



Universität Trier

Fachbereich VI

Angewandte Geoinformatik

Bachelorarbeit

Kostenberechnung basierend auf einer Netzwerkanalyse, zur Auswertung potenzieller Kindergarten-Standorte der Stadt Trier.

Erstgutachter: Univ.-Prof. Dr. Thomas Udelhoven

Zweitgutachter: Dr. Johannes Stoffels

vorgelegt von: Juliane Franziska Dengg

Schöndorfer Str. 35, 54292 Trier

Matr.Nr.: 1093617, s6judeng@uni-trier.de

Trier, September 2015

Inhaltsverzeichnis

1	Einleitung	3
1.1	Zielsetzung	3
1.2	Aufbau	3
I	Theoretischer Teil	5
2	Datenmaterial	5
2.1	Geodaten	5
2.1.1	Kindergarten-Standorte	6
2.1.2	Demographische Daten	6
2.1.3	Netzwerkdaten	6
2.2	Übersicht	8
3	Methode	9
3.1	Kostenberechnung	9
3.2	Netzwerkanalyse	9
3.2.1	Graphentheorie	9
3.2.2	Netzwerke	11
3.3	ArcGIS-Network-Analyst	12
3.3.1	Grundlagen	12
3.3.2	Network-Dataset	12
3.4	Start-Ziel-Kostenmatrix	16
3.4.1	Grundlagen	16
3.4.2	Dijkstra-Algorithmus	16
3.5	Network-Dataset aus OSM-Daten	17
3.5.1	OSM2NetworkDataset	18
3.5.2	ArcGIS Editor for OSM	19
II	Praktischer Teil	20
4	Anwendung	20
4.1	Datenaufbereitung	20
4.1.1	Demographische Daten	20
4.1.2	Generierung des Network-Dataset mit OSM2NetworkDataset	20

4.2	Berechnung der Start-Ziel-Kostenmatrix im Network-Analyst	20
4.3	Datenverknüpfung	23
4.4	Zuordnung der Haushalte unter Berücksichtigung der Demographie	24
4.4.1	Kandidaten	25
5	Ergebnisse und Auswertung	29
5.1	Zuordnung	29
5.1.1	Auswahl der Kindergärten	29
5.1.2	Kostenfunktion	29
5.2	Identifizierung des besten Kandidaten	31
6	Diskussion	35
6.1	Bewertung des Verfahrens	35
6.1.1	Netzwerkanalyse mit OSM-Daten	35
6.1.2	Zuordnung	36
6.2	Optimierungsmöglichkeiten	36
6.2.1	Datengrundlage mit höherer Auflösung	36
6.2.2	Iteratives Verfahren zur 2D-Standortplanung	36
7	Ausblick und Fazit	38
8	Abbildungs-, Tabellen- und Literaturverzeichnis	39
9	Eidesstattliche Erklärung	43

1 Einleitung

Wo sind geeignete Standorte für Kindergärten in Trier und wo überschneiden sich Einzugsgebiete? Wie verringern sich die Wegkosten für die einzelnen Haushalte mit einem neuen Kindergarten-Standort bzw. wie erhöhen sich die Kosten, wenn ein Standort wegfällt? Ausgehend von diesen Fragen entwickelte sich im Zeitraum vom 01.07.2015 bis 30.09.2015 die vorliegende Bachelorarbeit im Fachbereich Angewandte Geoinformatik der Universität Trier, unter der Betreuung von Herrn Prof. Udelhoven, als Erstgutachter, und Herrn Dr. Stoffels, als Zweitgutachter.

1.1 Zielsetzung

Ziel dieser Arbeit ist ein Verfahren darzulegen, mit dem die Änderung eines Zustandes durch Erschließung neuer Standorte oder dem Wegfallen bestehender Standorte erfasst und auf Effizienz überprüft werden kann. Mit dieser Methode soll unter potenziellen Standorten den Geeignetsten zu identifizieren.

Dieses Verfahren soll auf folgenden Anwendungsfall angepasst werden:

- Im ersten Schritt sollen basierend auf einer Netzwerkanalyse geeignete Standort-Kandidaten der Stadt Trier identifiziert werden
- Unter den Kandidaten soll der Standort ermittelt werden, welcher am effizientesten die Wegkosten aller Haushalte minimiert.
- In der Untersuchung sollen demographische Informationen über die Anteile der Kinder in den Haushalten und die Aufnahme-Kapazitäten der Kindergärten berücksichtigt werden.

Die Intention hinter dem Verfahren ist, eine Hilfestellung zur Standortwahl unter Berücksichtigung standortabhängiger Daten.

Die gewonnenen Erkenntnisse in den Ergebnisdaten sollen auch für Menschen ohne Fachkenntnisse, durch kartographische Visualisierung anschaulich und leicht verständlich dargestellt werden. Weiter sollen Optimierungsmöglichkeiten für das umgesetzte Verfahren festgestellt und kommuniziert werden.

1.2 Aufbau

Die vorliegende Arbeit teilt sich in einen theoretischen und einen praktischen Teil. Der theoretische Teil informiert über das Datenmaterial und die Methodik zur Lösung der Fra-

gestellungen. Das zweite Kapitel gibt eine Übersicht über die verwendeten Daten und deren Eigenschaften. Im nächsten Kapitel folgt eine Zusammenstellung von Methoden, die zur Findung eines Lösungsweges zur genannten Fragestellung angewendet werden könnten. Hier werden grundlegende Kenntnisse zum Verfahren der Netzwerkanalyse, basierend auf der Graphentheorie und der Struktur eines Netzwerks, vermittelt. Im weiteren Kapitel wird die Funktionsweise des Network Analysts von Esri beschrieben und das hierbei verwendete Network-Dataset thematisiert. Danach wird die Netzwerkanalyse zur Berechnung von Wegkosten vorgestellt. Im praktischen Teil wird die Umsetzung der Methoden erläutert. Die daraus resultierenden Ergebnisse und deren kartographische Visualisierung, sowie deren Auswertung werden im anschließenden Kapitel aufgeführt. Es folgt eine Diskussion zum angewendeten Verfahren und zu Optimierungsansätzen. Abschließend folgt ein Fazit mit Ausblick.

Teil I

Theoretischer Teil

2 Datenmaterial

Potenzielle Standorte neuer Kindergärten in der Stadt Trier werden in dieser Arbeit hinsichtlich ihrer räumlichen Auswirkung auf die demographische Struktur und auf die bestehenden Kindergärten analysiert. Im ersten Schritt einer Standortanalyse muss beurteilt werden, mit welchen Daten die zu untersuchenden Standorte in Verbindung stehen. Die „Kunden“ bzw. die Zielgruppe von Kindergärten sind Kinder im Alter zwischen drei und sechs Jahren (KiTa.de, o.J.). Diese Analyse bezieht sich auf die Erreichbarkeit der Kindergärten und soll bei der Erstellung einer Kostenfunktion die Anzahl der Kinder je Haushalt und die Aufnahme-Kapazität der Kindergärten berücksichtigen.

Aus diesen Kriterien ergibt sich das benötigte Datenmaterial:

- Die Kindergarten-Standorte der Stadt Trier,
- potenzielle Standorte für einen neuen Kindergarten,
- die räumliche Verteilung der Haushalte im Untersuchungsgebiet mit demographischer Information zur Anzahl der Kinder zwischen drei und sechs Jahren
- und ein Netzwerk zur Erreichbarkeitsanalyse.

Das Untersuchungsgebiet bezieht sich auf die Stadt Trier in Rheinland-Pfalz. Da eine räumliche Analyse durchgeführt werden soll, werden Daten mit Raumbezug benötigt.

2.1 Geodaten

Als Geodaten wurden frei verfügbare Daten aus dem OpenStreetMap-Projekt (OpenStreetMap, o.J.) benutzt, die über die Geofabrik (Geofabrik, 2014) bezogen wurden. Das OpenStreet-Map (OSM) Projekt wurde 2004 mit dem Ziel gegründet, eine freie Weltkarte zu erzeugen und die geographischen Daten lizenzkostenfrei bereitzustellen. Die Geofabrik ist ein freier Download-Server, welcher OSM-Daten in verschiedenen Kompressionsformaten zur Verfügung stellt. Die Vektordaten, die in Form von Shapefiles angeboten werden, sind mit Esri kompatibel und können direkt mit ArcGIS oder einem anderen GIS verarbeitet werden. OSM-Daten basieren auf dem Referenzsystem *World Geodetic System*

1984 (WGS 84) (Wiki, 2015a). Für diese Arbeit wurde mit der ZIP-Datei *rheinland-pfalz-latest.shp.zip* ein Ausschnitt aus den OSM-Daten zu Rheinland-Pfalz heruntergeladen.

2.1.1 Kindergarten-Standorte

Die räumlichen Daten zu den einzelnen Kindergärten der Stadt Trier wurden aus einem OSM-Datensatz zu Gebäuden extrahiert. Die Daten liegen als Vektorgrafik (.shp) vor und beinhalten Attribute zu Kindergartennamen, Gebäudetyp und einer OSM-ID. Da im Rohmaterial keine Information zur Kapazität der einzelnen Kindergärten vorliegt, mussten diese recherchiert und als weiteres Attribut hinzugefügt werden. Auch sind nicht alle Gebäude als Kindergarten gekennzeichnet, sodass noch weitere Standorte mit Hilfe von Google Maps (Google, o.J.) recherchiert wurden. Insgesamt gingen 32 Kindergarten-Standorte als Ausgangspunkte in die Netzwerkanalyse mit ein (siehe Abbildung 1).

2.1.2 Demographische Daten

Haushalte wurden mit dem Shapefile zu Gebäuden aus den OSM-Daten dargestellt. Da die Gebäude nicht nur Wohn-, sondern auch Dienstleistungs-, Industrie-, Kirchen- und andere Gebäudetypen enthalten, mussten die relevanten Typen selektiert und extrahiert werden. In den Attributen zu den Gebäuden fehlen demographische Informationen. Benötigt wurde die absolute Anzahl der Kinder je Haushalte. Die zur Verfügung stehenden demographischen Daten waren jedoch weniger hoch aufgelöst und lagen nur für die einzelnen Stadtbezirke vor. Durch den Anteil der Kinder je Haushalt in jedem Stadtbezirk musste der Idealfall approximiert werden (siehe 4.1.1).

2.1.3 Netzwerkdaten

Das Zitat "Your network will only be as good as the underlying data" (White, 2013) beschreibt, wie wichtig die Datengrundlage für ein gut funktionierendes Netzwerk ist. Das Netzwerk in dieser Arbeit basiert auf OSM-Daten im OSM-XML-Format. Der Detaillierungsgrad dieser Daten kann regional variieren. Städtische Bereiche sind dabei sehr gut erfasst, während periphere Gebiete sich in der Vergangenheit als lückenhaft erwiesen haben (OpenStreetMap, o.J.), (Ramm, 2009, S.3).

Die OSM-XML-Datei (.osm) nutzt ein XML-Format und basiert auf einer Liste von Instanzen aus den Grundelementen (data primitives) *nodes*, *ways* und *relations*, sowie deren *tags* (Wikib, 2015b). *Nodes* definieren Punkte im Raum, *ways* stellen Linien-Features und Polygongrenzen dar und *relations* werden benutzt, um topologische Beziehungen der Ele-

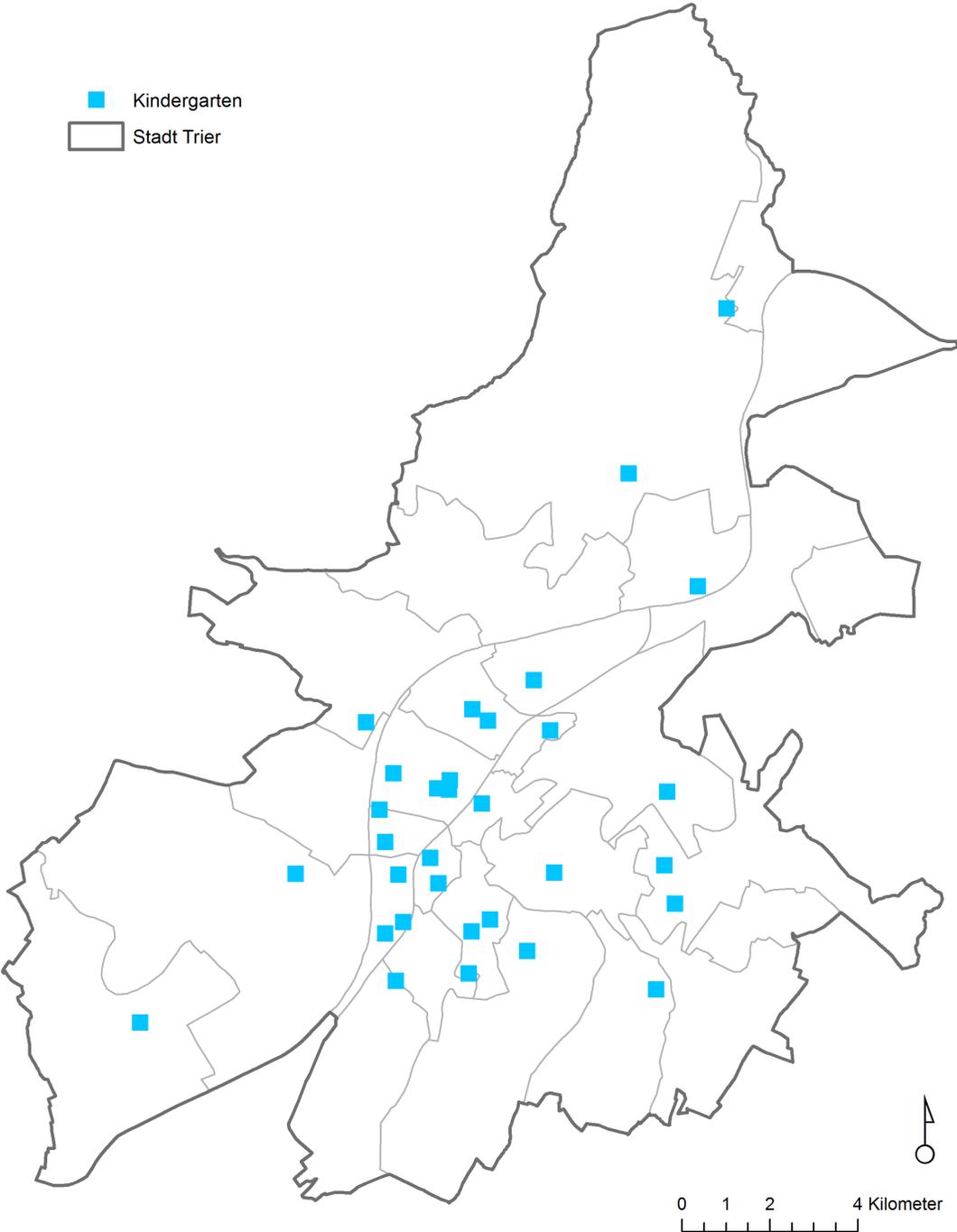


Abbildung 1: Verwendete Kindergarten-Standorte der Stadt Trier

mente untereinander zu beschreiben. Mit *tags* werden die Eigenschaften dieser Elemente beschrieben. Ein Tag besteht aus einem Key-Value-Paar, wobei sowohl der Schlüssel, als auch der zugehörige Wert als Freitext angegeben werden können. Das Beispiel "highway = residential" definiert ein Way-Element, das als Straße ("highway") definiert ist und als Ortsstraße ("residential") funktioniert bzw. als solche klassifiziert wurde. Diese Key-Value-Paare werden als *Map Features* bezeichnet und definieren in einem Netzwerk Verkehrswege, Verkehrsbeschränkungen oder Barrieren (Wikic, 2015c).

Die OSM-Rohdaten können nicht direkt als Netzwerk in einem GIS eingesetzt werden, sondern die Daten aufbereitet werden. Um diese beispielsweise in einer Netzwerkanalyse von Esri zu verwenden, müssen diese in ein kompatibles s.g. *Network-Dataset* Format konvertiert werden. (siehe 3.5). Eine Konfigurationsdatei im XML-Format, die Parameter zur Definition eines Netzwerkes enthält, wird im Konvertierungsprozess benötigt. Diese wurde aus einem unter 52north.org (52north, 2015) erreichbaren Archiv bezogen.

2.2 Übersicht

Die folgende Tabelle gibt eine Übersicht über die verwendeten Eingangsdaten.

Tabelle 1 Die verwendeten OSM-Daten

Dateiname	Datentyp	Dateigröße (KB)	Raumbezug	Quelle
Gebäude (Kindergärten/Haushalte)	.shp	111.458,0	GCS WGS 1984	Geofabrik
Stadtbezirke	.shp	102,0	GCS WGS 1984	Universität Trier
OSM-XML-Datei	.osm	180.000,0	GCS WGS 1984	Geofabrik

3 Methode

3.1 Kostenberechnung

DIE SUCHE NACH EINEM BESTEN STANDORT FÜR EIN GEPLANTES OBJEKT, KANN AUF DER BEDINGUNG BASIEREN, DIE ABSOLUTEN WEGKOSTEN DER POTENZIELLEN INTERESSENTEN ZUR GEPLANTEN ANLAGE ZU MINIMIEREN (Bill, 1999, S.103).

Die hier zitierte Bedingung bildet die Grundlage für die Untersuchungen in dieser Arbeit. Der untersuchte Fall gestaltet sich jedoch komplexer als die Suche der absoluten Wegkosten zu einer einzelnen Anlage, da neben der vorhandenen Verkehrsinfrastruktur, weitere standortspezifische Faktoren mit berücksichtigt werden müssen. Die Ermittlung des besten Standorts geschieht hier basierend auf der Minimierung der Wegkosten aller Interessenten zur ihnen jeweils zugeordneten Anlage. Zur Bestimmung der gesamten Wegkosten für alle Haushalte muss jedem Haushalt ein Kindergarten zugeordnet werden. Die Zuordnung soll unter Berücksichtigung der Kapazitäten der Kindergärten stattfinden und eine möglichst geringe Fahrzeit der Haushalte gewährleisten. Wegkosten lassen sich mit Netzwerkanalysen, basierend auf einem Netzwerk berechnen.

3.2 Netzwerkanalyse

Netzwerkanalysen spielen in der Geoinformatik eine wichtige Rolle und werden insbesondere zur Standortbewertung und im Transportwesen eingesetzt. Die Basis bilden hierbei Abstraktionen von real existierenden Netzwerken aus unterschiedlichen Anwendungsgebieten, z.B. für Verkehrssysteme, wie Straßen- oder Schienennetze, Leitungsnetze oder Telekommunikationsnetze (Lange, 2013, S.37). Die Netzwerkanalyse in dieser Arbeit bezieht sich auf ein Straßennetz.

3.2.1 Graphentheorie

Die mathematische Grundlage zur Modellierung eines Netzwerkes bildet die Graphentheorie. Ein Graph G ist formal definiert als eine Menge an Knoten V (engl. *vertices*) und Kanten E (engl. *edges*). Es gilt $G = \{V, E\}$. Eine Kante $u \in E$ verbindet ein Knotenpaar (A, B) mit $A, B \in V$ (WeickerWeicker2013).

Man unterscheidet gerichtete und ungerichtete Graphen. Ein ungerichteter Graph kann entlang der Kanten in beide Richtungen durchlaufen werden, die Knoten einer Kante sind

hier daher vertauschbar. In formaler Notation: $\{u, v\} \in E$.

Ein Graph heißt gerichtet, wenn jede Kante eine definierte Richtung besitzt, das Knotenpaar also geordnet ist (Bill, 1999, S.28). Eine gerichtete Kante $(u, v) \in E$ mit $u, v \in V$ stellt sich bildlich als Pfeil von u nach v dar. Der Knoten u fungiert hier als Start- und v als Endknoten. Die Kanten (u, v) und (v, u) können in einem gerichteten Graphen parallel existieren, z.B. um Straßen mit zwei Fahrtrichtungen abzubilden. Für ein Straßennetz mit Einbahnstraßen oder unterschiedlichen Fahrtzeiten für Hin- und Rückwege, müssen beispielsweise gerichtete Graphen zur Modellierung verwendet werden (WeickerWeicker2013).

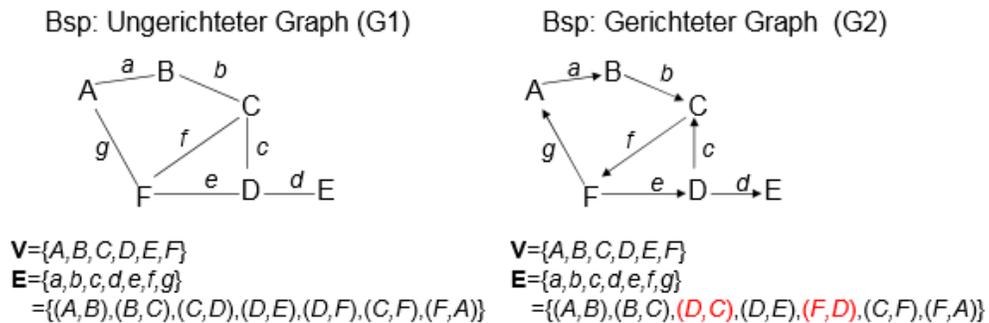


Abbildung 2: Geordnete und ungeordnete Knotenpaare in Graphen (Quelle: MPI-Magdeburg, o.D.)

Kanten und Knoten eines Graphen können numerische Werte als "Gewichtung" zugeordnet werden (ebd., S.28). Beispielsweise können Kantenelementen in einem Straßennetz Fahrtzeiten oder Streckenlängen zugeordnet sein.

Graphen lassen sich durch Adjazenz-, Inzidenzmatrizen oder Adjazenzlisten abbilden. Ein Graph G wird in einer $n \times n$ Adjazenzmatrix $A(G)$ (mit $n = |V|$) so gespeichert, dass der Eintrag von $a_{i,j}$ von $A(G)$ den Wert 1 besitzt, wenn eine Kante von i nach j existiert, ansonsten den Wert 0. Adjazenzmatrizen spiegeln die Nachbarschaften der Knoten wieder (Turau, 2004, S.25). Auch knoten- und kantengewichtete Graphen können in bewerteten Adjazenzmatrizen dargestellt werden. Bei einem kantengewichteten Graphen wird die Gewichtung $B[i, j]$ der Kante von i nach j , in der Position $b_{i,j}$ repräsentiert (ebd., S.28).

Netzwerkanalysen, basierend auf topologischen strukturierten Datensätzen, können durch graphentheoretische Algorithmen direkt und ohne geometrische Berechnungen gelöst werden (Bill, 1999, S.27). Der Graph ist somit eine für den vorliegenden Anwendungsfall sehr geeignete Abstraktion eines Netzwerks, mit dem der Berechnungsaufwand deutlich verringert werden kann.

3.2.2 Netzwerke

Netzwerkanalysen basieren auf planaren Graphen (Bill, 1999, S.99). "Ein Graph heißt planar, falls es eine Darstellung in der Ebene gibt, bei der sich die Kanten höchstens in ihren Endpunkten schneiden" (Turau, 2004, S.23). In der Praxis erfolgen Netzwerkanalysen mit gerichteten Graphen (Lange, 2013, S.37). Wie in 3.2.1 beschrieben, dient der Graph als Abstraktionsschema, um Netzwerke zu modellieren. "Eine geometrisch-topologische Anordnung von Knoten und Kanten definiert ein Netzwerk" (Geoinformatik-Service, 2001) Ein Weg in einem Netzwerk erfolgt von Knoten zu Knoten, über Kanten. "Eine Kante verbindet zwei Knoten, dadurch werden diese beiden Knoten adjazent. In diesen Knoten beginnen oder enden i.d.R. mehrere Kanten. Diese sind somit miteinander inzident" (Bartelme, 2005, S.75). Knoten repräsentieren Verknüpfungs-, Abbiege- oder Endpunkte, wie Kreuzungen oder Haltestellen, in einem Verkehrsnetz. Netzwerkkanten modellieren z.B. knotenverbindende Straßen, Eisenbahn- oder Schifffahrtslinien (Lange, 2013, S.37f, Bartelme, 2005, S.75). Die Knoten- und Kanten-Elemente tragen fachspezifische Informationen zu geometrischen Beziehungen oder Flussbeziehungen im Netzwerk, wie z. B. Weglängen, Widerstände, Leitungsdurchmesser, etc., aus denen in einer Analyse spezifische Wegkosten abgeleitet werden können (Bill, 1999, S.100). Die Bewertung von Kanten erfolgt i.d.R. durch Weglängen zwischen zwei Knoten in einer Längen- oder Zeiteinheit. Es werden sogenannte Impedanzen (*lat. impedire "hemmen"* Duden, o.J.), Widerstandswerte, modelliert (Lange, 2013, S.37f). Wichtige Fragestellungen basierend auf einer Netzwerkanalyse sind die Suche nach dem günstigsten Weg, dem besten Standort und die Suche nach günstigsten Reiserouten über mehrere Standorte (Bill, 1999, S.100).

Der beste Weg kann der metrisch oder topologisch günstigste Weg sein. Der topologisch günstigste Weg sucht die geringste Anzahl an Knoten und Kanten. Der metrisch günstigste Weg minimiert nach der Summe der gewichteten Kanten. Durch die Anwendung von Algorithmen der kombinatorischen Optimierung, wie z. B. der Tripel-Algorithmus oder der Dijkstra Algorithmus, können kürzeste Wege in einem Netz zwischen einem Start- und Zielpunkt ermittelt werden (ebd., S.30). Wegstrecken, Fahrzeiten oder sonstige Gewichtungsfunktionen werden durch diese Algorithmen minimiert. Der Dijkstra Algorithmus wird in Kapitel 3.4.2 genauer beschrieben.

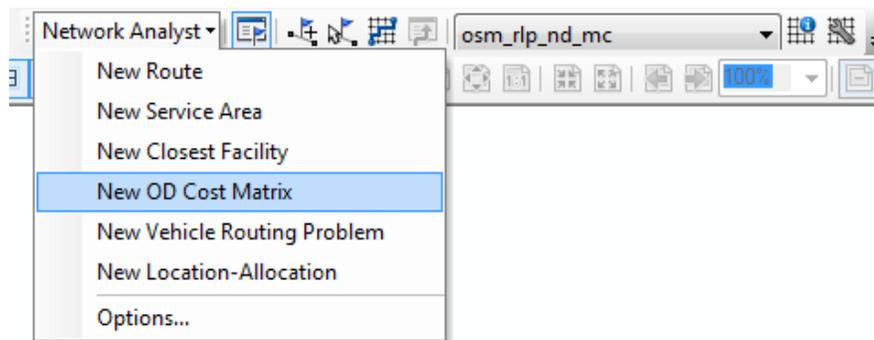


Abbildung 3: Analysen des Network-Analyst-Tools

3.3 ArcGIS-Network-Analyst

3.3.1 Grundlagen

Der Network-Analyst in ArcGIS von Esri ist eine Erweiterung mit Werkzeugen für unterschiedliche Fragestellungen der Erreichbarkeitsanalyse. Er bietet Methoden zur Findung von Routen, nächsten Einrichtungen, zur Generierung von Einzugsgebieten, Berechnung einer Start-Ziel-Kostenmatrix, zum Reisenden-Problem (*Vehicle Routing Problem*) und zu Zuordnungs-Problemen (*Location-Allocation*).

Zur Anwendung einer Analyse wird ein Netzwerkanalyse-Layer erstellt. Dieser Analyse-Layer stellt ein Framework zur Konfiguration eines Netzwerkproblems zur Verfügung. Für eine Berechnung müssen Eingangsdaten und Analyseigenschaften definiert werden (Esri, o.J.a). Die Analysen des Network-Analysts basieren auf speziellen Netzwerkformaten. Esri unterscheidet zwischen geometrischen Netzwerken, für Versorgungs- und Flussnetze, und dem sogenannten *Network-Dataset*, welches sich für Verkehrsnetze eignet (Esri, 2012b).

3.3.2 Network-Dataset

Ein Network-Dataset ist die Grundlage, auf dem die Analyse des Network-Analysts stattfindet. Mit Network-Datasets werden Verkehrsnetzwerke, wie Straßen, Fußwege oder Eisenbahnen, modelliert (ESRIc). Es ist ein Beispiel für ein gerichtetes Netzwerk, in dem für jede Kante eine Bewegungsrichtung definiert ist und dabei eine einzelne Ressource transportiert wird.

Das Network-Dataset besteht aus einer Menge von topologisch verbundenen Netzwerkelementen, die durch Konnektivitätsregeln eine einheitliche Netzwerkinfrastruktur aus der Realität modellieren (IQ, 2012). Die Netzwerkelemente sind Kanten-, Knoten- und Kantenübergangs-*Feature-Classes*, welche als sogenannte Feature-Quellen in das Network-

Dataset eingehen. Feature-Classes beinhalten jeweils räumliche Objekte vom gleichen Geometrietyp und mit den selben Eigenschaften. Optional sind dabei die sogenannten *Kantenübergangs-Feature-Classes*. Sie speichern Informationen, die die Bewegung zwischen zwei oder mehreren Kanten beeinflussen, z.B. Abbiegevorschriften. *Kanten- und Knoten-Feature Classes* sind für das Network-Dataset dagegen essenziell. Die Kanten-Feature-Class stellt Verbindungen aus Straßen, Bahnlinien oder Fußwege dar, Knoten-Feature-Classes modellieren in einem Verkehrsnetz Endpunkte, Verknüpfungspunkte oder Barrieren, beispielsweise Baustellen (Esri, 2012c).

Komplexere Netzwerke können mit multimodalen Network-Datasets abgebildet werden, welche verschiedene Verkehrsnetze, wie Straßen- und U-Bahnnetz, miteinander kombinieren. Hier beteiligen sich mehrere Kanten-Feature-Classes als Kanten-Feature-Quellen (Butler, 2008, S.266f). Mit 3D Network Datasets lässt sich das Routing in der dritten Dimension darstellen, z.B. in einem Gebäude (Esri, 2012d).

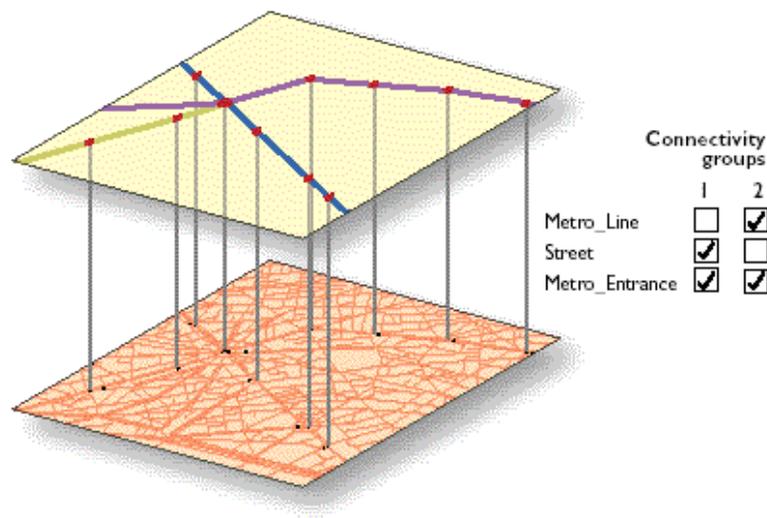


Abbildung 4: Darstellung der Konnektivität zwischen Metro Linien (oberer Layer) und einem Straßen-Netzwerk (unterer Layer). (Quelle: Esri, 2012e)

Konnektivität

Das Verhalten und die topologischen Beziehungen der Netzwerkelemente werden durch Konnektivitätsgruppen und -regeln definiert, die Verknüpfungen von Kanten- und Knotenelementen definieren. Jeder Kanten-Quelle wird eine Konnektivitätsgruppe zugewiesen, Knoten-Quellen können dagegen in einer oder mehreren Gruppen enthalten sein. Die Konnektivitätsgruppen entscheiden, welche Netzwerkelemente sich miteinander verbinden können. Zwei Kanten-Quellen (z.B. eine Straße und einer U-Bahnlinie), die unter-

schiedlichen Konnektivitätsgruppen zugeordnet sind, können nicht miteinander verknüpft werden. In diesem Fall kann nur ein Knoten (z.B. eine U-Bahnstation), der in beiden Konnektivitätsgruppen teilnimmt, diese Kanten miteinander verbinden (siehe Abbildung 4). Dieser Fall betrifft multimodale Netzwerke.

Konnektivitätsregeln definieren die Art der Verbindungen für Netzwerkelemente. Die Ver-

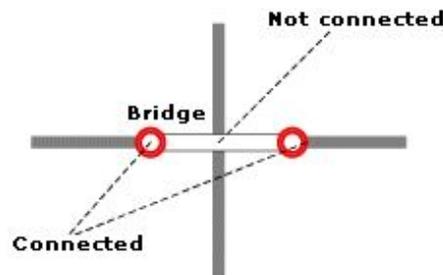


Abbildung 5: Verbindung zweier Punkte zur Abstraktion einer Brücke im Netzwerk (Quelle: Esri, 2012e)

knüpfungen verschiedener Quellelemente innerhalb einer Konnektivitätsgruppe werden durch zu spezifizierende Regeln gesteuert. Verknüpfungen können nur an Linien-Endpunkten erfolgen. Durch eine Verknüpfung lässt sich die Überschneidung zweier Kanten vermeiden, für die kein gemeinsamer Schnittpunkt existiert. Bei der Endpunkt-Konnektivität "werden aus Linien-Features Kanten, die nur an lagegleichen Endpunkten verbunden sind" (Esri, 2012e). Diese Regel ermöglicht es, Überquerungsobjekte zu modellieren. Beispielsweise kann eine Endpunkt-Konnektivität für Brücken eingestellt werden, damit diese sich nicht mit unterführenden Straßen verbinden und nur an ihren Endpunkten mit anderen Kanten-Features verknüpft werden (siehe Abbildung 5). Hierbei müssen Brücken und Straßen derselben Konnektivitätsgruppe zugeordnet sein (ebd.). Straßen können eine "Beliebiger-Stützpunkt"-Konnektivität (*Any Vertex Connectivity*) haben, d.h. Straßen-Features werden an lagegleichen Stützpunkten miteinander verbunden (Zimmermann, 2010, S.7).

Netzwerkattribute

Den Netzwerkelementen sind bestimmte Attribute zugeordnet. Die Netzwerkattribute definieren die Eigenschaften eines Network-Datasets, nach denen es durchlaufen wird. "Beispiele für Attribute sind die Zeit, die für einen bestimmten Straßenabschnitt benötigt wird, Angaben darüber, welche Straßen für welche Fahrzeuge beschränkt sind, Geschwindigkeiten für bestimmte Straßen und Einbahnstraßen" (Esri, 2012f). Die grundlegenden Eigenschaften sind der Name, Verwendungstyp, Einheiten, Datentyp und standardmäßige Verwendung. Der Name benennt das Attribut. Der Verwendungstyp gibt an, welche In-

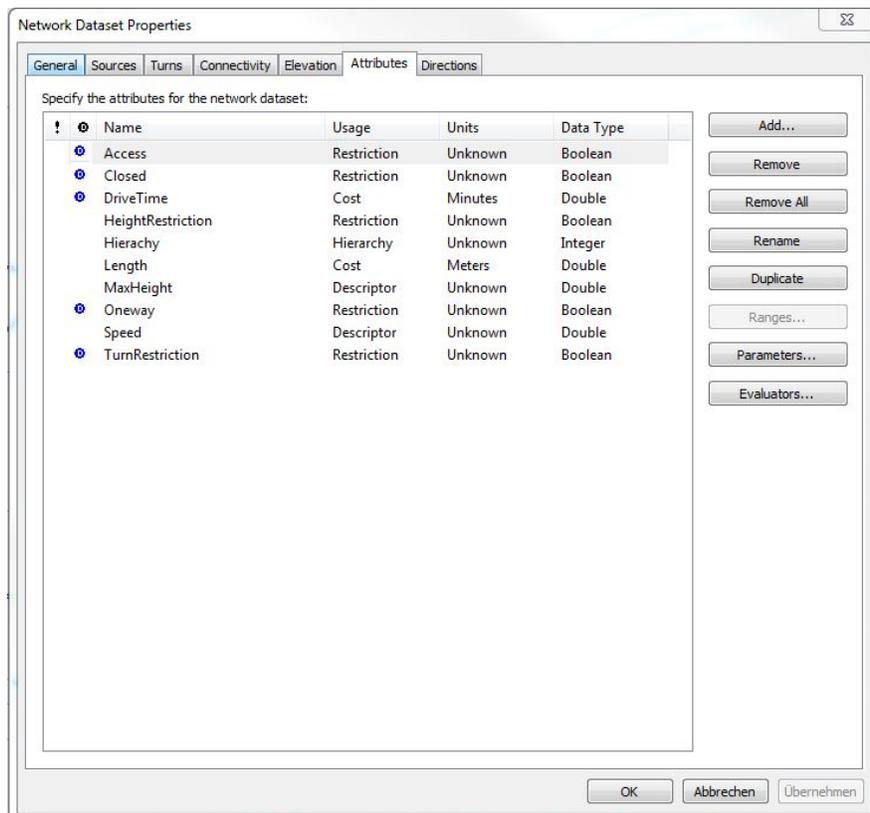


Abbildung 6: Eigenschaften eines Network-Datasets

formationen durch das Attribut modelliert wird: Kosten, Einschränkungen, Deskriptoren oder Hierarchien. Bei den Einheiten für eine Kostenattribut kann es sich um Entfernung- oder Zeiteinheiten handeln (z.B. Meter oder Minuten). Für die anderen Verwendungstypen ist keine Einheit definiert (unbekannte Einheit). Als Datentypen können Boolesche, Integer, Float oder Double Werte ausgewählt werden, wobei Kostenattribute sich immer auf numerische Werte vom Typ Integer, Float oder Double beziehen. Einschränkungen enthalten immer Boolesche Werte und Hierarchien Integer Werte. Attributdefinitionen sind beispielsweise die Fahrtzeit in Minuten, verwendet als Kosten und vom Typ Double oder die Geschwindigkeit als Deskriptor vom Typ Double (siehe Abbildung 6). Die Kostenattribute, wie Fahrtzeit oder Bedarf (z.B. "die Menge des Mülls, der von einer Straßen abgeholt wird" (Esri, 2012f), werden zum Berechnen und Modellieren von Impedanzen verwendet. Der Begriff "Impedanz" wird in diesem Zusammenhang als Kosteneinheit verwendet. Die Kosten können entlang einer Kante proportional aufgeteilt werden, sodass halbe Kosten bei halben Kantendurchlauf berechnet werden. Zur Suche des besten Weges werden in einer Netzwerkanalyse häufig die Impedanzen minimiert (ebd.). Der Verwendungstyp Einschränkung gibt an, ob ein Netzwerkelement passierbar oder unpassierbar ist (Butler,

2008, S.268). Hierbei kann es sich um Einbahnstraßen oder Barrieren durch Baustellen handeln. Mit Hierarchien können Präferenzen für bestimmte Netzwerkelemente konfiguriert werden. Wenn zum Beispiel Straßenklassen Autobahnen und Landstraßen separieren, können durch Angabe einer Hierarchie, beim Durchlaufen des Straßennetzes in einer Besten-Weg-Suche, Landstraßen vermieden werden (Esri, 2012f).

3.4 Start-Ziel-Kostenmatrix

3.4.1 Grundlagen

Mit einer Start-Ziel-Kostenmatrix können Wegkosten in einem Verkehrsnetz, zu einer festgelegten Impedanz, von mehreren Startpunkten zu mehreren Zielen berechnet werden. Das Ergebnis ist eine $m:n$ Relation zwischen Start- und Zielpunkten, die in einer Kostenmatrix dargestellt wird. Der Netzwerkanalyse-Layer besteht aus Sublayern zu Eingabe- und Ausgabedaten sowie für Einschränkungen, die optional in Form von Linien-, Punkt-, oder Polygon-Feature-Classes hinzugeladen werden können. In den Analyseigenschaften kann festgelegt werden, nach welcher Impedanz in der Analyse minimiert werden soll, z.B. die Fahrtzeit in Minuten. Außerdem können die Restriktionsattribute, wie Einbahnstraßen, aktiviert werden, die in der Analyse beachtet werden sollen. Der Suchbereich im Netzwerk kann, durch einen Standardgrenzwert für die Impedanz, eingeschränkt werden. Auch die Anzahl der zu suchenden Ziele kann mit einer Zahleneingabe begrenzt werden (Esri, 2012g).

3.4.2 Dijkstra-Algorithmus

Der Algorithmus zur Start-Ziel-Kostenberechnung basiert auf dem Dijkstra-Algorithmus (Esri). Der Dijkstra Algorithmus ist ein populärer Algorithmus zur Berechnung kürzester Pfade in einem Graphen mit einer amortisierten Laufzeit von $O(n * \log n + m)$ mit $n = |V|, m = |E|$. Dijkstra's Algorithmus löst das Bester-Weg-Problem entlang eines gewichteten Graphen (ohne negative Kantengewichte), anhand einzelner Eingabe-Quellen für Start und Ziel. Um die kürzeste Route von einem Startpunkt s zu einem Zielpunkt d zu finden, betrachtet Dijkstra's Algorithmus Menge an Knoten S zusammen mit ihrer Entfernung zu s über den kürzesten Weg. Der Algorithmus entfernt den Knoten mit den geringsten Gesamt-Wegkosten aus S und fügt alle noch nicht betrachteten Nachbarn zusammen mit ihrem zugehörigen Wegkosten (nach s) zu S hinzu. Der Algorithmus wiederholt sich bis der Zielknoten d zu S hinzugefügt oder der gesamte Graph durchlaufen wurde (Esri, o.J.b).

Der aufgeführte Pseudocode (Algorithm 1) zeigt den genauen Algorithmus, der in diesem Fall die Abstände (als aufsummierte Gewichte) aller Knoten über den günstigsten Weg zum Startknoten s zurückliefert. Dies ist die Information, die auch für den in dieser Arbeit betrachteten Anwendungsfall relevant ist.

Algorithm 1 Dijkstra Algorithmus

```

procedure DIJKSTRA(Knoten  $V$ , Kanten  $E \subset V \times V$ , Gewichte  $weight$ , Startknoten
 $s \in V$ )
  for all  $v \in V$  do
     $dist[v] \leftarrow \infty$ 
     $parent[v] \leftarrow NIL$ 
  end for
   $dist[s] \leftarrow 0$ 
   $Q \leftarrow V$ 
  while  $Q \neq \emptyset$  do
     $u \leftarrow Extract - Min(Q)$ 
    for all  $e = (u, v) \in E$  do
      if  $dist[v] > dist[u] + weight[e]$  then
         $dist[v] \leftarrow dist[u] + weight[e]$ 
         $parent[v] \leftarrow u$ 
      end if
    end for
  end while
  return  $dist$ 
end procedure

```

3.5 Network-Dataset aus OSM-Daten

Das Netzwerk in dieser Arbeit basiert auf einem OSM-Datensatz im OSM-XML-Format (siehe 2.1.3). Diese Daten lassen sich jedoch nicht direkt in einer Netzwerkanalyse des Network-Analyst von Esri verwenden, sondern müssen zu einem Network-Dataset konvertiert werden. "Die manuelle Anpassung der OSM-Daten, um ein Netz zu generieren und im GIS zu nutzen, ist relativ aufwändig. Deshalb sollte möglichst eine automatisierte Lösung geschaffen werden, um die Daten zu konvertieren (IQ, 2012)". Zur automatisierten Konvertierung der OSM-Daten zu einem ArcGIS-Network-Dataset wurden für diese Arbeit zwei Varianten betrachtet: die *OSM2NetworkDataset*-Applikation von Eva Peters (52north, 2015) und der *ArcGIS Editor for OpenStreetMap* von Esri (White, 2015).

3.5.1 OSM2NetworkDataset

OSM2NetworkDataset wurde im Rahmen einer Bachelorarbeit von Eva Peters entwickelt und bietet ein Open-Source-Werkzeug zur automatisierten Konvertierung von OSM-Daten in ein Network-Dataset von Esri. Parameter zur Generierung eines Verkehrsnetzes können hierbei nach Bedürfnissen angepasst werden (IQ, 2012). Mit diesem Konfigurations-Werkzeug können Verkehrsnetzwerke für beliebige Fahrzeugtypen und für bestimmte Länder erzeugt werden (Teucher, o.J. Folie 8).

Neben einer OSM-XML-Datei wird eine Konfigurationsdatei im XML-Format als Eingangsdatei benötigt. Diese XML-Datei beinhaltet Parameter, die ein Netzwerk, aus Straßen, Fahrradwegen oder Fußwegen, für ein bestimmtes Land beschreiben. Zu unterschiedlichen Parametern werden z.B. Transportart, Hierarchie, Einschränkungen, Fahrzeit und Maximalgeschwindigkeit definiert. Für eine Netzwerkanalyse werden die Parameter `<access barrier restrictions>` (Zugangsbeschränkungen durch Barrieren) und `<access highway restrictions>` (Zugangsbeschränkungen für Straßen) benötigt. Die Access Barrier Restrictions regeln, welche Barrieren für für welche Verkehrsteilnehmer unpassierbar sind. Die Access Highway Restrictions steuern, welche Straßen von welchen Verkehrsteilnehmern benutzt werden dürfen.

Die Parameter `<access barrier restrictions interpretation>` und `<access highway restrictions interpretation>` interpretieren die OSM-Tags zu Einschränkungen durch Barrieren und für Verkehrswege. Damit wird definiert welche Wege für den jeweiligen Verkehrsteilnehmer zugänglich bzw. zulässig sind (Zimmermann, 2010, S.43). Das Key-Element ist der Value-Eintrag eines OSM-Tags (`v`) und das Value-Element wird durch "yes" oder "no" bestimmt. "yes" bedeutet z.B., dass eine Barriere passierbar ist und "no" heißt, dass eine Barriere beschränkt bzw. unpassierbar ist. Der Value eines OSM-Tags wird somit als Key verwendet und "der Value definiert dessen Wirksamkeit". Hierzu ein Beispiel (52north, 2015):

OSM-Tags:

```
<tag k="barrier" v="gate"/>
<tag k="access" v="private"/>
```

Parameter:

```
<listElementNoYes key="private" value="no"/>
```

Interpretation: Die Barriere ('gate') ist beschränkt ('private' => 'no').

OSM2NetworkDataset generiert das Network-Dataset mit einer Any Vertex Connectivity (siehe 3.3.2), da bei Überquerungsobjekten in OSM-Daten generell kein gemeinsamer Knoten existieren darf, sodass keine Abbiegemöglichkeit an solchen Objekten besteht

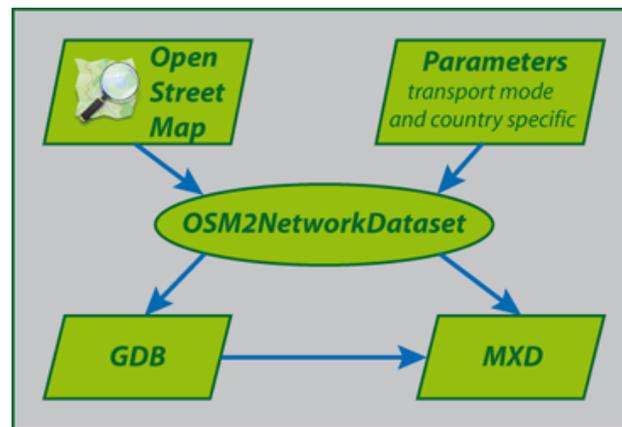


Abbildung 7: Workflow zur Erstellung eines Network-Datasets (Quelle: 52north, 2015)

(Zimmermann, 2010, S.41). Die Applikation ist lizenziert nach der *Apache Licence*, Version 2.0, und wurde in Java 6 implementiert. Das Programm erstellt ein Kartendokument im MXD-Format und eine Geodatabase (.gdb), die ein *Feature-Dataset* enthält. Dieses beinhaltet das Network-Dataset und Feature-Classes zu Straßen, Barrieren und Abbiegeeinschränkungen, welche im Network-Dataset als Feature-Quellen referenziert werden (52north, 2015) (siehe 3.3.2).

3.5.2 ArcGIS Editor for OSM

Der ArcGIS Editor for OpenStreetMap (OSM Editor) ist ein freies, Open-Source Add-On für ArcGIS for Desktop (Esri, o.J.b). Damit werden Werkzeuge zur Verarbeitung von OSM-Daten, die OpenStreetMap-Toolbox, in ArcMap bereitgestellt. Mit dem Werkzeug *Create OSM Network Dataset* aus dieser Toolbox, können die OSM-Daten aus einem OSM-Feature-Dataset, zu einem Network-Dataset transformiert werden. Die Regeln zur Generierung des Netzwerks basieren hier ebenfalls auf einer Konfigurationsdatei im XML-Format. Das generierte Network-Dataset beinhaltet die Netzwerkelemente als Feature-Classes zu Barrieren, Abbiegungen, Verbindungen und Straßen (White, 2013).

Teil II

Praktischer Teil

4 Anwendung

4.1 Datenaufbereitung

4.1.1 Demographische Daten

Basierend auf demographischen Daten zu den Stadtbezirken Trier, wurde ein *Score* zur Gewichtung aller Haushalte (*HH*) berechnet. Dieser stellt die Anteile der Kinder im Alter zwischen 3-6 Jahren je Haushalt zu einem bestimmten Stadtbezirk dar (siehe Abbildung 8).

$$childScore_{Bezirk} = \frac{\sum Kinder_{Bezirk}}{\sum Haushalte_{Bezirk}} \quad (1)$$

Der berechnete Score für jeden Stadtbezirk wurde nun über eine Räumliche Verknüpfung mit dem *Spatial Join*, den einzelnen Haushalten zugewiesen. Hierbei werden die Attribute von zwei Feature-Classes über die räumliche Lage miteinander verknüpft und in einer neu erzeugten Feature-Class gespeichert (Esri). Für die einzelnen Haushalte gilt demnach:

$$childScore_n = childScore_{Bezirk(n)} \forall n \in HH \quad (2)$$

4.1.2 Generierung des Network-Dataset mit OSM2NetworkDataset

Das verwendete Network-Dataset wurde mit dem ArcGIS Editor for OSM erstellt, da es bei der Konvertierung mit OSM2NetworkDataset zu Problemen mit dem vorliegenden Datensatz gekommen ist.

4.2 Berechnung der Start-Ziel-Kostenmatrix im Network-Analyst

Im ersten Schritt wird das Network-Dataset aus OSM-Daten in das Dokument geladen. Dieses Netzwerk muss in der Werkzeugleiste des Network-Analysts ausgewählt sein. Im nächsten Schritt wird ein Analyse-Layer zur Start-Ziel-Kostenmatrix (OD Cost Matrix) angelegt (siehe Abbildung 9). In die jeweiligen Sublayer werden Start (Origin) und

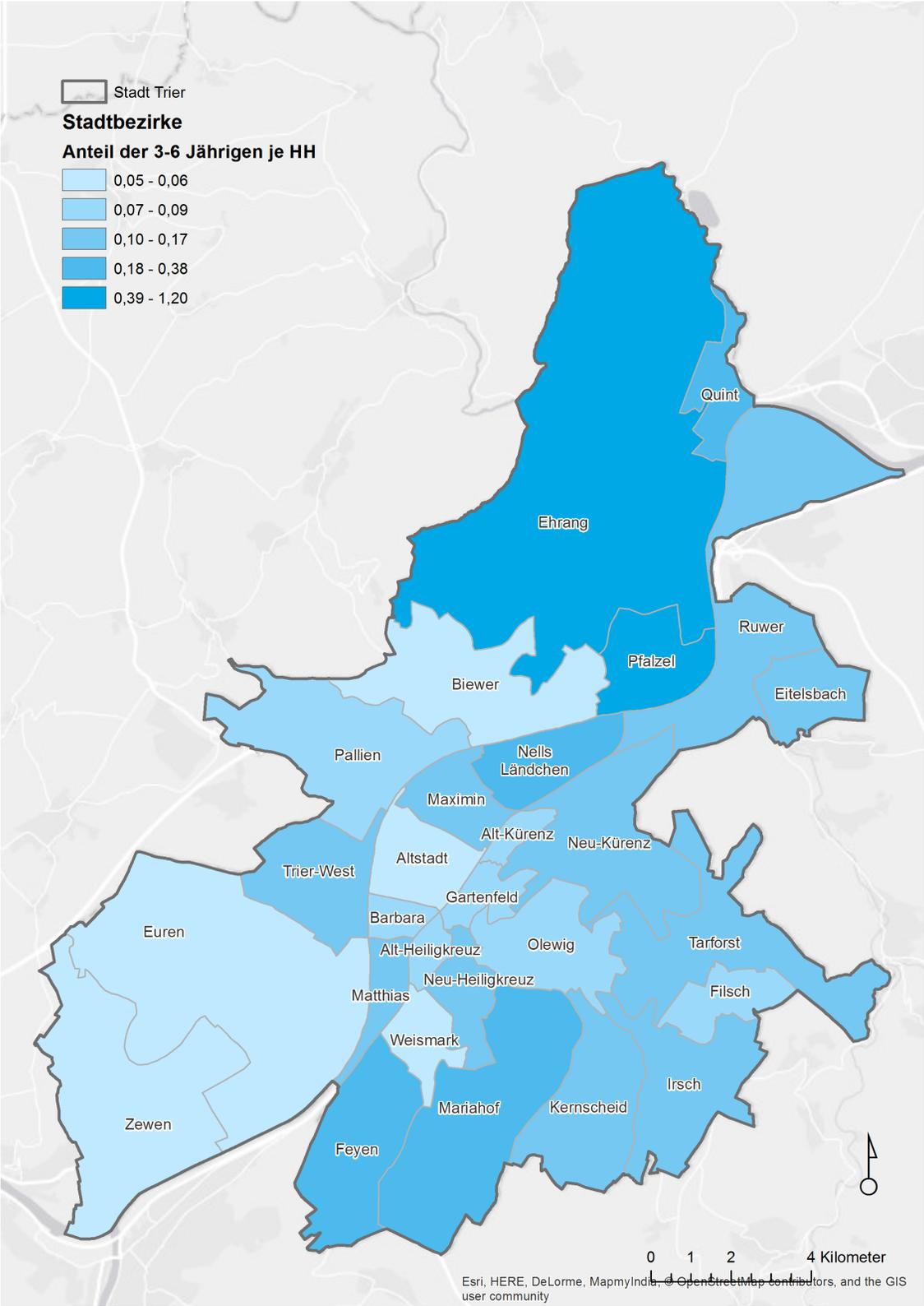


Abbildung 8: Anteile der 3-6 Jährigen je Haushalte, je Stadtbezirk

Zielpunkte (Destinations) über das Network-Analyst Fenster geladen. Die Kindergarten-Feature Class werden den Origins zugewiesen und die Haushalte den Destinations. Wichtig ist hierbei im Feld Name die vorhandenen IDs für Kindergärten und Haushalte mitzuführen, um spätere Verknüpfungen über Joins zu ermöglichen. Dieser Schritt ist nötig, da der Network-Analyst eigene ID für Origins und Destinations generiert, wodurch die Referenz auf die Eingangs-Feature Classes verloren geht. Mit Joins lassen sich Datentabellen, basierend auf einem eindeutigen gemeinsamen Wert eines Feldes, einem sogenannten Schlüssel, verbinden. (Esri, 2013).

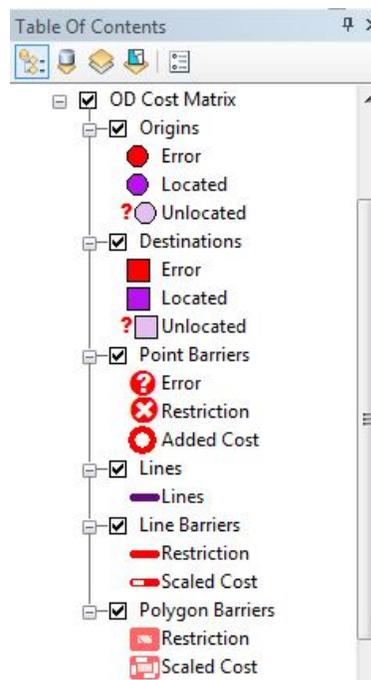


Abbildung 9: Analyselayer der Start-Ziel-Kostenmatrix

In den Eigenschaften zur Generierung der Start-Ziel-Kostenmatrix wurde als Impedanz die Fahrtzeit (DriveTime) in Minuten angegeben (siehe Abbildung 10). In der Analyse werden somit die Kosten nach der Fahrtzeit minimiert und die schnellsten Wege zwischen Kindergärten und Haushalten ermittelt. Als Einschränkungen (Restrictions) wurden nur Einbahnstraßen (Oneway) und Abbiegevorschriften (TurnRestrictions) aktiviert. Access und Closed wurden nicht ausgewählt. Diese Auswahl ist auf Grund von Fehlern in den vorhandenen OSM Daten nötig, durch die zu vielen Haushalten wegen fehlender Erreichbarkeit keine Kosten berechnet werden konnten. Die Option 'Destinations to Find' wurde auf 'All' gesetzt, damit die Kostenberechnung von den einzelnen Kindergärten zu jeweils allen Haushalten erfolgt. Wendungen an Knoten (U-Turns at Junctions) wurde erlaubt.

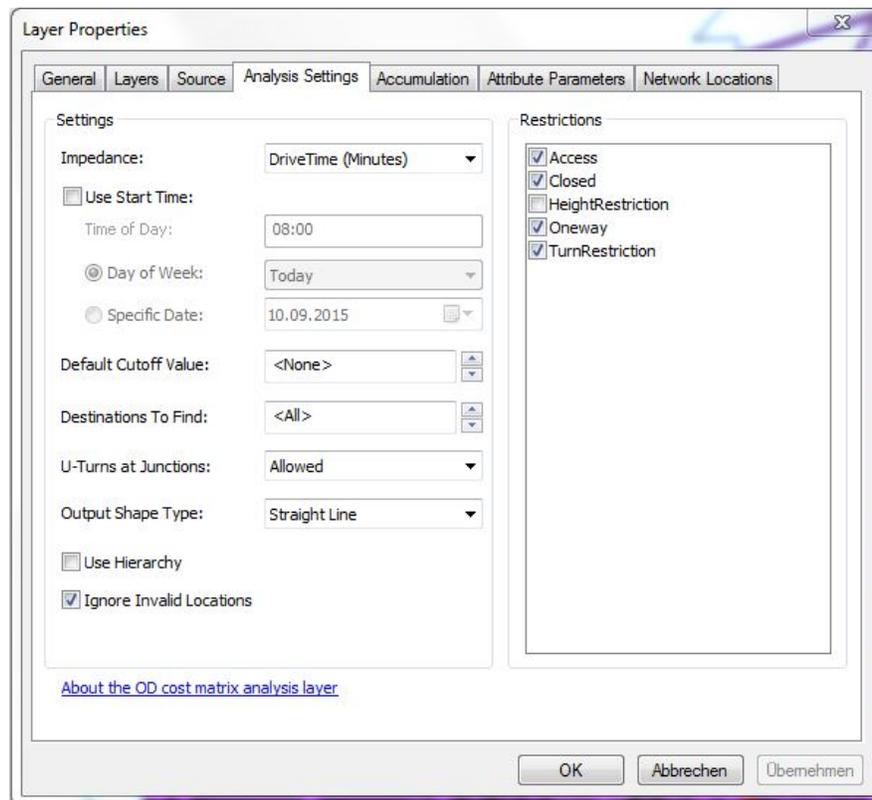


Abbildung 10: Analyseeigenschaften der Start-Ziel-Kostenmatrix

Nach der Konfiguration der Eigenschaften wurde die Start-Ziel-Kostenmatrix berechnet und in dem Sublayer 'Lines' ausgegeben. Die Berechnung benötigt für den vorliegenden Datensatz von 21930 Haushalten und 32 Kindergärten ca. eine Minute auf einem modernen Desktop-PC. Die Kosten werden in der Attributtabelle zu Lines in dem Feld 'Total DriveTime' dargestellt. Abbildung 11 zeigt einen Ausschnitt aus der Attributtabelle, welche nach aufsteigenden Kosten sortiert vorliegt.

4.3 Datenverknüpfung

Nach Berechnung der Kostenmatrix, wurde diese über mehrere Joins mit den Eingangsdaten verknüpft, mit dem Ziel den Anfangs berechneten *childScore* und die Kindergartenkapazitäten mit der Kostenmatrix zusammen zu führen. Da in der Kostentabelle (dem Ausgabeformat der Kostenmatrix) kein direkter Schlüssel für diese Attributtabelle vorhanden ist, musste ein Zwischenschritt durchgeführt werden. Über die Schlüssel OriginID/ObjectID und DestinationID/ObjectID wurden mit zwei Joins die Attribute aus Origin und Destinations der Kostenmatrix hinzugefügt. Da diese nun über die OSM-ID's der Kinder-

Lines						
ObjectID	Shape	Name	OriginID	DestinationID	DestinationRank	Total DriveTime
1547289	Polyline	41569523 - 41170601	71	22452	1	0,117386
1547290	Polyline	41569523 - 47322679	71	22480	2	0,128355
1547291	Polyline	41569523 - 130421029	71	25489	3	0,132631
1547292	Polyline	41569523 - 47322680	71	22481	4	0,140204
1547293	Polyline	41569523 - 130421033	71	25490	5	0,146397
1547294	Polyline	41569523 - 47322237	71	22479	6	0,158285
1547295	Polyline	41569523 - 91842117	71	22856	7	0,167679
1547296	Polyline	41569523 - 47322236	71	22478	8	0,172406
1547297	Polyline	41569523 - 128083803	71	24536	9	0,173329
1547298	Polyline	41569523 - 128083860	71	24590	10	0,182078
1547299	Polyline	41569523 - 130421027	71	25488	11	0,184269
1547300	Polyline	41569523 - 128083801	71	24534	12	0,193834
1547301	Polyline	41569523 - 128083806	71	24538	13	0,195155
1547302	Polyline	41569523 - 130421026	71	25487	14	0,203176
1547303	Polyline	41569523 - 128083859	71	24589	15	0,203956
1547304	Polyline	41569523 - 128083828	71	24560	16	0,206772
1547305	Polyline	41569523 - 128083789	71	24529	17	0,207277
1547306	Polyline	41569523 - 130421024	71	25486	18	0,216508
1547307	Polyline	41569523 - 128083815	71	24547	19	0,216552
1547308	Polyline	41569523 - 128083810	71	24542	20	0,219524
1547309	Polyline	41569523 - 128083782	71	24525	21	0,219857
1547310	Polyline	41569523 - 128083794	71	24531	22	0,227683
1547311	Polyline	41569523 - 128083843	71	24574	23	0,229368
1547312	Polyline	41569523 - 128083798	71	24533	24	0,230693
1547313	Polyline	41569523 - 325094909	71	43827	25	0,231645
1547314	Polyline	41569523 - 41170599	71	22451	26	0,235515
1547315	Polyline	41569523 - 128083836	71	24568	27	0,238
1547316	Polyline	41569523 - 128083814	71	24546	28	0,241787
1547317	Polyline	41569523 - 128083851	71	24581	29	0,242173
1547318	Polyline	41569523 - 128083830	71	24562	30	0,243169
1547319	Polyline	41569523 - 128083856	71	24586	31	0,246177
1547320	Polyline	41569523 - 128083816	71	24548	32	0,247795
1547321	Polyline	41569523 - 128083813	71	24545	33	0,25005

Abbildung 11: Ausschnitt aus der Kostentabelle

gärten und Haushalte verfügte konnten über zwei weitere Joins deren Attribute ebenfalls zur Kostenmatrix hinzugefügt werden. Im Anschluss wurde die angereicherte Start-Ziel-Kostenmatrix zu einer CSV-Datei exportiert, welche im weiteren Schritt, der Zuordnung nach Kapazität, als Eingangsdaten benötigt wurde.

4.4 Zuordnung der Haushalte unter Berücksichtigung der Demographie

Der Algorithmus zur Zuordnung der Haushalte zu einem Kindergärten, unter Berücksichtigung des Scores (siehe 4.1.1) und der Kapazität, wurde in Python umgesetzt und beinhaltet folgende Schritte:

1. Jedem Kindergarten ist jeweils ein Stapel Referenzen zu Haushalten, sortiert nach Distanz (Impedanz), zugewiesen (aufsteigend)
2. Der anteilmäßig am wenigsten belegte Kindergarten bekommt den in seinem Stapel obersten, noch nicht zugeordneten Haushalt zur Menge der ihm zugeordneten Haushalte (HH_{KG}) hinzugefügt. Dem Haushalt wird die Distanz zu diesem Kindergarten als $driveTime_{HH}$ zugewiesen. Dieser und die bereits zugeordneten Haushalte werden vom Stapel entfernt.

3. Anschließend Sortierung des Kindergarten-Stapels(KG) nach Belegung, die für jeden Kindergarten (KG) nach folgender Formel berechnet wird:

$$Belegung_{KG} = \frac{\sum_{i \in HH_{KG}} childScore_i}{Kapazitaet_{KG}} \quad (3)$$

4. Wiederhole Schritt 2 und 3, bis alle HH zugeordnet sind (Referenzstapel leer)

In Formel 3 wird deutlich, dass die Berechnung der Belegung und damit die Zuordnung sowohl abhängig von der Kapazität der einzelnen Kindergärten $Kapazitaet_{KG}$ als auch vom aus den demographischen Daten abgeleiteten Score $childScore_n$ mit $n \in HH$.

4.4.1 Kandidaten

In einer Visualisierung der Fahrtzeitkosten der einzelnen Haushalte zu den ihnen jeweils zugewiesenen Kindergärten hoben sich deutlich Bereiche mit längeren Fahrtzeiten hervor (Abbildung 12). In diesen Bereichen lässt sich in einer weiteren Visualisierung der Zuordnung außerdem ein "buntes" Punktmuster erkennen (Abbildung 13), welches auf Lücken zwischen den direkten Einzugsgebieten hindeuten, wo keine eindeutige Präferenz für einen bestimmten Kindergarten erkennbar ist. Auf dieser Informationsbasis wurden drei Kandidaten für einen weiteren Kindergarten ausgewählt (siehe Abbildung 14), die in einem weiterem Verfahren auf ihre Auswirkung hin untersucht wurden (5.2). Allen Kandidaten wurde eine gleiche Kapazität von 50 Plätzen für Kinder zugewiesen.

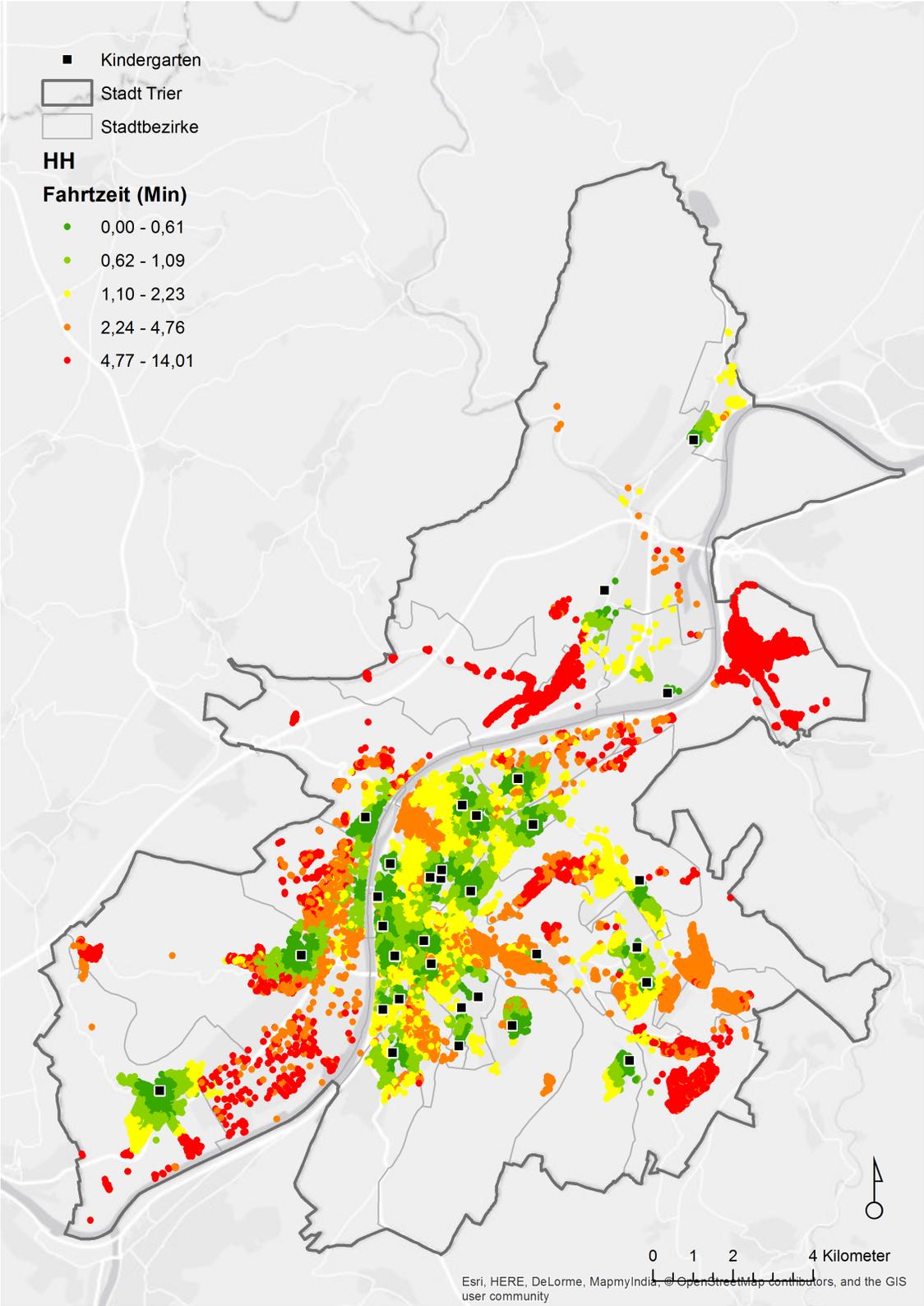


Abbildung 12: Distanzen der Haushalte zu den jeweils zugeordneten Kindergärten

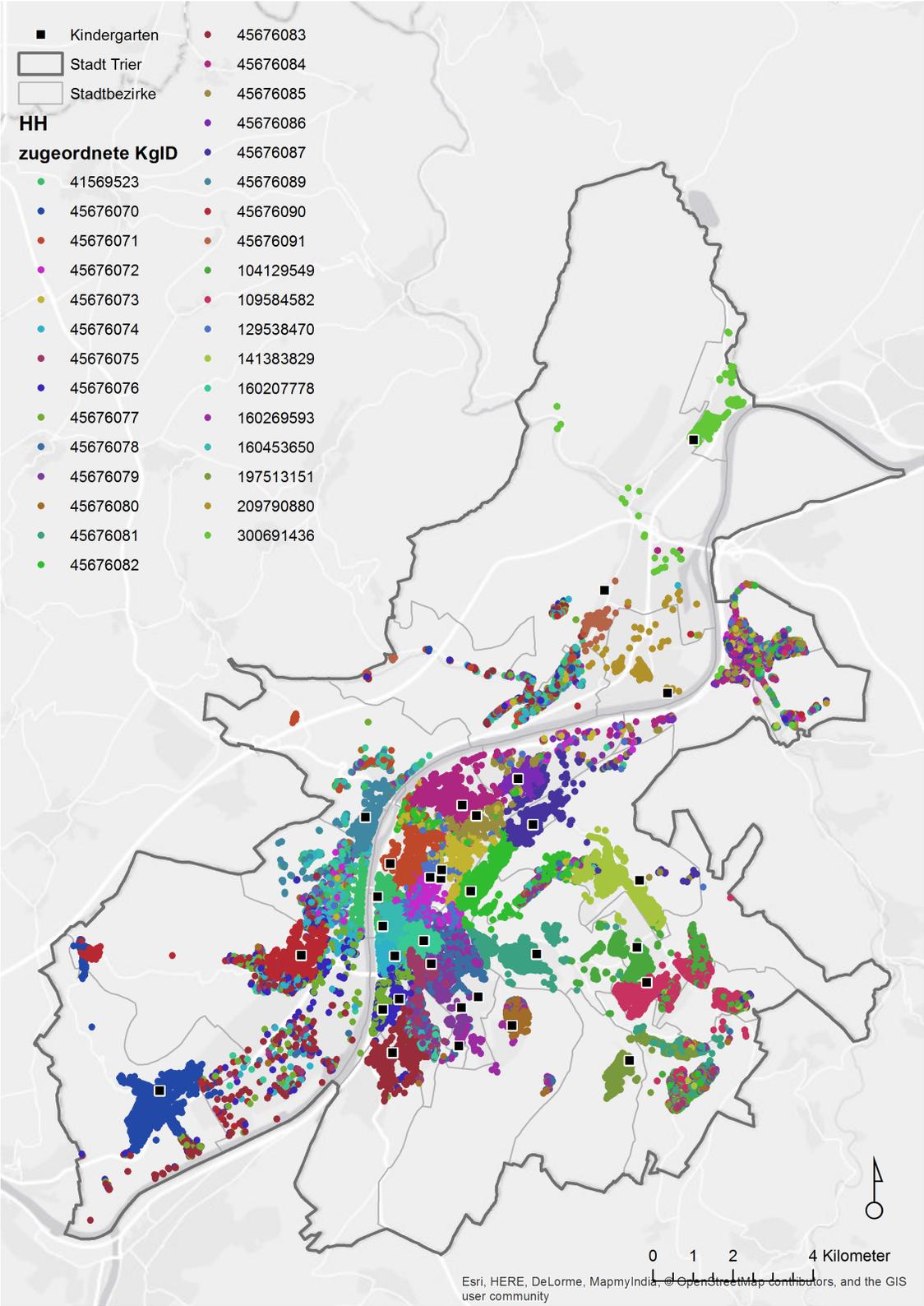


Abbildung 13: Zuordnung der Haushalte zu den Kindergärten

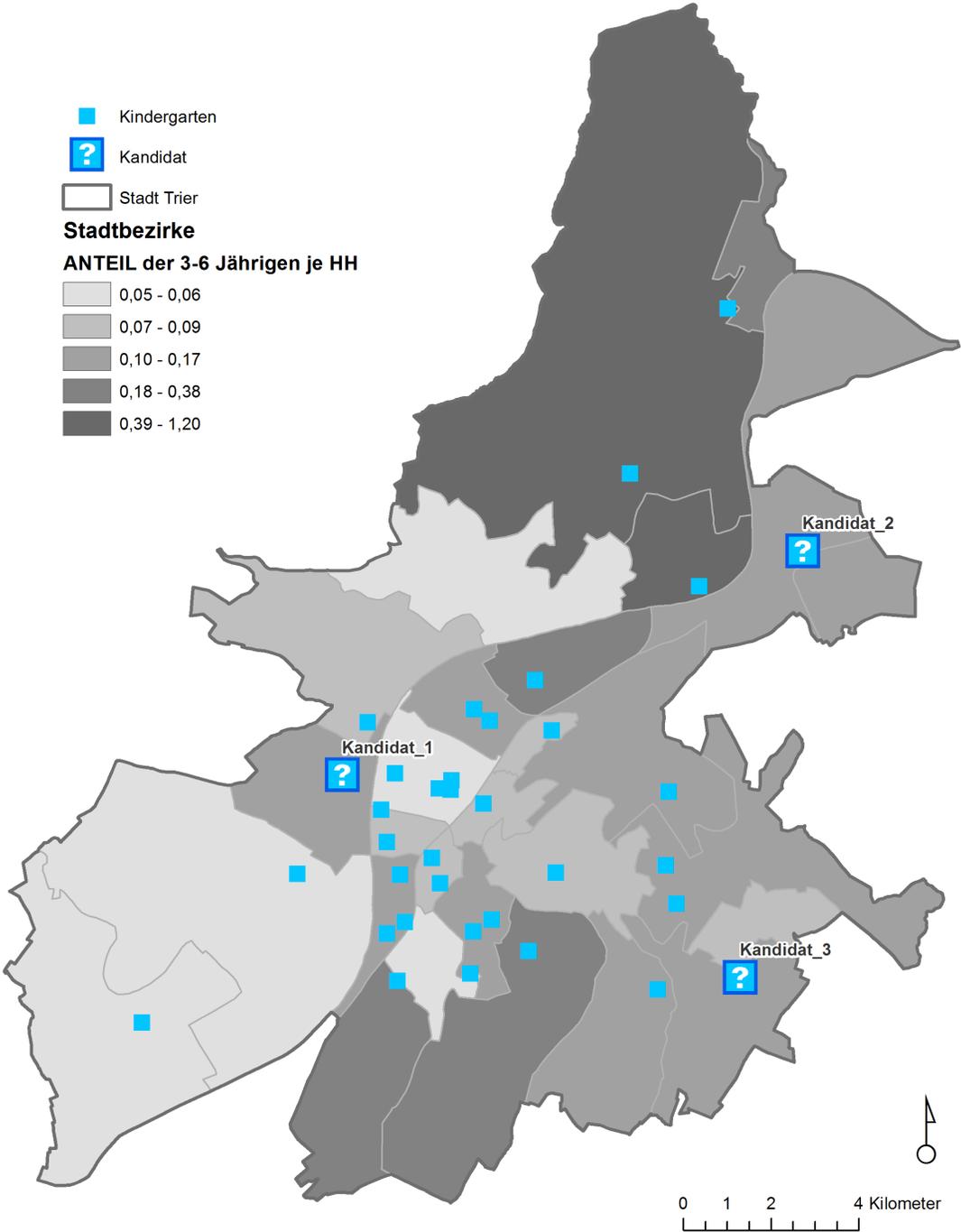


Abbildung 14: Kandidaten-Standorte

5 Ergebnisse und Auswertung

5.1 Zuordnung

Die Visualisierung der Zuordnung, der Haushalte zu einzelnen Kindergärten, zeigt eine Clusterung der Haushalte um die zugewiesenen Kindergärten (siehe Abbildung 15). Bereiche, wo benachbarte Haushalte verschiedenen Kindergärten zugeordnet werden, bedeuten ein sich überschneidendes Einzugsgebiet. Hier befinden sich "Lücken" zwischen den Einzugsgebieten der Kindergärten. Aus diesem Grund entsteht in Überschneidungsgebieten ein "buntes" Punktmuster. Diese Lücken könnten mit weiteren Standorten erschlossen werden.

5.1.1 Auswahl der Kindergärten

Die Darstellung der Fahrtkosten in Abbildung 12 zeigt, dass mit der Zuordnung die Fahrtkosten der Haushalte minimiert wurden. Die zu einem Kindergarten am nächstgelegenen Haushalte wurden diesem auch zugewiesen. Deutlich stechen die Bereiche mit höherer Fahrtzeit hervor (rote Punkte), die zustande kommen, wenn sich kein Kindergartenstandort in nächster Nähe befindet.

Das Verhältnis zwischen Kapazität der Kindergärten und deren Belegung ist nach der Zuordnung untereinander ausgeglichen (siehe Abbildung 16). Dies war zu erwarten, da der Algorithmus aus 4.4. eine einheitliche Verteilung nach Kapazität gewährleistet. Alle Kindergärten weisen eine Überbelegung von durchschnittlich 24,05 % mit einer Standardabweichung von 0,26 % auf, der Median beträgt 23,98 %.

5.1.2 Kostenfunktion

Um die tatsächliche Gewichtung eines einzelnen Haushalts für die Kostenfunktion zu bestimmen, wird die errechnete Fahrtzeit zum nächsten Kindergarten mit dem errechneten *childScore* des Haushalts multipliziert:

$$costScore_{HH} = driveTime_{HH} * childScore_{HH} \quad (4)$$

Dies führt zu einer besonders starken Gewichtung wenn Haushalte in kinderreichen Stadtteilen liegen und gleichzeitig eine große Entfernung zum zugeordneten Kindergarten haben. Abbildung 17 visualisiert den jeweils berechneten *costScore* der Haushalte.

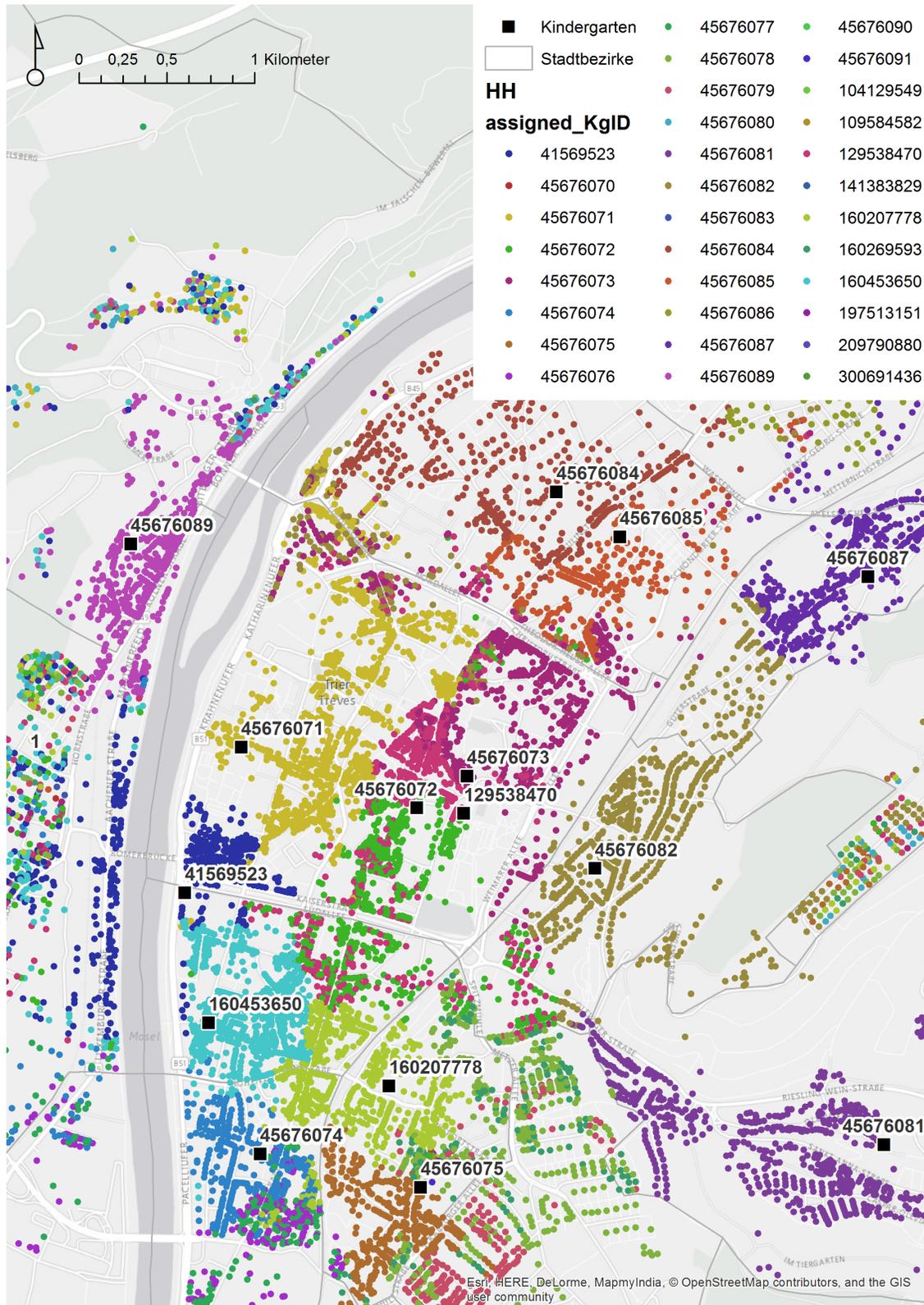


Abbildung 15: Zoom der Zuordnungsansicht

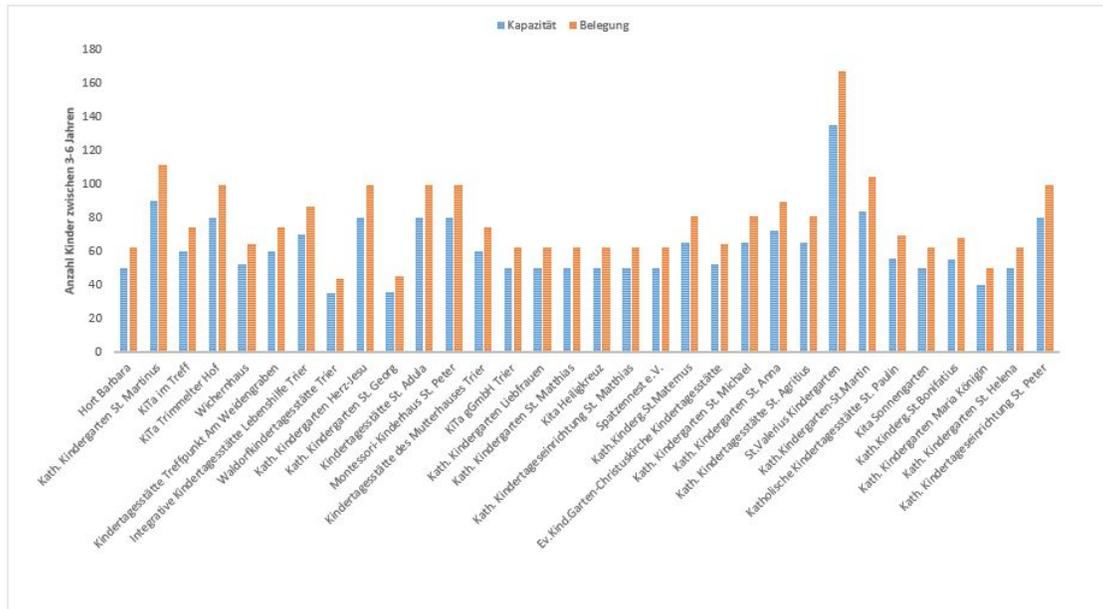


Abbildung 16: Verhältnis zwischen Kapazität und Belegung der Kindergärten

5.2 Identifizierung des besten Kandidaten

Zur Bestimmung des Kandidaten, der am besten dazu beiträgt, die gesamten Wegkosten für alle Haushalte zu minimieren, fungiert eine Vergleichsparameter. Hier wurden das arithmetische Mittel, der Median und die Standardabweichung für alle drei Kandidaten als Parameter der deskriptiven Statistik zur Auswertung herangezogen.

Das arithmetische Mittel \bar{x} repräsentiert den Wert, der die Summe der quadrierten Abstände zwischen den Variablenwerten minimiert" (Udelhoven, 2012). Im hier vorliegenden Fall:

$$\overline{costScore} = \frac{1}{|HH|} \sum_{n \in HH} costScore_n \quad (5)$$

Der Median teilt eine sortierte Wertemenge in gleich große Hälften, die sog. 50%-Quantils ($Q_{0.5}$) und minimiert "die Summe der Beträge der Abstände" (Bill, 1999, S.106). Dieser Parameter ist bei der Schätzung der zentralen Tendenz weniger beeinflusst von Ausreißern, als das arithmetische Mittel.

Die Standardabweichung σ erfasst die Streuung von Werten zum Mittelwert, innerhalb einer Verteilung.

In den Kostentabellen der Kandidaten wurden die Parameterberechnungen jeweils auf den $costScore$ der Kostenfunktion (4) bezogen.

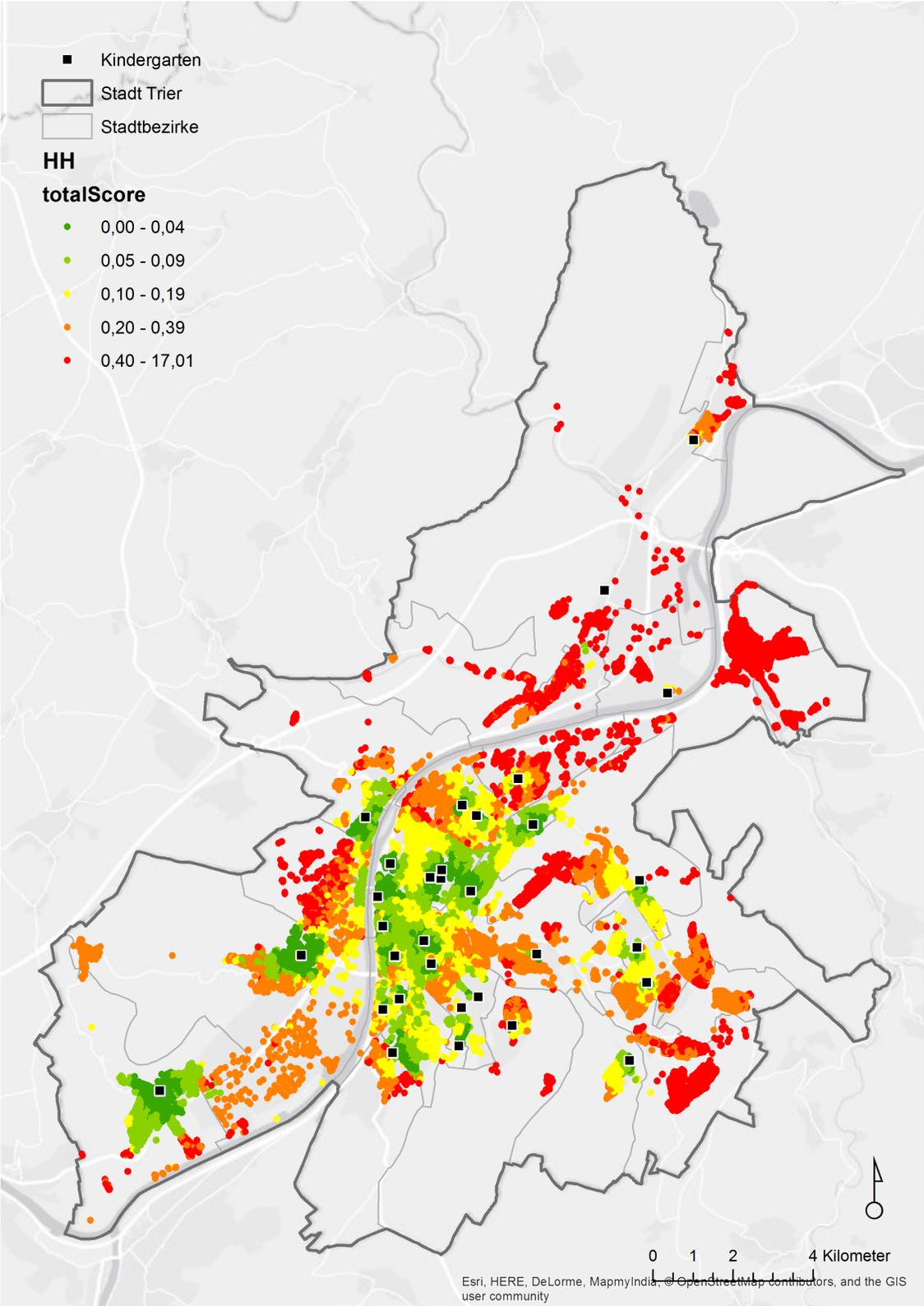


Abbildung 17: Darstellung der Kostenfunktion für einzelne Haushalte

Tabelle 2 Berechnung inklusive Ausreißerwerten

Kandidat	\bar{x}	$Q_{0.5}$	σ
1	1.2205	0.1364	121.2392
2	1.2599	0.1291	123.9646
3	1.2462	0.1369	122.5862

Tabelle 3 Berechnung exklusive Ausreißerwerten

Kandidat	\bar{x}	$Q_{0.5}$	σ
1	0.3127	0.1364	0.839
2	0.3108	0.1290	0.9212
3	0.3180	0.1368	0.87275

Durch den Vergleich der Werte von Tabelle 2, welche von Ausreißern (19 Haushalte) beeinflusst werden, und Tabelle 3, welche die Ausreißer nicht miteinbezogen hat, wird deutlich, dass der Mittelwert stark von Ausreißern beeinflusst wird. Die Berücksichtigung der Ausreißer erhöht deutlich die Standardabweichung. Daher wird in Tabelle 2 der Median als Vergleichsparameter betrachtet. Kandidat 2 weist hier die minimalen Kosten auf und gilt somit als bester Kandidat für einen neuen Kindergarten-Standort (siehe Abbildung 18).

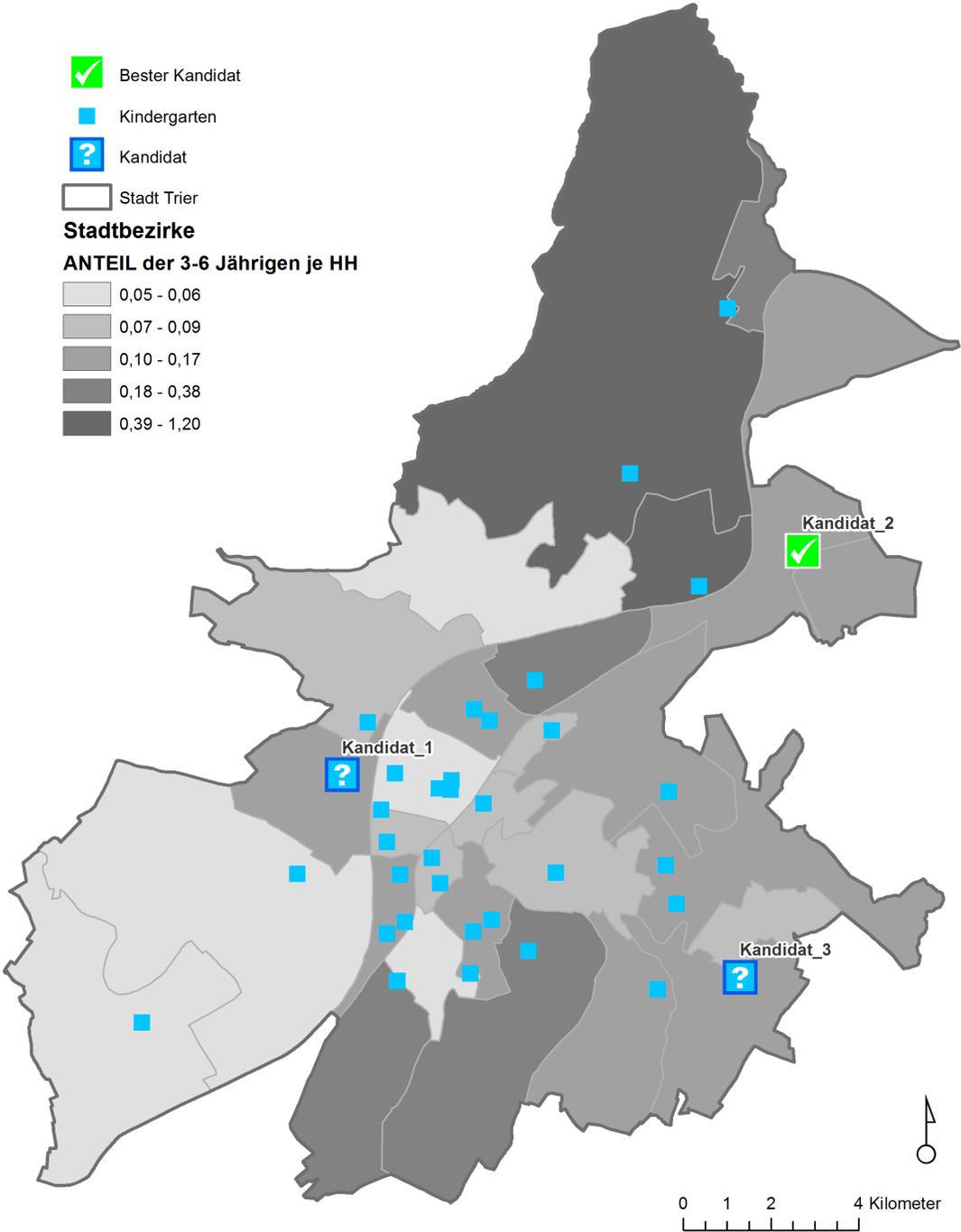


Abbildung 18: Bester Kandidat

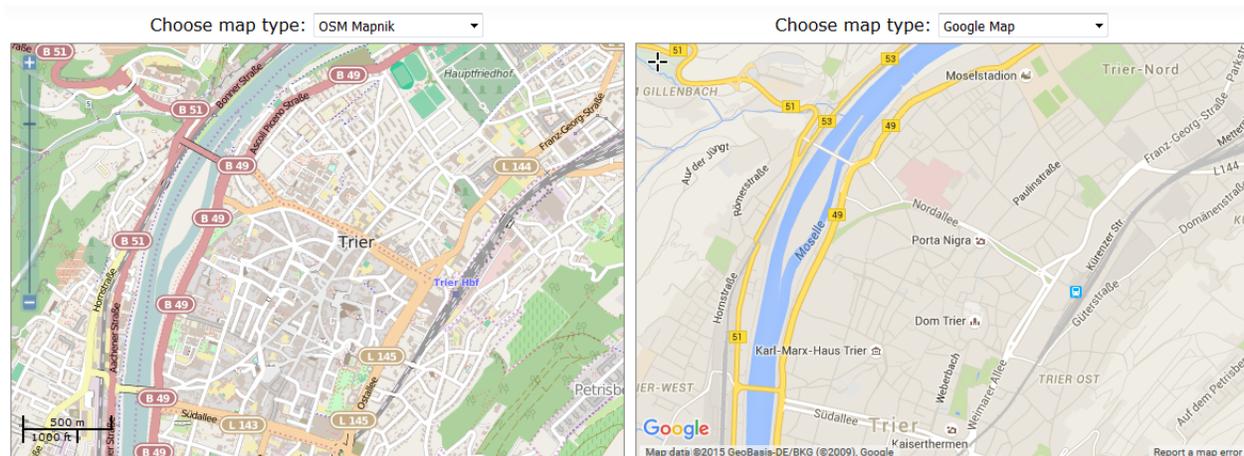


Abbildung 19: OSM vs. GoogleMaps (Quelle: Geofabrik, o.D.)

6 Diskussion

6.1 Bewertung des Verfahrens

6.1.1 Netzwerkanalyse mit OSM-Daten

Mit dem Network-Dataset, basierend auf OSM-Daten, konnte eine Netzwerkanalyse mit dem Network Analyst durchgeführt werden. Eine Erleichterung bei der Generierung eines Network-Datasets aus OSM-Daten ist dabei die automatisierte Datenkonvertierung mit dem ArcGIS-Editor for OSM (3.5). Da das OpenStreetMap-Projekt eine Gemeinschaftsprojekt auf freiwilliger Basis ist und keiner zentralen Kontrolle unterliegt (Zimmermann, 2010, S.16), kann es vorkommen, dass Regionen weniger gut erschlossen sind als andere. Auch bezüglich Informationen, die sich auf ein Verkehrsnetz beziehen. Die in OSM-Daten verwendeten Map-Features weisen z.B. noch Umsetzungsbedarf zu zeitlichen Beschränkungen, und richtungsabhängiger Höchstgeschwindigkeiten auf (ebd., S.74). Diese konnten in den Netzwerkeigenschaften nicht konfiguriert werden. Das Straßennetzwerk für das Untersuchungsgebiet Trier scheint gut ausgearbeitet zu sein, hier wäre ein Vergleich mit auf Vollständigkeit überprüften Netzwerken interessant. Mit dem Tool *MapCompare*, welches von der Geofabrik bereitgestellt wird, kann man OSM-Daten für bestimmte Ausschnitte mit GoogleMaps vergleichen (siehe Abbildung 19). Der Network-Analyst bietet eine intuitives benutzerfreundliches Interface und konnte die Start-Ziel-Kostenmatrix ohne langem Rechenaufwand berechnen. Jedoch müsste umständliche Joins durchgeführt, um die Ergebnisdaten mit den Eingangsdaten zu verknüpfen (4.3).

6.1.2 Zuordnung

Die Zuordnung der Haushalte zu den Kindergärten nach dem Algorithmus in Python (4.4) lieferte ein gutes Resultat. Auch als Vorakquise zur Auswahl neuer Standorte konnte das Ergebnis genutzt werden (4.4.1). Mit diesem Algorithmus wurden alle Haushalte zugeordnet, sodass die Kindergärten überfüllt wurden. Je nach Fragestellung könnte diese Zuordnung auch anhand von Kapazitäten limitiert werden.

6.2 Optimierungsmöglichkeiten

6.2.1 Datengrundlage mit höherer Auflösung

Ein Untersuchungsergebnis basiert neben dem eigentlichen Verfahren zu einem Großteil auf der Qualität und Auflösung der verwendeten Daten. Um die Realität zu modellieren, müssen diese einen hohen Grad an Genauigkeit aufweisen. Die Eingangsdaten, die in dieser Arbeit verwendet wurden, sind eine Annäherung an die Realität. Demographische Daten, bezüglich der Anzahl der Kinder im Kindergarten-Alter je Haushalt wurden durch die Anteile der Kinder je Haushalt eines Stadtbezirks approximiert. Besser aufgelöste demographische Daten würden das Ergebnis optimieren.

6.2.2 Iteratives Verfahren zur 2D-Standortplanung

In dieser Arbeit basiert das Lösungsverfahren zur Standortfindung auf einem Testverfahren. Denkbar wäre ein Verfahren, welches automatisiert einen besten Standort definiert, ohne Kandidaten im Vorfeld zu bestimmen. Unter der Nutzung bivariater Statistik in der Methode 2D-Medianzentrum, nach G. BAHRENBURG u.a. (1985), lässt sich ein Ansatz zur automatisierten Bestimmung eines Standortes im zweidimensionalen Raum definieren (Bill, 1999, S.106).

"Gegeben sei eine Menge von Punkten $P_i(x_i, y_i) \forall i \in \{1, \dots, n\}$, die auf den zu bestimmenden Standort einen Einfluss mit dem Gewicht p_i ausüben. Im einfachsten Falle üben alle Punkte P denselben Einfluss auf den gesuchten Standort aus, so dass $p_i = 1 \forall i \in \{1, \dots, n\}$ gilt. Mit Hilfe eines iterativen Ansatzes wird das zweidimensionale Medianzentrum bestimmt:

1. Man lege ein regelmäßiges Gitter über die Fläche, auf der die Einflusspunkte P liegen. Die gewählte Gitterweite legt die Genauigkeit der Standortbestimmung fest.

2. Für jeden Gitterpunkt $G_j(x_j, y_j)$ bestimme man die gewichtete Summe der Abstände zu allen Punkten P_i nach

$$f(G_j) = \sum_{i=1}^n \sqrt{((x_i - x_j)^2 + (y_i - y_j)^2)} * p_i \quad (6)$$

3. Man wähle den Gitterpunkt G_j , für den $f(G_j)$ am kleinsten ist.
4. Um diesen Punkt G_j lege man eine feineres Gitter (z.B mit der halben Gitterweite) und wiederhole den Vorgang ab Schritt 2 solange, bis der gesuchte Standort in der gewünschten Genauigkeit festgelegt ist. Als endgültigen Standort wähle man denjenigen Punkt G_j , der im feinsten Gitternetz die geringste Abstandssumme zu den Punkten P_i aufweist (Bill, 1999, S.107)."

Diese Methode ist primär für punktförmige Einflussfaktoren konzipiert. Das Verfahren lässt sich auf den Anwendungsfall dieser Arbeit übertragen. Im zweiten Schritt, Formel (6), ist jedoch statt der Summe der euklidischen Abstände, die hier verwendete Kostenfunktion (4.4) anzuwenden. Anstatt eines Gitters würden sich im ersten Schritt auch Punkte aus dem Straßennetz anbieten.

7 Ausblick und Fazit

Das angewendete Verfahren eignet sich grundsätzlich zur Planung von Standorten, bei denen die Kapazität ein wichtiges Kriterium darstellt. Es könnte beispielsweise auch auf Schulen übertragen werden, unter Verwendung von öffentlichen Verkehrsmitteln und Gehwegen. Als Grundlage für die Netzwerkanalyse müsste hier ein multimodales Netzwerk eingesetzt werden, in dem unterschiedliche Verkehrsmittel berücksichtigt werden. Für Wahllokale wäre dieses Verfahren ebenfalls denkbar, um für Wahlberechtigte die Wegkosten zu minimieren.

Auch die Planung von neuen Geschäftsstandorten beispielsweise für den Einzelhandel kann mit der vorgestellten Vorgehensweise optimiert werden, auch wenn Kapazitäten hier eine weniger große Rolle spielen. In einem solchen Szenario bietet sich ebenso die Einbeziehung von auf Zielgruppen zugeschnittenen demographischen Daten an.

Die rasante Entwicklung des OSM-Projekts lässt auf eine Verfeinerung der bisherigen Daten erhoffen. Damit würde auch die Qualität von Netzwerken, basierend auf OSM-Daten, gesteigert werden.

8 Abbildungs-, Tabellen- und Literaturverzeichnis

Tabellenverzeichnis

1	Die verwendeten OSM-Daten	8
2	Berechnung inklusive Ausreißerwerten	33
3	Berechnung exklusive Ausreißerwerten	33

Abbildungsverzeichnis

1	Verwendete Kindergarten-Standorte der Stadt Trier	7
2	Veranschaulichung Graphen (Quelle: MPI-Magdeburg, o.D.)	10
3	Analysen des Network-Analyst-Tools	12
4	Konnektivität (Quelle: Esri, 2012e)	13
5	Brückenkonnektivität (Quelle: Esri, 2012e)	14
6	Eigenschaften eines Network-Datasets	15
7	Workflow (Quelle: 52north, 2015)	19
8	Anteile der 3-6 Jährigen je Haushalte, je Stadtbezirk	21
9	Analyselayer der Start-Ziel-Kostenmatrix	22
10	Analyseeigenschaften der Start-Ziel-Kostenmatrix	23
11	Ausschnitt aus der Kostentabelle	24
12	Distanzen der Haushalte zu den Kindergärten	26
13	Zuordnung der Haushalte zu den Kindergärten	27
14	Kandidaten-Standorte	28
15	Zoom der Zuordnungsansicht	30
16	Verhältnis zwischen Kapazität und Belegung der Kindergärten	31
17	Darstellung der Kostenfunktion für einzelne Haushalte	32
18	Bester Kandidat	34
19	OSM vs. GoogleMaps (Quelle: Geofabrik, o.D.)	35

Literatur

52north (2015). *OSM2NetworkDataset*. Stand: 05.03.2015.

URL: <http://52north.org/communities/geoprocessing/osm2nds/index.html>.

Bartelme, N. (2005). *Geoinformatik, Modelle, Strukturen, Funktionen*. Berlin, Heidelberg (Springer Verlag).

Bill, R. (1999). *Grundlagen der Geo-Informationssysteme. Analysen, Anwendungen und neue Entwicklungen*. 2. Auflage. Bd. Band 2. Wichmann.

Butler, J.A. (2008). *Designing Geodatabases for Transportation*. (ESRI Press) Redlands (California).

Duden (o.J.). *Impedanz, die*. Zugriff: 29.09.2015.

URL: <http://www.duden.de/suchen/dudenonline/impedanz>.

Esri (2012b). *Arten von Netzwerkanalyse-Layern*. Stand: 07.10.2012.

URL: <http://help.arcgis.com/de/arcgisdesktop/10.0/help/index.html#na/004700000032000000>.

— (2012c). *Was ist ein Netzwerk-Dataset?* Stand: 7/10/2012.

URL: <http://help.arcgis.com/de/arcgisdesktop/10.0/help/index.html#/004700000007000000>.

— (2012d). *Was ist Network Analyst?* Stand: 7/10/2012.

URL: <http://help.arcgis.com/de/arcgisdesktop/10.0/help/index.html#na/004700000001000000/>.

— (2012e). *Konnektivität*. Stand: 7/10/2012.

URL: <http://help.arcgis.com/de/arcgisdesktop/10.0/help/index.html#na/004700000009000000/>.

— (2012f). *Netzwerkattribute*. Stand: 7/10/2012.

URL: <http://help.arcgis.com/de/arcgisdesktop/10.0/help/index.html#na/00470000000m000000/>.

— (2012g). *Start-Ziel-Kostenmatrix-Analyse*. Stand: 7/10/2012.

URL: <http://help.arcgis.com/de/arcgisdesktop/10.0/help/index.html#na/004700000004r000000/>.

— (2013). *Verbinden von Attributen einer Tabelle mit einer anderen Tabelle*. Stand: 9/12/2013.

URL: <http://resources.arcgis.com/de/help/main/10.1/index.html#/005s0000002q000000>.

Esri (o.J.a). *Algorithms used by Network Analyst*.

URL: http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Algorithms_used_by_Network_Analyst.

— (o.J.b). *ArcGIS Editor for OpenStreetMap*. Zugriff: 29.09.2015.

URL: <http://www.esri.com/software/arcgis/extensions/openstreetmap>.

Geofabrik.

— (2014). *Daten*. Zugriff:29.09.2015. Geofabrik GmbH Karlsruhe.

URL: <http://www.geofabrik.de/de/data/>.

Geoinformatik-Service (2001). *Netzwerk*. Stand: 29.10.2001.

URL: <http://www.geoinformatik.uni-rostock.de/einzel.asp?ID=1242>.

Google (o.J.). *Google Maps*. Zugriff: 29.09.2015.

URL: <https://www.google.de/maps>.

IQ, GIS (2012). *Von der Basemap zur Analyse: OpenStreetMap in ArcGIS (2/2)*. Stand: 16.08.2012.

URL: <http://gisiq.blogs.esri.de/2012/08/von-der-basemap-zur-analyse.html>..

KiTa.de (o.J.). *KiTa / Kindertagesstätte*. Zugriff: 29.09.2015.

URL: <http://www.kita.de/wissen/kinderbetreuung/kindertagesbetreuung/kita-kindertagesstaette>.

Lange, N. de (2013). *Geoinformatik in Theorie und Praxis*. Berlin [u.a.] (Springer Spektrum). MPI-Magdeburg.

OpenStreetMap (o.J.). *FAQs*. Zugriff: 29.09.2015.

URL: http://www.openstreetmap.de/faq.html#was_ist_osm.

Ramm F. und Topf, J. (2009). *OpenStreetMap - Die freie Weltkarte nutzen und mitgestalten*. Berlin (Lehmanns Media).

Teucher, M. (o.J.). „Modelle und Methoden der wissenschaftlichen Visualisierung von Geodaten“. In: *Visualisierung von frei verfügbaren Daten*. Universität Trier.

Turau, V. (2004). *Algorithmische Graphentheorie*. 2. Auflage. München, Wien (Oldenbourg Verlag).

Udelhoven, T. (2012). „Modul: Statistik I: Statistische Grundlagen für Bio- und Geowissenschaften“. Magisterarb. Universität Trier.

White, C. (2013). *Create a Network Dataset from OSM Data*. Stand: 4.11.2013.

URL: <https://github.com/Esri/arcgis-osm-editor/wiki/Create-a-network-dataset-from-osm-data>.

White, C. (2015). *ArcGIS Editor for OSM, 10.2.x Desktop*. Stand: 3. April 2015.

URL: <http://www.arcgis.com/home/item.html?id=16970017f81349548d0a9eead0ebba39>.

Wiki, OpenStreetMap (2015a). *Software comparison/Converting to WGS84*. Stand: 31.08.2015.

URL: http://wiki.openstreetmap.org/wiki/Software_comparison/Converting_to_WGS84.

Wikib, OpenStreetMap (2015b). *OSM XML*. Stand: 30.04.2015.

URL: http://wiki.openstreetmap.org/wiki/OSM_XML.

Wikic, OpenStreetMap (2015c). *Elements*. Stand: 30.08.2015.

URL: <http://wiki.openstreetmap.org/wiki/Elements>.

Zimmermann, E. (2010). „OpenStreetMap in ArcGIS: Automatisierte Datenaufbereitung zu Netzwerkanalysen“. Bachelorarbeit. Hochschule Karlsruhe - Technik und Wirtschaft, Fakultät für Geomatik, Fachbereich Kartographie und Geomatik.

9 Eidesstattliche Erklärung

Erklärung zur Bachelorarbeit

Hiermit erkläre ich, dass ich die Bachelorarbeit / Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die aus fremden Quellen direkt oder indirekt übernommenen Gedanken als solche kenntlich gemacht habe. Die Arbeit habe ich bisher keinem anderen Prüfungsamt in gleicher oder vergleichbarer Form vorgelegt. Sie wurde bisher nicht veröffentlicht.

Ort, Datum:

Unterschrift: