bevölkerung können eine Park-and-Ride Anlage innerhalb von 2,5 bis 5 Minuten erreichen. Damit sind die Park-and-Ride Anlagen für drei Viertel der Fläche bzw. der Bevölkerung innerhalb von fünf Minuten mit dem Pkw erreichbar. Dieses Ergebnis zeugt von der sehr guten Erschließung der Region mit Bahnstationen.

Bei der Interpretation der Fahrzeiten muss jedoch beachtet werden, dass der Straßengraph die Fahrzeit im unbelasteten Netz wiedergibt und Verzögerungseffekte durch Stau in der Morgenspitzenstunde, Wartezeiten an Ampeln (siehe Kapitel 6.4 osm2po und Kapitel 9.4 Thematische Genauigkeit der Kantengeschwindigkeit), Parkplatzsuche am Bahnhof sowie Zu- und Abgangswege zum Bahnhof bzw. Parkplatz nicht berücksichtigt werden.

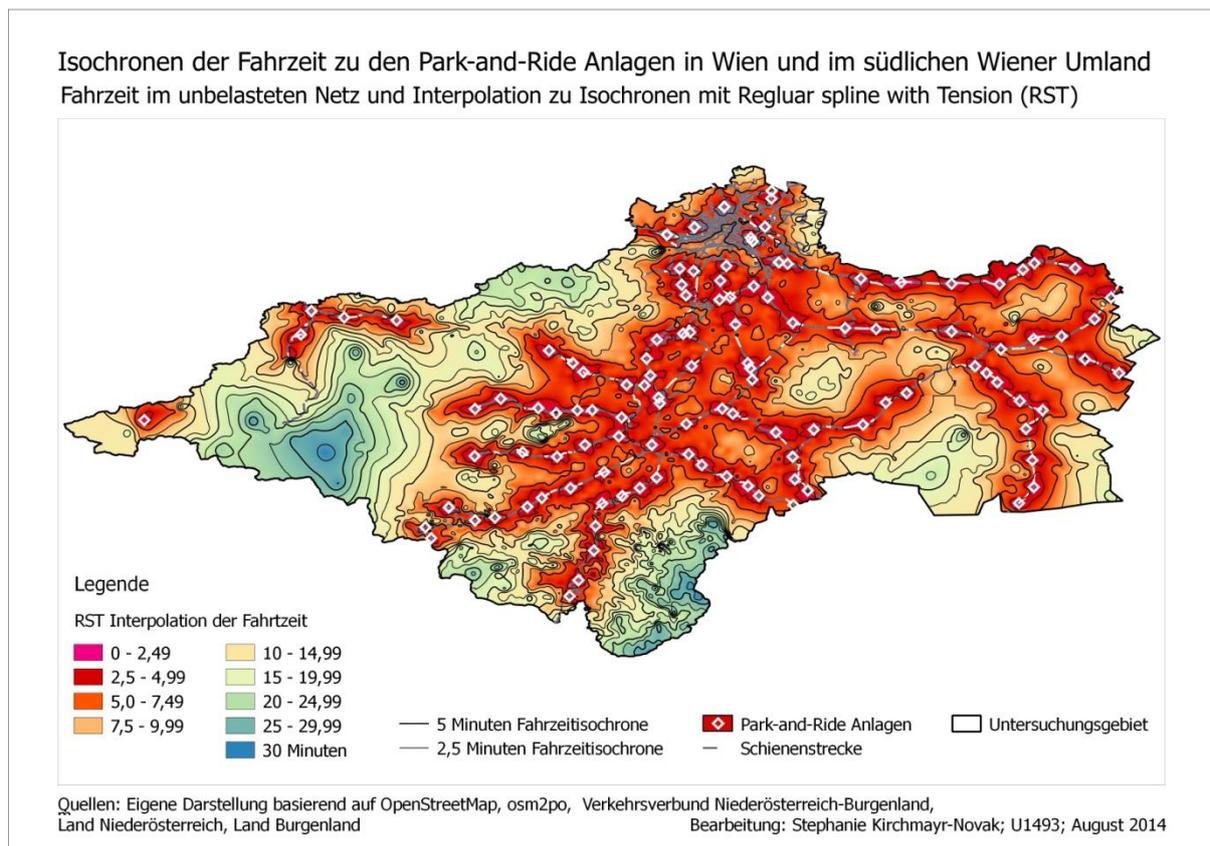
In Verkehrsmodellen sind diese Adaptionen des unbelasteten Verkehrsgraphs im Rahmen der Kalibrierung des Verkehrsgraphs ausschlaggebend für die Ergebnisgüte. Insbesondere die Berücksichtigung von Wartezeiten an Lichtsignalanlagen sollte umgesetzt werden um realitätsnähere Fahrzeiten und Fahrrouten zu erhalten. In der gegenständlichen Arbeit entfällt diese Art der Kalibrierung, da sie den Rahmen der Arbeit sprengen würde. Es ist jedoch darauf hinzuweisen, dass gerade die Berücksichtigung von Kapazitätsengpässen in der Morgenspitzenstunde der Pendlerverkehre das Ergebnis stark verändern könnte.

Tabelle 30: Durch Park-and-Ride Anlagen erschlossene Fläche und Bevölkerung nach Fahrzeitklassen

Fahrzeit bis zur Park-and-Ride Anlage	Insgesamt (inkl. Wien)			davon südliches Wiener Umland		
	Erschlossene Fläche in ha	Erschlossene Fläche in ha	Anteil in %	Erschlossene Bevölkerung 2010*	Erschlossene Bevölkerung 2010*	Anteil in %
0 - 2,49 Minuten	26.220	20.330	49,5	668.700	355.000	54,1
2,5 - 4,99 Minuten	23.260	9.710	23,6	882.600	150.000	22,9
5,0 - 7,49 Minuten	16.080	5.690	13,8	649.400	86.300	13,2
7,5 - 9,99 Minuten	2.580	1.750	4,3	73.700	29.000	4,4
10 - 14,99 Minuten	3.400	2.780	6,8	61.800	28.800	4,4
15 - 19,99 Minuten	730	670	1,6	8.400	5.400	0,8
20 - 24,99 Minuten	170	170	0,4	1.300	1.300	0,2
25 - 29,99 Minuten	0	0	0,0	0	0	0,0
Insgesamt	72.440	41.100	100,0	2.345.900	655.800	100,0

* Vereinfachte Berechnung der erschlossenen Bevölkerung anteilige Umlegung der Siedlungsfläche je Fahrzeitklasse
 Quelle: eigene Berechnungen

Abbildung 30: Einzugsgebiete der Park-and-Ride Anlagen in Wien und im südlichen Wiener Umland



Quelle: Eigene Darstellung basierend auf OSM und osm2po

11 Schlussfolgerungen und Ausblick

Die Ergebnisse der Arbeit zeigen, dass für die Konvertierung von OpenStreetMap-Daten in einen Straßengraph Open Source Software und die für ihre Anwendung erforderliche Dokumentation in ausreichender Qualität zur Verfügung stehen. Auch ist es möglich, mit Hilfe direkter intrinsischer Verfahren Aussagen über die Datenqualität des so erzeugten Straßengraphs abzuleiten. Der im Rahmen der Arbeit vorgestellte Evaluierungsprozess stößt jedoch dort an seine Grenzen, wo ein routenfähiger Referenzstraßengraph für die umfassende Evaluierung des vollständigen Straßengraphen erforderlich ist. Dies trifft auf die Bewertung der Lagegenauigkeit ebenso zu, wie auf die Bewertung der thematischen Genauigkeit der Geschwindigkeitsangaben, Abbiegerelationen und Einbahnregelungen. Letztere wurden auch einer extrinsischen visuellen Kontrolle ausgewählter Teilbereiche unterzogen. Diese diente in erster Linie dazu, sich mit den Stärken und Schwächen des Verkehrsgraphs vertraut zu machen. Ergänzend zu intrinsischen Bewertungsverfahren wurden die aus der Literatur bekannten extrinsischen Evaluierungsansätze vorgestellt. Insgesamt zeigt sich, dass für die vollständige Bestimmung der Datenqualität nutzergenerierter, freier Daten jedenfalls offizielle – freie oder kostenpflichtige Referenzdatensätze – erforderlich sind.

Generell zeigt sich, dass der auf Basis der OpenStreetMap erzeugte Straßenverkehrsgraph eine sehr gute Grundlage für Netzwerkanalysen ist, der die Verkehrswege in Wien und im südlichen Wiener Umland vollständig abzubilden in der Lage ist. Bevor der Straßenverkehrsgraph jedoch für anspruchsvollere Anwendungen herangezogen wird, wird empfohlen drei Eingangsparameter bei der Konvertierung der OpenStreetMap-Daten zu adaptieren. Zunächst sollte die Standardgeschwindigkeit für Autobahnauffahrten in der Konvertierungsdatei von osm2po von 30 km/h auf 50 km/h angehoben werden. Darüber hinaus wird empfohlen, die Fahrzeit um eine durchschnittliche Wartezeit an Lichtsignalanlagen zu erhöhen, die im ersten Schritt mit einer Minute pro Lichtsignalanlage festgelegt werden kann. Schließlich sollte vor jeder Konvertierung anhand des Datensatzes und der Anforderungen des Anwendungsfalls überprüft werden, ob die Aufnahme von Straßen „einfachsten Ausbaustand“, in der Regel asphaltierte Wirtschaftswege, erforderlich ist oder entfallen kann.

Die Anwendung von nutzergenerierter Geoinformation (Volunteered Geographic Information) und Open Source Software erfordert die umfassende Befassung mit den jeweiligen Datengrundlagen bzw. den Spezifika der Open Source Software und deren laufenden Adaptationen. Der Aufwand alle relevanten Informationen für jeweils aktuellen Bearbeitungsstand der Software zusammenzutragen, sollte dabei nicht unterschätzt werden, da über eine größere Anzahl an Quellen (Foren, Mailinglisten, Anwenderblogs, etc.) verteilt sind, als bei kommerziellen Produkten. Ein wesentliches Element der Open Source Software, die user community – also die Anwender und Entwickler der Software – ist im Falle der benutzten Werkzeuge sehr aktiv. Für alle aufgetretenen Fragen konnten Funktionserläuterungen oder Anwendungsbeschreibungen gefunden werden. Auch persönliche Anwenderfragen in den Foren wurden innerhalb weniger Tage von den Programmierern bzw. anderen Anwendern beantwortet. Diese Unterstützung der „Nutzergemeinschaft“ erhöht den Wert der freien Software substantiell.

Wie eingangs erwähnt, sollte der erstellte und geprüfte Straßenverkehrsgraph vor weiteren, anspruchsvolleren Routinganwendungen, z.B. Berechnungen von Erreichbarkeitspotenzialen, noch adaptiert werden. Um die Verkehrsflüsse ausreichend genau reproduzieren zu können, sollten die für die Region repräsentativen Aufschläge für Ampelwartezeiten, Zugangszeiten zum jeweiligen Bahnhof bzw. Stellplatzsuchzeiten genauso berücksichtigt werden, wie eine stärkere räumliche Differenzierung der Standardgeschwindigkeiten bei mangelnder Information über zulässige Höchstgeschwindigkeiten. Gerade bei der Modellierung von Pendlerverkehren sind darüber hinaus auch die Kapazitätsbeschränkungen im Straßennetz, gerade während der Morgenspitzenstunde als Zeitaufschläge auf die Kantengeschwindigkeit im unbelasteten Straßennetz zu berücksichtigen.

12 Literatur

- AURENHAMMER, F. (1991): Voronoi Diagrams – A Survey of a Fundamental Geometric Data Structure. Institute für Informationsverarbeitung Technische Universität Graz. www.cimec.org.ar/~ncalvo/aurenhammer-voronoi.pdf (Zugriff: 12.11.2014).
- BARRON C., NEIS P., ZIPF, A. (2013): iOSMAnalyzer - Ein Werkzeug für intrinsische OSM Qualitätsuntersuchungen. In: Strobel, J. et al. (Hg.) (2013): Angewandte Geoinformatik 2013. Berlin/Offenbach: Herbert Wichmann Verlag, VDE Verlag GmbH.
- BRASSEL K., BUCHER F., STEPHAN E., VCKOVSKI A. (1995): Completeness. In Guptill S C, Morrison J L (eds) Elements of spatial data quality. Oxford, Elsevier Science: 81–108.
- BRINKHOFF, T. (2005): Geodatenbanksysteme in Theorie und Praxis. Einführung in objektrelationale Geodatenbanken unter besonderer Berücksichtigung von Oracle Spatial. Herbert Wichmann Verlag, Heidelberg.
- CAPRIOLI, M., SCOGNAMIGLIO, A., STRISCIUGLIO, G. and TARANTINO, E. (2003): Rules and Standards for spatial data quality in GIS environments. 21st International Cartographic Conference. Cartographic Renaissance. ICC 10 - 16 August 2003. http://icaci.org/files/documents/ICC_proceedings/ICC2003/Papers/222.pdf (Zugriff: 14.11.2014).
- CHEN, H. (2011): Entwicklung von Verfahren zur Beurteilung und Verbesserung der Qualität von Navigationsdaten. Dissertation. Universität Stuttgart, Institut für Photogrammetrie.
- CHRISMAN, N. (1991): The error component in spatial data. In: Maguire, D.J., Goodchild, M.F and Rhind, D.W. (eds.), Geographical Information Systems: Principles and Applications. Volume 1, p. 165-174.
- DIN PAS 1071 (2007): Modell für die Beschreibung der Qualität von Geodatenprodukten. Deutscher Dachverband für Geoinformationen, Beuth Verlag, Berlin.
- EEA European Environment Agency (2010): Corine Land Cover 2006 seamless vector data. European EEA standard re-use policy: unless otherwise indicated, re-use of content on the EEA website for commercial or non-commercial purposes is permitted free of charge, provided that the source is acknowledged (<http://www.eea.europa.eu/legal/copyright>). Copyright holder: European Environment Agency (EEA). <http://www.eea.europa.eu/data-and-maps/data/clc-2006-vector-data-version#tab-metadata> (Zugriff: 14.11.2014).
- ELMASRI R.; NAVATHE, S. (2002): Grundlagen von Datenbanksystemen. 3., überarb. Aufl. . - München : Pearson Studium, 2002.
- FORSCHUNGSGESELLSCHAFT MOBILITÄT (2011): Projekt VeGIS Anforderungen an einen Verkehrsweegegraph für die Verwendung in Verkehrsmodellen. Version 1.3, (Zugriff: 14.03.2011)
- FREE SOFTWARE FOUNDATION (2008). What is free software? <http://www.fsf.org/about/what-is-free-software> (Zugriff: 14.12.2014).
- GOODCHILD, M. (2007): Citizens as sensors: the world of volunteered geography. In: GeoJournal (2007) 69:211–221
- GOODCHILD, M., HUNTER, G. (1997): A simple positional accuracy measure for linear features. International Journal of Geographical Information Science 11, 299-306.
- HUNTER, G. (1999): New Tools For Handling Spatial Data Quality: Moving from Academic Concepts to Practical Reality. URISA Journal Vol., 11, No. 2. Summer 1999
- GRASER, A. (2011): Drive time isochrones – an example using Finnish Airports. February 12, 2011. <http://anitagraser.com/2011/02/12/drive-time-isochrones/> (Zugriff: 12.11.2014)
- GRASER, A., STRAUB, M., DRAGASCHNIG, M. (2013): Towards an open source analysis toolbox for street network comparison: Indicators, tools and results of a comparison of OSM and the official Austrian Reference Graph. Under review for TGIS special issue FOSS4G2013.
- GUPTILL, S. C., MORRISON J. L., EDS. (1995): Elements of Spatial Data Quality (The International Geographic Association). Elsevier, Oxford, 250 S.

- HAKLEY, M. (2010): How good is volunteered geographical information? A comparative study of Openstreetmap and ordnance survey datasets. *Environment & Planning B: Planning and Design* 37, 682-703
- HOFIERKA, J., GALLAY, M., KANUK, J. (2013): Spatial Interpolation of Airborne Laser Scanning Data with Variable Data Density. In: Buchroithner et al. (2013): *Proceedings of the 26th International Cartographic Conference*. Dresden, Germany, 25–30 August 2013
- HUNTER, G. (1999): New tools for handling spatial data quality: moving from academic concepts to practical reality. *URISA Journal* 11, 25-34.
- ISO/TC 211 N3086 (2011): Draft text for DIS, final ISO/CD 19157, Geographic information - Data quality. ISO/TC 211/WG 9/19157 Editing Committee. <http://www.isotc211.org/protdoc/211n3086/>
- JERMANN, J. (2004): GIS-basiertes Konzept zur Modellierung von Einzugsbereichen auf Bahn-Haltestellen. Dissertation. ETZ Zürich, November 2004
- JOHNSTON, K., VER HOEF, J., KRIVORUCHKO, K., LUCAS, N. (2001): *Using ArcGIS Geostatistical Analyst*. Environmental Systems Research Institute, 2001 Redlands , CA.
- JOOS, G. (2000): Zur Qualität von objektstrukturierten Geodaten. Schriftenreihe des Studienganges Geodäsie und Geoinformation der Universität der Bundeswehr Münschen. Heft 66/2000.
- KLAMPFERER, D. (2014): Open Government Data. Motive und Barrieren der öffentlichen Verwaltung. Masterarbeit zur Erlangung des akademischen Grades Master of Arts Business. FH Oberösterreich. Linz, 2014.
- KRAMPE, S. (2009): Entwicklung eines Lkw-Empfehlungsnetzes für die Region Frankfurt RheinMain auf Basis von OpenStreetMap. In: Strobl, J., Blaschke, T.; Griesebner, G.; (Hrsg.): *Angewandte Geoinformatik 2013*. Beiträge zum 21. AGIT-Symposium Salzburg.
- LANGE, N. (2013): *Geoinformatik in Theorie und Praxis*. 3. Auflage. Springer Spektrum. Springer Verlag Berlin Heidelberg, 2002, 2006, 2013
- LONGLEY P., GOODCHILD, M, MAGUIRE, D., RHIND, D. (2005): *Geographic Information Systems and Science*. Second Edition. John Wiley and Sons Ltd.
- MAGISTRAT DER STADT WIEN (2013): *Masterplan Verkehr Wien 2003, Evaluierung 2013*.
- MAGISTRAT DER STADT WIEN (2014a): *Smart City Rahmenstrategie*.
- MAGISTRAT DER STADT WIEN (2014b): *Stadtentwicklungsplan STEP 2025*
- MAX PLANK GESELLSCHAFT (2003): Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen vom 22. Oktober 2003. <http://openaccess.mpg.de/Berliner-Erklaerung> (14.11.2014).
- MITAS L., MITASOVA H. (1999) Spatial Interpolation. In: P.Longley, M.F. Goodchild, D.J. Maguire, D.W.Rhind (Ed s.), *Geographical Information Systems: Principles, Techniques, Management and Applications*, Geoinformation International, Wiley: 481-492
- MITASOVA, H., MITAS, L. (1993) Interpolation by Regularized Spline with Tension: I. Theory and implementation. *Mathematical Geology*, 25, 641-655.
- MÖLLER, C. (2012): *osm2po*. <http://www.osm2po.de/> (Zugriff: 15.10.2013)
- MÖLLER, C. (2013a): How Does Directed and has_reverse_cost Arguments when set to [true, false] and [true, true] Affect Routing? <http://gis.stackexchange.com/questions/48071/how-does-directed-and-has-reverse-cost-arguments-when-set-to-true-false-and> (Zugriff: 15.01.2013)
- NEIS, P. (2014): *The OSM Contributor Activity Report – Edition 2014*. <http://neis-one.org/2014/08/osm-activity-2014/> (Zugriff: 14.11.2014)
- NEIS, P., ZIELSTRA, D. & ZIPF, A. (2011): The Street Network Evolution of Crowdsourced Maps: OpenStreetMap in Germany 2007–2011. In: *Future Internet*, 4 (1), 1-21.
- NETELER M., MITASOVA, H. (2008) *Open Source GIS: A GRASS GIS Approach*. Springer Verlag, New York

- O'REILLY, T. (2005): What is web 2.0: Design patterns and business models for the next generation of software, O'Reilly blog, oreilly.com/web2/archive/what-is-web-20.html (Zugriff: 13.02.2013).
- OGC (2010): OpenGIS Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option. Version: 1.2.1. Status: Corregendum http://portal.opengeospatial.org/files/?artifact_id=25354 (Zugriff: 14.11.2014)
- OPEN KNOWLEDGE FOUNDATION (2014): Conformant Licences. <http://opendefinition.org/licenses/> (Zugriff: 14.11.2014)
- POSTGIS 1.4.2 Manual <http://postgis.refractor.net/documentation/manual-1.4/index.html> (Zugriff: 19.07.2014).
- POSTGIS PROJECT STEERING COMMITTEE (o. D) PostGIS 1.4 Manual. SVN Revision (11337). <http://postgis.net/stuff/postgis-1.4.pdf> (Zugriff: 14.11.2014).
- RAMM, F., TOPF, J. (2010): OpenStreetMap. Die freie Weltkarte nutzen und mitgestalten. 3. Auflage. Lehmanns Media, Berlin.
- RAYMOND, E.S. (2002): The cathedral and the bazaar, version 3.0. <http://www.catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar/index.html> (Zugriff: 13.02.13).
- REHRL, K., Edlinger K.-M., Friedwagner, A., Hahn, B., Langthaler, T., Wagner, A., Wimmer, M. (2012): Evaluierung von Verkehrsgraphen für die Berechnung von länderübergreifenden Erreichbarkeitspotenzialen am Beispiel von OpenStreetMap. In: Strobl, J., Blaschke, T. & Griesebner, G. (Hrsg.) (2012): Angewandte Geoinformatik 2012. ISBN 978-3-87907-520-1.
- REY, S.J. (2008): Show me the code: Spatial analysis and open source. School of Geographical Sciences, Arizona State University. Online at <http://mpra.ub.uni-muenchen.de/9260/> MPRA Paper No. 9260, posted 24. June 2008 01:29 UTC
- RITTLER C. (2011): Kordonenerhebung Wien in den Jahren 2008–2010 i.A. d. Planungsgemeinschaft Ost.
- SCHOLZE-STUBENRECHT, W. (Red.), Wermke, M. (Hrsg.) (1997): Der Duden in zwölf Bänden : das Standardwerk zur deutschen Sprache. 5. Duden - Fremdwörterbuch. Mannheim, Wien (u.a.). Dudenverlag 1997
- TAKUBO A., PATRUSHEV, A., HOSSAIN, A., GONZALEZ, C., KASTL, D., POTTS, D., TECHER, D., MIYAWAKI, E., THURKOW, F., JUNOD, R., FENOY, G., MAHADEOKAR, J., LENG, J., BEHNCKE, K., KUMAR, K., NAGASE, K., BASA, M., WIESENHAAN, M., ISLAM, R., WOODBRIDGE, S., HOUSSEMAN, S., PASCHKE S. (2013): pgRouting Manual, Release 2.0.0.-rc1 (bf13fd7 v2.0.0.-rd1) <http://pgrouting.org/> (Zugriff: 15.10.2013)
- THALLER, D. (2009): Die Open-Source-Plattform "OpenStreetMap", eine Konkurrenz für Geodatenhersteller? Gezeigt im Vergleich mit Geodaten der Firma TeleAtlas, Mastersthesis, Universität Wien.
- VAN OORT, P. (2005): Spatial data quality: from description to application. Publications on Geodesy 60. ISBN 90 6132 295 2. ISSN 0165 1706
- VEREGIN (1999): Data quality parameters. In: Longley P, Goodchild M, D Maguire, D Rhind (1999): Geographical Information Systems. Volume 1 Principles and Technical Issues. Second Edition. John Wiley & Sons, Inc.
- WITSCHKO, T., KAUFMANN, T. (2005): Modellierung und Bewertung der Datenqualität in Prozessen der Geodatenverarbeitung. In: Strobl, J., Blaschke, T. & Griesebner, G. (Hrsg.) (2005): Angewandte Geoinformatik 2005. Beiträge zum AGIT-Symposium 2005, Salzburg 2005, S. 790 - 799.
- ZIMMERMAN, E. (2010): OpenStreetMap in ArcGIS: Automatisierte Datenaufbereitung für Netzwerkanalysen. Bachelor-Thesis an der Hochschule Karlsruhe - Technik und Wirtschaft, Fakultät für Geomatik. Durchgeführt bei ESRI Deutschland GmbH.