



TECHNISCHE
UNIVERSITÄT
WIEN

Diplomarbeit

Einsatzmöglichkeiten von nutzergenerierten
Mobilitätsdaten in der
Stadt- und Regionalforschung
gezeigt am Beispiel von Fahrrad-Tracking-Daten in Wien

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Diplom-Ingenieurs

unter der Leitung von

Univ.Ass. Dipl.-Ing. Robert Kalasek

E280 - Department für Raumentwicklung, Infrastruktur- und Umweltplanung

Fachbereich Stadt- und Regionalforschung

eingereicht an der Technischen Universität Wien

Fakultät für Architektur und Raumplanung

von

Florian Pühringer

Wien, am 28.03.2017

Kurzfassung

Das gesellschaftliche Leben wird derzeit vom technologischen Fortschritt im Bereich der mobilen Endgeräte geprägt. Niemals zuvor war die von der Weltbevölkerung produzierte Datenmenge größer als heute. Die verschiedensten Sensoren machen die Aufzeichnung von nahezu jeder menschlichen Interaktion möglich, was auch für Themenfelder aus den Regionalwissenschaften neue Entwicklungsmöglichkeiten eröffnet. Vor allem die Geokodierung - also die koordinative Verortung - räumlicher Phänomene und die Möglichkeit, große Stichproben von Nutzerverhalten direkt dynamisch zu erfassen, stellt eine interessante und innovative technologische Perspektive dar, Datengrundlagen für Forschungs- und Planungsprojekte zu sammeln. Vor diesem Hintergrund sucht diese Arbeit nach planerischen Anknüpfungspunkten, wobei vor allem die Analyse von auf Tracking-Daten basierenden Bewegungsmustern in den Vordergrund gestellt wird.

Der Zugang und das Verständnis von „Raum“ aus der Stadt- und Regionalforschung bilden den theoretischen Rahmen der vorliegenden Diplomarbeit. Die Auseinandersetzung mit dem Begriff und dessen Dimensionen ist eine der elementaren Voraussetzungen für räumliche Analysen. Daher soll auch der neue und zukunftsweisende Typ der Tracking-Daten aus den verschiedenen räumlichen Zugängen beleuchtet werden.

Im Kern der Arbeit folgt die praktische Auswertung und Analyse von Fahrrad-Tracking-Daten aus Wien, die von *Bike Citizens* - einem österreichischen Start-Up im Marktsegment von app-basierten Mobilitätslösungen - zur Verfügung gestellt wurden. Dabei handelt es sich um eine Fahrrad-Routing-App für Smartphones, die in der Lage ist, mit dem Fahrrad zurückgelegte Wege detailliert aufzuzeichnen. Die Analyse erfolgt mit in den Regionalwissenschaften etablierten Werkzeugen, allen voran geographischen Informationssystemen. Dabei sollen die Möglichkeiten und Potentiale für den Einsatz derartiger georeferenzierter Daten aufgezeigt, aber auch auf Grenzen in der Aussagekraft eingegangen werden. Anhand verschiedener räumlicher Maßstabsebenen wurden die Daten mithilfe von Hands-On Beispielen auf ihre planerische Eignung überprüft. Die Anwendungsfelder reichen dabei von Geschwindigkeitsanalysen einzelner Netzabschnitte über Auswertungen zu Routenwahl und Routensplit hin zu Einzugsbereichsanalysen von Points of Interest. Ableitend aus den Erkenntnissen werden verschiedene Aspekte hinsichtlich Aussagekraft und Datencharakteristik dargelegt, sowie Rahmenbedingungen und Einsatzmöglichkeiten aufgezeigt. Zudem werden die Tracking-Daten mit Beispielen klassischer Mobilitätsdatenerhebungen, wie *Österreich Unterwegs 2014*, verglichen. Der Fokus liegt dabei auf den Prozessen und der jeweiligen Datengüte sowie der räumlichen Aussagekraft.

Die Einordnung der Tracking-Daten in den Kontext der Stadt- und Regionalforschung, anhand der vorab definierten Raumzugänge, leitet zum Abschluss der Arbeit über. Durch die Eigenheiten der Tracking-Daten, wie beispielsweise deren starke Verräumlichung, ergeben sich innovative Nutzungsmöglichkeiten, welche auch im theoretischen Rahmen spannende Aspekte aufzeigen.

Abstract

Social life is currently strongly influenced by technological progress in the field of mobile devices. Never before was the amount of data produced by the world population greater than it is today. The various sensors provide the means for recording almost any human interaction. This also indicates new fields of application in urban and regional research. Geocoding spatial phenomena and the opportunity to capture large samples of user behavior directly represent an interesting and innovative technological perspective to collect data for research and planning projects. This work is looking for planning connections to point out meaningful *big data* implementations. The analysis focuses on movement patterns based on tracking data.

The approach and understanding of „space“ form the theoretical framework of urban and regional research. The analysis of this concept and its dimensions is one of the fundamental prerequisites for spatial analysis. Therefore, the future-oriented type of tracking data from the different spatial accesses should also be analyzed.

The core of the work is the practical evaluation and analysis of bicycle tracking data covering the Vienna city area. The data was provided by *Bike Citizens* – a bicycle routing app for smartphones, which is able to record cycling trajectories in detail. The analysis is using tools established in urban and regional research, primarily geographic information systems. In this context, the potentials for the use of this kind of data in urban and regional research are to be pointed out, but also the limitations in the meaningfulness will be discussed. At different spatial levels, the data were examined by means of hands-on examples for their planning suitability. These examples range from speed analyses of individual network sections via evaluations of route choices to accessibility analyzes of points of interests. Concluding the findings of this exploration and analysis, various aspects of meaningfulness and data characteristics are discussed. In a next step, the tracking data are compared with previously shown examples of mobility data surveys, such as „Österreich Unterwegs 2014“. The focus is on the processes and the respective data quality.

The classification of the tracking data in the context of urban and regional research, based on the predefined different spatial concepts, leads to the conclusion of the work. Due to the characteristics of the tracking data, such as their strong spatialization, innovative possibilities of use arise that can also lead to interesting insights for the underlying theoretical framework.

Inhaltsverzeichnis

Kurzfassung	2
Abstract	3
1 Einleitung	9
1.1 Motivation	11
1.2 Ziel der Arbeit - Forschungsfragen	12
1.3 Aufbau der Arbeit	13
2 Raumverständnis	15
2.1 Physisch-materieller / geographischer Raum	16
2.2 Relationaler Raum	17
2.3 Sozialer Raum	18
3 Nutzergenerierte Daten	20
3.1 Allgemeine Bedeutung von Daten in der Stadt- und Regionalforschung	20
3.2 Grundprinzip der nutzergenerierten Daten	22
3.3 Verräumlichte nutzergenerierte Inhalte	22
3.4 Datenzugänglichkeit - Open(Geo-)Data	24
3.5 Technische Grundlagen - GPS-Tracking	27
3.6 Datenschutz und rechtliche Rahmenbedingungen	29
4 Mobilitätsdatenerhebung - State of the Art und aktuelle Entwicklungen	33
4.1 Von der aktiven Befragung... - Klassische Mobilitätsdatenerhebung gezeigt am Beispiel Österreich Unterwegs 2014	34
4.1.1 Hintergrund und Ziele	34
4.1.2 KOMOD-Studie	34
4.1.3 Methodik der Erhebung	35
4.1.4 Methodik der Datenaufbereitung	37
4.1.5 Ergebnisse	38
4.2 ... zur passiven Beobachtung - Neue Formen der Datenerhebung	40
4.2.1 Technologieeinsatz bei Mobilitätserhebungen	40
4.2.2 Konkrete Beispiele	42
5 Einsatzmöglichkeiten nutzergenerierter Mobilitätsdaten - Ein Hands-On Beispiel	47
5.1 Allgemeines zum Radverkehr in Wien	47
5.2 Methodische Vorgehensweise	49
5.3 Der Bike Citizens Datensatz	51
5.4 Datenaufbereitung und -verarbeitung	54
5.4.1 Map Matching	56
5.4.2 Inhaltliche Anreicherung	61

5.4.3	Benutzte Software	63
5.5	Daten Check-Up	64
5.5.1	Deskriptive Merkmale	64
5.5.2	Repräsentativität	65
5.5.3	Datengüte	70
5.5.4	Unterscheidung nach Fahrtzweck	71
5.6	Makro-Ebene	74
5.6.1	Visualisierung	74
5.6.2	Radverkehr und Wetter	76
5.6.3	Radrouten	77
5.7	Meso-Ebene	82
5.7.1	Routensplit	82
5.7.2	Einzugsbereichsanalyse	85
5.7.3	Verkehrsspinne	89
5.8	Mikro-Ebene	91
5.8.1	Geschwindigkeitsanalyse	91
5.8.2	Analyse von Kreuzungsbereichen	93
5.9	Fazit und Diskussion	97
5.9.1	Defizite	97
5.9.2	Potentiale	99
6	Tracking-Daten in der Stadt- und Regionalforschung - Resümees	101
6.1	Vergleich der Datencharakteristik	101
6.2	Bewertung der Einsatzmöglichkeiten von Tracking-Daten	103
6.3	Rahmenbedingungen für den Einsatz Tracking-Daten	106
6.4	Tracking-Daten aus raumtheoretischer Perspektive	107
6.5	Ausblick	110
7	Literaturverzeichnis	111
8	Abbildungsverzeichnis	115
9	Tabellenverzeichnis	117
10	Anhang	118

Besten Dank an

Robert Kalasek, für die gute Betreuung sowie für die zahlreichen kritischen und gleichzeitig produktiven Inputs.

Andreas Friedwagner und das restliche Verracon Team, für die Idee zum Thema, die Kontakt-herstellung zu Bike Citizens und die vielen Anregungen.

das *Bike Citizens* Team aus Graz, für den zur Verfügung gestellten Datensatz sowie an Steffen Wirth für die angenehme Kommunikation.

Rudolf Giffinger, für die Idee zur theoretischen Einbettung des Themas.

Johannes Suitner, für die Ratschläge zur Gliederung der Arbeit.

Martin Berger, für die Einschätzung der verkehrsplanerischen Relevanz der Daten.

meine Familie und insbesondere meine Eltern, für das Ermöglichen des Studiums und die Unterstützung in den letzten Jahren.

Sonja, für den Rückhalt während der Ausarbeitung und die zahlreichen Vorschläge zu Korrekturen.

alle weiteren Freunde und Kollegen, für die spannenden und hilfreichen Diskussionen zum Thema.

digital traces.



Aus Gründen der besseren Lesbarkeit wird in der vorliegenden Arbeit bei Begriffen auf die Anführung von maskulinen und femininen Formen verzichtet. Die maskuline Schreibweise, auch jene des Plurals, bezieht stets Frauen und Männer mit ein.

Hintergrundbild Seite 7: [Eigene Darstellung / Bike Citizens Tracks]

1 Einleitung

Der digitale Wandel hat die Gesellschaft in den letzten Jahren nachhaltig verändert - das Smartphone ist zum ständigen Begleiter geworden. Es hilft Wege zu finden, Öffnungszeiten zu ermitteln, Neuigkeiten mitzuteilen und über das Leben der Mitmenschen immer am Laufenden zu bleiben. Der Nebeneffekt, dass bei all diesen Aktionen durch die in Smartphones verbauete Vielzahl an Sensoren Daten generiert, aufgezeichnet sowie weiterverarbeitet werden und diese Daten mittlerweile mächtige Informationsquellen geworden sind, wurde nicht nur von Weltkonzernen bereits erkannt. Auch in der lokalen Start-Up-Szene, einem Begriff, der ebenfalls in den letzten Jahren an Bedeutung gewonnen hat, haben sich derartige datenbasierte Geschäftsmodelle bereits etabliert.

Niemals zuvor war die von der Weltbevölkerung produzierte Datenmenge in den verschiedensten Bereichen größer als heute. Die Menge der digitalen Daten lag schon im Jahr 2013 bei ungefähr 4,4 Zettabytes - wobei sich diese Zahl etwa im Abstand von zwei Monaten verdoppelt und somit im Jahr 2020 bereits 44 Zettabytes groß sein wird.¹ Auch der Trend rund um *Internet of Things* hat die Entwicklung weiter verstärkt. Er bezeichnet dabei ein Konzept, bei dem immer mehr miteinander verknüpfte, intelligente Sensoren dem Menschen in allen erdenklichen Lebenssituationen zur Seite stehen und ihm helfen, die optimale Lösung oder den optimalen Weg zu finden.² Dabei ist es für die Funktionsweise der jeweiligen Anwendung meist nötig, die automatisch entstandenen Daten zumindest kurzfristig abzuspeichern.

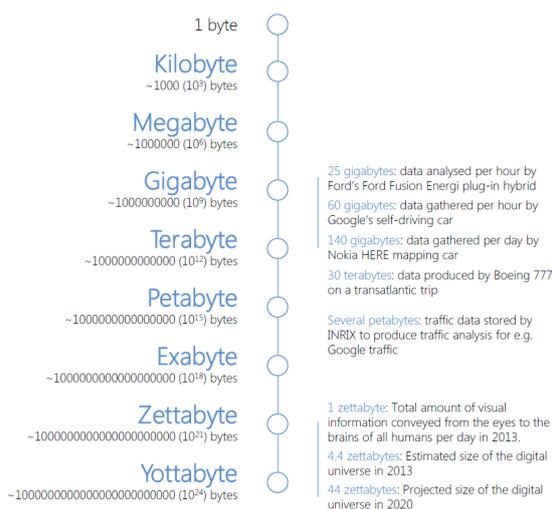


Abbildung 1.1:
Übersicht der Datenmengen zur besseren Einschätzung der genannten Größen. [ITF u. OECD, 2015, S.10]

Ein Großteil dieser unvorstellbaren Menge an Daten lässt dabei Rückschlüsse auf das menschliche Verhalten zu, wie beispielsweise WiFi Anmeldungen, Nutzungsdaten von Haushaltsgeräten, Medien auf sozialen Netzwerken oder auch Daten, die von modernen Fahrzeugen generiert und gespeichert und in der Regel auch an Unternehmen übermittelt werden. Vor allem im Bereich des motorisierten Individualverkehrs sind Sensoren nicht nur in modernen Fahrzeugen längst Standard, auch Autobahnen, Kreuzungen oder Parkhäuser „sammeln“ Daten, welche in einem nächsten Schritt analysiert werden um mögliche Erkenntnisse daraus zu erlangen. So entwickelt beispielsweise der US-Konzern Inrix Modelle und Simulationen, bei denen Daten aus Fahrzeugen für eine intelligente Stra-

¹vgl. [ITF u. OECD, 2015, S.9]

²vgl. [Freier u. Marco, 2014, S.5]

ßensystemplanung in Echtzeit herangezogen werden. Dies beschränkt sich aber nicht nur auf die Position oder die Geschwindigkeit des Fahrzeugs, sondern reicht von der Scheibenwischeraktivität und dem Status der Lichtanlage für Echtzeitprognosen des Wetters bis zur automatischen Navigation zum nächsten Parkplatz am Ziel, den das System aufgrund eines gerade wegfahrenden Fahrzeugs soeben identifiziert hat. Das Resultat aus immer leistungsfähigeren und günstigeren Endgeräten sowie Sensoren verbunden mit der Möglichkeit, große Mengen an Daten mit geringen Kosten zu speichern und auch mit hohen Bandbreiten zu übermitteln, sind die Grundbausteine der Konzepte rund um den Begriff *Big Data*. Die Technologie ist dabei soweit, dass die Daten oft schon in Echtzeit - wie anhand des Beispiels von Inrix - weiterverarbeitet werden können.³

Anders als Anwendungen im Bereich von *Internet of Things* kann die Datengenerierung mithilfe von Smartphones als überwiegend bewusst beschrieben werden. Der Nutzer entscheidet selbst, welche Anwendungen er weshalb installiert. Die Verbreitung von Smartphones schreitet zügig voran und das klassische „Handy“ wird sukzessive ersetzt. So nutzten laut dem Mobile Communications Report 2015⁴ bereits 86% aller Handynutzer (dies sind in Österreich 77% der Gesamtbevölkerung) ein Smartphone. Zusätzlich geben 36% der mobilen Internetnutzer an, dass dessen Nutzung mittlerweile fest in ihrem Leben integriert sei. 52% nutzen mobiles Internet bewusst und selektiv - 11% stehen der mobilen Internetnutzung kritisch gegenüber. Die Nutzung als Telefon ist bei vielen Smartphone Besitzern zur Nebensache geworden. Mit Ausstattungsmerkmalen, die in ihrer Leistung teilweise schon mit einfachen Notebooks oder PCs vergleichbar sind, bietet sich eine unüberschaubar große Anzahl an Anwendungsmöglichkeiten. Die Geräte besitzen dabei nicht nur immer schnellere Prozessoren, auch etliche Mikrosensoren sind bei vielen Modellen schon fix verbaut: Barometer, Beschleunigungssensor, Bluetooth, Elektromagnetischer Sensor, Fingerabdrucksensor, GPS, Gyroskop, Helligkeitssensor, Magnetometer, Mikrofon, Näherungssensor, NFC oder Pulsmesser.⁵

All diese Prozesse werden auch als *digitale Transformation* betitelt: Eine Entwicklung, die einen gesellschaftlich nachhaltig-wirksamen Wandel beschreibt, dem auch Unternehmen und Behörden ausgesetzt sind und daher versuchen sich anzupassen.

Spätestens wenn es um georeferenzierte Daten geht, stellt sich die Frage nach einer möglichen Verschmelzung dieser Entwicklung mit der Raumplanungswelt. „Raumplanung ist seit ihren Anfängen auf Datengrundlagen angewiesen, die es ihr ermöglichen, den Raum und seine Dynamik soweit zu verstehen, dass steuernde Eingriffe hinsichtlich Wirksamkeit und Wirkung abgeschätzt werden können. Es gibt in Raumplanungsarchiven zahlreiche Dokumente, die die mühselige Datenbeschaffung noch vor der Zeit automatisierter Verkehrszählungen, der Fernerkundung oder der elektronischen Aufbereitung von Volkszählungen belegen.“⁶ Die Bedeutung von Datengrundlagen hat sich in der Planung im Laufe der Zeit kaum verändert, jedoch haben sich die Erhebungsmethoden sowie die konkrete Ausprägung und Charakteristika immer weiterentwickelt. Auch die dargestellte Entwicklung des digitalen Wandels durch *Big Data* wird hier möglicherweise Prozesse revolutionieren können. Vor allem die innovative Möglichkeit, große Stichproben von tatsächlich beobachteten räumlichen Interaktionen zur Verfügung

³vgl. [ITF u. OECD, 2015, S.10]

⁴vgl. [MobileMarketingAssociation, 2016, 26]

⁵Ein Überblick der Sensoren inklusive detaillierter Erklärungen findet sich auf: <http://www.zeit.de/digital/mobil/2014-05/smartphone-sensoren-iphone-samsung/>

⁶[Koll-Schretzenmayr, 2013, S.118]

zu haben, könnte die Grundlagen der Planung und damit auch deren Ergebnisse nachhaltig verbessern. Diese Daten ermöglichen dabei Anwendungen, die auf Basis von konventionellen statistischen Erhebungen nicht durchführbar wären.

Die vorliegende Arbeit beschäftigt sich mit der Verschmelzung von *Big Data*, digitaler Transformation und der Raumplanungswelt und erforscht, inwieweit ein großer geokodierter Datensatz im planerischen Kontext neue Einblicke und innovative Analysen ermöglicht. Kern dieser Diplomarbeit ist ein Hands-On Beispiel, in dem Fahrrad-Tracking-Daten als Beispiel für nutzergenerierte Mobilitätsdaten mithilfe einer maßstabsbasierten Ebenenbetrachtung auf ihre Einsatzmöglichkeiten hinsichtlich Fragestellungen aus der Stadt- und Regionalforschung untersucht werden. Der zur Verfügung gestellte Datensatz stammt von einem österreichischen Start-Up im Marktsegment von app-basierten Mobilitätslösungen - *Bike Citizens* aus Graz. Neben den Einsatzmöglichkeiten der Daten zielt die Arbeit auf das Aufzeigen von Grenzen und Rahmenbedingungen ab. Durch eine Vielzahl an praxisorientierten Analysen wird einerseits die Eignung als Datenbasis für die Forschung und Planung dargestellt und andererseits Knowhow für den Einsatz zukünftiger *Big Data*-Anwendungen geschaffen.

1.1 Motivation

Die Idee zum Thema der Diplomarbeit habe ich Mag. Andreas Friedwagner zu verdanken. Er ist Geschäftsführer bei Verracon GmbH, einem Planungsbüro in Wien, bei dem ich seit Frühjahr 2015 als studentischer Mitarbeiter tätig bin. Die Idee beruhte auf der Überlegung, dass immer größere Datenmengen, vor allem durch mobile Endgeräte, erzeugt werden und verfügbar sind und diese Daten auch für die Planung von bisher unterschätzter Relevanz sein könnten.

Auch aufgrund meines persönlichen Interesses in den Themenbereichen mobile Technologien und App-basierte Lösungen konnte ich mich schnell mit dem Thema identifizieren und fasste den Entschluss, ein Konzept in diese Richtung zu erarbeiten. Mag. Andreas Friedwagner konnte wenige Zeit später über die Mobilitätsagentur Wien und Dipl.-Ing. Martin Blum den Kontakt zu *Bike Citizens* nach Graz herstellen, die mit der Zusage, einen Datensatz zur Verfügung zu stellen, das Projekt endgültig auf Schiene brachten. Mit diesen Daten konnte der empirische Teil der Arbeit vollzogen werden, wodurch auch mein persönlicher Wunsch erfüllt wurde, mein in der studentischen Laufbahn entdecktes großes Interesse an geographischen Informationssystemen in die Arbeit zu integrieren.

1.2 Ziel der Arbeit - Forschungsfragen

Die Arbeit untersucht, inwieweit nutzergenerierte Mobilitätsdaten in der Stadt- und Regionalforschung als Datengrundlage dienen können. Als Beispiel für diesen Datentyp dienen zur Verfügung gestellte Fahrrad-Tracking-Daten. Der Schwerpunkt der Auswertung liegt nicht in der Schaffung von empirisch belegten Ergebnissen von Detailfragen sondern in einem Aufzeigen von Anwendungsfeldern.

1. Welche Einsatzmöglichkeiten von Fahrrad-Tracking-Daten in der Stadt- und Regionalforschung können aus den Hands-On Beispielen abgeleitet werden?

Die Entscheidung, Einsatzmöglichkeiten aufzuzeigen, begründet sich mit den Eigenschaften des zur Verfügung gestellten Datensatzes, der in der Raumplanungswelt ob der Ausprägung seines Inhalts durchaus als innovativ bezeichnet werden darf. Die evidente Beantwortung konkreter Fragestellungen mithilfe der Daten hätte als Vorbereitung einer intensiven Auseinandersetzung der Eigenheiten und Eigenschaften der Fahrrad-Tracking-Daten bedurft - diese Abhandlung füllt nun einen Großteil des Inhalts dieser Arbeit.

2. Welche Eigenheiten und Rahmenbedingungen müssen beim Einsatz der Fahrrad-Tracking-Daten in der Stadt- und Regionalforschung beachtet werden?

Eine Auseinandersetzung mit den Charakteristika und Eigenschaften der Daten ist wichtig, um mögliche Defizite in der Aussagekraft zu entdecken. Diese Frage soll einerseits aus den Ableitungen der Ergebnisse des Hands-On Beispiels und andererseits durch den Vergleich mit anderen Mobilitätsdaten(-erhebungen) beantwortet werden. Ziel ist es hier, klare Unterschiede in der Datengüte und Datencharakteristik herauszuarbeiten, um eine Einordnung des zur Verfügung gestellten Datensatzes zu vollziehen.

3. Wie können die Fahrrad-Tracking-Daten aus unterschiedlichen raumtheoretischen Perspektiven von Nutzen sein?

Ein wesentliches Element in der Regionalwissenschaft sind die Zugänge und Verständnisse rund um den Begriff des Raums. Das Bewusstsein dieser Zugänge und das Wesen des Raumes zu erfassen, sind elementare Bausteine in komplexen, räumlich-analytischen Fragestellungen.

Die Eigenschaften, Ausprägungen und Anwendungsarten der Fahrrad-Tracking-Daten werden anhand der in Kapitel 2 definierten Raumzugänge und -verständnisse eingeordnet. Aufgrund der mit dem Datentyp verbundenen, innovativen Möglichkeit, räumliche Interaktionen direkt durch nutzergenerierte Daten abzubilden, können diese auch im theoretischen Rahmen spannende Aspekte aufzeigen.

1.3 Aufbau der Arbeit

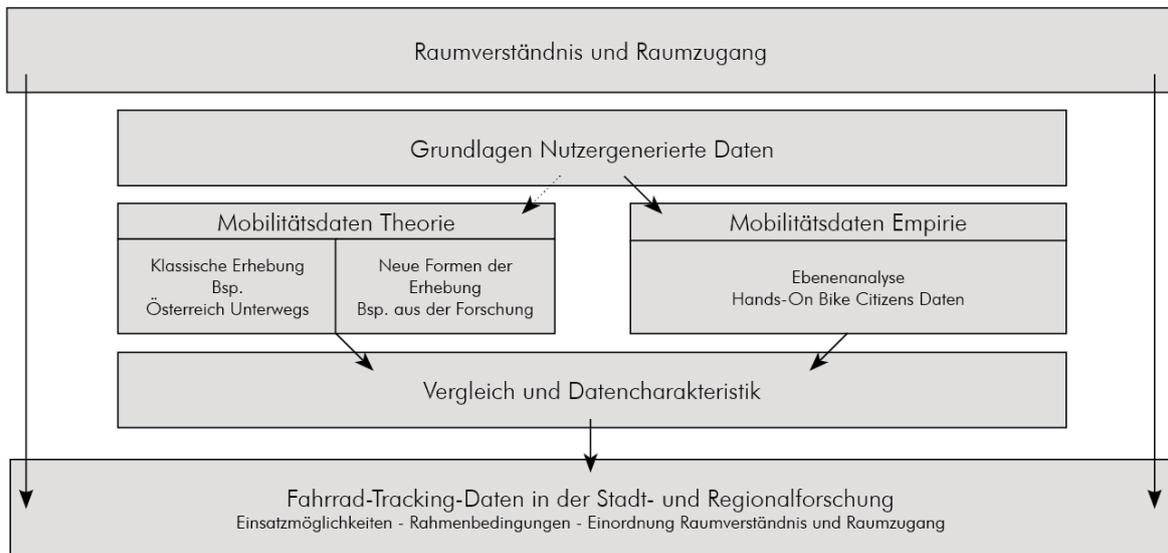


Abbildung 1.2: Aufbau der Arbeit. Als Beispiel für Nutzergenerierte Daten dienen die zur Verfügung gestellten Fahrrad-Tracking-Daten. [Eigene Darstellung]

Die Arbeit gliedert sich im Wesentlichen in fünf inhaltliche Blöcke. Abgestimmt auf das Ziel der Arbeit soll vorab das Raumverständnis und der Raumzugang in der Theorie erklärt und definiert werden. Die raumtheoretischen Perspektiven bilden dabei den Rahmen der Arbeit.

Im zweiten Block wird mit der Welt der nutzergenerierten Daten die zweite wichtige Säule der Arbeit beleuchtet. Dabei wird vorab auf die generelle Bedeutung von Daten für die Regionalwissenschaften eingegangen. Danach werden neben Konzepten zur Datengenerierung durch Nutzer auch Aspekte wie die Zugänglichkeit der Daten - vor allem hinsichtlich der aktuellen *Open Data*-Bewegung - betrachtet. Vorbereitend auf das Hands-On Beispiel im Kern der Arbeit werden technische Grundlagen zur GPS-gestützten Datenerfassung und datenschutzrechtliche Rahmenbedingungen dargelegt.

Im Verlauf der Arbeit soll der Fokus immer weiter auf den Sektor der Mobilität verlagert werden. Diese Entscheidung begründet sich auf den für die praktische Auswertung zur Verfügung gestellten Fahrrad-Tracking-Daten aus Wien und der Tatsache, dass im Bereich der Mobilität bereits ein Grundstock an Literatur und Forschung zu neuen Formen der Datenerfassung und -verarbeitung vorhanden ist. Die Ergebnisse sollen anhand der Ergebnisse des Hands-On Beispiels als auch durch Verweise und Zitate von Fachliteratur gestützt werden. Ein weiterer Grund für diese Entscheidung ist die in den letzten Jahren stattgefundene Entwicklung im Bereich der Mobilität, welche vor allem die Städte vor große Herausforderungen stellt. Die Erarbeitung von Mobilitätslösungen ist somit auch in der Stadt- und Regionalforschung ein gegenwärtiges Anwendungsgebiet. Durch die technischen Entwicklungen ergeben sich neue analytisch-methodische Herausforderungen aber auch Erkenntnispotentiale.

Der dritte Block beinhaltet nähere Erläuterungen zu Erhebungen und Mobilitätsdaten. Mit *Österreich Unterwegs* aus dem Jahr 2014 soll ein relativ aktuelles State of the Art Beispiel für eine umfassende repräsentative Datenerhebung gezeigt werden. Auch dieses hat keinen Anspruch, die gesamte Erhebung im Detail darzulegen. Es geht um die Vermittlung der grundsätzlichen Abläufe, um einen kurzen Einblick in die Prozesse und Eigenheiten der Erhebung zu bekommen. Im Kontrast dazu folgen Einblicke in neue Entwicklungen im Bereich der Datenerhebung, sowohl auf der Prozessebene als auch durch die kurze Präsentation ausgewählter Forschungsprojekte.

Im Hauptteil der Arbeit folgt die empirische Auswertung von nutzergenerierten Daten, welche anhand eines Hands-On-Blocks gezeigt wird. Dabei geht es weniger um die Vermittlung konkreter Ergebnisse, der Leser soll vielmehr ein Gefühl dafür bekommen, welche Aufgabenstellungen mit den Daten lösbar sind und wo es zu eventuellen Herausforderungen kommen könnte. Bei den zur Verfügung gestellten Daten handelt es sich um nutzergenerierte Mobilitätsdaten aus Wien, genauer georeferenzierte Fahrrad-Tracking-Daten, die mittels der App *Bike Citizens* in regelmäßigen kurzen Intervallen aufgezeichnet werden. Die Methodik der Hands-On-Analyse wird später noch im Detail vorgestellt.

Der letzte Block hat zum Ziel, die Erkenntnisse rund um die Eigenschaften und Charakteristika der Daten aus dem Hands-On Beispiel anderen Erhebungen gegenüberzustellen und sie in ihren Merkmalen und Attributen zu unterscheiden sowie in weiterer Folge die Forschungsfragen zu beantworten. Es soll genauere Erkenntnis darüber gewonnen werden, wie der Typ der *Bike Citizens* Daten aus regionalwissenschaftlicher Sicht einzuordnen ist.

Kurzfasit

Für die Arbeit wichtige Erkenntnisse werden in einer Box am Kapitelende zusammengefasst.
--

2 Raumverständnis

Der Zugang zum Raumverständnis kann als eine der elementaren Voraussetzungen der Stadt- und Regionalforschung gesehen werden. Dies legt schon der Inhalt der Disziplin nahe, beschäftigt sich die Forschung doch meist mit Phänomenen, die in verschiedener Art und Weise auch verräumlicht sind. Nicht selten ist der „Raum“ auch selbst von Erkenntnisinteresse.

Die Entscheidung, diese Thematik gleich zu Beginn der Arbeit zu besprechen, begründet sich mit dem vorbereitenden Verständnis der später dargestellten Inhalte zu Nutzergenerierten Daten und Mobilitätsdatenserhebungen. Zudem soll dies die Relevanz des Raumverständnis für die Einordnung der Ergebnisse betonen. Anzumerken ist auch, dass je nach Fachdisziplin innerhalb der Raumplanung andere Sichtweisen des Raums nötig und vorhanden sind. Durch die Betrachtungsweise der unterschiedlichen Raumverständnisse kann später die Einordnung und Beurteilung der Daten für die jeweilige Disziplin leichter erfolgen.

Die Frage nach einer Definition von Raum und der Klarstellung, was dieser Begriff eigentlich meint und bedeutet, ist nicht neu und gilt auch im Bereich der Geographie als elementar. Die Frage wird unter anderem im Werk von Peter Weichhart „Entwicklungslinien der Sozialgeographie“ aus dem Jahr 2008 intensiv diskutiert, indem er seine langjährigen Erfahrungen und wissenschaftlichen Beiträge kompakt zusammenfasst. Wichtig ist Weichhart die Umformulierung der Frage nach dem Raum. „Wenn wir einfach fragen: Was ist Raum?, dann riskieren wir durch diese sprachrealistische Fragehaltung, dass die Antwort schlicht in einer metaphysischen Spekulation besteht. [...] Es ist daher vernünftiger, schlicht und einfach die Verwendungsweisen von Begriffen zu analysieren.“⁷

Er entscheidet sich aus diesem Grund gegen die Beantwortung der konkreten Frage nach einer Definition des Raums und führt stattdessen die verschiedenen Zugangsweisen, Deutungen und Wahrnehmungen von Raum im Rahmen einer Bestandsaufnahme auf. Sein Ziel ist es, aufbauend auf den sehr unterschiedlichen Bedeutungsweisen, eine Art Inventar zu erstellen - also eine Übersicht der verschiedenen Zugänge zum Raum, wie es am Anfang dieser Arbeit schon als elementarer Inhalt gefordert wurde. Weichhart sammelt insgesamt sieben verschiedene Zugänge und setzt sich dabei intensiv mit ihren Bedeutungen und Wechselwirkungen auseinander. Zugunsten einer leichteren Einordnung der Ergebnisse dieser Diplomarbeit und der Übersichtlichkeit für den Leser werden diese sieben Zugänge auf drei wesentliche zusammengefasst. Diese sind:⁸

der physisch-materielle Raum, der relationale Raum und der soziale Raum.

⁷[Weichhart, 2008, S.76]

⁸nach [Giffinger, 2016]

Im Folgenden werden die Charakteristika und Sichtweisen der unterschiedlichen Raumzugänge kurz erklärt und Anwendungsfelder anhand von Beispielen aus den Regionalwissenschaften dargelegt. Außerdem werden mithilfe von Abbildungen aus den unterschiedlichen Raumzugängen die Erklärungen bildhaft unterstützt, um die raumtheoretischen Konzepte möglichst greifbar darzustellen.

2.1 Physisch-materieller / geographischer Raum

Der physisch-materielle Raum ist wahrscheinlich der ontologisch greifbarste Raumtyp. Er dient als Behältnis und ist Träger von Eigenschaften wie Strukturen und anderen Charakteristika. Der Raum dient dabei „als Häferl“⁹, in das man etwas hineingeben kann. Für Newton war dieser Raum eine Art dreidimensionaler Container, in den alles materielle eingebettet ist.¹⁰ Diese Ansicht und Definition von Raum hat eine lange Historie. Ergänzend zum physisch materiellen Raum soll im selben Verständnis auch der Begriff des geographischen Raums zum Einsatz kommen. Dieser meint nicht mehr als eine Gebietsabgrenzung oder eine Adressangabe. Je nach Typus kann diese eher ungenau (zum Beispiel Mittelmeerraum, Gebirgsraum) bis sehr genau (exakte Koordinaten) sein. Für Weichhart ist das Charakteristische am geographischen Raum, dass es sich um einen Ausschnitt der materiellen Welt handelt.

In den Regionalwissenschaften spielt dieser Raumtyp eine Rolle, wenn es beispielsweise um Entfernungen, Abstände oder oberflächenbezogene Eigenschaften geht. Er ist die Voraussetzung für Positionierung auf der Erdoberfläche, Grundbuchfestlegungen, Trassen- und Grenzverortungen oder auch für die Navigation¹¹. Als Behältnis ist er Träger von Eigenschaften. Die materielle Komponente zeigt dabei sowohl quantitative als auch qualitative Ausprägungen.

Im Sinne der Mobilitätsforschung ist der Behältnis-Raumtyp zweckmäßig etabliert, da viele Daten nur auf aggregierter Basis innerhalb bestimmter Grenzen - wie beispielsweise Bezirken oder Zählsprenkeln - vorhanden sind. Die Bewegungen werden dabei aggregiert für verschiedene Behältnisse dargestellt. Als Beispiel hierfür dient die Pendlerstatistik: Zwar ist bekannt, wieviele Personen von Gemeinde A in die Gemeinde B pendeln, jedoch nur auf Basis einer Matrix ohne konkrete räumliche Information. Dazu wird eine politische Karte mit Verwaltungsgrenzen - den Behältnissen oder „Häferln“ - mit den Informationen der Matrix gefüllt.

⁹vgl. [Weichhart, 2008, S.77]

¹⁰vgl. [Weichhart, 2008, S.77]

¹¹vgl. [Giffinger, 2016]

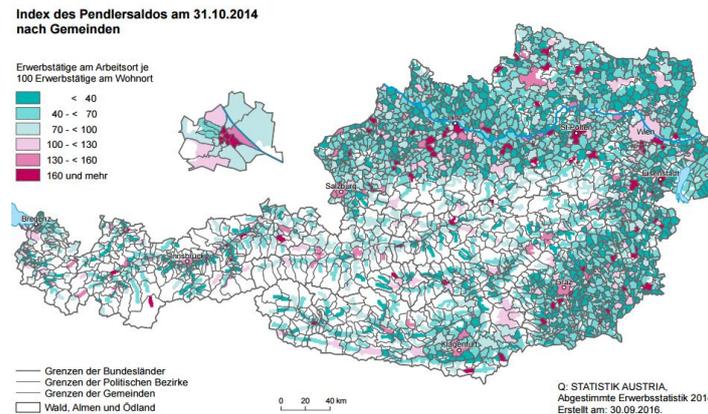


Abbildung 2.1: Pendlersaldo als Beispiel für eine Darstellung des physisch-materiellen / geographischen Raums.
[Statistik Austria]

2.2 Relationaler Raum

Der relationale Raum gestaltet sich im Vergleich schon deutlich abstrakter - seine Bedeutung ist aber gerade im Bereich der Mobilitätsforschung für das Raumverständnis unerlässlich. Bei diesem Konzept steht der Raum nicht für etwas materiell-existierendes, sondern für „immaterielle Relationen und Beziehungen oder für etwas Gedachtes.“¹² Dieser Typ Raum charakterisiert sich als eine Ordnungsrelation ko-existierender Dinge (Ausgehend von Leibnitz) und kommt ohne Container (Häferl) aus. Der Raum stellt dabei eine Struktur dar, innerhalb derer Dinge ein- und angeordnet werden können. Das zugrunde liegende Raster ist dabei vom jeweiligen Betrachter frei gewählt und dient letztendlich dazu, Dinge in geordnete Strukturen zu bringen. „Das Konzept geht von der Vorstellung aus, dass Raum ausschließlich durch die Beziehungen und die Relationalität der physisch-materiellen Dinge zueinander konstituiert wird. Wenn also aus einem solchen Raum die Dinge herausgenommen werden, dann bleibt schlicht und einfach gar nichts übrig. Denn der Raum entsteht erst durch die zwischen den Dingen und Körpern existierenden Lagerrelationen. Ohne Dinge gibt es keinen Raum.“¹³

Konkrete Beispiele aus der Stadt- und Regionalforschung für physisch-materiellen Kontext im relationalen Raum sind Erreichbarkeitsanalysen, Zeitaufwände, Transportkosten, oder Topologien. Relationales Raumverständnis ist unabdingbar bei Studien im Umfeld des Bereichs der Mobilität. Es ist zudem Voraussetzung für alle (geographischen) Informationssysteme.

Als Beispiel für eine relationale Raumabbildung soll hier eine Darstellung des Flugverkehrs der Vereinigten Staaten dienen. Gut zu erkennen sind dabei die Verbindungen (Relationen) zwischen den Metropolen der USA. Je intensiver die Farbdarstellung, desto mehr Verbindungen sind vorhanden. Entfernt man aus dieser Darstellung die Relationen, dann bleibt - auch gemäß den Rahmenbedingungen des relationalen Raums - nichts übrig. Bei derartigen Darstellungen verwendete Verwaltungsgrenzen dienen dabei rein der Orientierung des Betrachters, nicht der Darstellung des Inhalts.

¹²vgl. [Weichhart, 2008, S.78]

¹³[Weichhart, 2008]



Abbildung 2.2: Flugverkehrskarte als Beispiel für ein Abbild von Relationen im Raum. [Wildfiretoday, 2017]

Weichhart geht aber weiter und untermauert, dass dieses Verständnis von Raum (vielleicht unbewusst) auch gemeint ist, wenn Fußballtrainer oder Theaterregisseure von Raum sprechen. Diese Vermutung macht er bildhaft an Beispielen fest:

Regisseure wollen durch die „Raumgestaltung einen Spannungsbogen aufbauen“ (also einen Effekt durch die Anordnung von Dingen im Raum kreieren) oder ein Fußballtrainer meinte, sein Team habe nur aufgrund des „besseren Raumaufbaus“ (also den Relationen zwischen Spielern, Ball und Toren) gewonnen. Von Relevanz ist also nicht nur das Vorhandensein von Relationen, sondern auch deren Anordnung zueinander.¹⁴

Es ist schon hier vorwegzunehmen, dass dieser Typus vor allem im empirischen Teil der Arbeit für das tiefere Verständnis und die abstrahierende Deutung von besonderer Bedeutung sein wird.

2.3 Sozialer Raum

Um das Raumverständnis und die Zugänge zu komplettieren, ist die Einbeziehung des Sozialen in ein Konzept unumgänglich. Weichhart bezeichnet diesen Typ als sozial konstituierten und konstruierten Raum. Dieser wird durch Handlungspraxis begründet und stellt in mehrfacher Hinsicht ein gesellschafts- und entwicklungsbedingtes soziales Konstrukt dar.¹⁵ Dieses ist gleichzeitig Ressource und Constraint.¹⁶ Der relationale Raum dient für den sozialen Raum aber oftmals als Bezugsbasis und steht somit in enger Verbindung mit letzterem. Aber auch intensive Wechselwirkungen zwischen physisch-materiellem und geographischem Raum sind nicht von der Hand zu weisen. Die Ausstattung des „Häferls“ und die sozialen Interaktionen beeinflussen sich gegenseitig maßgebend.

Für Weichhart ist auch der Prozess der Regionalisierungen ein Beispiel dieser Wechselwirkungen. Als Beispiel für normativ-politische Regionalisierungen zeigt er anhand des Beispiels des

¹⁴vgl. [Weichhart, 2008, S.79]

¹⁵vgl. [Weichhart, 2008, S.326]

¹⁶vgl. [Giffinger, 2016]

Salzburger Raumordnungsgesetzes 1977, dass „bestimmte Normen Gültigkeit besitzen, welche die Praxis der produktiv-konsumtiven Regionalisierung beeinflussen.“ So ist die heutige Problematik der Zersiedlung maßgeblich auf eine Ausnahmegenehmigung für das Bauen im Grünland des genannten Raumordnungsgesetzes zurückzuführen.¹⁷

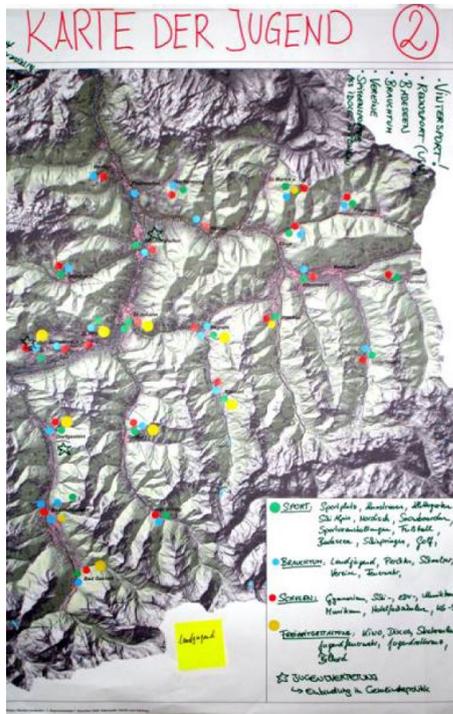


Abbildung 2.3:
„Karte der Jugend“ - [Unterlagen VO
Forschungs- und Planungsdesign 2017]

Als konkretes Beispiel für eine Darstellung des sozialen Raums soll ein Zwischenergebnis aus einem Workshop zum regionalen Entwicklungskonzept Pongau dienen. Sogenannte Akteurslandkarten sind eine etablierte Form in Workshopprozessen und helfen, aktors- bzw. zielgruppenspezifische, räumliche Sichtweisen zu verstehen. Dargestellt ist hier die „Karte der Jugend“ - in der für die Jugend wichtige Punkte der Region von den Workshopteilnehmern verortet wurden. Die Analyse lässt sich für viele weitere regionale Akteursgruppen wiederholen. Die Überlagerung dieser Ergebnisse ergibt dann für mehrere Gruppen wichtige Orte in der Region. „Gesellschafts- und entwicklungsbedingte Konstruktion des Raums bedingt aus sozialen Interaktionen ist eine Voraussetzung für aktors- und zielgruppenorientierte Planungsansätze.“¹⁸ Aber auch das Beispiel „mobility2know_4_ways2go“¹⁹ kann in diesem Kontext genannt werden. Hier wurde untersucht, inwiefern Soziale Lage, Wohnstandort und Soziales Milieu das Mobilitätsverhalten beeinflussen, wobei des Mobilitätsverhalten letztendlich wieder die soziale Interaktion im Raum ist.

¹⁷vgl. [Weichhart, 2008, S.327]

¹⁸vgl. [Giffinger, 2016]

¹⁹<https://www2.ffg.at/verkehr/projekte.php?id=721&lang=de&browse=programm>

3 Nutzergenerierte Daten

Neben den raumtheoretischen Zugängen sind für das Thema der Arbeit grundlegende Definitionen und Konzepte rund um den Begriff der nutzergenerierten Daten wichtig. Vorweg wird in diesem Kapitel die generelle Bedeutung von Daten in den Regionalwissenschaften sowie Defizite bisheriger Datenbestände beleuchtet. Danach folgt die grundlegende Systematik der nutzergenerierten Daten inklusive einigen Aspekten zur Zugänglichkeit. Vorbereitend auf das Hands-On Beispiel werden auch technische Grundlagen im Bereich des GPS-gestützten Trackings gezeigt. Die datenschutzrechtlichen Rahmenbedingungen in Österreich schließen das Kapitel der Nutzergenerierten Daten ab.

3.1 Allgemeine Bedeutung von Daten in der Stadt- und Regionalforschung

Die Stadt- und Regionalforschung beschäftigt sich traditionell mit städtischen und regionalen Entwicklungen, Strukturen und Prozessen und ist in ihren Methoden quantitativ geprägt. Dadurch entsteht eine große Bedeutung von Datengrundlagen, die für die Qualität der Ergebnisse und der Erreichung von Evidenz maßgebend ist. Eine breite und repräsentativ gesicherte Datenbasis bietet Ausgangspunkte für analytische Fragestellungen und führt zu einer standhaften Argumentationslinie. Zudem erleichtern anschaulich visualisierte Daten im letzten Schritt von Projekten die Kommunikation von Projekthalten und -ergebnissen. Von besonderem Interesse sind hier natürlich alle (unmittelbar oder mittelbar) raumbezogenen Daten, die für die Identifikation von Ursache- und Wirkungsbeziehungen auf räumlicher Ebene unverzichtbar sind. Die Analysewerkzeuge dieser Strukturen und Prozesse sind dabei meist Geographische Informationssysteme.

Eine starke Abhängigkeit bestand lange Zeit auch von klassischen Datenquellen aus öffentlichen Verwaltungsträgern. Dabei mussten in Forschungsprojekten immer wieder Datendefizite hinsichtlich der inhaltlichen Breite und Tiefe als auch bezüglich der Aktualität der Daten in Kauf genommen werden. Die Aktualität der Daten ist deswegen oft unzureichend, da diese meist nur einen sehr kleinen Querschnitt abbilden und die Erhebungen nur in großen Intervallen wiederholt werden. Es kann also nur eine sehr kurze Momentaufnahme verarbeitet werden - und dieser aufgenommene Moment liegt meist auch noch weit in der Vergangenheit. Zum Beispiel liegen im Bereich der Mobilitätsforschung die letzten Ergebnisse einer österreichweiten Erhebung fast 15 Jahre zurück.

Ein weiteres Defizit betrifft die räumliche Auflösung der Daten. Daten werden aus dem physisch-materiellen bzw. geographischen Raum in historisch-politisch gewachsenen administrativen Strukturen extrahiert. Die Dimension dieser Strukturen war aus pragmatischen Gründen wie die Strukturierung, Erfassung und Aktualisierung der Daten eher groß. Unter-

halb der Gemeindeebene war das Datenangebot grundsätzlich eingeschränkt, was für viele Anwendungen eine unzureichende Granularität darstellt. So argumentiert beispielsweise auch Wonka, wenn er meint „Nur dann, wenn sich die Daten auf Gebietseinheiten beziehen, die alle die gleiche Größe haben und keinen Grenzänderungen unterworfen sind, wie das bei einer Rastergliederung der Fall ist, sind sie universell einsetzbar.“²⁰ für eine Abkehr vom Konzept historisch-administrativ abgegrenzter Raumbehälter hin zu einer rasterbasierten Statistik. Dies hat zur Folge, dass vermehrt auch Daten auf Rasterbasis angeboten werden - eine etwa seit 10 Jahren existierende Datenvariante die auf der adressbasierten Statistik basiert. „Diese (Rasterzellen, Anm.) homogenisieren das räumliche Erscheinungsbild, Absolutwerte sind gleich Dichtewerten zu interpretieren, und die Integration mit beliebigen anderen Datensätzen bis hin zur Fernerkundung ist effektiv und algorithmisch einfach möglich.“²¹ Jedoch lässt sich durch den Ansatz der rasterbasierten Statistik das Heterogenitätsproblem nicht restlos in den Griff bekommen. Vor allem kleine Einheiten können heterogen sein, zumal die Abgrenzung nicht nach Homogenitätskriterien sondern aufgrund eines geometrischen Prinzips passiert.

Zu den genannten schwierigen Rahmenbedingungen kommen behördliche Datenschutzinterpretationen, die zur Vererrung von Information auf der Detailebene führen. „Generell ist festzuhalten, dass adressbezogene und damit koordinatengebundene, statistische Erhebungen und entsprechend georeferenzierte Datenbanken als Grundlage für flexible und problemgerechte Aggregation wünschenswert sind und als Grundlage zur Erarbeitung von Planungs- und Entscheidungsunterlagen enormes Potenzial aufweisen.“²² Auch im Bereich der Mobilitätsdaten - welchen in dieser Arbeit große Aufmerksamkeit zukommt - ist die Ausgangslage ähnlich. Als Beispiel können hier Pendlermatrizen genannt werden, die von den geschilderten Defiziten ebenso betroffen sind.

Aus diesen Rahmenbedingungen begründen sich die Wünsche nach neuen Quellen für räumliche Daten. In den letzten Jahren hat hier vor allem die Entwicklung von freien Datenbeständen - im weiteren *Open Data* genannt - neue Perspektiven aufgezeigt, die im Kapitel 3.3 näher vorgestellt werden. Aber auch der digitale Wandel und die Entwicklungen rund um nutzergenerierte Daten können nachhaltige Alternativen hervorbringen. Durch die Datengenerierung kann eine neue Basis für innovative Projekte erschlossen werden. Besonders spannend ist dieser Aspekt dann, wenn es sich um georeferenzierte Daten handelt die über einen besonders langen Zeitraum kontinuierlich erhoben werden.

Die Anwendungsmöglichkeiten sind ähnlich wie *Big Data* selbst: vielfältigst.

²⁰ [Wonka, 2010, S.155]

²¹ [Strobl, 2005, S.166]

²² [Strobl, 2005, S.155]

3.2 Grundprinzip der nutzergenerierten Daten

Person of the year: YOU. Yes, you.
You control the Information Age.
Welcome to your world.

*Titelcover des Time Magazine vom
25. Dezember 2006.*

Um das Prinzip der nutzergenerierten Daten verstehen zu können, ist es wichtig einige Aspekte rund um das Thema Web 2.0 zu beleuchten. Dieser Begriff bezeichnet einen Trend, der davon geprägt ist, dass die Nutzer selbst in den Vordergrund rücken und Inhalte zur Verfügung stellen sowie über das Internet interagieren. Auch wenn die Entwicklung zum Zeitpunkt der Entstehung dieser Arbeit schon sehr weit fortgeschritten ist, prägten die technischen Möglichkeiten und die neue Philosophie des Web 2.0 ein außergewöhnlich weites Spektrum an Inhalten und Disziplinen. Prominentestes Beispiel dafür ist wohl Wikipedia. Meist wird vom Betreiber nur mehr eine Plattform zur Verfügung gestellt, mithilfe derer die Nutzer Inhalte erstellen und bearbeiten können.²³ Eine komplette Liste von Beispielen lässt sich wahrscheinlich nicht erstellen, anzumerken ist aber, dass unter den 25 meist besuchten Websites 2016 (global) mit LinkedIn.com, Instagram, Twitter, Yahoo, Wikipedia, Facebook und Youtube gleich mehrere zu finden sind, die auf diesem Grundprinzip basieren.²⁴ Lediglich die Bezeichnung des Konzepts hat sich über die Jahre verändert - von Web 2.0 damals zu Social Media heute. Die technischen Aspekte rund um das Thema werden im Folgenden nicht weiter beleuchtet, da für diese Arbeit viel mehr das Konzept dahinter von Bedeutung ist.

Die Idee der nutzergenerierten Inhalte ist aber keineswegs erst mit der weiten Verbreitung des Internets aufgekommen. In Zeitungen ist beispielsweise das Konzept des Leserbriefs schon lange ein fixer Bestandteil - und auch eine Form, wie Nutzer (in diesem Fall Leser) eigene Inhalte in der Plattform (der Zeitung) veröffentlichen können.

Im Folgenden werden einige Aspekte und Konzepte rund um nutzergenerierte Daten vorgestellt. Die Bandbreite reicht dabei von grundlegenden Begriffen über die Zugänglichkeit von Daten bishin zu datenschutzrechtlichen Rahmenbedingungen. Zudem werden noch technische Vorgehensweisen rund um die GPS-gestützte Erhebung von nutzergenerierten Daten beleuchtet.

3.3 Verräumlichte nutzergenerierte Inhalte

Die Entwicklungen der letzten Jahre haben sich auch im Bereich der (Internet-)Geographie und Kartographie längst etabliert. Auch bei diesbezüglichen Plattformen sind von Nutzern freiwillig bereitgestellte Informationen das Grundkonzept. Dieses Prinzip wird auch als „Crowdsourcing“ in der Literatur angeführt: „Crowdsourcing ist eine interaktive Form der Leistungserbringung, die kollaborativ oder wettbewerbsorientiert organisiert ist und eine große Anzahl extrinsisch oder intrinsisch motivierter Akteure unterschiedlichen Wissensstands unter Ver-

²³vgl. [OECD, 2016]

²⁴vgl. [Alexa, 2016]

wendung moderner IuK²⁵-Systeme auf Basis des Web 2.0 einbezieht. Leistungsobjekt sind Produkte oder Dienstleistungen unterschiedlichen Innovationsgrades, welche durch das Netzwerk der Partizipierenden reaktiv aufgrund externer Anstöße oder proaktiv durch selbsttätiges Identifizieren von Bedarfslücken bzw. Opportunitäten entwickelt werden.“²⁶

In Hinblick auf Räumlichkeit lassen sich nutzergenerierte Daten folgenderweise kategorisieren.²⁷

- non-spatial user-generated content
 - Daten und Inhalte ohne expliziten Raumbezug. Beispiele hierfür sind die Inhalte von Wikipedia Artikeln oder Twitter Tweets.
- spatially implicit user-generated content
 - Daten und Inhalte, die einen räumlichen Bezug besitzen, aber der Inhalt im Vordergrund steht. Als Beispiel ein Twitter Tweet, der die von einem mobilen Endgerät mitgesendete GPS-Position enthält oder ein hochgeladenes Foto auf Flickr, welches sich mithilfe der in der Datei mit abgespeicherten Koordinaten ebenfalls georeferenzieren lässt. Passive Verräumlichung der Inhalte.
- spatial user-generated content
 - Zweck und Inhalt haben eindeutigen Raumbezug, beispielsweise ein Eintrag in OpenStreetMap. Aktive Verräumlichung der Inhalte.

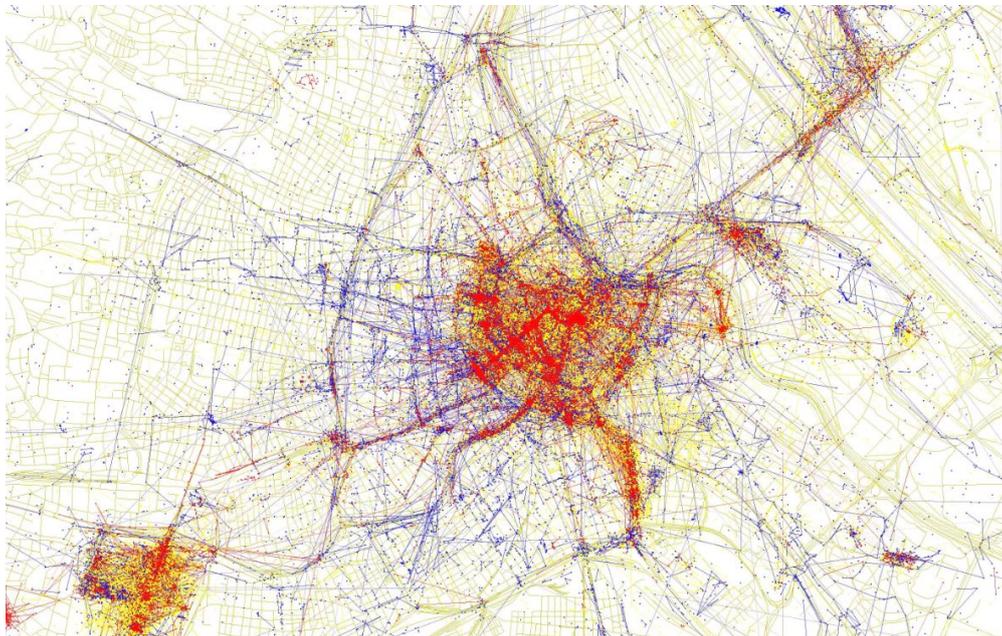


Abbildung 3.1: Ein Beispiel für spatially implicit user-generated content: Verortete Aufnahmestandorte von Fotos der Bilder-Plattform Flickr.com von Wien. Deutlich erkennbar sind touristische HotSpots wie Schönbrunn im Südwesten. [Fischer, 2017]

²⁵Internet- und Kommunikationstechnik

²⁶vgl. [Lessmann u. a., 2008]

²⁷nach [Haklay, 2010]

Als prominentestes Beispiel für spatial user-generated content dient OpenStreetMap. Die Idee dahinter ist, mithilfe einer sehr großen Zahl an Freiwilligen einen großen Bestand an frei nutzbaren Geodaten zu sammeln und diese unter dem Leitbild des *Open Data* Konzepts allen zur Verfügung zu stellen. Ein Grundpfeiler der Datenbeschaffung sowie der Datenkontrolle ist dabei das „local knowledge“ der Nutzer. Im Dezember 2016 lag die Zahl dieser OpenStreetMap (OSM) Autoren bei ca. 3,2 Millionen, die Datenbank verfügt über knapp 5,5 Milliarden gespeicherte GPS-Punkte.²⁸ Der Abdeckungsgrad der Karte ist mittlerweile so weit fortgeschritten, dass OpenStreetMap in vielen Anwendungen und Projekten als Basisdatensatz und Kartengrundlage verwendet werden kann. Zur Qualität der Daten gibt es bereits ein breites Feld an Arbeiten. Auf einige davon soll hier kurz verwiesen werden: [Zipf, 2009] und [Krammer, 2014]

Der Ansatz hinter OpenStreetMap wird in der Literatur auch als „Freiwilligengeographie“²⁹ oder noch verbreiteter „Volunteered Geographic Information“³⁰ bezeichnet.

Grundsätzlich ist für diese Arbeit von Relevanz: Bei nutzergenerierten Inhalten handelt es sich um einen Datentyp, bei dem Inhalte von den Nutzern selbst erzeugt werden. In dieser Systematik kann weiter nach Typen der Datengenerierung unterschieden werden. Während bei OpenStreetMap Nutzer aktiv Inhalte beitragen, wird das Hands-On Beispiel der *Bike Citizens* Daten im Kapitel 5 zeigen, dass es auch Anwendungen gibt, bei denen nutzergenerierte Daten passiv generiert werden können und der Nutzer eigentlich einen anderen Zweck verfolgt.

3.4 Datenzugänglichkeit - Open(Geo-)Data

„Open knowledge is any content, information or data that people are free to use, re-use and redistribute — without any legal, technological or social restriction.“
- Open Knowledge Foundation 2015.

Die Konzepte rund um frei verfügbare Datenquellen haben sich in den letzten fünf bis zehn Jahren entwickelt. „Offene Daten sind sämtliche Datenbestände, die im Interesse der Allgemeinheit ohne jedwede Einschränkung zur freien Nutzung, zur Weiterverbreitung und zur freien Weiterverwendung frei zugänglich gemacht werden. Bei *Open Data* handelt es sich nicht ausschließlich um Datenbestände der öffentlichen Verwaltung, auch privatwirtschaftlich agierende Unternehmen, Hochschulen und Rundfunksender sowie Non-Profit-Einrichtungen produzieren entsprechende Beiträge“³¹. Trotz *Open Data* können jedoch gewisse Bedingungen zu berücksichtigen sein - beispielsweise die Nennung der Quelle oder des Urhebers oder Einschränkungen in der Publikation.

Wie Von Lucke in seiner Definition der *Open Data* Materie festhält, meint diese ein sehr breites Feld von Daten, wobei dabei nicht nur nach der veröffentlichenden Institution differenziert werden kann, sondern auch nach dem Inhalt des jeweiligen Datenbestandes. Für die Planung ist der Bereich rund um Open-Geo-Data von höchster Relevanz.

²⁸[OpenStreetMapStats, 2016]

²⁹vgl. [Lechner, 2011, S.18]

³⁰[Goodchild, 2007]

³¹[von Lucke, 2011, S.5]

Lange Zeit war der Einsatz von geographischen Informationssystemen (GIS) Experten vorbehalten, Grund dafür war die notwendige Expertise in der Bedienung. Eingesetzt wurden diese Systeme vorrangig in der staatlichen Verwaltung - naheliegend, dass auch die Datenzufuhr von staatlicher Stelle mehrheitlich übernommen wurde. In Österreich ist in diesem Fall das Bundesamt für Eich- und Vermessungswesen (BEV) der größte staatliche Geodatenproduzent.³² Die Erstellung dieser Datenbestände passierte weitgehend im Zuge der Digitalisierung von analogen Daten und war dementsprechend kostenaufwendig.³³

Die Entwicklungen des Web 2.0 haben auch hier nachhaltige Spuren hinterlassen. GIS-Systeme haben sich weiterentwickelt, sind kostengünstiger bzw. teilweise kostenlos und verlangen nach einheitlichen Standards und großen frei verfügbaren Datenmengen.³⁴

„In Zeiten des WebGIS und des Web 2.0 entstehen für Planer beinahe täglich neue Daten und Instrumente, die frei kombiniert werden können. Diese sollten bereits heute von Planern nicht nur erkannt, sondern auch aufgegriffen, zusammengesetzt und weiterentwickelt werden - denn es kann heute bei der Entwicklung neuer Konzepte nicht mehr darum gehen, das Rad neu zu erfinden. Wie ein Eichhörnchen kann der Planer sich auf die Suche nach bestehendem Material begeben und dieses zusammen tragen. Mit wenig Aufwand kann so innerhalb kürzester Zeit eine spezifische Sammlung an Informationen und Instrumenten angehäuft werden.“³⁵ Diese Thematik stellt aber nicht nur für die Planung und Forschung wichtige zukünftige Wünsche dar, sondern ist durchaus auch auf die Privatwirtschaft übertragbar - auch Start-Ups im Dienstleistungsbereich könnten von einer großen Geodatenverfügbarkeit profitieren.

Ergebnisse sind neben klassischen Datenbeständen, wie die Landbedeckungsklassen oder Verwaltungsgrenzen, auch eher speziellere Anwendungen wie beispielsweise OpenRailwayMap (<http://www.openrailwaymap.org/>) wo sehr detailliert Eisenbahnstrecken samt Höchstgeschwindigkeiten und Signalanlagen ersichtlich sind. Für die breitere Öffentlichkeit wahrscheinlich relevanter ist die OpenCycleMap (<https://www.opencyclemap.org/>) wo Nutzer weltweit Radwege erfassen, ergänzen und abrufen können. Beide Beispiele basieren auf der Idee von OpenStreetMap und der damit Hand in Hand gehenden Vor- und Nachteile sowie den Unterschieden zu proprietären/kommerziellen Daten. „Diese Unterschiede betreffen die Ebenen Urheberschaft (z.B. amtliche Statistik vs. private Erfassung), Lizenz- und Nutzungsbedingungen (sehr eingeschränkt vs. Public Domain), Kosten (kostenintensiv vs. kostenlos) und schließlich auch Verfügbarkeit (beschränkt auf Grund des Datenschutzes vs. unbeschränkt) und Qualität (garantierte Qualität vs. Vertrauen in die Kompetenz der Mitwirkenden).“³⁶ Es sind somit auch Beispiele für spatial user-generated content und damit nutzergenerierte Daten.

³²vgl. [Kalasek u. Weninger, 2015, S.18]

³³vgl. [Greve, 2002, S.121]

³⁴vgl. [Kalasek u. Weninger, 2015, S.17]

³⁵vgl. [Körnig-Pich u. a., 2010, S.181], nach [Kalasek u. Weninger, 2015, S.17]

³⁶[Kalasek u. Weninger, 2015, S.15]

Exkurs - Open Government Data

Neben den genannten privaten Initiativen im Rahmen der *Open Data* Bewegung etabliert sich diese sukzessive auch bei administrativen Daten. „Das Bundeskanzleramt, die Städte Wien, Linz, Salzburg und Graz gründeten am 13. Juli 2011 gemeinsam die „Cooperation Open Government Data Österreich“, kurz „Cooperation OGD Österreich“. Bund, Länder, Städte und Gemeinden wollen in Kooperation mit Communities, Wissenschaft, Kultur und der Wirtschaft die Basis für die Zukunft von Open Government Data in Österreich legen. Durch die Einigung auf gemeinsame Standards sollen effektive Rahmenbedingungen geschaffen werden, die zum Nutzen aller Interessensgruppen sind.“³⁷ Auf der *Open Data* Plattform <https://www.data.gv.at/> finden sich immer mehr frei verfügbare Datensätze, die von Verwaltungseinrichtungen der Allgemeinheit zur Verfügung gestellt werden. Darunter sind auch eine Menge georeferenzierter Daten zu finden. Im aktuellen Regierungsprogramm Österreichs³⁸ ist zudem vermerkt, dass „bis Ende 2020 alle Daten der öffentlichen Verwaltung nach Möglichkeit als *Open Data* zur Verfügung stehen. Das erhöht einerseits die Transparenz und schafft auch die Basis für Geschäftschancen für innovative Unternehmen und Startups. [...] Zudem soll die Einbindung von Datenbeständen aus dem privaten Sektor auf freiwilliger Basis forciert werden.“³⁹ Aber auch auf der EU-Ebene wurde erkannt, dass die Entwicklung eines lukrativen und know-how-intensiven Wirtschaftszweigs im Bereich der Geodatenverarbeitung eine möglichst umfangreiche und hochwertige Datenbasis braucht. In der INSPIRE⁴⁰-Richtlinie wird dazu festgehalten: „Die Probleme bei der Verfügbarkeit, Qualität, Organisation, Zugänglichkeit und gemeinsamen Nutzung von Geodaten betreffen in gleicher Weise zahlreiche Bereiche der Politik und Information und nahezu alle Verwaltungsebenen. Ihre Lösung erfordert Maßnahmen für den Austausch, die gemeinsame Nutzung, die Zugänglichkeit und die Verwendung von interoperablen Geodaten und Geodatendiensten über die verschiedenen Verwaltungsebenen und Sektoren hinweg. Deshalb sollte in der Gemeinschaft eine Geodateninfrastruktur geschaffen werden.“⁴¹ Aus der im Kapitel 2.3 geschilderten Datenabhängigkeit der Stadt- und Regionalforschung und den damit verbundenen bisherigen Defiziten gibt es große Hoffnungen im Bereich von Open Government Data. Allerdings bleibt die offene Frage nach der Datenqualität, Repräsentativität und Vollständigkeit, welche zukünftig auch hier abgesichert sein muss.

Kurzfasit

Neben der Einteilung von (Nutzergenerierten) Daten nach ihrer räumlichen Wirkung ist auch die Zugänglichkeit entscheidend. Datenkonzerne wie Google oder auch Automobilfirmen können selbst schon auf eine große Menge an nutzergenerierten Daten zugreifen, die aber für die Forschung und Planung - wenn überhaupt - nur gegen hohe Gebühren verfügbar sind. Die Entwicklung im Bereich von OGD stellt eine Quelle zur Anreicherung der Datenvielfalt von nutzergenerierten Daten dar.

³⁷ vgl. <https://www.data.gv.at/infos/cooperation-ogd-oesterreich/>

³⁸ Februar 2017

³⁹ Arbeitsprogramm der Bundesregierung 2017/2018 S.17ff: <https://goo.gl/RRKCIU>

⁴⁰ Infrastructure for Spatial Information in the European Community

⁴¹ [Parlament u. Rat, 2007]

3.5 Technische Grundlagen - GPS-Tracking

Beim GPS-Tracking wird mithilfe eines GPS-Empfängers in einem bestimmten Zeitintervall eine fortlaufende Positionsbestimmung über eine Satellitenverbindung durchgeführt. Die so lokalisierten Punkte enthalten dabei nicht nur die exakten GPS-Koordinaten, sondern auch einen Zeitstempel. Vorbereitend auf die Auswertung der *Bike Citizens* Daten als Hauptteil der Arbeit sollen hier technische Voraussetzungen und Probleme im Bereich des automatisierten GPS-gestützten Trackings zur Erhebung nutzergenerierter Daten mithilfe von Smartphones aufgezeigt werden.

Grundsätzlich sind im Bereich des GPS-gestützten Trackings folgende Probleme und Herausforderungen bekannt:⁴²

- Zufallsfehler: Bei wiederholten Positionsbestimmungen ist davon auszugehen, dass es unausweichlich zu Positionierungsfehlern kommt. Dies kann verschiedene physikalische Gründe haben, die zu einer Störung des Signals führen. Gründe liegen dabei beispielsweise beim Mehrwegeempfang, wo das Signal abgelenkt wird oder erst über Umwege wieder zum Empfänger kommt. Dies kann beispielsweise durch Brechung des Signals in der Ionosphäre passieren.
- Positionierungslücken: Die exakte Positionsbestimmung über GPS funktioniert nur bei Sichtbarkeit von mindestens drei Satelliten. Dies führt vor allem in bebauten Gebieten, aber auch in Gegenden mit hohen Objekten (Wälder oder Bergtäler) zu Problemen und Ungenauigkeiten.
- Kaltstart Problematik: Durch die Bewegung in der Erdumlaufbahn muss bei jedem Start eines GPS-Empfängers die Suche nach Satelliten von neuem gestartet werden. Dies führt zu einer so genannten Kaltstart-Problematik, welche zur Folge hat, dass insbesondere kurz nach der Aktivierung des Geräts besonders starke Ungenauigkeiten auftreten.

Die in den letzten Jahren weiter fortgeschrittene Verbreitung der Smartphones bietet mittlerweile eine Alternative zu klassischen GPS-Geräten: Die Personen sind bereits mit einem GPS-fähigen Gerät ausgestattet und es bedarf nur mehr einer softwareseitigen Lösung zur Wegeaufzeichnung. Zudem bieten Smartphones den Vorteil, dass der Kaltstart-Problematik mithilfe von ungefährender Positionsbestimmung über GSM/UMTS bzw. umliegenden WLAN-Zellen entgegen gewirkt werden kann - zweiteres nur mit aktiver Internetverbindung. Jedoch sind die verbauten GPS-Module in Smartphones nicht mit jenen spezialisierter GPS-Logger vergleichbar - vor allem in puncto Genauigkeit.⁴³

Speziell unter schlechten Empfangsbedingungen - wie beispielsweise konstruktionsbedingt oft in öffentlichen Verkehrsmitteln - kommen Smartphones schnell an ihre Grenzen. So schreibt schon die KOMOD-Studie im Jahr 2011: „Anhand von Referenzmessungen konnte zum Beispiel nachgewiesen werden, dass in einem Railjet-Zug der ÖBB zwischen Wien und Salzburg mit dem damals aktuellen iPhone 3GS und 4 keine Positionierung durchgeführt werden konnte. Mit einem QStarz BT-Q1000X Empfänger ist die Aufzeichnung einer solchen Fahrt, ausgenommen der Tunnelabschnitte, möglich“.⁴⁴ Einige Versuche auf der Strecke zwischen Wien und

⁴²vgl. [Fellendorf u. a., 2011, S.152 ff]

⁴³vgl. [Bierlaire u. a., 2013, S.1]

⁴⁴[Fellendorf u. a., 2011, S.163]

Linz im Dezember 2016 und Jänner 2017 zeigten, dass die Positionsbestimmung auf einem aktuellen Android Gerät (Xiaomi Mi5) auf freien Abschnitten mit einiger Wartezeit möglich ist, die Lokalisierung aber nur über das Mobilfunknetz durchgeführt wird. Eine zeitgleich gestartete Applikation zur genauen Positionsbestimmung rein über GPS konnte sich während der gesamten Fahrt nicht mit den geforderten drei Satelliten verbinden. Es zeigt sich, dass die Signalabschirmung durch Fahrzeuge als besondere Herausforderung bei der Aufzeichnung von Bewegungsabläufen hinzukommt.

Die im Hands-On Beispiel genutzten Daten von *Bike Citizens* beruhen technisch gesehen auf einer im Sekundenintervall durchgeführten Datenaufzeichnung der aktuellen GPS-Position. Diese Bewegungsspuren werden in einem nächsten Schritt miteinander verknüpft, was als Endergebnis den Fahrtverlauf - auch Track genannt - erkennen lässt. Da die Lokalisierung über die Smartphone-internen Sensoren passiert, zeigen auch dort die hier genannten Störfaktoren ihre Auswirkung. So sind beispielsweise beim überbauten und unter der Straße geführten Radweg auf der Praterbrücke zwischen Donauinsel und Prater grobe Ungenauigkeiten zu erkennen:

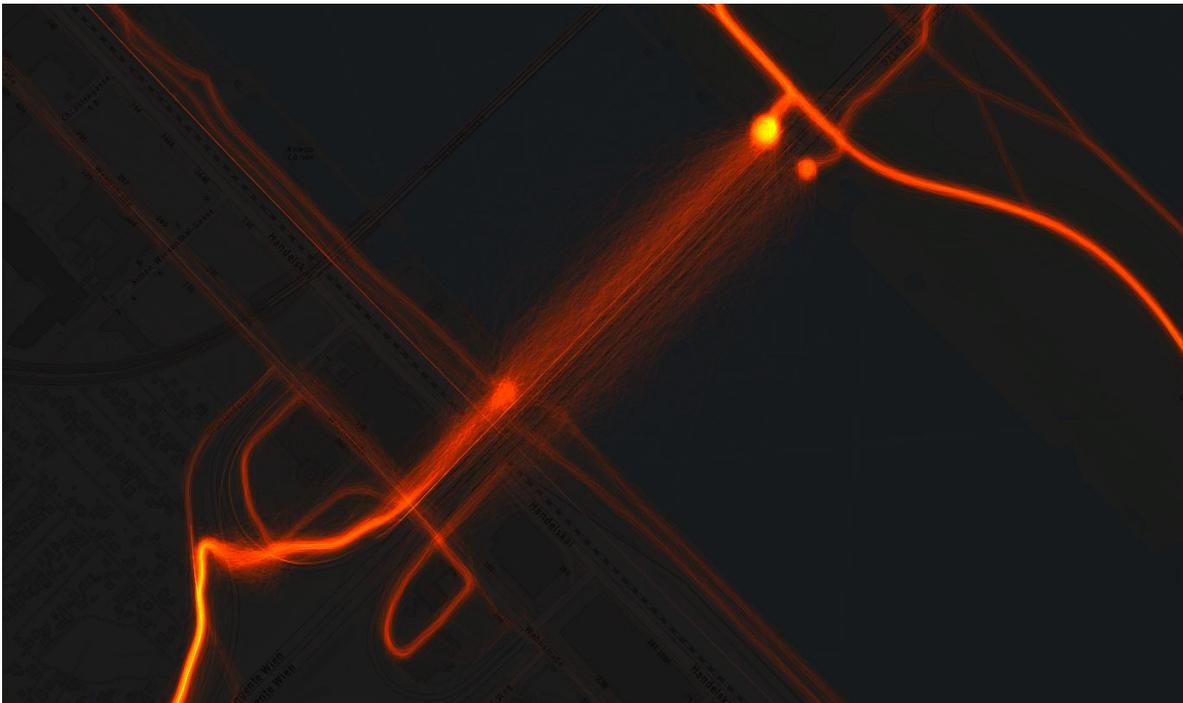


Abbildung 3.2: Bike Citizens Tracks im Bereich der Praterbrücke: Deutlich sind die Ungenauigkeiten der Tracks verglichen mit den Straßenabschnitten auf der Donauinsel (nördlich) und im zweiten Bezirk (südlich) erkennbar. Sichtbar sind auch die spiralförmigen Radauffahrten.
[Eigene Darstellung / Bike Citizens]

Diese Ungenauigkeiten haben in späterer Folge auch Auswirkungen auf die Analysen und Anwendungen. In erster Linie problematisch sind ungenaue Positionen für die Zuordnung der gemessenen Punkte und Tracks zu einzelnen Routen und Trassen. So ist beispielsweise die Darstellung der aktuellen Verkehrssituation in Google Maps nur möglich, wenn die Positionssignale der KFZ-Nutzer, klar von denen der übrigen mobilen Nutzer getrennt, den Straßenab-

schnitten zugewiesen werden können. Auch für die Abbildung von Frequenzen innerhalb von Teilabschnitten ist die Kantenzuweisung einzelner Fahrten unabdingbar. Ein auf möglichst genauen Positionen basierender Datensatz ist somit wünschenswert - aber durch die dargestellten Einschränkungen nicht immer möglich, weshalb der Umgang mit Ungenauigkeiten in Analysen berücksichtigt werden muss.

3.6 Datenschutz und rechtliche Rahmenbedingungen

Die Innovationen und technischen Möglichkeiten rund um die Themen Web 2.0, *Big Data* und neue Datenerhebungsstrategien werden oft von Bedenken im Bereich des Datenschutzes begleitet. Mit den vorgestellten Technologien und Prozessen ist die Aufzeichnung von Mobilitätsmustern möglich, welche auch eine Ableitung auf persönliche Aktivitäten zulassen. Spätestens wenn diese Aufzeichnungen auch Einzelpersonen zuordenbar sind, stellen sich klarerweise erhebliche Fragen zum Schutz der Privatsphäre. Wie sicher sind persönliche Daten im Netz? Was darf mit den Daten gemacht werden? Aber auch im Bereich der klassischen Mobilitätsenerhebungen gilt es, personenbezogene Daten im Rahmen der gesetzlichen Rahmenbedingungen zu behandeln. Ebenso werden im späteren Hands-On Beispiel nutzergenerierte Mobilitätsmuster analysiert und ausgewertet, weshalb vorab eine Auseinandersetzung mit den österreichischen datenschutzrechtlichen Grundlagen unerlässlich scheint.

Grundsätze in Österreich

In Österreich wird Datenschutz im Datenschutzgesetz 2000 (DSG2000) mit der Datenschutznovelle 2010 geregelt. Dies sieht grundsätzlich vor, dass die eine natürliche Person betreffenden personenbezogenen Daten schutzwürdig sind und nicht beliebig verwendet werden dürfen (§§1 - 3 DSG2000). Im Artikel 2 werden von Abschnitt 1 bis Abschnitt 11 folgende Inhalte geregelt: Verwendung der Daten nach den Grundsätzen der Verfassungsbestimmungen, notwendige Datensicherheitsmaßnahmen, notwendige Publizitätserfordernisse, Rechte bzw. Rechtsschutz der betroffenen Personen, Kontrollorgane, Ausnahmen und Einschränkungen für bestimmte, besondere Verwendungszwecke, zusätzliche Erfordernisse für automatisierte Einzelentscheidungen und Strafbestimmungen.

Wesentlich sind für den täglichen Gebrauch personenbezogener Daten die Artikel 2 Abschnitte 2-5. Zusammengefasst dürfen personenbezogene Daten nur insoweit verwendet werden⁴⁵:

- nur für festgelegten, (dem Betroffenen vorher bekannt gegebenen) eindeutigen und rechtmäßigen Zweck.
- wenn sie für den Zweck der Datenanwendung wesentlich sind und nicht über diesen Zweck hinausgehen.
- wenn sie so verwendet werden, dass sie für den Verwendungszweck im Ergebnis sachlich richtig und soweit nötig auf dem neuesten Stand sind.

⁴⁵nach [Wolf, 2016, S.41]

- wenn sie nur solange in personenbezogener Form aufbewahrt werden, als dies für die Zweckerreichung, für die sie ermittelt wurden, erforderlich ist. Danach sieht das Gesetz eine physische Löschung vor - das heißt, es darf keine Möglichkeit zur Wiederherstellung geben.
- soweit Zweck und Inhalt der Datenanwendung von den gesetzlichen Zuständigkeiten oder rechtlichen Befugnissen des jeweiligen Verwenders der Daten gedeckt sind und die schutzwürdigen Geheimhaltungsinteressen der Betroffenen nicht verletzen.
- sie dürfen an Dritte nur dann übermittelt werden, wenn sie aus einer zulässigen Datenanwendung stammen und der Empfänger (Dritte) dem Übermittelnden seine ausreichende gesetzliche Zuständigkeit oder rechtliche Befugnisse glaubhaft gemacht hat.
- wenn sie durch den Zweck und den Inhalt der Übermittlung die schutzwürdigen Geheimhaltungsinteressen der Betroffenen nicht verletzen
- die Eingriffe in das Grundrecht auf Datenschutz nur im erforderlichen Ausmaß und mit den gelindesten Mitteln erfolgen.

Zusätzlich werden im §4 Z2 DSG2000 besonders sensible und daher auch besonders schutzwürdige Daten natürlicher Personen genannt. Diese dürfen nur unter Ausnahmebestimmungen verwendet werden (geregelt in §9). Dazu gehören:

- rassische und ethnische Herkunft⁴⁶
- politische Meinung
- Gewerkschaftszugehörigkeit
- religiöse oder philosophische Überzeugung
- Gesundheit
- Sexualleben

Auskunftsrecht, Recht auf Richtigstellung und Widerspruchsrecht

Das Auskunftsrecht

Im §27 DSG2000 wird das Auskunftsrecht geregelt. Dieses sieht vor, dass Anfragen zu Datenverwendung innerhalb von acht Wochen beantwortet werden müssen (auch bei Negativauskunft). Dieses Recht steht jeder Person einmal im Jahr zu.

Als Antwort auf die Auskunftsanfrage muss der Person mitgeteilt werden:

- um welche Daten die Person betreffend es sich handelt.
- für welchen Zweck diese Daten verwendet werden.
- woher die Daten stammen.
- wie lange die Daten gespeichert bleiben.

⁴⁶Originalwortlaut §4 DSG2000

Das DSG2000 sieht auch vor, dass die Antwort an die auskunftsuchende Person in verständlicher Sprache zu erfolgen hat. Bei Verweigerung oder falscher Auskunft dient die Datenschutzbehörde als Ansprechpartner für Beschwerden.

Das Recht auf Richtigstellung oder Löschung

§28 DSG2000 regelt die Richtigstellung oder Löschung von Personendaten. Dazu muss ein Antrag gestellt werden, welcher innerhalb von acht Wochen zu beantworten ist. Etwaige Übermittlungsempfänger der Daten hat der Verwender der Daten über die Richtigstellung oder Löschung der Bestreitung zu informieren.⁴⁷

Das Widerspruchsrecht

Jeder Betroffene hat das Recht, gegen die Verwendung seiner Daten wegen Verletzung überwiegender schutzwürdiger Geheimhaltungsinteressen, die sich aus einer besonderen Situation ergeben, Widerspruch zu erheben. Bei Vorliegen dieser Voraussetzungen sind die Daten innerhalb von acht Wochen zu löschen. Bei öffentlich zugänglichen Anwendungen kann der Betroffene jederzeit auch ohne Begründung einen Widerspruch einlegen. Die Daten sind dann ebenfalls spätestens nach acht Wochen zu löschen. Beide Regelungen kommen nur bei nicht gesetzlich angeordneten Datenaufnahmen zur Anwendung.

Datenschutzbehörde

Die Datenschutzbehörde ist nach §35 DSG2000 eine unabhängige Kontrollstelle. Der Leiter der DSB ist unabhängig und weisungsfrei und auf fünf Jahre durch den Bundespräsidenten bestellt. Jedermann kann sich wegen einer behaupteten Verletzung seiner Rechte oder ihn betreffender Pflichten an diese Behörde wenden. Zur Herstellung des rechtmäßigen Zustands kann die DSB Empfehlungen aussprechen. Wird diese innerhalb der gesetzten Frist nicht befolgt, so drohen je nach Art des Verstoßes:⁴⁸

- Strafanzeige nach §51 (Gewinn- oder Schädigungsabsicht) oder §52 (Verwaltungsstrafgesetz)
- Klage vor dem zuständigen Gericht (bei schwerwiegenden Verstößen durch Auftraggeber des privaten Bereichs)

Die Strafen reichen von einer Freiheitsstrafe von bis zu einem Jahr (bei Gewinn oder Schädigungsabsicht) bis zu 25.000 Euro bei Verwaltungsübertretung, wobei hier nach Schwere der Übertretung abgestuft wird. Wer Daten nicht fristgerecht beauskunftet, richtigstellt oder löscht, riskiert beispielsweise eine Strafe von bis zu 500 Euro. Bei Verletzung von Offenlegungs- oder Informationspflichten oder Sicherheitsmaßnahmen für Daten riskiert man eine Strafe von bis zu 10.000 Euro.⁴⁹

⁴⁷vgl. [Wolf, 2016, S.42]

⁴⁸vgl. [Wolf, 2016, S.42]

⁴⁹vgl. [Wolf, 2016, S.43]

Entwicklungen - EU Verordnung

Die EU hat am 24. Mai 2016 die neue Datenschutz-Grundverordnung erlassen. Diese ersetzt eine Richtlinie aus dem Jahr 1995, die in jedem Mitgliedsland in eine eigene Rechtsordnung transformiert wurde. Die neue Verordnung ist dabei wesentlich konkreter und soll den Datenschutz auf EU-Ebene auf einen einheitlichen Standard bringen.

Die Eckpfeiler der Verordnung sind dabei folgende Punkte:⁵⁰

- Recht auf Vergessenwerden
- Verarbeitung der Daten nur nach ausdrücklicher Einwilligung der betroffenen Person
- Recht auf Datenübertragbarkeit (an einen anderen Dienstleister)
- Recht der Betroffenen, bei Verletzung des Schutzes der eigenen Daten darüber informiert zu werden
- Datenschutzbestimmungen müssen in klarer und verständlicher Sprache erläutert werden
- Bei Verstößen wird härter durchgegriffen. Im Fall eines Unternehmens werden Strafen von bis zu 4% des Jahresumsatzes des vergangenen Geschäftsjahres eingefordert

In Grundzügen werden diese Eckpfeiler der neuen Verordnung, die ab Frühjahr 2018 in Kraft treten wird, bereits jetzt im österreichischen Datenschutzgesetz berücksichtigt. Spätestens bis dahin müssen die Bestimmungen in nationales Recht umgesetzt sein. Auffallend sind die neuen höheren Strafsätze für Unternehmen, welche bei entweder bis zu 4% des Jahresumsatzes oder bis zu 20 Mio. Euro liegen - je nach höherer Summe.

Wichtige Gesetze im Zusammenhang mit Mobilitätserhebungen

Die Erfassung und auch die Veröffentlichungen von Mobilitätsdaten durch breite Erhebungen betreffen in Österreich nicht nur das Datenschutzgesetz. Folgende Liste soll zusammenfassen, welche gesetzlichen Rahmenbedingungen zu beachten sind:⁵¹

- Datenschutzgesetz (Schutz personenbezogener Daten) BGBl I 1999/165 zuletzt geändert durch BGBl I 2009/133 und 135
- Bundesstatistikgesetz (Veröffentlichungspflichten Statistik Austria) BGBl I Nr. 163/1999 zuletzt geändert durch BGBl I Nr. 125/2009
- Informationsweiterverwendungsgesetz (Weitergabe von Daten öffentlicher Stellen) BGBl. I Nr. 135/2005)
- Telekommunikationsgesetz (Standortdaten dürfen nur für die Besorgung eines Kommunikationsdienstes ermittelt oder verarbeitet werden, BGBl I Nr. 70/2003 zuletzt geändert durch BGBl I Nr. 65/2009

⁵⁰vgl. [EuropäischesParlament, 2016]

⁵¹vgl. [Fellendorf u. a., 2011, S.80]

4 Mobilitätsdatenerhebung - State of the Art und aktuelle Entwicklungen

Die Disziplin der Raumplanung und Raumordnung bringt als Querschnittsmaterie ein sehr breites Spektrum an Anforderungen für die Datenverarbeitung mit sich. Dabei sind aber nicht nur die einzelnen Fachbereiche der Disziplin heterogen, auch die benötigten Daten müssen mit der jeweiligen Disziplin und der Aufgabenstellung oder dem Projekt harmonieren. Eine Tatsache eint die Querschnittsmaterie aber stark: Für Evidenz und optimierte Prozesse werden Daten benötigt. Die Relevanz von Daten speziell für die Stadt- und Regionalforschung wurde schon in Kapitel 3.1 aufgezeigt.

Daten spielen dabei in den verschiedensten Phasen von Forschungs- oder Planungsprojekten eine entscheidende Rolle: Angefangen bei der Bestandsaufnahme in der Planungsregion oder der Ausgangslage eines Forschungsprojekts, über die Findung von optimalen Maßnahmen oder dem Aufbau maßgeschneiderter Modelle, bis hin zur argumentativen und stützenden Grundlage für die Ergebnisse. Daten haben in der Planung verschiedenste Anwendungsmöglichkeiten - wohlgemerkt, dass der Begriff der Daten hier ein ähnlich großes Spektrum aufzeigen kann wie die Raumplanung selbst. Wie schon anhand des Aufbaus der Arbeit erklärt, soll der Fokus immer weiter in Richtung Mobilität wandern. Dies trifft auch auf den Begriff der Daten zu, welcher im Folgenden im Kontext von Mobilitätsdaten zu verstehen ist.

Die in den Punkten 4.1. bis 4.4. geschilderten Inhalte haben ihren Ursprung in einem Seminar der österreichischen Forschungsgesellschaft für Straße, Schiene und Verkehr, welches ich am 21. und 22. April 2016 dienstlich besuchen durfte. Das Seminar trug den Titel „Evolution versus Revolution der Erhebung und Anwendung von Mobilitätsdaten.“ Neben spannenden Vorträgen passend zum Thema dieser Diplomarbeit bot sich dort auch die Möglichkeiten mit einigen Experten aus den Themenbereichen in Kontakt zu treten. Auch der Aufbau und die Konzeption der Arbeit ist von diesem Seminar inspiriert.

4.1 Von der aktiven Befragung... - Klassische Mobilitätsdatenerhebung gezeigt am Beispiel Österreich Unterwegs 2014

4.1.1 Hintergrund und Ziele

Mobilitätsdaten der österreichischen Bevölkerung stellen eine wesentliche Entscheidungsgrundlage für Planungen und Projekte dar. Sie sind Berechnungsbasis für wesentliche Kennzahlen wie den Modal Split. Der Handlungsbedarf spitzte sich in den letzten Jahren zu, weil zwar in den einzelnen Bundesländern immer wieder Teilerhebungen erfolgt sind, die letzte österreichweite Befragung zum Mobilitätsverhalten aber aus dem Jahr 1995 stammt. Vor dem Hintergrund des technologischen Fortschritts, des Strukturwandels und den daraus resultierenden, sich stark verändernden Rahmenbedingungen ein sehr langer Zeitraum. Zudem wurde die Erhebung *Österreich Unterwegs* vorgestellt und intensiv diskutiert.

4.1.2 KOMOD-Studie

Bereits im Vorfeld von *Österreich Unterwegs* sollte eine Konzeptstudie die Rahmenbedingungen für eine österreichweite Mobilitätsdatenerhebung festlegen. „Neueste und bewährte Erhebungstechnologien und -verfahren sowie Finanzierungs- und Organisationsmodelle werden mit ihren Vor- und Nachteilen analysiert und modular zu einem tragfähigen Erhebungskonzept zusammengestellt. [...] Vorgesehen ist dabei eine breit angelegte Einbindung aller relevanten Akteure des „Verkehrswesens“ wie Bundesländer, Städte, Statistik Austria, Interessensvertretungen, Verkehrsanbieter, Forschungseinrichtungen, Verkehrsplaner, etc., um die Anforderungen dem Bedarf anzupassen“⁵² - so die Einleitung im Endbericht zur Konzeptstudie Mobilitätsdaten Österreichs (KOMOD).

Die Studie beschäftigt sich dabei anfangs mit den Anforderungen der Erhebung und Analyse. Dabei werden vorhandene Mobilitätsdatenerhebungen untersucht sowie Meinungen von Experten, Stakeholdern und Workshopergebnisse berücksichtigt. Außerdem werden die konkreten Inhalte von Erhebung, Methoden und Komponenten, Technologieeinsatz, Entwicklung von Modulen, Organisations- und Finanzierungsmodelle sowie Anforderungen des Datenschutzes behandelt.

Zudem wurden im Rahmen der Ergebnisse von KOMOD folgende Ziele vereinbart⁵³:

- Schaffung einer einheitlichen, aktuellen und qualitativ hochwertigen, bundesweiten Datengrundlage zum Mobilitätsverhalten der Österreicher,
- Etablierung eines methodischen Standards, um bessere Vergleichbarkeit von Erhebungen unterschiedlicher Institutionen und Zeitpunkte zu ermöglichen,
- weitreichende Verfügbarkeit der Daten,⁵⁴
- Nutzen von Synergien mit anderen Gebietskörperschaften,

⁵²vgl. [Fellendorf u. a., 2011, S.11]

⁵³[Kirnbauer, 2016]

⁵⁴Die (Roh-)Daten sind mittlerweile nach dem *Open Data* Prinzip frei verfügbar unter <https://goo.gl/KL1Bhh> abrufbar. Zur Nutzung ist allerdings die Unterzeichnung einer Datennutzungserklärung notwendig.

Österreich Unterwegs ist somit nicht nur eine Datenerhebung, sondern vielmehr ein Prozess und Leitprojekt rund um die Zurverfügungstellung von Mobilitätsdaten in Österreich. Diese finden später ein durchaus breites Feld von Nutzergruppen. Dazu zählen neben Verwaltungsorganen (Bund, Länder, Gemeinden) auch Universitäten und Forschungseinrichtungen, Ingenieurbüros und Verkehrsplaner, Verkehrsdienstleister, Verkehrsverbünde und Infrastrukturgesellschaften. Neben diesen wurden in der KOMOD-Studie, der Vorstudie zu *Österreich Unterwegs* von Experten auch noch Unternehmen der Mobilitätsindustrie, Energieversorger, Regionalverbände, Tourismus und Freizeitwirtschaft, NGOs des Verkehrssektors (zB ÖAMTC, ARBÖ, VCÖ,...) und viele weitere genannt.⁵⁵

4.1.3 Methodik der Erhebung

Österreich Unterwegs wurde als ganzjährige Haushaltserhebung angelegt und zwischen Herbst 2013 und Herbst 2014 durchgeführt. Diese bestand aus einem Wegeprotokoll für zwei aufeinanderfolgende Stichtage und einem Bogen zu Mobilitätsfragen den ganzen Haushalt betreffend. Die Befragten wurden nicht alle zur gleichen Zeit kontaktiert, um auch Daten im Jahreszeitenvergleich analysieren zu können.

Die Stichprobe wurde per Zufallsauswahl aus dem Zentralen Melderegister gezogen und entspricht allen Anforderungen einer österreichweiten repräsentativen Stichprobe. Insgesamt wurden 65.080 Haushalte kontaktiert.



Abbildung 4.1: Ablauf des Prozesses rund um Österreich Unterwegs. [Riegler, 2016, S.14]

Die Grafik zeigt den Ablauf des Prozesses *Österreich Unterwegs* als Ganzes - ausgehend von Vorarbeiten im KOMOD-Projekt bis hin zur Datenbereitstellung.

Wie Riegler⁵⁶ ausführt, wurde als Teil des „KOMOD-konformen Erhebungsdesigns“ in einem mehrstufigen Rückkopplungsprozess ein Methodenmix aus schriftlich-postalischer, telefonischer und webbasierter Befragung entwickelt, um durch eine möglichst große Rücklaufquote die Qualität der Erhebung zu sichern. Die aufgesplitteten Rücklaufwege zielten dabei auf

⁵⁵vgl. [Fellendorf u. a., 2011, S.63]

⁵⁶vgl. [Riegler, 2016]

verschiedene Personengruppen in der Bevölkerung ab. Während man jüngere und technologie-affine Bevölkerungsgruppen mit der Möglichkeit der webbasierten Fragebogen-Beantwortung zu motivieren versuchte, blieb eher technologie-unerfahrenen Personen immer noch die Möglichkeit der klassischen Beantwortung per Post oder Telefon.

Elementarer Bestandteil des Prozesses rund um *Österreich Unterwegs* war auch die Qualitätssicherung:

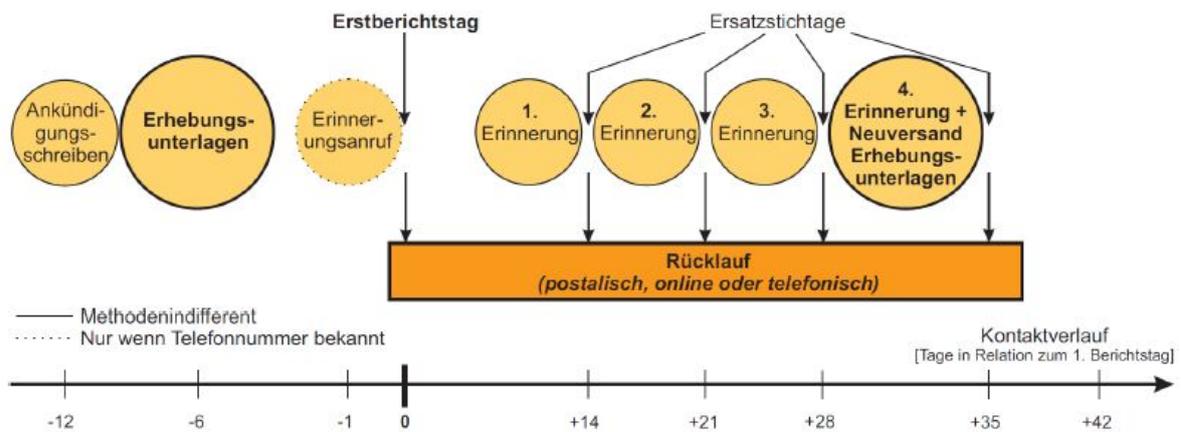


Abbildung 4.2: Schematische Darstellung des Befragungsablaufs mit Fokus auf die Qualitätssicherung durch Erinnerungen. [Riegler, 2016, S.14]

Innerhalb von sechs Wochen wurde durch aktive und wiederholte Kontaktaufnahme versucht, alle Haushalte zur Teilnahme zu bewegen. Dies führte zum Resultat, dass mit 28% knapp mehr als jeder vierte Haushalt an der Befragung teilnahm. Aufgrund der in der Vorstudie festgelegten Qualitätskriterien der Antwortbögen konnten am Ende nur 26% als verwertbar eingeordnet werden.

Bezüglich der Rücklaufquote gilt es noch anzumerken, dass mit knapp 75% die Mehrheit der Haushalte den klassischen Papierfragebogen zur Übermittlung der Antworten wählte. Die Telefoninterviews (9%) und Online-Fragebögen (16%) spielten nur eine untergeordnete Rolle, wobei die Qualität der Antworten (vor allem bei den Telefoninterviews) deutlich über der der klassischen Papierfragebögen lag.

Für das weitere Verständnis der Erhebung ist noch der Fragebogen selbst von Bedeutung. Jeder der ausgewählten Haushalte bekam im Format DIN A3 einen Haushaltsfragebogen und einen personenbezogenen Fragebogen. In zweitem mussten für eine Person eines Haushalts an zwei aufeinander folgenden Tagen alle Wege im Detail eingetragen werden. Trotz der laufenden Qualitätssicherung ist hier die Fehleranfälligkeit durch vergessene Wege oder auch durch Überschätzung von Distanzen und Zeit durchaus gegeben. Deswegen war die Datenaufbereitung auch von enormer Bedeutung.

② Personenbogen: erster Berichtstag



Wochentag: Datum: Bitte Zuordnungsnummer für Person im Haushalt eintragen: Bitte Vorname/ Namenskürzel eintragen:

Waren Sie an diesem Tag außer Haus? ja nein → bitte zum zweiten Berichtstag (auf der Rückseite)

Ein Hinweis zu Beginn: Ein Weg ist hier immer eine Strecke von einem Ausgangspunkt zum Ziel. Nach einem Zwischenstopp beginnt ein neuer Weg. Vergessen Sie bitte nicht, auch die Wege nach Hause als eigene Wege einzutragen!

Wo begann Ihr erster Weg?
zu Hause anderer Ort → Postleitzahl/ Gemeinde: Straße:

	Erster Weg	Zweiter Weg	Dritter Weg	Vierter Weg	Fünfter Weg	Sechster Weg	Siebenter Weg
Wann begann der Weg? Stunde: <input type="text"/> Minute: <input type="text"/> Uhr: <input type="text"/>							
Zu welchem Zweck haben Sie den Weg unternommen? Bitte nur eine Angabe!	zur Arbeit <input type="checkbox"/> dienstlich/geschäftlich <input type="checkbox"/> Schule/Ausbildung <input type="checkbox"/> Bringen/Holen/Begleitung von Personen <input type="checkbox"/> Einkauf <input type="checkbox"/> private Erledigung <input type="checkbox"/> privater Besuch <input type="checkbox"/> sonstige Freizeit <input type="checkbox"/> nach Hause <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zur Arbeit <input type="checkbox"/> dienstlich/geschäftlich <input type="checkbox"/> Schule/Ausbildung <input type="checkbox"/> Bringen/Holen/Begleitung von Personen <input type="checkbox"/> Einkauf <input type="checkbox"/> private Erledigung <input type="checkbox"/> privater Besuch <input type="checkbox"/> sonstige Freizeit <input type="checkbox"/> nach Hause <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zur Arbeit <input type="checkbox"/> dienstlich/geschäftlich <input type="checkbox"/> Schule/Ausbildung <input type="checkbox"/> Bringen/Holen/Begleitung von Personen <input type="checkbox"/> Einkauf <input type="checkbox"/> private Erledigung <input type="checkbox"/> privater Besuch <input type="checkbox"/> sonstige Freizeit <input type="checkbox"/> nach Hause <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zur Arbeit <input type="checkbox"/> dienstlich/geschäftlich <input type="checkbox"/> Schule/Ausbildung <input type="checkbox"/> Bringen/Holen/Begleitung von Personen <input type="checkbox"/> Einkauf <input type="checkbox"/> private Erledigung <input type="checkbox"/> privater Besuch <input type="checkbox"/> sonstige Freizeit <input type="checkbox"/> nach Hause <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zur Arbeit <input type="checkbox"/> dienstlich/geschäftlich <input type="checkbox"/> Schule/Ausbildung <input type="checkbox"/> Bringen/Holen/Begleitung von Personen <input type="checkbox"/> Einkauf <input type="checkbox"/> private Erledigung <input type="checkbox"/> privater Besuch <input type="checkbox"/> sonstige Freizeit <input type="checkbox"/> nach Hause <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zur Arbeit <input type="checkbox"/> dienstlich/geschäftlich <input type="checkbox"/> Schule/Ausbildung <input type="checkbox"/> Bringen/Holen/Begleitung von Personen <input type="checkbox"/> Einkauf <input type="checkbox"/> private Erledigung <input type="checkbox"/> privater Besuch <input type="checkbox"/> sonstige Freizeit <input type="checkbox"/> nach Hause <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zur Arbeit <input type="checkbox"/> dienstlich/geschäftlich <input type="checkbox"/> Schule/Ausbildung <input type="checkbox"/> Bringen/Holen/Begleitung von Personen <input type="checkbox"/> Einkauf <input type="checkbox"/> private Erledigung <input type="checkbox"/> privater Besuch <input type="checkbox"/> sonstige Freizeit <input type="checkbox"/> nach Hause <input type="checkbox"/> anderes, und zwar: <input type="text"/>
Welche Verkehrsmittel haben Sie benutzt? Wenn Sie mehrere Verkehrsmittel genutzt haben, geben Sie alle an. Geben Sie bitte auch an, wenn Sie ein Stück zu Fuß oder einen kompletten Fußweg zurückgelegt haben.	zu Fuß <input type="checkbox"/> Fahrrad <input type="checkbox"/> Pkw als Lenker/in <input type="checkbox"/> Pkw als Mitfahrer/in <input type="checkbox"/> Moped/Motorrad <input type="checkbox"/> Stadt-/Regionalbus <input type="checkbox"/> U-Bahn/Straßenbahn <input type="checkbox"/> S-Bahn/Eisenbahn <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zu Fuß <input type="checkbox"/> Fahrrad <input type="checkbox"/> Pkw als Lenker/in <input type="checkbox"/> Pkw als Mitfahrer/in <input type="checkbox"/> Moped/Motorrad <input type="checkbox"/> Stadt-/Regionalbus <input type="checkbox"/> U-Bahn/Straßenbahn <input type="checkbox"/> S-Bahn/Eisenbahn <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zu Fuß <input type="checkbox"/> Fahrrad <input type="checkbox"/> Pkw als Lenker/in <input type="checkbox"/> Pkw als Mitfahrer/in <input type="checkbox"/> Moped/Motorrad <input type="checkbox"/> Stadt-/Regionalbus <input type="checkbox"/> U-Bahn/Straßenbahn <input type="checkbox"/> S-Bahn/Eisenbahn <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zu Fuß <input type="checkbox"/> Fahrrad <input type="checkbox"/> Pkw als Lenker/in <input type="checkbox"/> Pkw als Mitfahrer/in <input type="checkbox"/> Moped/Motorrad <input type="checkbox"/> Stadt-/Regionalbus <input type="checkbox"/> U-Bahn/Straßenbahn <input type="checkbox"/> S-Bahn/Eisenbahn <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zu Fuß <input type="checkbox"/> Fahrrad <input type="checkbox"/> Pkw als Lenker/in <input type="checkbox"/> Pkw als Mitfahrer/in <input type="checkbox"/> Moped/Motorrad <input type="checkbox"/> Stadt-/Regionalbus <input type="checkbox"/> U-Bahn/Straßenbahn <input type="checkbox"/> S-Bahn/Eisenbahn <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zu Fuß <input type="checkbox"/> Fahrrad <input type="checkbox"/> Pkw als Lenker/in <input type="checkbox"/> Pkw als Mitfahrer/in <input type="checkbox"/> Moped/Motorrad <input type="checkbox"/> Stadt-/Regionalbus <input type="checkbox"/> U-Bahn/Straßenbahn <input type="checkbox"/> S-Bahn/Eisenbahn <input type="checkbox"/> anderes, und zwar: <input type="text"/>	zu Fuß <input type="checkbox"/> Fahrrad <input type="checkbox"/> Pkw als Lenker/in <input type="checkbox"/> Pkw als Mitfahrer/in <input type="checkbox"/> Moped/Motorrad <input type="checkbox"/> Stadt-/Regionalbus <input type="checkbox"/> U-Bahn/Straßenbahn <input type="checkbox"/> S-Bahn/Eisenbahn <input type="checkbox"/> anderes, und zwar: <input type="text"/>
Wo lag das Ziel? Vermerken Sie möglichst die genaue Adresse. Wenn Sie diese nicht kennen, notieren Sie eine kurze Beschreibung.	Postleitzahl/Gemeinde: <input type="text"/> Adresse/Beschreibung: <input type="text"/>						
Wann sind Sie angekommen? Stunde: <input type="text"/> Minute: <input type="text"/> Uhr: <input type="text"/>							
Wie weit war der Weg? ca. <input type="text"/> km							
Haben Sie weitere Wege zurückgelegt? ja <input type="checkbox"/> nein <input type="checkbox"/>							ja <input type="checkbox"/> nein <input type="checkbox"/> Anzahl weiterer Wege: <input type="text"/>

Bitte füllen Sie auch die Rückseite für Ihren zweiten Berichtstag aus!

Abbildung 4.3: Personenbezogener Fragebogen Österreich Unterwegs. [Website bmvit]

4.1.4 Methodik der Datenaufbereitung

Fehler und Verzerrungen sind bei großen Erhebungen nichts Außergewöhnliches. Bei *Österreich Unterwegs* wurden zum Großteil zwei verschiedene Verfahren verwendet, um diesen entgegen zu steuern:⁵⁷

- Datenimputation fehlender Angaben (nicht, fehlerhaft oder unvollständig berichtete Wege) auf Basis inhaltsorientierter Regeln, zum Beispiel:
 - Teilung von Wegen,
 - Bereinigung und Zusammenführen von Wegetappen zu Wegen,
 - Imputation fehlender Zwischenwege,
 - Imputation fehlender Heimwege am Ende des Berichtstages.
- Reduzierung bzw. Behebung von, durch das Stichprobendesign und durch unterschiedliche Antwortbereitschaft verursachte, Ziehungs- und Stichprobenfehlern oder Verzerrungen mittels Gewichtung und Hochrechnung

⁵⁷vgl. [Herry u. Tomschy, 2016]

Während die Datenimputation die Bereinigung von Fehlern und daraus resultierend die Schaffung eines korrekten Datensatzes zum Ziel hatte, lag der Schwerpunkt bei der Gewichtung und Hochrechnung in der Maximierung der Repräsentativität der Studie: „Das Wesen und das Ziel der Gewichtung ist es, eine in Bezug auf alle relevanten Mobilitätsindikatoren und auf die Soziodemographie der ProbandInnen, sowie auf die definierten zeitlichen und räumlichen Einheiten, repräsentative Stichprobe herzustellen. Das heißt, dass die Verteilung der relevanten Mobilitätsindikatoren und Merkmale der befragten Personen in der Stichprobe jener in der Grundgesamtheit entspricht. Es ist also darauf zu achten, dass durch die Gewichtung zeitliche, räumliche, soziodemographische und sonstige mobilitätsrelevante Verzerrungen in der Stichprobe bestmöglich ausgeglichen werden.“⁵⁸

Für die Gewichtung wurde ein iteratives Verfahren herangezogen, dessen Schritte sich vereinfacht wie folgt darstellen: Zuerst wurde eine Zielverteilung der Tageswegehäufigkeit basierend auf den Daten von Stichtag 1 ermittelt. Dabei wurde nach mehreren Merkmalen differenziert: Bundesland, Raumtyp, Wochentag, Jahreszeit, Geschlecht, Alter, ... Basierend auf dieser Zielverteilung wurde die Tageswegehäufigkeitsverteilung für den zweiten Stichtag berechnet. Im dritten Schritt passierte die Gewichtung des zweiten Stichtags nach räumlichen, zeitlichen und soziodemographischen Faktoren. In einem weiteren Schritt wurde dies auch für den ersten Stichtag durchgeführt. Dies erfolgte sowohl auf der Ebene der Berichtstage (zB Anzahl der Wege pro Tag) als auch auf der Wegebene (gefahrte Wege, Wegekettten, Wegelängen) Die einzelnen Gewichte wurden auf verschiedenen Ebenen bei jedem Durchlauf neu berechnet und angepasst. So erfolgte eine stückweise Annäherung auf die final verwendeten Gewichtungsfaktoren. Mit diesen Faktoren lassen sich aber lediglich relative Analysen durchführen - also beispielsweise die Tageswegehäufigkeit pro Person an einem Werktag. Um auch absolute Berechnungen zu bekommen, wurden die Daten mithilfe der ermittelten Gewichte hochgerechnet.

Diese Gewichte sind zur Nachvollziehbarkeit auch in den frei verfügbaren Daten hinterlegt und können für eigene Berechnungen verwendet werden.

4.1.5 Ergebnisse

Die Ergebnisse von *Österreich Unterwegs* sind nach Zustimmung einer Lizenzvereinbarung frei verfügbar. Im folgenden sollen diese kurz in ihrem Aufbau beleuchtet werden - auf inhaltliche Ergebnisse wird verzichtet.

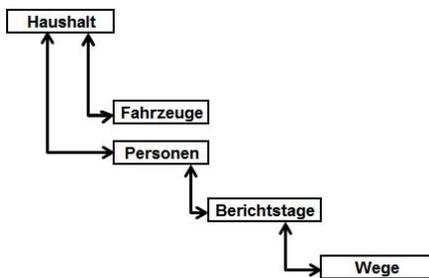


Abbildung 4.4: Ergebnisstruktur [Österreich Unterwegs 2014]

Die zur Verfügung gestellten Daten folgen dabei diesem Aufbau: Ausgehend von der Haushaltsbefragung wurden Datensätze für zur Verfügung stehende Fahrzeuge und Personen je Haushalt erstellt. Darunter folgt die Ebene der Berichtstage. Hier finden sich im wesentlichen groben Daten zum jeweiligen Berichtstag (Fragebogen Grafik 4.2) wie Jahreszeit oder Wochentag. Auf der Ergebnisebene der Wege sind diese dann im Detail sichtbar.

⁵⁸[Herry u. Tomschy, 2016, S.18]

Die Granularität der Daten liegt dabei auf Gemeindeebene. Dies heißt, dass die Daten vor der Veröffentlichung aggregiert worden sind und somit keine Rückschlüsse auf Einzelverhalten möglich ist. Zusätzlich ist eine Zuordnung der Daten zu einem Raumtyp möglich - wie beispielsweise peripheren Bezirk oder Großstadt. Die Daten sind sowohl als csv-File als auch als SPSS-Datenblatt verfügbar.

Kurzfasit

Obwohl in der KOMOD-Studie durchaus innovative Ansätze aufscheinen ist *Österreich Unterwegs 2014* eine klassische Haushaltsbefragung. Dabei wurden Informationen aus dem physisch-materiellen und geographischen Raum abgefragt, die nun als tabellarische Information vorliegen. Über Zuordnungskennzahlen ist zwar eine Verräumlichung möglich, allerdings nur wieder zurück zu räumlichen (Verwaltungs-)Einheiten im physisch-materiellen und geographischen Raum. Zwar können mithilfe der Daten auch Relationen abgebildet werden - dies aber auch nur zwischen Gebietskörperschaften, was nicht im Sinne eines relationalen Raumverständnisses ist. Die Aggregation der Daten lässt keine eigene von Verzerrungen befreite Bearbeitung zu. Der einzig innovative Aspekt bleibt die freie Nutzung als Open Government Data.

4.2 ... zur passiven Beobachtung - Neue Formen der Datenerhebung

Die bereits in der Einleitung geschilderten, sich verändernden Rahmenbedingungen rund um *Big Data* und Digitalisierung machen auch vor dem Bereich der Mobilitätsdatenerhebungen nicht Halt. Der große Trend kann wie in der Überschrift zusammengefasst werden - weniger aktive Befragung und mehr passive Beobachtung. Dies führt zu größeren Stichproben und auch Langzeitbeobachtungen werden möglich - im Gegenzug müssen Überlegungen zur Datengüte, Repräsentativität und zum Datenschutz angestellt werden, die oftmals im Detail um einiges komplexer sind, als diese vorab erscheinen. Vor allem im Bereich der Datenqualität und Repräsentativität offenbaren sich bei vielen technologischen Ansätzen noch Schwächen im Detail.

Im Folgenden sollen einige Beispiele automatisierter Datenerhebungen als Vorstufe zum späteren Hands-On Beispiel gezeigt werden. Einleitend dazu werden die wichtigsten Inhalte betreffend technologischer Datenerhebungen der schon zuvor erwähnten KOMOD-Studie erläutert werden.

4.2.1 Technologieeinsatz bei Mobilitätserhebungen

Auch wenn die Nachfrage nach dem Einsatz technologischer Hilfsmittel im Erhebungsdesign nicht quantifizierbar ist, bleibt beim Lesen des KOMOD-Berichts das persönliche Gefühl, dass die Lager hier wohl etwas gespalten waren - zwischen Befürwortern der klassisch-bewährten Methoden und Befürwortern der modern-innovativen Methoden. Seit der letzten Erhebung 1995 hat sich nicht nur das Verkehrsverhalten der Österreicher verändert, auch die Technologie im Bereich der GPS- und GSM-basierten Erhebungsmethoden hat sich enorm weiterentwickelt.

Große Teile der KOMOD-Studie handeln davon, warum die Entscheidung bei *Österreich Unterewegs* schlussendlich auf ein eher klassisches Erhebungsdesign fiel. In der über 400-Seiten umfassenden Studie war es ein allgemeiner Wunsch von Experten, den Einsatz von technologiegestützten Erhebungsmethoden zu forcieren und auch ein Testfeld für neue Erhebungstechnologien aufzubauen.⁵⁹ Auch bei der Frage nach der Einschätzung von verschiedenen Methoden zur Eignung für die Erhebung des Mobilitätsverhaltens wurde die Wegedatenaufzeichnung mittels GPS-Geräten bei der Auswertung der Expertenfragebögen mit durchschnittlich 6,4 von 10 Punkten bewertet - der vierthöchste Wert. Die Kombination verschiedener Ansätze fand mit 8,2 Punkten den größten Zuspruch.

Deswegen wurde der Technologieeinsatz genau analysiert und beschrieben - spätestens hier ist jedoch anzumerken, dass die Studie im Jahr 2011 veröffentlicht wurde und einige Beispiele zeigen, wie schnell sich die Technologien rund um Smartphones in den letzten fünf Jahren weiterentwickelt haben und wie schwer es ist, zuverlässige Prognosen betreffend technologischer Entwicklungen zu erstellen. Hier nur ein kurzes Beispiel: In der Studie wurde die Problematik der verschiedenen Plattformen im Bereich der Smartphones thematisiert und diskutiert und welche davon zukünftig den Markt beherrschen werden. So wurde davon ausgegangen, dass

⁵⁹vgl. [Fellendorf u. a., 2011, S.60]

bis zum Jahr 2015 die Windows 7 Plattform die iOS Plattform von Apple in den Marktanteilen überholen wird. Tatsächlich ist laut aktuellen Zahlen Ende 2016 Windows auf mobilen Endgeräten beinahe bedeutungslos.⁶⁰

Die Unterschiede bzw. Vor- und Nachteile von Smartphones gegenüber GPS-Empfängergeräten wurden bereits in den Grundlagen thematisiert. Im Folgenden sollen die Vor- und Nachteile in der Anwendung sowie Konzepte zur Erhebung diskutiert werden.

Im Bereich der Mobilitätserhebungen können grundsätzlich folgende Formen unterschieden werden:

Tabelle 4.1: Einteilung der Tracking-Systematik⁶¹

	Fahrzeug	Persönlich
Passiv	Passives Fahrzeug-Tracking	Persönliches passives Tracking
Aktiv	Aktives Fahrzeug-Tracking	Persönliches aktives Tracking

Diese Einteilung folgt einer ganz klaren Logik: Das GPS Gerät wird entweder an einem Fahrzeug fix montiert oder von der Person getragen und zeichnet die zurückgelegten Wege entweder automatisch (also passiv ohne Zutun der Person) oder nach Aktivierung auf.

Die fixe Montage des Geräts an oder in einem Fahrzeug hat zwar den Vorteil, dass beispielsweise an mehreren Tagen alle Fahrten eines PKWs aufgezeichnet werden können. Der Nachteil liegt aber ganz klar in der Erfassung von Wegen, bei denen ein Verkehrsmittelwechsel stattfindet.

Beim persönlichen Tracking bleibt das GPS-Gerät immer bei der Person selbst. Je nachdem, ob die Erhebung der passiven oder aktiven Systematik folgt, ist durch die Person ein kurzes Aktivieren beim Wegantritt notwendig. Durch die lückenlose Aufzeichnung der GPS-Positionen - im Idealfall verknüpft mit weiteren Sensordaten wie Beschleunigungswerten - können in der Nachbearbeitung auch Verkehrsmittelwechsel erkannt und erhoben werden. Die Trennung der einzelnen Wege gestaltet sich aber oft noch schwierig. Durch längeren Stillstand (beispielsweise durch Empfangsprobleme in Gebäuden oder auch in U-Bahnen) ist die Wegeerkennung fehlerbelastet. Deswegen ist es state of the art, bei GPS-gestützten Erhebungen immer auch mit einer Nachbefragung zu arbeiten. (Siehe ⁶² und folgendes Kapitel zum Projekt *SmartSurvey*). Die Person kann dabei - meist über eine Webplattform - die aufgezeichneten und automatisch erkannten Wege kontrollieren und ergänzen.

Im Bereich der Systematik ist auch immer wieder die Belastung der Erhebungsperson von Relevanz. Im Idealfall sollte diese so gering wie möglich gehalten werden, um einerseits die Motivation zur Zusammenarbeit hoch zu halten und andererseits auch Fehler durch eventuell vergessene Aktivierung von vornherein möglichst zu vermeiden. Zusammenfassend können durch den Technologieeinsatz in Erhebungen folgende Probleme traditioneller Methoden vermindert werden.⁶³

⁶⁰Siehe <http://derstandard.at/2000046888593/Knapp-90-Prozent-Marktanteil-Android-fast-so-dominant-wie-Windows>

⁶²vgl. [Fellendorf u. a., 2011, S.156 ff]

⁶³vgl. [Fellendorf u. a., 2011, S.154] und eigene Überlegungen

- Personen über- oder unterschätzen die Wegelängen und Wegezeiten in Fragebögen,
- Personen berichten Wege bewusst oder unbewusst nicht,
- Personen vergessen kürzere Wege, was wiederum zu einer falschen Wegeanzahl führt,
- die erhobenen Daten können validiert werden, um die Qualität zu erhöhen und nicht ausschließlich auf ausgefüllte Fragebögen zu vertrauen.

4.2.2 Konkrete Beispiele

Im Folgenden sollen drei Beispiele aus der Mobilitätsforschung diese neuen Trends verdeutlichen und bereits genutzte Möglichkeiten aufzeigen. Dabei spielt auch das bereits vorgestellte Konzept der nutzergenerierten Daten bereits eine wichtige Rolle - nichtzuletzt weil die Datenbasis der gezeigten Beispiele meist direkt durch Nutzerverhalten entstanden ist. Die Projekte sind also auch Beispiele, wie nutzergenerierte Daten bereits in Mobilitätsdatenerhebungen genutzt werden.

SmartSurvey - Mobilitätserhebungen mit Smartphones (AIT)

„AIT SmartSurvey bietet ein effizientes, qualitativ hochwertiges und kostengünstiges Service zur Sammlung von Mobilitätsinformationen mit Smartphones. Die zurückgelegten Wege und verwendeten Verkehrsmittel werden mittels Smartphone automatisch erhoben und in einem digitalen Mobilitätstagebuch gespeichert - weitaus genauer als bei einer schriftlichen Erfassung.“
- AIT

Mit SmartSurvey hat das AIT im Prinzip ein digitales Mobilitätstagebuch erschaffen. In Bezug auf Detaillierungsgrad und Umfang wurden dabei folgende Anforderungen berücksichtigt:⁶⁴

- möglichst lückenlose Erfassung aller Wege und Etappen der intermodalen Wegeketten, unabhängig von ihrer Länge und der Fortbewegungsform,
- detaillierte Informationen über Routen- und Verkehrsmittelwahl sowie aller multimodalen Umsteigeorte und -zeiten,
- Langzeiterhebung des Mobilitätsverhaltens über mindestens eine Woche, da sich Multimodalität und Mobilität innerhalb und außerhalb der, Alltagsroutine erst über einen längeren Zeitraum manifestieren,
- kontinuierliche Aktualisierung der Daten für die Analyse von Trends und Anpassungen des Mobilitätsverhaltens an neue Angebote.

⁶⁴vgl. [Widhalm, 2016, S.46]

Der Einsatz von Smartphones anstelle von GPS-Geräten bietet den Vorteil, dass ein Großteil der Teilnehmer der Studie bereits über ein technologisches Endgerät zur Datenerfassung verfügt. Smartphones bieten mit ihren zahlreichen Sensoren wie GPS, Accelerometer, Gyroskop und Kompass zahlreiche Erhebungsmöglichkeiten, welche wiederum eine automatisierte Auswertung erleichtern. So können die Algorithmen relativ genau ermitteln, ob der Teilnehmer gerade in einer U-Bahn, Straßenbahn oder in einem Autobus sitzt - drei Kategorien, die bei *Österreich Unterwegs* beispielsweise zusammengefasst als "Öffentlicher Verkehr" dargestellt werden und hier im Detail unterschieden werden können. Zusätzlich zur automatischen Wegeerkennung haben die Teilnehmer die Möglichkeit über eine Web-Plattform ihre zurückgelegten Wege zu sichten und bei Bedarf anzupassen oder zu bearbeiten.

Im Unterschied zu einer traditionellen Erhebung kommt dem Teilnehmer aber eine deutlich geringere Belastung zu, da das händische Ausfüllen eines aufwendigen Fragebogens entfällt und der Großteil der Daten automatisiert generiert wird. Gleichzeitig können mit *SmartSurvey* aber Erhebungen über längere Zeiträume und damit mehrere „Stichtage“ durchgeführt werden. Zudem enthalten die Informationen beispielsweise bei *Österreich Unterwegs* lediglich die Quelle und das Ziel jedes Weges - jedoch nicht die exakt zurückgelegte Strecke und Route. Wohnort und Arbeitsplatz werden von *SmartSurvey* automatisch erkannt - andere Wegezwecke können vom Teilnehmer zugeordnet werden. Die Wegeaufzeichnung passiert automatisch und wird auf dem Smartphone zwischengespeichert. Der Teilnehmer wird von Zeit zu Zeit aufgefordert, seine aufgezeichneten Wege mit dem Server zu synchronisieren. Erst dann können die Algorithmen Auswertungen durchführen.

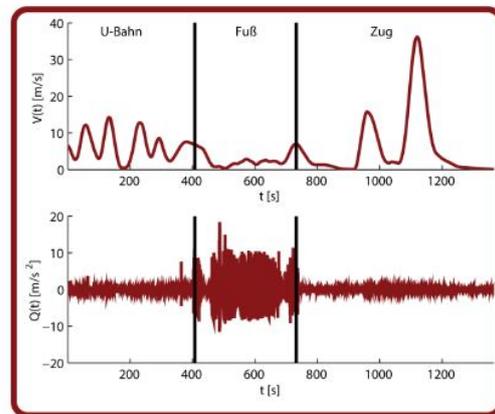


Abbildung 4.5: Sensorenauswertung SmartSurvey. Mithilfe der jeweiligen Geschwindigkeit und der Beschleunigung zum Zeitpunkt t berechnet der Algorithmus, mit welchem Verkehrsmittel der Teilnehmer unterwegs ist. [Widhalm, 2016]

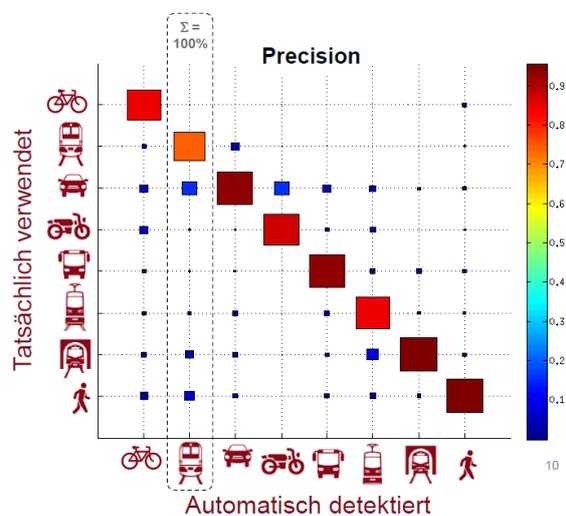


Abbildung 4.6: Verkehrsmitteldetektion SmartSurvey. Über eine Web-Plattform können die Teilnehmer die automatisch-generierten Daten einsehen und überprüfen - der Vergleich zeigt, dass der vom AIT entwickelte Algorithmus durchaus brauchbare Ergebnisse bringt und das Verkehrsmittel in den meisten Fällen richtig erkennt. [Widhalm, 2016]

Die *SmartSurvey* App des AIT wird laufend weiterentwickelt, um in Zukunft auch folgende Herausforderungen und Ziele zu erfüllen:

- Verbesserung der Qualität bei gleichzeitiger Reduzierung des Stromverbrauchs durch stärkere Integration von Karten- und Routinginformationen,
- Verringerung des Aufwands durch Unterstützung bei Eingaben auf Basis von historischen Angaben der Teilnehmer,
- Unterstützung von smarten Endgeräten (zB Smartwatch),
- Echtzeit Verkehrsmittelerkennung zur Integration in innovative Mobilitätslösungen.

Floating-Phone-Daten & Mobilitätsdaten im öffentlichen Verkehr (TU Graz)

Die TU Graz beschäftigte sich in einem Forschungsprojekt mit der Anwendbarkeit von Floating Phone Data (FPD) als Grundlage für Mobilitätsmodellierungen. „FPD basiert auf dem Ansatz, dass sich ein im Fahrzeuginnenraum (MIV oder ÖV) mitgeführtes Mobilfunkgerät während der Fahrt in verschiedenen Mobilfunkzellen entlang der Fahrtstrecke ein- bzw. ausbucht. Somit besteht die Möglichkeit, dass ein Mobilfunknetzbetreiber die Bewegungsdaten über Monitoringsysteme erfasst, anonymisiert und sie weiterverarbeitenden Einrichtungen zur Verfügung stellt.“⁶⁵ Grundlage für alle weiteren Analysen sind sogenannte Trajektorien, also mit Pfaden abgebildete Bewegungsbilder. Von Bedeutung ist dabei die anonymisierte IMSI (International Mobile Subscriber Identity), welche auf einer alle 24 Stunden erneuerten ID-Nummer basiert. Die Trajektorie im Rohzustand bildet dabei aber lediglich die vom Mobiltelefon verwendeten Sendestationen ab, also in einer Genauigkeit, die für die meisten Mobilitätsfragen eher unzureichend sein dürfte. Deswegen kommt in einem nächsten Schritt ein Map-Matching-Verfahren zum Einsatz. Dieses hat zum Ziel, die relativ ungenauen Mobilfunk-Trajektorien auf ein genaueres Netz zu projizieren - welches je nach Anwendung der Schienen- oder Straßeninfrastruktur entspricht.

Dieses Map-Matching ist ein erster Schritt eines Prozesses, welcher zum Ziel hat, zu unterscheiden, ob sich der Nutzer in einem Auto oder einem öffentlichen Verkehrsmittel befindet. Dabei werden folgende weitere Punkte analysiert:

- Ermittlung der Distanz der Mobilfunktrajektorie zur nächstgelegenen Infrastruktur,
- Vergleich mit Fahrplan: Korrelieren bestimmte Häufigkeitsverteilungen (zB Stundentakt) mit dem Fahrplan auf der jeweiligen Strecke? Bei ÖV eher Abbildung mit Höhen und Tiefen (viele Trajektorien zur nahezu identischen Zeit, da alle in gleichem Zug) und bei MIV eher gleichmäßige Verteilung,
- Gibt es markante Aufenthaltsorte im Verlauf der Trajektorie? (Beispielsweise an Umsteigeknoten)
- Ist die Reisegeschwindigkeit typisch für ein Verkehrsmittel?

⁶⁵[Cik, 2016, S.74]

All diese genannten Faktoren sowie die Ausgangsdaten (Mobilfunktrajektorie, Straßen- und Schienennetz und ÖV Fahrplan) werden in einem Tool automatisiert abgearbeitet. Somit sind beispielsweise weitreichende Analysen im schienengebunden Regional- und Fernverkehr möglich. Diese wurden am Beispiel der Erlaufalpbahn (Niederösterreich) veranschaulicht.

Neben Aussagen zu Fahrgastzahlen (welche wiederum mit tatsächlichen Zählungen verglichen wurden) konnten auch interessante Fahrgaststromanalysen durchgeführt werden. Durch vorab definierte Zielgebiete (Linz + Umgebung, Krems + St. Pölten, Wien) konnte das Tool ermitteln, wie sich die Fahrgäste der Erlaufalpbahn auf diese drei Ziele verteilen.

Aufgrund der angesprochenen Ungenauigkeit der Mobilfunktrajektorien funktionieren diese Analysen am besten in einem regional-analytischen Kontext auf der Makroebene. Für Analysen mit Fokus auf das Stadtgefüge wird die Methode wahrscheinlich weniger zufriedenstellende Ergebnisse liefern.

Da die Datenauswertung weitgehend automatisiert abläuft, lässt sich die Analyse auch einfach über mehrere Tage oder Wochen durchführen. So ist eine Beobachtung von Pendlerbewegungen über große Zeitperioden relativ einfach möglich, um auch saisonale Schwankungen abbilden zu können.

In einer weiteren Entwicklungsphase sollen nicht nur Quelle-Ziel Verbindungen abgebildet, sondern die einzelnen Wege und Trajektorien auch um eine Aktivitätenanalyse erweitert werden.

Floating Car Data

Wie schon an mehreren Stellen dieser Arbeit angedeutet, gibt es im Bereich des motorisierten Individualverkehrs bereits seit mehreren Jahren hochtechnologische Entwicklungen zur automatisierten Datengenerierung und Auswertung. Im Zusammenhang mit der inhaltlichen Ebene dieser Arbeit ist hier die Systematik rund um Floating Car Data (FCD) zu nennen. Das Grundprinzip bei FCD beruht dabei auf der Datengenerierung durch (GPS-)Sensoren im Fahrzeug

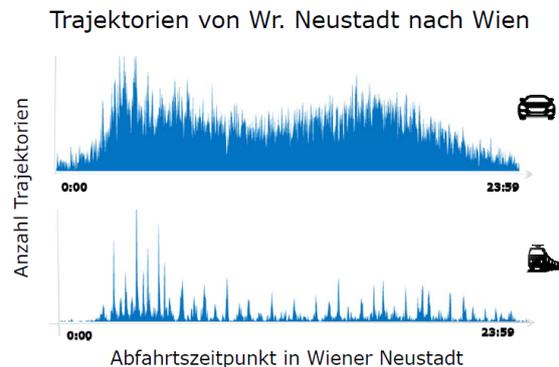


Abbildung 4.7: Vergleich der Anzahl der Trajektorien zum jeweiligen Abfahrtszeitpunkt in Wiener Neustadt. Durch die Überlappung mit Fahrplänen können die Trajektorien entweder dem ÖV oder dem IV zugewiesen werden. Zudem zeigt sich in der Verteilung ein deutlicher Unterschied. [Cik, 2016]

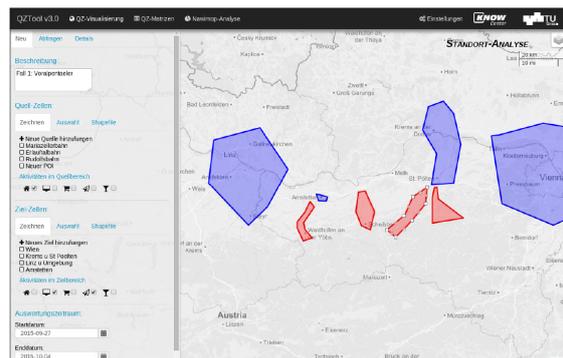


Abbildung 4.8: Im Tool der TU Graz besteht die Möglichkeit, Quell- und Ziel Gebiete festzulegen - jedoch nur auf relativ grober Maßstabsebene. Für Detailanalysen ist im Bereich der Mobilfunkdaten nur begrenztes Potential vorhanden. [Cik, 2016]

(zB Mobiltelefon oder Navigationssysteme) oder durch das Fahrzeug selbst. Diese ermittelten Standortdaten werden zur zentralen Auswertung an einen Server übermittelt und sind Basis für (Echtzeit-)Analysen zur Verkehrssituation und Reisegeschwindigkeit sowie natürlich zur Berechnung von Belastungen einzelner Streckenabschnitte. Für die Datenübermittlung ist ein GSM/UMTS/LTE Modul notwendig. Deshalb werden einige Projekte im Bereich FCD auch mithilfe von Smartphones durchgeführt (Zum Beispiel [Rehrl u. a., 2016]).

Eine Weiterentwicklung von FCD stellt extended Floating Car Data (xFCD) dar. „Dabei werden fahrzeugseitige Betriebsdaten wie beispielsweise Kraftstoffverbrauch, Motordrehrate oder Scheibenwischeraktivität aufgezeichnet. [...] Die xFCD-Daten bieten zudem neue Analyse- und Anwendungsmöglichkeiten, wie etwa die Bewertung der Energieeffizienz einzelner Streckenabschnitte oder des Einflusses von Schnee- oder Niederschlagsereignissen auf die segmentspezifische Verkehrsqualität, wodurch Entscheidungsprozesse für die Verkehrssteuerung unterstützt werden.“⁶⁶

Im Rahmen der Generierung von FCD Daten stellt sich immer die Frage nach der zugrundeliegenden Flotte an Fahrzeugen. Diese sind die Basis der Erhebung und auch der Repräsentativität. In Wien werden beispielsweise FCD-Sensoren aus Taxis für die Verkehrsanalysen verwendet.

Rund um den Begriff Floating Car Data gibt es bereits sehr viele praktische Projekte und wissenschaftliche Artikel. Da der Fokus der Arbeit auf dem Radverkehr liegt soll das Thema FCD zwar nicht unerwähnt bleiben, aber auch nicht weiter verfolgt werden. Es ist aber nicht auszuschließen, dass zukünftig auch FCD-Konzepte bei Fahrrädern zur Anwendung kommen - beispielsweise zur Diebstahlprävention- und verfolgung.

Kurzfasit

Das Kapitel 4.2 gibt Einblick in aktuelle Entwicklungen und Projekte im Bereich der Mobilitätserhebungen. Diese unterscheiden sich in ihrem Ansatz und Prozess deutlich von klassischen Erhebungen wie *Österreich Unterwegs*. Bei den meisten Ansätzen werden dabei nicht nur einfach Informationen aus dem Raum-Behältnis entnommen (wie bei I) sondern auch Analysen mithilfe des relationalen Raumzugangs gedacht und durchgeführt. Dabei spielen Grenzen keine Rolle mehr - der Raum bildet sich nur über seine relationalen Interaktionen ab. Dies ist beispielsweise bei FCD-Daten der Fall, aber auch bei *SmartSurvey*. Da letzteres als durchaus innovativer und verbreiteter Ansatz einer technologischen Mobilitätserhebung gesehen werden kann wird *SmartSurvey* auch im Kapitel 6.1 mit *Österreich Unterwegs* und den Ergebnissen aus dem empirischen Hands-On Beispiel einem genaueren Vergleich unterzogen. Alle drei Projekte können auch als Anwendungsfeld von nutzergenerierten Daten gesehen werden.

⁶⁶ [Krampe u. Pucher, 2016, S.62]

5 Einsatzmöglichkeiten nutzergenerierter Mobilitätsdaten - Ein Hands-On Beispiel

Die theoretischen Rahmenbedingungen zu nutzergenerierten Daten werden nun um eigene Hands-On Erfahrungen ergänzt. Die folgenden Kapitel sollen zeigen, welche Herausforderungen die bereitgestellten Fahrrad-Tracking-Daten mit sich bringen, welche Fragestellungen sich gut und weniger gut beantworten lassen und welche konkreten Ergebnisse im Rahmen dieser Diplomarbeit entstanden sind. Augenmerk liegt dabei auf einem Ausloten von Möglichkeiten und Grenzen in Anwendungsfeldern der Stadt- und Regionalforschung - wobei auch hier der Fokus auf den Bereich der Mobilitätsforschung gelegt wird. Was die Wünsche der Regionalwissenschaften an zukünftige Generationen nutzergenerierter Mobilitätsdaten sind, wird im Fazit und in den darauf folgenden Kapiteln Thema sein. In Vorbereitung auf das Hands-On Beispiel werden einige Grundlagen zum Radverkehr in Wien besprochen werden. Dies soll die Einordnung der später dargestellten Ergebnisse vereinfachen.

5.1 Allgemeines zum Radverkehr in Wien

Besonders in den Städten bietet das Fahrrad in der Alltagsmobilität große Vorteile. Dazu zählt vor allem die Schnelligkeit, die durch den geringen Platzbedarf und das dadurch ermöglichte Umfahren von Staubereichen, wegfallender Wartezeiten im Vergleich zum ÖV und das beschleunigte Abstellen (im Vergleich zum Parken eines PKWs) entsteht. Eine aktuelle Studie⁶⁷ zeigt, dass das Fahrrad bis ca. 4 Kilometer das schnellste Fortbewegungsmittel ist. Geschlagen könnte es in Zukunft nur von seiner Weiterentwicklung, dem Pedelec⁶⁸ werden. Die Durchschnittsgeschwindigkeit von Radfahrern im Stadtverkehr liegt aber bereits jetzt bei 15 bis 18km/h in Wien - die Wunschgeschwindigkeit bei 21 km/h⁶⁹. Deswegen ist es nicht nur planerischer Grundgedanke, sondern auch politisches Ziel, den Radverkehrsanteil am Modal Split weiter zu erhöhen.

Die dargestellten Grundlagendaten zum Radverkehr basieren auf Ergebnissen der Studie *Österreich Unterwegs 2014*. Die Wegedaten wurden nach Bundesland (=Wien) und dem Verkehrsmittel (=Fahrrad) selektiert. Diese Auswahl ergab eine Teilstichprobe von n=1373. Die Darstellung zeigt die zurückgelegten Distanzen kategorisiert. Betrachtet man die Daten absolut, so zeigen sich nach oben hin extreme Ausreißer (mit über 100km), weswegen der Mittelwert nur bedingt für Aussagen zur durchschnittlichen Distanz geeignet ist. Da Median und Modus aber mit 3 Kilometern übereinstimmen, lässt sich feststellen, dass ein typischer zurückgelegter

⁶⁷vgl. [Bmvit, 2016]

⁶⁸Pedal Electric Cycle - Fahrrad unterstützt beim Treten den Fahrer mit elektrischer Energie

⁶⁹vgl. [Bmvit, 2016]

Weg in Wien mit dem Fahrrad bei ungefähr 3 Kilometern liegt. Der österreichweite Durchschnitt liegt bei 3,5 Kilometern.

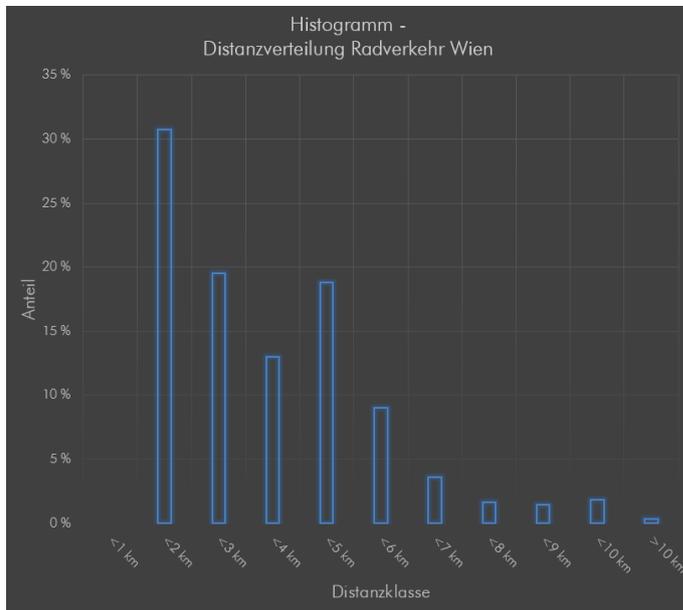


Abbildung 5.1:
Distanzverteilung Radverkehr Wien.
[Eigene Darstellung basierend auf Österreich Unterwegs 2014]

Österreich Unterwegs kommt aufgrund tiefergehender Analysen zu folgendem Schluss in Bezug auf die Nutzung des Fahrrads: „Beim Fahrradverkehr ist die Standardabweichung mit 25 Prozentpunkten durchschnittlicher Abweichung der Einzelwerte vom Mittelwert besonders hoch. Dies liegt daran, dass es sehr viele Personen gibt, die das Fahrrad gar nicht benutzen, und Personen, die an Werktagen relativ viele Wege mit dem Fahrrad zurücklegen. Zusammengefasst lässt sich sagen, entweder man ist ein häufiger Fahrradnutzer oder man benutzt es fast gar nicht.“⁷¹

In *Österreich Unterwegs* wurde auch erhoben, ob den Befragten ein Fahrrad zur Verfügung steht. Darauf gaben in Wien 51% an, ein eigenes Fahrrad zur Nutzung zur Verfügung zu haben. Österreichweit verfügen 68% der Haushalte über zumindest ein eigenes funktionstüchtiges Fahrrad, Wien liegt hier also unter dem Durchschnitt.

Abschließend sei zum Radfahren gesagt: Das Fahrrad hat - vor allem im urbanen Bereich - in seiner Konzeption als Verkehrsmittel riesige Vorteile - sowohl für den Nutzer (Geschwindigkeit, kostengünstig, körperliche Aktivität), als auch für die Stadt im Allgemeinen durch die weitere Verlagerung im Modal Split bei gleichzeitiger Abnahme des MIV. In Hinblick auf die Förderung und Attraktivierung von aktiver Mobilität hat das Fahrrad auch in der Forschung und Planung in den letzten Jahren stark an Stellenwert gewonnen.

Zur Bedeutung des Radverkehrs in der Wiener Mobilität liefert der Modal Split Antworten. Um das Fahrrad als Alltagsfortbewegungsmittel anstatt als Freizeitaktivität in den Vordergrund zu rücken, wurden nur Daten von Werktagen analysiert. Hier zeigt sich, dass der Radverkehr einen Anteil von 4% (n=35842) am gesamten Verkehr in Wien trägt. Ein Vergleich mit der Radfahr-Metropole Kopenhagen beispielsweise macht deutlich, dass verglichen mit den dortigen 30%⁷⁰ in Wien noch viel Potential vorhanden ist, auch wenn der Vergleich aufgrund der sehr unterschiedlichen Gegebenheiten - sowohl in kultureller Hinsicht als auch städtebaulich - so eigentlich nicht aussagekräftig ist. *Österreich Unterwegs*

⁷⁰vgl. http://www.epomm.eu/tems/result_city.phtml?city=227&list=1 - zuletzt aufgerufen am 20.1.2017

⁷¹vgl. [Tomschy u. a., 2016]

5.2 Methodische Vorgehensweise

Kapitel 5 - Hands-On Fahrrad-Tracking-Daten		
Vorbereitung	Durchführung	Ergebnisse
5.3 Der Bike Citizens Datensatz Inhaltlicher Aufbau, Datenschutz, Verfügbare Attribute	5.6 Makro Ebene Visualisierung, Radfahren und Wetter	5.9 Fazit und Diskussion Defizite und Potentiale
5.4 Datenaufbereitung und -verarbeitung Konvertierung, Map Matching, Informationsdichte	5.7 Meso Ebene Routensplit, Einzugsbereich, Verkehrsspinne	
5.5 Daten-Check-Up Repräsentativität, Datengüte, Nutzergruppen	5.8 Mikro Ebene Geschwindigkeit, Kreuzungen	

Abbildung 5.2: Aufbau Kapitel 5 [Eigene Darstellung]

Im folgenden Abschnitt werden die *Bike Citizens* Daten als Beispiel für nutzergenerierte Mobilitätsdaten auf ihre planerischen Potentiale überprüft und somit die Basis zur Zielerreichung der Arbeit gelegt. In den Kapiteln 5.3 bis 5.5 werden notwendige vorbereitende Schritte der Datenaufbereitung und Ergänzung gezeigt. Zudem soll der Datensatz vor dem Hintergrund der Überprüfung von Gütekriterien einem ersten Check-Up unterzogen werden.

Beim Kern des empirischen Teils und den inhaltlichen Analysen fiel die methodische Entscheidung auf eine Ebenenbetrachtung, wobei die Dimensionen selbst definiert wurden. Die Ebenen unterscheiden sich in ihrem Maßstab und bieten eine unterschiedliche räumlich-inhaltliche Granularität. Letzteres begründet sich mit der Menge zu verarbeitender Daten und den vorhandenen Ressourcen. Zudem kann die angestrebte Einschätzung der Potentiale und Herausforderungen der Tracking-Daten auch über passend gewählte beispielhafte Analysen erfolgen. Die Erstellung einer Ebenenlogik erleichtert die Kategorisierung und Einordnung räumlicher Analysen. Für das Verständnis der Ebenenbetrachtung ist ebenso die Interskalarität von Bedeutung. Diese meint, dass sich räumliche Phänomene in ihrer Wirkung nicht nur auf eine Ebene beschränken, sondern ihre Wirkung über mehrere entfalten.

Im Sinne eines Raumverständnisses, dass räumliche Prozesse und Phänomene nicht von Verwaltungseinheiten abgegrenzt werden, dienen dargestellte Bezirks- oder Stadtgrenzen lediglich der Orientierung des Betrachters.

- **Makro-Ebene:** Die Analysen der Makro-Ebene beschäftigen sich mit dem Großraum Wien als räumliche Einheit, wobei das Umland durch die Datenverfügbarkeit begrenzt wird. Ziel sind Aussagen zum Radverkehr allgemein (Statistische Auswertungen), Einfluss des Wetters auf den Radverkehr oder auch eine grobe Quell- und Zielverteilung der Tracks. Dies ist die größte Maßstabsebene der Analyse.

- **Meso-Ebene:** Die Meso-Ebene stellt eine räumliche Einheit zwischen größter und kleinster Betrachtungsebene dar. Sie dient als mittlere Maßstabsebene und bezieht sich in ihrer Ausprägung auf Stadtteile und größere Quartiere. Die genaue räumliche Definition ist aber von der Fragestellung abhängig und wird zielführend gewählt. Anwendungsbeispiele sind die Analyse von Routenentscheidungen oder auch Einzugsbereichsanalysen für bestimmte Points of Interest.
- **Mikro-Ebene:** Auf der kleinsten Ebene liegt der Fokus sehr tief im Stadtgefüge. Der Analyseausschnitt orientiert sich in der Größenordnung an einem Straßenabschnitt oder einem Kreuzungsbereich. Zu den Zielen gehören hier Aussagen zu Geschwindigkeiten bei Steigungen oder Zeitverlust bei verschiedenen Kreuzungsoptionen. In dieser Ebene werden vor allem die track-points eine große Rolle spielen.

Je nach Ebene werden Fragestellungen überprüft und mithilfe der Daten analysiert. Zudem soll jeweils eine Aussage zur Tauglichkeit der Daten für die jeweilige Analyse getroffen werden. Es werden Werkzeuge herangezogen, die in der Stadt- und Regionalforschung weit verbreitet sind. Ziel ist es auch, zu überprüfen, inwieweit sich nutzergenerierte Daten mit den gewohnten Raumplanungswerkzeugen verarbeiten lassen. Dazu zählen neben statistischen Tools wie SPSS auch GIS-Systeme wie ArcMap oder QGIS bis hin zu Excel und Datenbanktools.

Der Fokus der Hands-On Analyse liegt weniger in der Beantwortung von konkreten Fragestellungen, sondern vielmehr im Aufzeigen eines großen Spektrums an Anwendungsmöglichkeiten, welches Anknüpfungspunkte für weitere vertiefende Arbeiten bietet.

Übergeordnetes Ziel ist es dabei, eine Wissensbasis als Vorbereitung für die Beantwortung der Forschungsfragen zu kreieren. Primär sollen für die Fragen nach den **Einsatzmöglichkeiten** und nach den **Rahmenbedingungen** der Fahrrad-Tracking-Daten Lösungen produziert werden.

5.3 Der Bike Citizens Datensatz

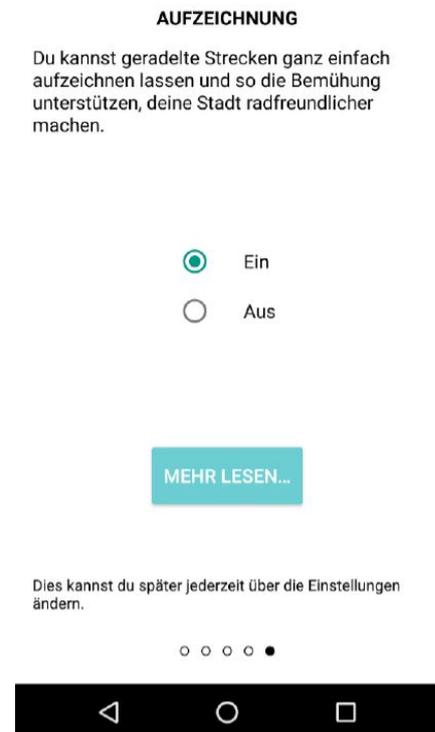


Abbildung 5.3:
Screenshot Bike Citizens App mit Opt-In
Auswahl zur Datenaufzeichnung

„Aus“. Diese Einstellung lässt sich nachträglich jederzeit verändern. Registrierte Nutzer haben weiters die Möglichkeit, aufgezeichnete Strecken einzusehen und erhalten zusätzlich eine persönliche Statistik. Laut *Bike Citizens*⁷³ entscheiden sich ungefähr 80% der Nutzer für eine Aufzeichnung der Wege. Dies variiert jedoch - je nachdem, ob die Option als Opt-In oder Opt-Out aufscheint.

Die datenschutzrechtlichen Rahmenbedingungen wurden bereits in Kapitel 3 ausgeführt. Im folgenden Wortlaut findet sich die Erklärung, die vom Nutzer akzeptiert werden muss, sobald dieser der Aufzeichnung der Fahrten zustimmt:⁷⁴

Für die folgenden Analysen wurden aufgezeichnete Fahrten der Fahrrad-Routing-App *Bike Citizens*⁷² im Großraum Wien verwendet. Die App kann kostenlos auf dem Smartphone (unter Berücksichtigung der technischen Anforderungen) installiert werden - die Karteninhalte sind dann aber meistens kostenpflichtig. Eine Ausnahme stellt Wien dar, wo über eine Partnerschaft mit der Mobilitätsagentur der Stadt die Karten und Inhalte kostenlos verfügbar sind.

Kernnutzen der App ist ein Routing-Service für Fahrradfahrten: Die App berechnet dabei aufgrund der eingestellten Parameter Radprofil (Citybike, Mountainbike, Rennrad) und Routingverhalten („Schnell“ - „Normal“ - „Gemütlich“) den passenden Weg. Dabei greift der Routing-Algorithmus aber nicht nur auf das vorhandene Straßen- und Radverkehrsnetz zu, sondern auch auf Erfahrungswerte von Radfahrkurieren. Zudem kann der Nutzer mithilfe von voreingestellten Strecken durch die Stadt auf Erkundungstour gehen - etwa um touristische Attraktionen zu erleben.

Bei der ersten Installation der App wird der Nutzer gefragt, ob seine Wege aufgezeichnet werden sollen. Voreingestellt ist nach Zufallsprinzip entweder „Ein“ oder

⁷²App erhältlich für

Android: <https://play.google.com/store/apps/details?id=org.bikecityguide>

iOS: <https://itunes.apple.com/at/app/bikecityguide/id517332958>

⁷³vgl. [BikeCitizens, 2017]

⁷⁴*Bike Citizens* App für iOS, aufgerufen am 10.01.2017

Wenn die Fahrtenaufzeichnung aktiviert ist, wird bei Benutzung der Navigation oder des Fahrradcomputers die Fahrt aufgezeichnet. Die gemachten Fahrten können im Menüpunkt Aktivität eingesehen werden. [...]

Warum es diese Funktion gibt

1) Viele finden es sehr motivierend, die mit dem Fahrrad zurückgelegten Fahrten rückblickend einsehen zu können und so z.B. auch anderen einfach zeigen zu können, wie gut das Fahrrad als urbanes Verkehrsmittel geeignet ist.

2) Bei hinreichend großer Datenbasis lassen sich viele wertvolle Informationen extrahieren. Zum Beispiel kann besser beurteilt werden, wo Verbesserungen der Infrastruktur am nötigsten wären oder wo Ampelschaltungen besser koordiniert werden sollten. Aus diesem Grund stellen wir die von App-Benutzern generierten Daten in anonymisierter Form für Städteplaner und Forschungsprojekte zur Verfügung.

3) Auch das *Bike Citizens*-Routing selbst kann mithilfe der aufgezeichneten Daten verbessert werden. So können etwa durchschnittliche Wartezeiten an Kreuzungen berechnet werden, was ein wichtiger Faktor bei der Routenwahl ist. Wartezeiten lassen sich auch aufgrund der statistischen Kartendaten schätzen, allerdings ist dies weniger genau.

Was wird gespeichert?

Entlang einer gefahrenen Strecke werden periodisch folgende Daten gesammelt: Position, Zeitstempel, aktuelle GPS-Genauigkeit, Meereshöhe, Richtung laut Kompass Sensor. Für jede Fahrt werden außerdem folgende Informationen gesendet: eine bei der Installation der App zufällig generierte Seriennummer (um verschiedene Fahrten zuordnen zu können), das Smartphone-Modell, Betriebssystemversion, App-Version, Uhrzeiten zu Beginn und am Ende der Fahrt, Uhrzeit

Wer bekommt die Daten?

Vor Weitergabe an Dritte werden aufgezeichnete Daten grundsätzlich anonymisiert, falls sie das nicht bereits sind. Wir geben grundsätzlich keine Daten an Dritte weiter, außer für Projekte, die direkt oder indirekt auch unseren Benutzern zugute kommen, etwa durch Verbesserung der Radinfrastruktur (z.B. Weitergabe von aggregierten Daten an Körperschaften, zum Beispiel Stadtverwaltungen, um Punkte, an denen ein Ausbau sinnvoll ist, zu identifizieren) oder durch Verbessern des Routing / der Navigation. [...] Es wird keine Weitergabe der Daten zu Werbezwecken geben.

Mithilfe dieser Erklärung wird dem Nutzer klargemacht, was nach der Zustimmung aufgezeichnet wird und wofür die Daten später verwendet werden - so wie es auch die datenschutzrechtlichen Rahmenbedingungen verlangen. Interessant ist dabei auch die Rahmenbedingung, dass die Daten nur für Projekte genutzt werden, die direkt oder indirekt auch den Benutzern zugute kommen.

Der zur Verfügung gestellte Datensatz umfasst alle aufgezeichneten Fahrten der *Bike Citizens* App Nutzer von 1. März 2015 bis 31. August 2016. Insgesamt sind dies 146.947 Fahrten im Großraum Wien. Durch die zunehmende Verbreitung der App zeigt der Monatsvergleich einen stark steigenden Trend. Im August 2016 hatte die App 2271 aktive Nutzer.⁷⁵

Im Folgenden werden einige Aspekte zum technischen Aufbau und Inhalt des Datensatzes dargelegt.

⁷⁵vgl. [BikeCitizens, 2017]

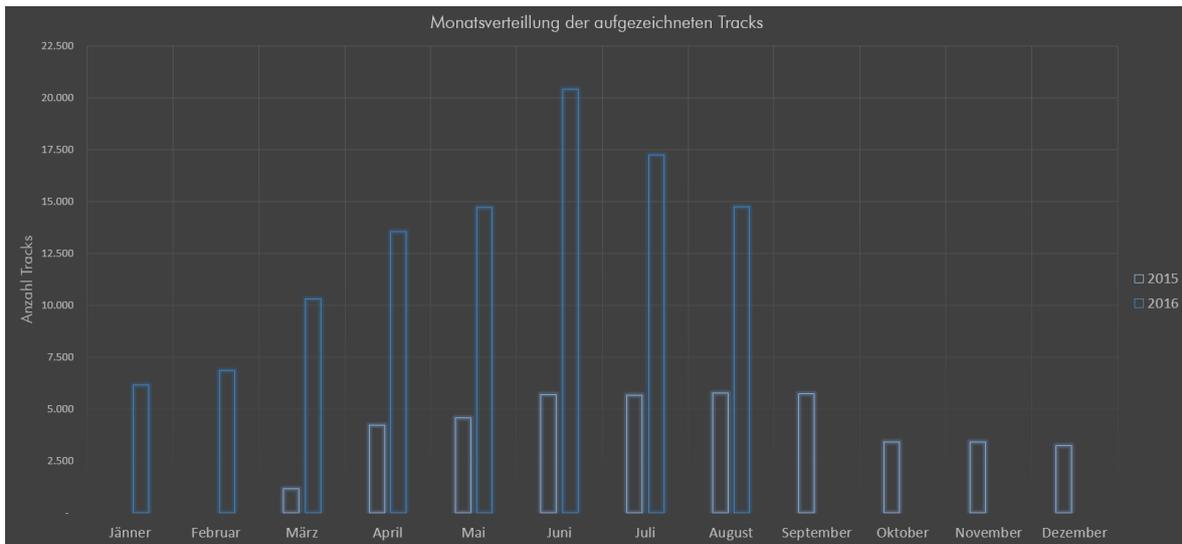


Abbildung 5.4: Anzahl der Tracks im Monatsvergleich [Eigene Darstellung]

Der Datensatz werden in einer .gpx-Datei übermittelt, ein Format welches in der Planung und im Bereich der Geographischen Informationssysteme zurzeit noch nicht weit verbreitet. Gpx ist für den Austausch von GPS-Daten konzipiert, basiert auf dem XML-Standard und vereint Inhalte auf drei Ebenen:⁷⁶ Gpx steht dabei für GPS Exchange Format.

- **Waypoint:** Einzelne (meist vorab) abgespeicherte Punkte (z.B. zur Navigation)
- **Route:** geplantes Routing, Sortierte Abfolge von Punkten, Folge von Wegpunkten
- **Track:** Liste aufeinanderfolgender Punkte, welche durch ein GPS-Gerät aufgezeichnet wurde.



Abbildung 5.5: Skizze der .gpx-Elemente: die graue Strecke ist die Route, die größeren blauen Punkte die Waypoints und punktiert der Track. [Eigene Darstellung]

Die Ebenen Waypoint und Route werden in der .gpx Datei von *Bike Citizens* nicht genutzt. Die Ebene Track, in der alle verfügbaren Daten gespeichert sind, unterteilt sich in zwei Bereiche:

- **tracks** enthält vollständige Fahrten in Form von Polylinien. Diese Linien-Features lassen sich nicht weiter unterteilen und enthalten ausschließlich Informationen zur kompletten Fahrt, nicht zu Teilbereichen. Folgende Attribute sind zusätzlich zum geographischen Verlauf enthalten:

⁷⁶vgl. Website der gpx-Entwickler: <http://www.topografix.com/gpx.asp> (Zul. aufgerufen am 11.1.2017)

- Zeitstempel Start
- Zeitstempel Ende
- Dauer der Fahrt
- zurückgelegte Distanz
- **track-points:** sind die Basis der tracks, also jene Punkte, welche die App periodisch jede Sekunde ermittelt und abspeichert. Die Aneinanderreihung der track-points ergibt die eigentliche Fahrt. Die track-points enthalten zusätzlich zur Position noch folgende weitere Attribute:
 - ID, über die Punkte zu einem Track zusammengefasst werden können
 - Zeitstempel
 - GPS-Genauigkeit (in Metern)
 - Fortlaufende Nummerierung

Kurzfasit

Für die Analyse stehen zwei verschiedene Datensätze zur Verfügung: Zum einen Daten vom Typ Line Feature, welcher die Geometrie aller knapp 150.000 Fahrten beinhalten und zum anderen Point-Features, welche die zu jeder Sekunde abgespeicherte Position jeder Fahrt darstellen. Aus Datenschutzgründen wurden seitens *Bike Citizens* alle tracks am Start sowie am Ziel um 100 Meter verkürzt. Daher ist keine absolut exakte Aussage zu Start- und Zielpunkt mehr möglich.

5.4 Datenaufbereitung und -verarbeitung

Das zur Verfügung gestellte .gpx-File stellte mit über 21 Gigabyte im Office-Bereich verbreitete Standard-PC-Infrastruktur vor Probleme - ein Öffnen mit Software wie ArcMap, QGis aber auch Notepad+ ist nicht ohne weiteres möglich. (anzumerken noch, dass ArcMap das .gpx Format generell nicht unterstützt)

Mithilfe der OSGeo4W Erweiterung, einem Batch-Programm welches über die Windows-Konsole bedient wird und eine Vielzahl von Geo-OpenSource Algorithmen beherrscht kann der Datensatz verarbeitet werden. Teil dieser Erweiterung ist auch GDAL (Geospatial Data Abstraction Library), welches unter seinen vielen Befehlen und Funktionen das Tool ogr2ogr beherbergt. Damit können auf einfache Weise (Geo-)Daten in verschiedene Dateisysteme konvertiert werden. Mithilfe des folgenden Befehls ist eine Konvertierung der .gpx Datei in ein .shp Shapefile möglich:

Textblock 5.1: ogr2ogr Syntax für die Konvertierung in ein .shp Format

```
for %N in (C:\work\diplomarbeit\bc_data\*.gpx) DO
ogr2ogr -t_srs epsg:4326
C:\work\diplomarbeit\bc_data\%~nN-wp.shp %N waypoints
-skipfailures
for %N in (C:\work\diplomarbeit\bc_data\*.gpx) DO
ogr2ogr -t_srs epsg:4326
C:\work\diplomarbeit\bc_data\%~nN-rt.shp %N routes
-skipfailures
for %N in (C:\work\diplomarbeit\bc_data\*.gpx) DO
ogr2ogr -t_srs epsg:4326
C:\work\diplomarbeit\bc_data\%~nN-tr.shp %N tracks
-skipfailures
pause
```

Der Befehl führt für alle drei Ebenen der .gpx Datei die Konvertierung durch, wobei letztendlich nur die Ebene *Tracks* Inhalte beherbergt. Während die Polylinien ohne Probleme in eine .shp Datei konvertiert werden können, ist diese Methode für die Punkte ob der Menge an Daten nicht zu empfehlen. Stattdessen bietet sich aufgrund der großen Datenmenge eine Datenbanklösung an. Im GIS-Bereich ist hier die PostgreSQL Erweiterung PostGis verbreitet, in diese auch mithilfe von ogr2ogr importiert werden kann:

Textblock 5.2: ogr2ogr Syntax für den Datenbankimport

```
'ogr2ogr -append -f PostgreSQL
"PG:dbname=gis user=puehflo password=xyz" bc_data_160921-at-vienna.gpx'
```

Das Ergebnis des Datenbankimports ist die Tabelle „track_points“ mit 133.587.593 Zeilen (=Punkte). Das sind je Track im Durchschnitt knapp über 900 abgespeicherte GPS-Positionen. Selbst im Datenbankformat belegen nur die track-points ca. 22 Gigabyte Speicherplatz. Über den QGIS DB-Manager ist ein Import in ein GIS möglich - wobei ob der Datengröße weiterhin erhebliche Wartezeiten auftreten. So dauert der Export eines Ausschnitts der Punkteebene im Ausmaß von ca 250x250 Metern in eine .shp Datei knapp eine Stunde.

Um im weiteren Verlauf der Datenanalyse effektiver arbeiten zu können, wurden die mitgelieferten Attribute erweitert. Da der Zeitstempel sowie die Beschreibung zur Distanz und Dauer im Textformat hinterlegt sind, war hier eine Aufsplittung unumgänglich. So wurde der track-Datensatz um die Attribute date (yyyymmdd), month, distance, duration, weekday, start_hr (Stunde des Starts) erweitert. Mithilfe von einfachen SQL⁷⁷-Queries können die Daten zum jeweiligen Analysezzweck passend gefiltert werden.

Ergebnis sind zwei verschiedene Datenschichten: Einmal die Tracks als Line-Features, welche nur Informationen zur gesamten Fahrt beinhalten sowie die einzelnen GPS-Positionen als Point-Features, die einen detaillierten Einblick für Teilbereiche ermöglichen.

⁷⁷Structured Query Language: Datenbanksprache zum Bearbeiten und Abfragen innerhalb einer Datenbankstruktur

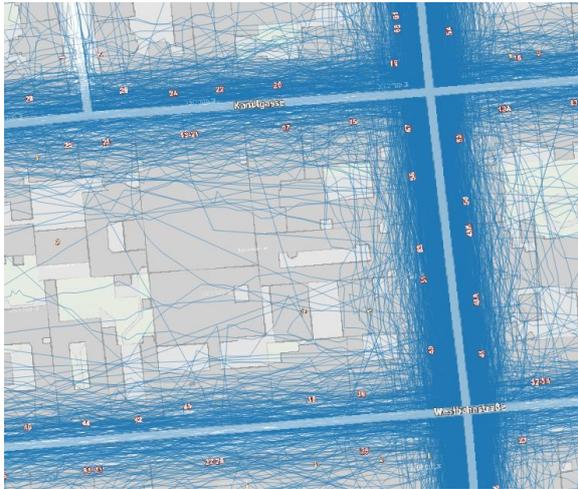


Abbildung 5.6:
Tracks als Line Features [Eigene Darstellung, Bike Citizens, Kartengrundlage: www.basemap.at]

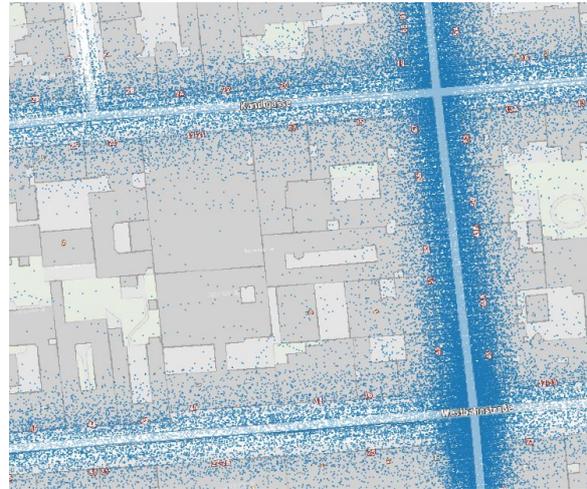


Abbildung 5.7:
GPS-Positionen als Point Features [Eigene Darstellung, Bike Citizens, Kartengrundlage: www.basemap.at]

5.4.1 Map Matching

Wie aus den Abbildungen 5.6 und 5.7 ersichtlich sind die *Bike Citizens* Daten nicht in kantenbasierter Form vorhanden. Aufgrund von Erfahrungen im Bereich graphenbasierter Netzwerkanalysen war schon nach dieser Feststellung klar, dass für tieferegehende Analysen eine Kantenzuweisung wünschenswert wäre. Dies ist notwendig, um beispielsweise Aussagen zu Frequenzen und Auslastungen einzelner Abschnitte zu bearbeiten. Auch ein Abgleich mit der befahrenen Infrastruktur (entspricht der Frage, wie groß der Anteil der Fahrten auf Radwegen ist) ist ohne Map-Matching kaum möglich. Auch für Aussagen zur Auslastung (hinsichtlich Kapazitätsgrenze) von Radinfrastrukturen sind kantenbasierte Daten nötig.

Der Begriff „Map Matching“ meint einen Prozess, der den Abgleich von gemessenen Positionen mit einem vorliegenden digitalen Vektordatenbestand vollzieht. Die Vorgehensweise ist vor allem aus Navigationsgeräten bekannt: Durch die fehlende Exaktheit und speziell Fehler bei der Ermittlung der GPS-Position kann es zu Abweichungen von der tatsächlichen Position kommen. Ein Navigationsgerät kennt die Fahrtrichtung und die Infrastruktur (= Straße = digitale Karte) und „snappt“ (verschiebt) die GPS-Position auf diese. Der Benutzer bekommt davon nichts mit, denn für ihn bleibt das „virtuelle Auto“ auf dem Navigationsgerät immer auf der momentan befahrenen Straße.

Strava, ein kalifornisches Unternehmen, betreibt eine gleichnamige APP und Plattform, mit der Nutzer ihre zurückgelegten Rad- und Laufmeter GPS-gestützt aufzeichnen können. Das Unternehmen verfolgt dabei schon länger Projekte, diese Daten für Forschungs- und Planungszwecke zu benutzen und zu vermarkten. *Strava Metro makes riding, running and walking in cities better. Millions of people upload their rides and runs to Strava every week via their smartphone or GPS device. Metro anonymizes and aggregates this data and then partners with departments of transportation and city planning groups to improve infrastructure for bicyclists and pedestrians.*⁷⁸

⁷⁸[Strava, 2016]



Abbildung 5.8: Ausschnitt des Strava-Datensatzes für Paris. Anhand der HeatMap wird die Zahl der einzigartigen Nutzer je Kanten-element visualisiert. [Eigene Darstellung / Strava Metro]

Voraussetzung dafür ist das Map Matching, welches Strava automatisiert durchführt. Strava Metro⁷⁹ bietet für GIS-Tools optimierte Beispiel-Tracks verschiedener Städte an - diese enthalten jeweils den Beobachtungszeitraum von einer Woche. Dabei stehen kantengenaue Detailinformationen für Auswertungen zur Verfügung. So kann innerhalb der Strava Daten nicht nur die Frequenz je Kante dargestellt werden, auch eine Differenzierung von wievielen Nutzern diese Frequenz stammt ist möglich. Zudem sind schon Werte für verschiedene Zeitscheiben, Richtungsabhängige Unterscheidungen für Kanten oder durchschnittliche Reisezeiten je Kante vorhanden. Auch für Quell-Ziel-Analysen gibt es

bereits vorgefertigte Polygondatensätze. Erwähnenswert sind auch die Detailinformationen auf Kreuzungsebene, welche ebenfalls für Auswertungen zur Verfügung stehen. Mithilfe der Strava-Daten können bereits Netzwerkanalysen aufgrund von realen (beobachteten) Fahrzeiten durchgeführt werden.

Die erweiterten Anwendungsmöglichkeiten durch kantenbasierte Daten sind enorm. Etliche Fragestellungen lassen sich sonst nicht beantworten. Die *Bike Citizens* Daten liegen leider nicht kantenbasiert vor, weshalb eine tiefere Auseinandersetzung mit Map Matching erforderlich ist. Zudem soll die Möglichkeit bewertet werden, die Daten als Anwender selbst auf ein Netz zu transferieren. Map Matching ist aber keines Falls trivial. Dies wird im Folgenden verdeutlicht.

Querschnittsmodell

Ein erster - relativ simpler - Zugang ist die Erstellung eines Querschnittsmodells: An jeder Kante des Zielnetzes werden Mittelpunkte generiert, an denen wiederum in einem nächsten Schritt ein Querschnitt gelegt wird. Die Dimension des Querschnitts sollte jeweils passend zum Straßenabschnitt gewählt sein, was bei einem heterogenen Straßennetz erste Komplikationen mit sich bringt. Als Zielnetz stehen dabei OpenStreetMap oder GIP⁸⁰ Österreich zur Verfügung.

⁷⁹[Strava, 2016]

⁸⁰GraphenIntegrationsPlattform

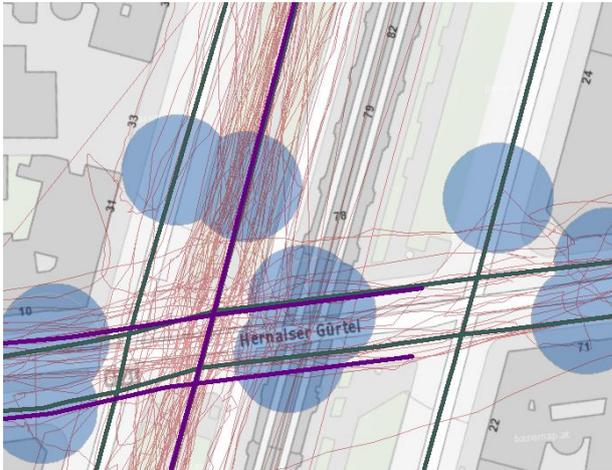


Abbildung 5.9: Beispiel Gürtel. [Eigene Darstellung]

Eine kurze Veranschaulichung anhand des Gürtels in Wien: Dieser besteht innerhalb eines Graphen aus je einer Kante je Richtung für den MIV und einer Kante für den Radweg (violett), der sich im Beispiel neben der südlich führenden Richtungsfahrbahn befindet. Es sollen alle Tracks (hellrot) mithilfe von Querschnitten (dunkelblaue Buffer) auf die richtige Netzkante übertragen werden. Schnell wird erkenntlich, dass diese Zuordnung mithilfe dieses Prinzips ohne ein dafür perfekt vorbereitetes Netz nicht funktioniert, da es sonst unausweichlich zu Überlagerungen kommt, die in weiterer Folge

zu quantitativen Verzerrungen führen. Ebenso die alleinige Nutzung des OGD-Datensatzes zum Radwegenetz hätte nicht den erwünschten Effekt, da in Bereichen ohne Radinfrastruktur, die aber trotzdem im Mischerverkehr befahren werden können, keine Kanten vorhanden sind. Überlegungen die Tracks zum jeweils nächstgelegenen Querschnitt zu transferieren stellten sich als äußerst komplex und ressourcenaufwendig dar.

An dieser Stelle sei auch noch angemerkt, dass ein Export eines passenden Netzes mit Radfahrinfrastruktur sowie mit dem Rad befahrbaren Straßenabschnitten aus OpenStreetmap aufgrund teilweise inkonsistent vergebener Attribute scheiterte.

Workaround Rasteranalyse

Da auf der Vektorebene keine akzeptable Lösung für das Problem gefunden werden konnte, soll ein Workaround aus der Rasterwelt Abhilfe bringen. Beim Map Matching geht es aus Anwendersicht um die Berechnung der Belastung von Kanten - oder anders formuliert - wieviele Fahrten finden auf einem x-beliebigen Teilabschnitt des Radnetzes statt. Beim Workaround ist es zunächst wichtig, die richtige Rastergröße zu wählen - in Abstimmung mit dem durchschnittlichen Wiener Straßenquerschnitt. Natürlich kann diese auch kleiner sein, was wiederum die Rechenleistung erhöht.

Ziel der Rasteranalyse ist es, alle Tracks, die eine Zelle berühren oder anschneiden, zu zählen und diesen Wert in die Rasterzelle zu schreiben. So kann eine Liniendichte pro Zelle berechnet werden und der Vergleich der Zellen zeigt, wo große Frequenzen auftreten und wo nicht. Zusätzlich stellt eine Überlagerung mit der Radinfrastruktur (bei richtig gewählten Buffern oder einem Fuzzy-Ansatz) kein Problem dar.



Abbildung 5.10:
Testausschnitt Rasteranalyse mit 5m Rasterzellengröße
[Eigene Darstellung]

Das Ergebnis der Rasteranalyse liefert auf den ersten Blick zufriedenstellende Ergebnisse, der Fehler liegt aber im Detail: Das ArcGis Tool Line Density, welches in einer definierten Nachbarschaft einer Zelle die Dichte-Berechnung durchführt, hat keine Möglichkeit für diese Berechnung die Anzahl der Linien zu berücksichtigen.

Stattdessen wird für die Berechnung die Längeneinheit pro Fläche verwendet.⁸¹ Diese Problematik liegt nicht nur den ArcGis Dichteberechnungen zugrunde, auch anderen GIS-Systemen fehlt ein derartiger Algorithmus. Dies lässt den Schluss zu, dass die Berechnung grundsätzlich nicht trivial ist. Das Ergebnis ist also keine Frequenz oder Belastung, sondern eine Angabe von gefahrenen Metern je

Hektar, was in diesem Fall völlig unbrauchbar ist. Zudem werden in dieser Analyse ungenaue Tracks durch ihre Charakteristik eines Zick-Zack-Verlaufes stärker gewichtet als korrekte Tracks, die gerade verlaufen.

Ist bei ausreichender Rechenleistung die Möglichkeit vorhanden mit der Punktebene Berechnungen durchzuführen, so wäre der Work-Around mit der Rasteranalyse brauchbar, da Algorithmen zum Zählen von Punkten innerhalb von Zellen vorhanden sind.

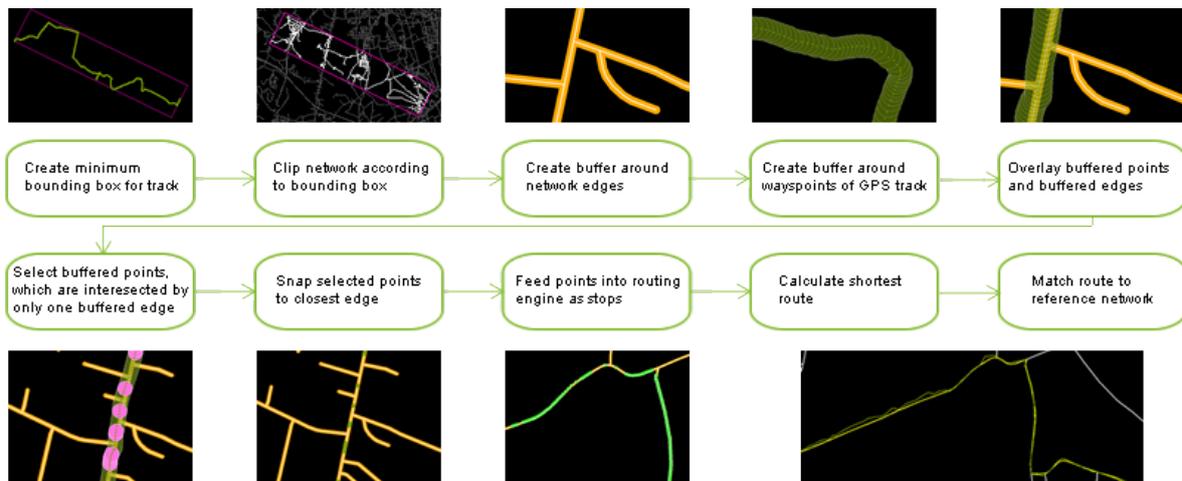
Bereits bekannte Map Matching Ansätze

Zur Map Matching Problematik gibt es bereits einen Grundstock an Literatur. Viele Ansätze sind aber auf spezielle Anwendungen zugeschnitten und können nicht ohne Adaption übernommen werden.

Der Ansatz, der auf dem Blog „Gicycle“⁸² (geographical information support for cyclists) vorgestellt wurde, wäre einfach umsetzbar, hat aber auch einige Kritikpunkte. Die Vorgehensweise soll hier kurz erläutert werden:

⁸¹vgl. <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-line-density-works.htm>

⁸²vgl. [Gicycle, 2016]



g

Abbildung 5.11: Der Aufbau des Map Matching Konzepts von Gicycle [Gicycle, 2016]

Zuerst wird der zu analysierende Track aus Performancegründen von den restlichen Tracks extrahiert. Auch das Zielnetz wird mithilfe eines Clips auf ein Minimum reduziert. Um die Straßenbreite besser abzubilden, werden die Kanten des Netzes mithilfe eines Buffers verbreitert. Auch um die track-points selbst wird ein Buffer gelegt, um eventuelle Ungenauigkeiten in der GPS-Bestimmung zu kompensieren. Die beiden Buffer werden in einem nächsten Schritt miteinander verschritten und nur die points ausgewählt, die eindeutig einer Kante zugeordnet werden können. Übrig bleiben jene Punkte, die beispielsweise an Kreuzungsbereichen liegen und somit Überschneidungen mit mehreren Kanten aufweisen. Die eindeutig zugewiesenen Punkte werden in einen Routing-Algorithmus geladen, welcher in gängigen GIS-Programmen zur Standardausstattung gehört. Der Routing Algorithmus berechnet die kürzeste Route und berücksichtigt dabei alle ausgewählten Punkte. Diese Berechnung wird anschließend in das Zielnetz übertragen und der nächste Track folgt. Fehler bringt dieser Ansatz dann, wenn einige points von massiven Fehlern belastet sind und so das Routing-Ergebnis Kanten inkludiert, die eigentlich gar nicht befahren wurden.

Neben diesen Ansätzen gibt es in der Literatur noch unzählige weitere. Erwähnt werden sollen hier [Straub u. Graser, 2015] sowie [Quddus u. a., 2007], die sich beide intensiv mit der Materie auseinandersetzen.

Der Ansatz von Gicycle scheiterte an der Masse der Punkte und der dazu benötigten großen Rechnerleistung. Verglichen mit einer Bearbeitungszeit der Analysen der Mikroebenen von teilweise 48h wäre der Ressourceneinsatz hier unverhältnismäßig gewesen. Einen Teilbereich zu modellieren ist zudem nicht im Sinne dieser Arbeit.

Tool aus der Ökologie

Mit „Movement Ecology Tools for ArcGis“ wurde ein Werkzeug getestet, das eigentlich aus der Ökologie stammt. Es wurde für die Analyse von Tracking-Daten von Elefanten entwickelt. Aus räumlicher Sicht stellt sich hier auch die Herausforderung der Umlegung von Trajektorien auf

Rasterzellen, also eine durchaus vergleichbare Ausgangslage. Das Toolset kann nach Installation in ArcGis als Erweiterung geladen werden und bietet eine breite Palette von Funktionen, die in einer Dokumentation⁸³ detailliert beschrieben werden. Viele davon wurden konzipiert, um Lebensbereiche und Heimatreservoirs von Tieren auf Basis von Tracking-Daten automatisiert auswerten zu können. Auch die Möglichkeit der Berechnung von Korridoren ist gegeben.

In der Anwendung der *Bike Citizens* Daten scheitern die meisten Tools an der großen Menge an Punkten. Auch die Ausrichtung auf ein spezielles Datenformat stellte sich als Herausforderung heraus. Praktikabel in der Anwendung zeigte sich aber das mitgelieferte Filter-Tool, welches eine einfache Selektion nach Datum oder Dauer der Tracks ermöglicht und im späteren Verlauf der Analyse noch öfter zum Einsatz kam.

Kurzfasit

Aus Sicht der Regionalwissenschaften soll hier der Wunsch nach kantenbasierten Daten (als Ergebnis des Map Matchings) festgehalten werden. Dies ist die Voraussetzung für zahlreiche, spannende Fragestellungen und analytische Herangehensweisen. Der Versuch das Map Matching als Anwender durchzuführen, ist ohne verfügbare und geeignete Tools nicht zu empfehlen, da der Ressourcenaufwand - auch mit dem Aspekt einer möglichen Nachbearbeitung zur Qualitätsverbesserung - in keiner Relation steht und deswegen auch im Rahmen dieser Arbeit scheiterte. Das Map Matching nur für einen Teilbereich durchzuführen, widerspricht der Sicht des relationalen Raums. Analysen sollen nur durch die Reichweite des Datensatzes, nicht durch willkürliche Rahmen aufgrund fehlender Ressourcen begrenzt sein.

Ausblick Bike Citizens

Laut *Bike Citizens*⁸⁴ wird in Kürze ein Web-Tool vorgestellt, welches über ein Map-Matching Verfahren die Tracks auf den Kanten der OpenStreetMap zeigt und dabei für jedes Element verschiedene Informationen (zB Fahrzeiten und Frequenzen) abgerufen werden können. Für zukünftige Forschungsprojekte wäre es wünschenswert, wenn die Tracking-Daten dann auch graphenbasiert verwendet werden könnten.

5.4.2 Inhaltliche Anreicherung

Wetter und Klimadaten mit dem Radverkehr in Bezug zu setzen, ist in der Forschung keinesfalls neu und stellen einen spannenden Forschungszweig dar. Daher war es naheliegend, die vorliegenden Daten thematisch mit Aussagen zur jeweils bestehenden Wetterlage anzureichern um zusätzliche analytische Auswertungen durchzuführen und im weiteren Verlauf auch tiefere Hypothesen zu validieren oder zu falsifizieren.

⁸³<https://goo.gl/rrUDRp>

⁸⁴[BikeCitizens, 2017]

Leider ist das *Open Data*-Konzept noch nicht bei der ZAMG (österreichische Zentralanstalt für Meteorologie und Geodynamik) angekommen - die Daten werden nur nach Bezahlung herausgegeben. Glücklicherweise zeichnet ein Kollege bereits seit Anfang 2016 Wiener Wetterdaten über eine freie API-Schnittstelle der Website <https://openweathermap.org> auf und hat mir diese Daten dankenswerterweise zur Verfügung gestellt.

Dabei sind folgende Informationen für jede Stunde für das Wiener Zentrum vorhanden:

- Wetterbeschreibung⁸⁵, unterteilt in
 - Clear
 - Clouds
 - Drizzle
 - Fog
 - Mist
 - Rain
 - Snow
 - Thunderstorm
- Temperatur in Grad Celsius, auf zwei Kommastellen genau
- Luftfeuchtigkeit in Prozent
- Windgeschwindigkeit in Metern pro Sekunde

Ergebnis der Zuweisung ist, dass jede aufgezeichnete Fahrt auch über detailreiche Wetterattribute verfügt. Allerdings war aufgrund der Datenverfügbarkeit eine Zuweisung für tracks im Jahr 2015 leider nicht möglich, weshalb sich alle weiteren Aussagen im Zusammenhang der *Bike Citizens* Daten und dem Wetter ausschließlich auf das Jahr 2016 beziehen.

Aufgrund von Übertragungskomplikationen sind teilweise keine Wetterdaten verfügbar. Sofern es sich bei dem Datenfehler um eine Unterbrechung der Übertragung von weniger als zwei aufeinanderfolgenden Stunden handelt, wurden die Daten von der Stunde davor bzw. Stunde danach fort- oder rückgeschrieben. Bei fehlenden oder fehlerhaften Daten über einen Zeitraum von mehr als drei Stunden wurden die betroffenen Tracks für die Analyse nicht weiter berücksichtigt und somit aus dem Wetterdatensatz entfernt. Dies betraf 1146 aufgezeichnete Fahrten von insgesamt 104028 aus dem Jahr 2016 (ca. 1%).

⁸⁵Details zu den Kategorien können auf der Website <https://openweathermap.org/weather-conditions> eingesehen werden

Weitere Datengrundlagen

Neben dem Wetterdatensatz wurden für die Auswertung weiters folgende Datengrundlagen verwendet:

Tabelle 5.1: Verwendete Datengrundlagen und deren Quelle

Datensatz	Quelle
Administrative Grenzen Wien	http://data.gv.at
Adressdaten Österreich	http://www.bev.gv.at/
Radwegenetz Wien	http://data.gv.at
Straßengraph Wien	http://data.gv.at
Basemap Österreich	http://basemap.at

5.4.3 Benutzte Software

Es wurde die grundsätzliche Entscheidung getroffen, für die Analyse der *Bike Citizens Tracks* Werkzeuge zu verwenden, die bereits aus dem Studium bekannt und daher auch im Bereich der Stadt- und Regionalforschung verbreitet sind. Zusätzlich wurde Software verwendet, die bereits aus praktischen Erfahrungen neben dem Studium bekannt war.

Der zur Verfügung stehende Werkzeugkasten umfasst die folgenden Produkte:

- **ESRI Arcmap Version 10.2.** Dient für alle GIS-Operationen und Berechnungen, egal ob auf Vektor- oder Rasterebene.
- **Arcmap Modelbuilder:** Bietet die Möglichkeit, Prozessabläufe auf Basis graphischer Benutzeroberfläche analog zu Flussdiagrammen zu kreieren. Dient der Entwicklung von wiederverwendbaren Werkzeugen, aber auch der Strukturierung und Archivierung der Arbeitsprozesse in ArcMap
- **QGIS Version 2.18. Las Palmas de G.C.:** Wurde in erster Linie zur Visualisierung von Berechnungen und zum weiteren Export als Kartendarstellung verwendet. Außerdem wurden über den integrierten DB-Manager die track-points geladen, gefiltert und exportiert.
- **SPSS Version 24:** Diente für alle statistischen Auswertungen und zur Erstellung von Diagrammen und Histogrammen
- **PostGIS Version 2.3.:** Erweiterung für PostGreSQL zur Einbindung von geographischen Objekten. War für die Organisation der track-points zuständig.
- **Microsoft Excel 2016:** Standardtool in der Raumplanung zur computergestützten Datenanalyse und -organisation

Da später auch eigene Werkzeuge, die mithilfe des ArcMap-internen Modelbuilder erstellt wurden, vorgestellt werden, soll dieser im Folgenden kurz genauer beleuchtet werden. Die Grundidee des Modelbuilder ist ein Flussdiagramm, welches Daten, Prozesse und Parameter

beinhaltet. Die Verknüpfung dieser Inhalte führt zu einem (mehr oder weniger) komplexen Analysemodell, welches dann auf Knopfdruck als Gesamtes ausgeführt werden kann.



Abbildung 5.12: Grundkonzept eines Ablaufprozesses im Modelbuilder. [Eigene Darstellung]

Im Gegensatz zum manuellen Ausführen der einzelnen Tools in ArcMap haben Modelle den Vorteil, dass sie jederzeit angepasst werden können (zB bei Änderungen von Parametern, während die Grundstruktur gleich bleibt) und die Ergebnisse auch jederzeit reproduzierbar sind. In der Perfektion führen Modelle aus dem Modelbuilder zu eigenen Werkzeugen, die über die ArcMap-interne Toolbox aufgerufen werden können und natürlich auch vielseitig einsetzbar sind.

Der Modelbuilder bietet zudem auch die Möglichkeit der Integration von Iteratoren. „Iteration, häufig auch Wiederholung genannt, bedeutet, dass ein Prozess immer wieder wiederholt wird, wobei ein gewisser Grad an Automatisierung zum Einsatz kommt.“⁸⁶ Vor allem im Bereich größerer Datenmengen und Stichproben ist der Einsatz von Iteratoren oft unerlässlich, wenn für jeden einzelnen Fall die gleichen Berechnungen durchgeführt werden müssen. Der Iterator erkennt dabei aufgrund einer eindeutigen Identifizierung (zB über eine Track-ID) die einzelnen Fälle und führt die Berechnung automatisiert für jeden dieser Fälle aus, wobei die Ergebnisse am Ende des Prozesses gesammelt in eine Ausgabedatei gespeichert werden können.

5.5 Daten Check-Up

Als Vorbereitung für die empirische Datenauswertung ist eine inhaltliche Überprüfung der aufbereitenden Daten nötig. Der Fokus liegt dabei auf Aspekten der Datengüte, der Repräsentativität sowie Eigenheiten, die bei den weiteren Hands-On Beispielen berücksichtigt werden sollten.

5.5.1 Deskriptive Merkmale

Vor der Auswertung und Darstellung der folgenden Inhalte wurden alle Tracks auf fehlende Daten - sogenannte Missing Values - überprüft. Diese zeigen sich durch eine Nullausprägung bei den Attributen Distanz oder Zeit. Insgesamt wurden 464 Datensätze - das sind immerhin nur 0,3% - für die weiteren Analysen als Missing Values deklariert und daher in keiner Form berücksichtigt.

⁸⁶vgl. [ESRI, 2017]

Aufgrund der Übersichtlichkeit wurde die Entscheidung getroffen, die deskriptive Analyse vorrangig am Attribut der Distanz zu zeigen:

Tabelle 5.2: Deskriptive Statistik des Attributs Distanz (in Metern)

N Gültig	146483
N Fehlend	464
Mittelwert Distanz	4792
Standardfehler des Mittelwerts	13,95
Median	3124
Standardabweichung	5340,52
Varianz	28521216,57
Minimum	1
Maximum	148490

Auffallend ist die extrem große Streuung der Zahlen - sowohl erkenntlich an der Standardabweichung als auch an der Varianz. Das Minimum von 1m lässt zudem darauf schließen, dass die Deklaration des Wertes 0 als Missing Value wahrscheinlich noch nicht ausreicht und auch bei weiteren sehr niedrigen Werten Datenfehler vorliegen. Anders als im Bereich des Minimums handelt es sich beim Maximum - immerhin 148 Kilometern - um keine Datenfehler sondern tatsächlich gefahrene Tracks. Über eine Stichprobenanalyse wurden diese Tracks gesichtet und überprüft. Meist sind es Relationen zwischen dem Bereich um Korneuburg und der Gegend um den südlichen Neusiedlersee. Natürlich ist das Argument nicht ganz unbegründet, dass diese Relation im Begriff des *Radverkehrs in Wien* nur bedingt inkludiert ist. Andererseits hat die Stadt Wien im Bereich der offiziellen Radrouten - siehe dazu Kapitel 5.4.4 - durchaus auch Investitionen in Langstreckenrelationen gesetzt. Deswegen können diese Tracks zwar durchaus als Ausreißer im Fokus der deskriptiven statistischen Analyse gesehen werden - haben aber in der räumlichen Relation durchaus ihre Daseinsberechtigung.

5.5.2 Repräsentativität

Grundsätzlich stellt sich die Frage, wie gut bilden die Daten die Wirklichkeit - also den gesamten Fahrradverkehr in Wien - eigentlich ab? Die Antwort dieser Frage lässt sich mit statistischen Methoden annähern. Das Verkehrsplanungsbüro Nast erhebt im Auftrag der Stadt Wien schon seit 2011 mithilfe von elf automatisierten Zählstellen⁸⁷ verteilt an strategischen Punkten des Wiener Radwegenetzes Zähldaten. Diese automatisierten Anlagen erfassen - so gut als möglich - jeden vorbeifahrenden Radfahrer, vermerkt mit der aktuellen Uhrzeit und dem aktuellen Wetter werden die Werte abgespeichert.

⁸⁷Eigentlich sind 12 Zählstellen aktiv, jedoch wurden für die Auswertung die zwei nach Fahrtrichtung getrennten Stationen beim Opernring zu einer zusammengefasst

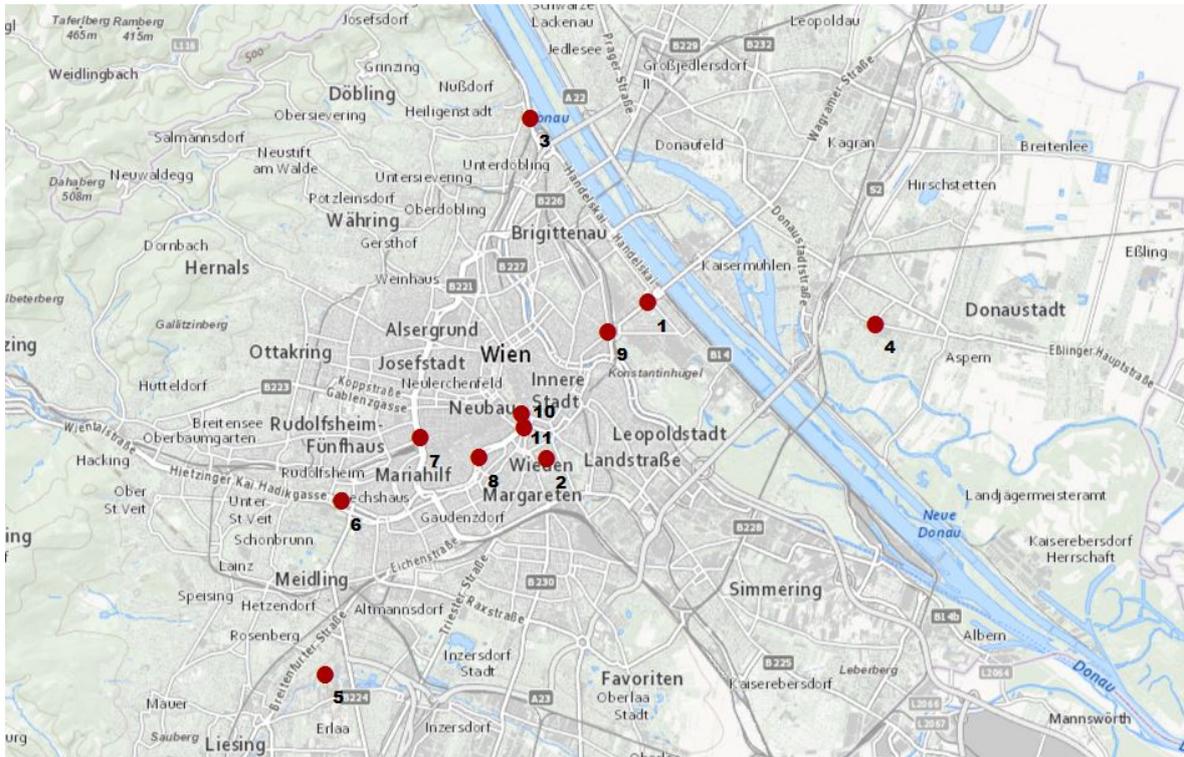


Abbildung 5.13: Verortung der Radverkehrsählstellen (Rote Punktsignatur) im Großraum Wien. Über die angegebene ID können die einzelnen Stellen mithilfe der nachfolgenden Tabelle identifiziert werden. [Eigene Darstellung; Kartengrundlage: www.basemap.at]

Mithilfe einer Querschnittsanalyse in ArcMap wurden nun in einem nächsten Schritt die tangierenden *Bike Citizens* Tracks gezählt. Die Länge des Querschnitts betrug dabei jeweils ungefähr die 5-fache Breite des Straßenprofils. Dies war notwendig, um auch etwaige GPS-Ungenauigkeiten berücksichtigen zu können.

Die gezählten Tracks je Querschnitt wurden monatlich ausgewertet und mit den Zählstellen-daten⁸⁸ gegenübergestellt. Da die Daten der Zählstellen nur als DTV24⁸⁹ vorliegen, mussten auch die *Bike Citizens* Daten in dieses Format konvertiert werden. Die folgende Tabelle zeigt den Zählstellenvergleich exemplarisch für Werktage - die Tabellen für Samstag und Sonntag befinden sich im Anhang.

⁸⁸[nast, 2016]

⁸⁹Anzahl der Radfahrer je 24 Stunden

Tabelle 5.3: Anteile der Bike Citizens Tracks an den Zählstellenwerten an Werktagen. Hervorgehoben jeweils der maximale Wert. [Eigene Darstellung]

ID / Ort	Jan16	Feb16	Mar16	Apr16	May16	Jun16	Jul16	Aug16	Mittelwert
1 Lassallestraße	0,10%	0,23%	0,47%	0,58%	0,56%	0,61%	0,54%	0,50%	0,45%
2 Argentinierstraße	0,24%	0,26%	0,44%	0,45%	0,66%	0,78%	0,75%	0,57%	0,52%
3 Donaukanal	0,05%	0,10%	0,08%	0,08%	0,10%	0,17%	0,18%	0,25%	0,13%
4 Langobardenstraße	0,00%	0,40%	0,62%	0,42%	0,71%	0,55%	0,69%	0,52%	0,49%
5 Liesingbach	0,09%	0,03%	0,15%	0,35%	0,41%	0,62%	0,49%	0,45%	0,32%
6 Wienzeile	0,16%	0,31%	0,43%	0,37%	0,51%	0,57%	0,44%	0,50%	0,41%
7 Neubaugürtel	0,17%	0,25%	0,26%	0,37%	0,37%	0,50%	0,33%	0,44%	0,34%
8 Margaritensteg	0,29%	0,41%	0,41%	0,40%	0,30%	0,40%	0,45%	0,40%	0,38%
9 Praterstern	0,08%	0,15%	0,39%	0,49%	0,45%	0,51%	0,53%	0,48%	0,39%
10 Operngasse	0,36%	0,39%	0,44%	0,46%	0,44%	0,52%	0,56%	0,50%	0,46%
11 Opernring	0,49%	0,47%	0,54%	0,48%	0,46%	0,49%	0,44%	0,39%	0,47%
Mittelwert	0,18%	0,27%	0,38%	0,40%	0,45%	0,52%	0,49%	0,46%	0,40%

Um die relationalen Anteile und die Größenordnung der Zahlen besser einordnen zu können, hilft ein beispielhafter Blick auf die absoluten Werte. Aufgrund der Konvertierung der Daten in das Format DTV24 enthalten die Ergebnisse aus der eigenen Querschnittsanalyse Kommazahlen, um hier keine Verzerrungen zu zeigen. Dies ist aufgrund des Formats notwendig, da die monatlichen Ergebnisse auf einen 24-stündigen Mittelwert heruntergebrochen wurden. Was der relationale Vergleich schon deutlich gemacht hat, zeigt sich dabei auch beim Vergleich der Absolutwerte. An vielen Zählstellen ist die Anzahl der *Bike Citizens* Fahrten sehr klein. Werte unter 1 entstehen, da an einigen Tagen kein einziger *Bike Citizens* App-Nutzer die jeweilige Zählstelle passierte.

Tabelle 5.4: Vergleich der absoluten Werte an den Zählstellen, differenziert jeweils nach ZS (Zählstellenwert) und BC (Bike Citizens Tracks). [Eigene Darstellung]

	Jan16		Feb16		Mar16		Apr16		May16		Jun16		Jul16		Aug16	
	ZS	BC	ZS	BC	ZS	BC	ZS	BC	ZS	BC	ZS	BC	ZS	BC	ZS	BC
Lassallestraße	643	0,6	831	2,0	1181	5,6	2177	12,7	2337	13,2	3687	22,6	4189	22,8	3841	19,3
Argentinierstraße	1018	2,4	1193	3,1	1482	6,6	2457	11,0	2499	16,6	3142	24,5	2803	21,0	2668	15,2
Donaukanal	315	0,2	411	0,4	648	0,5	1308	1,0	1434	1,5	2022	3,5	2157	4,0	2122	5,3
Langobardenstraße	170	0,0	213	0,9	289	1,8	511	2,1	529	3,8	753	4,1	748	5,2	721	3,8
Liesingbach	116	0,1	160	0,0	257	0,4	505	1,8	548	2,2	721	4,5	729	3,6	685	3,1
Wienzeile	331	0,5	435	1,3	622	2,7	1190	4,4	1392	7,2	1773	10,1	2079	9,2	1992	9,9
Neubaugürtel	733	1,3	880	2,2	1122	2,9	1917	7,0	1932	7,2	2523	12,5	2322	7,6	2118	9,3
Margaritensteg	405	1,2	465	1,9	683	2,8	1230	4,9	1309	3,9	1702	6,9	1644	7,4	1541	6,2
Praterstern	1093	0,8	1350	2,0	1774	6,9	3041	15,0	3175	14,2	4538	23,4	4579	24,3	4303	20,8
Operngasse	1879	6,8	1930	7,6	2543	11,3	4655	21,4	5881	25,9	6920	36,3	5474	30,4	4899	24,7
Opernring	1994	9,8	2338	11,0	3163	17,0	5585	27,0	5933	27,5	7880	38,7	7321	31,9	6897	26,6

Selbst in Monaten mit vergleichsweise vielen aufgezeichneten Fahrten bilden die Daten von *Bike Citizens* nur einen sehr kleinen Teil des Wiener Radverkehrs ab. Die Dimension ist dabei so klein, dass teilweise an mehreren Tagen kein einziger App Nutzer eine Zählstelle passiert.

Die statistische Methode des Konfidenzintervalls (Schwankungsbreite) erlaubt die Abschätzung der Güte von Schätzungen für eine Grundgesamtheit in Abhängigkeit von der Stichprobengröße. Unter bestimmten Bedingungen können über diese Methode Aussagen zur erwarteten Schwankungsbreite getroffen werden. Dies erfolgt im Sinne von „mit einer bestimmten Wahrscheinlichkeit liegt der wahre Wert innerhalb eines bestimmten Intervalls.“ Dieses Inter-

vall ist das Konfidenzintervall.⁹⁰ Angenommen wird dabei, dass der berechnete Wert theoretisch zu 95% innerhalb dieses Intervalls liegt.

Voraussetzung ist aber, dass die Stichprobe das Phänomen der Normalverteilung aufweisen muss und eine Zufallsstichprobe ist. Da beide Bedingungen nicht gegeben sind, kann die folgende Analyse nur einen Hinweis auf potenzielle Schwankungsbreiten aufzeigen.

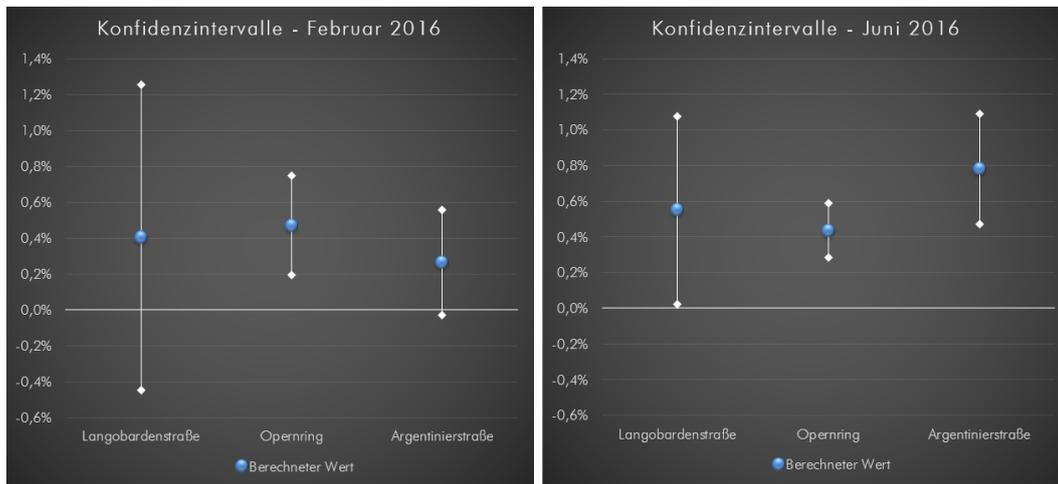


Abbildung 5.14: Darstellung der Konfidenzintervalle an ausgewählten Zählstellen in den Monaten Februar (vergleichsweise wenige Tracks) und Juni (vergleichsweise viele Tracks). Erkennbar sind besonders anhand der Zählstelle Langobardenstraße die besonders großen Intervalle. [Eigene Darstellung]

Der Darstellung 5.14 sind die Schwankungsbreiten der ausgewählten Zählstellen zu entnehmen. Besonders deutlich wird die Problematik am Beispiel der Langobardenstraße, wo das Intervall so groß ist, dass es bis in den negativen Bereich reicht. Aber auch die anderen Zählstellen zeigen vergleichsweise große Intervalle. Besonders durch den sehr kleinen Anteil der Tracks an den Zählstellenwerten erschwert diese Problematik die Repräsentativität nochmals deutlich. Die errechneten Werte sind somit vermutlich mit großer Unsicherheit behaftet. (Die komplette Tabelle zu den Konfidenzintervallen befindet sich im Anhang.)

Die Repräsentativität ist aber nicht nur von der Menge der Tracks abhängig. Es bleibt die Frage, ob die jahreszeitlichen Veränderungen durch die aufgezeichneten *Bike Citizens* Daten abgebildet werden und auch der räumlichen Verteilung des Wiener Radverkehrs entsprechen. Zur Überprüfung dieses Sachverhalts wurde die Methodik der Regressionsanalyse gewählt, wobei der Fokus klar auf den Korrelationsergebnissen liegt. Die Idee dahinter ist die Annahme, dass die Zahlenreihen der Zählstellenergebnisse mit den Zahlenreihen der *Bike Citizens* Tracks jeweils pro Zählstelle und Monat ein gleiches Verhalten zeigen - sprich korrelieren. Indikator ist hierfür der Korrelationskoeffizient, welcher möglichst hoch (=nah bei 1) liegen sollte. Ein möglichst hohes R-Quadrat (=ebenfalls nah bei 1) zeigt an, dass die Streuung gering ist und das Modell gut an die Daten angepasst ist. Da es sich hier nur um ein sehr simples Minimalmodell handelt, dient ein hohes R-Quadrat zur Absicherung des Korrelationskoeffizienten.

⁹⁰vgl. <http://www.univie.ac.at/soziologie-statistik/pflege/VL/VO%204%20Konfidenzintervall.pdf>

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.	95,0% Konfidenzintervalle für B	
		Regressionskoeffizient B	Standardfehler	Beta			Untergrenze	Obergrenze
1	(Konstante)	350,688	74,608		4,700	,000	202,373	499,003
	bc_data	183,367	5,510	,963	33,278	,000	172,413	194,321

a. Abhängige Variable: zaehl_data

Abbildung 5.15: Koeffizienten der Regressionsanalyse [Eigene Darstellung / SPSS]

Modellzusammenfassung^b

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Änderung in R-Quadrat	Statistikwerte ändern			Sig. Änderung in F
						Änderung in F	df1	df2	
1	,963 ^a	,928	,927	493,76113	,928	1107,412	1	86	,000

a. Einflußvariablen : (Konstante), bc_data
 b. Abhängige Variable: zaehl_data

Abbildung 5.16: Modellzusammenfassung der Regressionsanalyse [Eigene Darstellung / SPSS]

Zusammenfassung Repräsentativität

Verglichen mit den absoluten Zahlen, bilden die *Bike Citizens* Tracks einen verschwindend kleinen Teil des Radverkehrs ab. Dies liegt auch an der Zahl der aktiven Nutzer der App, welche im August 2015 bei 2271 lag. Dieser kleine Teil - so zeigt es die Regressionsanalyse - korreliert jedoch sehr stark mit den tatsächlichen Zahlen - und zwar über die Zeit und auch über die verschiedenen Zählstellen, also räumliche Verteilung. Sowohl der äußerst hohe Korrelationskoeffizient von 0,963, das korrigierte R-Quadrat von 0,927, als auch die makellose Signifikanz des Modells lassen daran wenig Zweifel. Die Analyse zeigt zudem, dass durch die größere Stichprobe in den Sommermonaten diese tendenziell besser für weitere Bearbeitungen geeignet sind.

Es kann also abschließend gesagt werden: Es handelt sich um eine sehr kleine Stichprobe, die aber sowohl über die zeitliche Achse, als auch über die Verteilung im Raum als durchaus repräsentativ eingestuft werden kann.

Der kritische Punkt ist aber, dass es sich um keine Zufallsstichprobe handelt. Das Ausmaß der Nichtzufälligkeit lässt sich aufgrund der fehlenden Info über die Zusammensetzung sowohl der *Bike Citizens* Nutzer, als auch jener die Zählstelle passierenden Radfahrern nicht ermitteln. Im Endeffekt stellt sich also auch die Frage, wie repräsentativ statistische Aussagen zum Radverkehr auf Basis der Zählstellendaten sind. Anzumerken ist aber noch, dass auch bei der Regressionsanalyse die kritisch geringen Stichproben und die Erkenntnisse aus den Konfidenzintervallen zu eventuellen Problemen führen können und daher die Ergebnisse nicht überbewertet werden sollen.

Kurzfasit

Die Repräsentativität kann aufgrund der Nichtzufälligkeit sowie der geringen Ausprägung der Stichprobe für den aktuell vorliegenden Datensatz nicht bestätigt werden. Mit Blick auf die zu erwartende Ausweitung des Benutzerkreises⁹¹ ist davon auszugehen, dass zumindest punktuell Stichprobengrößen erreicht werden können, die ein zufriedenstellendes Sicherheitsniveau gewährleisten. Diese Aussage bezieht sich aber nur auf die Menge und den Umfang der Stichprobe, nicht auf die Zusammensetzung dieser.

5.5.3 Datengüte

Zur Überprüfung der Datenqualität sind drei wichtige Gütekriterien vorhanden.⁹²

- Die **Objektivität** ist die Unabhängigkeit der Test-Ergebnisse. Das Ergebnis soll nicht durch unterschiedliche Personen in der Erfassung beeinflusst werden.
- Die **Reliabilität** meint die Zuverlässigkeit eines Tests. Diese ist dann gegeben, wenn bei wiederholten Messungen gleiche Ergebnisse zu erwarten sind.
- Die **Validität** gibt an, ob der Test auch das misst was er messen soll.

Die *Bike Citizens* Daten werden in einem theoretischen Rahmen aus den Standpunkten der verschiedenen Gütekriterien hinsichtlich des Ziels einen bestimmten Anteil des Radverkehrs in Wien abzubilden überprüft.

Die Objektivität der Testergebnisse kann in Hinblick auf die Erfassung bestätigt werden. Die Erhebung kann nicht durch einzelne Personen beeinflusst werden, da die Daten automatisiert generiert werden. Allerdings gibt es verschiedene andere Einflüsse auf den Anteil der App-Nutzer, die weder genau prognostiziert, noch bestimmt werden können. Die Objektivität des Datensatzes kann also nicht in vollem Umfang bestätigt werden. Da die Objektivität meist auch als Basis für die weiteren Gütekriterien gesehen wird, ist auch dort mit keinen guten Ergebnissen zu rechnen.

Die Reliabilität - also die Zuverlässigkeit - kann keinesfalls bestätigt werden, da die Nutzung der App sehr unterschiedlich ausfällt. Für den Radverkehrsanteil wäre es wünschenswert, wenn eine bestimmte Personengruppe die App täglich und regelmäßig nutzt. Dies ist jedoch nicht der Fall. Die Reliabilität kann also nicht bestätigt werden.

Die Validität ist in hohem Maße von der Reliabilität abhängig. Da über die Methode kein Einfluss über die Regelmäßigkeit der Nutzung bekannt ist, kann auch die Validität hinsichtlich der Messung eines Radverkehrsanteils in Wien nicht bestätigt werden. Die Daten entsprechen also nicht dem Gütekriterium der Validität.

⁹²vgl. <http://wirtschaftslexikon.gabler.de/Definition/testguetekriterien.html> - aufgerufen am 10.02.2017

Kurzfasit

Die Gütekriterien haben zusammen mit der Repräsentativität gezeigt, dass die folgenden Aussagen der Ergebnisse nur für den Anteil der *Bike Citizens* Tracks des aktuell vorliegenden Datensatzes Gültigkeit besitzen und keine Rückschlüsse auf den Gesamtverkehr möglich sind.

5.5.4 Unterscheidung nach Fahrtzweck

Der Zweck des Radfahrens ist für die Mobilitätsforschung aber auch für die Planung von entscheidender Bedeutung. Grundsätzlich kann hier zwischen dem Alltagsradverkehr und dem Freizeitradverkehr unterschieden werden. Im Sinne der Förderung der aktiven Mobilität soll natürlich vorrangig der Alltagsradverkehr angesprochen werden. Auch hinsichtlich repräsentativer Aspekte der Ergebnisse ist eine Kenntnis der Nutzergruppe nötig.

Diese definiert sich dahingehend, dass das Fahrrad in der Alltagsmobilität eines Individuums gegenwärtig ist und damit regelmäßig(e) Wege zurückgelegt werden. Der Freizeitradverkehr hingegen meint Wege, die vorrangig aus freizeithlichen Gründen zurückgelegt werden und das auch in einer gewissen Unregelmäßigkeit. Deswegen ist ein näherer Check der Daten bezüglich ihrer Ausprägungen von typischen Merkmalen von Freizeitfahrten vs. Alltagsfahrten notwendig. Die Wochentagsverteilung zeigt relativ gleich verteilte Mengen der Tracks. Lediglich an Samstagen werden tendenziell weniger Tracks aufgezeichnet. Diese Verteilung lässt vorerst keine exakte Einordnung des Datensatzes zu.

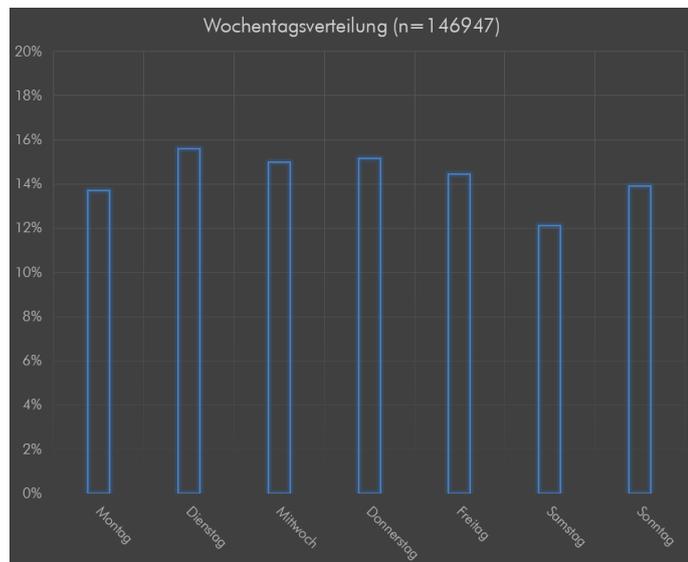


Abbildung 5.17:
Wochentagsverteilung der Bike Citizens Tracks
[Eigene Darstellung]

Die Stundenverteilung zeigt die höchsten Werte am frühen Abend zwischen 16 und 18 Uhr. Die Anteile in den Morgenstunden zwischen 6 und 9 Uhr sind zwar relativ konstant - dies aber nur bei einem Wert von ca. 5%. Die Stundenverteilung lässt zusammen mit der Wochentagsverteilung die Vermutung zu, dass der Datensatz gewissermaßen beide Nutzergruppen beinhaltet. Die mittlere Distanz je Fahrt ist ebenfalls ein Indiz für Nutzergruppen - folgt man der Annahme, dass Alltagsfahrten jenem Durchschnitt des Wiener Radverkehrs aus Erhebungen wie *Österreich unterwegs* (ca. 3km) entsprechen und bei Freizeitfahrten, die zum Zwecke des

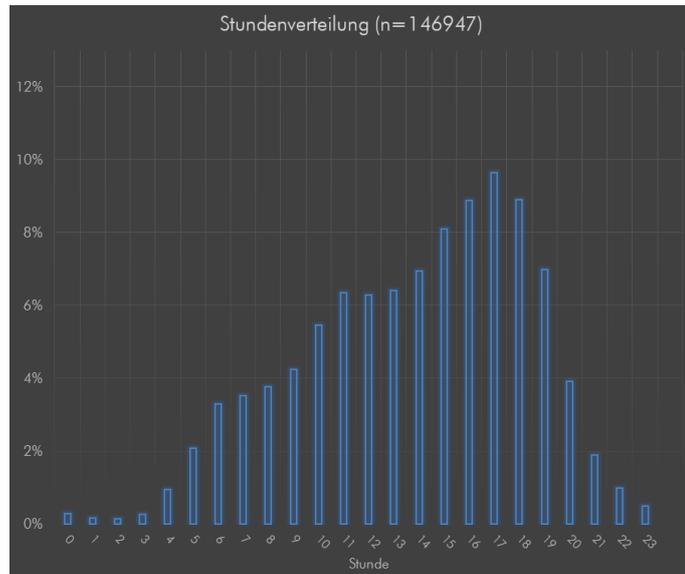


Abbildung 5.18: Stundenverteilung der Bike Citizens Tracks [Eigene Darstellung]

Radfahrens durchgeführt werden, deutlich längere Distanzen zurückgelegt werden. Die Darstellung der mittleren Distanz macht deutlich, dass vor allem bei den aktuelleren Daten ungefähr der erwähnte Durchschnittswert erreicht wird. Dies lässt darauf schließen, dass aktuelle Daten eher Alltagsfahrten beinhalten als ältere aus dem Jahr 2015.

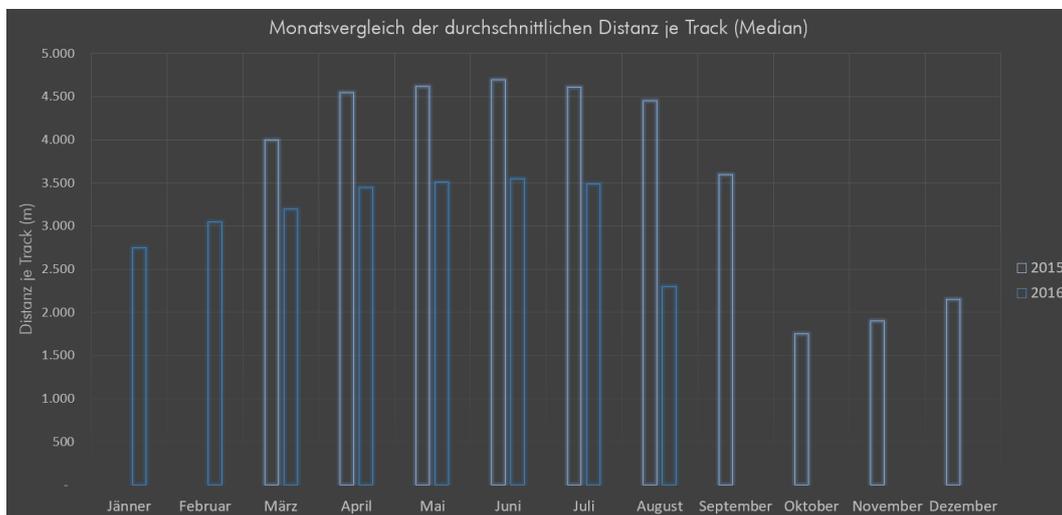


Abbildung 5.19: Distanzverteilung der Tracks über die Monate [Eigene Darstellung]

Die genauen Ursachen für die Abnahme bei aktuelleren Daten können nicht festgestellt werden, es gibt jedoch eine naheliegende Vermutung: Durch die zunehmende Nutzung der App hat sich die Menge der aufgezeichneten Tracks pro Monat stark gesteigert (siehe dazu Abbildung 5.2.). Bei den neuen Nutzern dürfte es sich demzufolge eher um solche handeln, die die App für kürzere Wege in der Stadt verwenden.

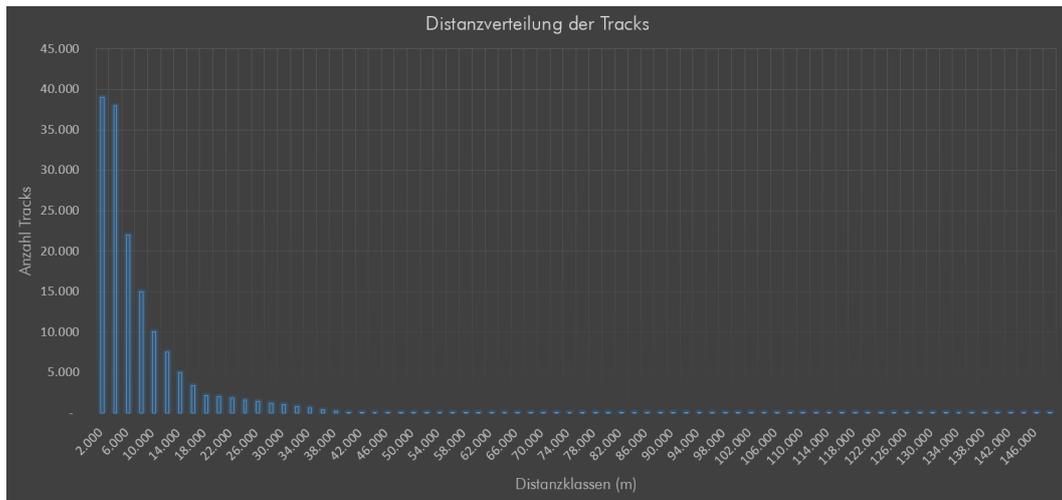


Abbildung 5.20: Häufigkeitsverteilung der Distanz [Eigene Darstellung]

Blickt man nun auf die Häufigkeitsverteilung der Distanz im Detail, so zeigt sich auch im Histogramm die in der deskriptiven Darstellung angesprochene große Streuung der Werte. Diese Darstellung begründet auch, weshalb der Median bei den meisten Aussagen zu Mittelwert und Durchschnitt geeigneter ist. In Hinblick auf die Nutzergruppen gibt die Grafik darüber Aufschluss, dass die Vermutung der Datensatz beinhaltet sowohl Alltags- als auch Freizeitradverkehr gestärkt wird. Eine erste Möglichkeit den Freizeitradverkehr zu extrahieren wäre dann vorhanden, wenn im Datensatz vermerkt wird, ob ein Nutzer gerade eine der angesprochenen touristischen Touren, die in der App hinterlegt sind, nutzt und abfährt.

Kurzfaszilit

Eine klare Trennung zwischen Alltagsradverkehr und Freizeitradverkehr lässt sich aus der Datenanalyse nicht feststellen. Viele Argumente sprechen dafür, dass beide Elemente etwa in gleicher Ausprägung vorhanden sind. Für eine klare Differenzierung ist eine Zuordnung der Datensätze zu Individuen nötig, um auch persönliche Aktivitätsmuster in eine Analyse einfließen zu lassen.

5.6 Makro-Ebene

Im folgenden Abschnitt werden die *Bike Citizens* Daten anhand der angekündigten Ebenenlogik analysiert. Die Makro-Ebene stellt dabei die größte Maßstabebene der Hands-On Datenanalyse dar. Die Inhalte orientieren sich an Aussagen für ganz Wien. Für Auswertungen dieser Ebene eignen sich unterschiedliche Methoden der räumlichen Analytik. Die gezeigten Ansätze reichen hier von der einfachen kartographischen Visualisierung, bei der die eigentliche analytische Arbeit auf der Betrachterseite geleistet wird, bis hin zu Aussagen zum Langstreckenradverkehr und der Infrastrukturplanung für diesen.

5.6.1 Visualisierung

Visualisierungen dienen in der Forschung und Planung nicht nur zur Darstellung von Gegebenheiten, sondern sind ebenso ein wichtiges Kommunikationsmittel. Die Vielfalt der räumlichen Darstellungen ist dabei nicht enden wollend und reicht von der einfachen politischen Land(Stadt-)Karte bishin zu relationalen Gegebenheiten im Raum - beispielsweise die Abbildung von lokalen Netzwerken über Akteurslandkarten. Mit nutzergenerierten Track-Daten wird die Kommunikation in der Planung sowie die Visualisierung - vor allem von Relationen im Raum - um eine weitere Ebene bereichert. Abgebildet werden Verflechtungen und Intensität der Inanspruchnahme von netzwerkartiger Infrastruktur.



Abbildung 5.21: Visualisierung der Bike Citizens Tracks im Jahr 2016. Je heller, desto höher ist die Frequenz. Die dargestellte Verwaltungsgrenze dient rein der Orientierung des Betrachters. [Eigene Darstellung / Bike Citizens Tracks]

Die spannende Komponente dieser weiteren Ebene ist, dass die Karte oder Abbildung des Raums nicht durch den Bearbeiter erstellt und gedacht wird, sondern die Daten rein durch das Verhalten der Nutzer im Raum entsteht. Und dieses Verhalten zeichnet wiederum das unverkennbare Muster der Stadt.

Bei der Wahl der Visualisierung sind dem Kartenersteller kaum mehr Grenzen gesetzt. Die dargestellten Visualisierungen wurden mithilfe von graphischen Methoden zur Dichtedarstellung erstellt. Dies spart die vorhergehenden Berechnungen mit Prozessen zur räumlichen Dichteberechnung wie *Kernel Density* oder *Point Density*. Sofern das Endprodukt lediglich der Visualisierung von Phänomenen dient stellt dies eine praktische Alternative dar.

Derartige Darstellungen können auch als *Heatmap* bezeichnet werden. Diese gibt es in der Kartographie in den verschiedensten Ausprägungen, sie folgen dabei alle dem Prinzip, dass über eine Farbskala quantitative Ausprägungen kartographisch dargestellt werden. Als Farbskala kann dabei beispielsweise jene der Temperaturskala verwendet werden - der Kreativität sind aber auch hier keine Grenzen gesetzt. Die beiden Abbildungen (5.6 sowie 5.7) bilden über die Farbskala die Nutzungsintensität ab. In einem ersten Schritt werden alle Punkte oder Linien in dieser Farbe dargestellt. Durch die große Menge an Daten kommt es unausweichlich zu Überschneidungen. Mithilfe des Grundprinzips der additiven Farbmischung werden Räume mit sehr vielen Datenüberlagerungen farblich intensiver dargestellt als jene mit weniger oder keinen.

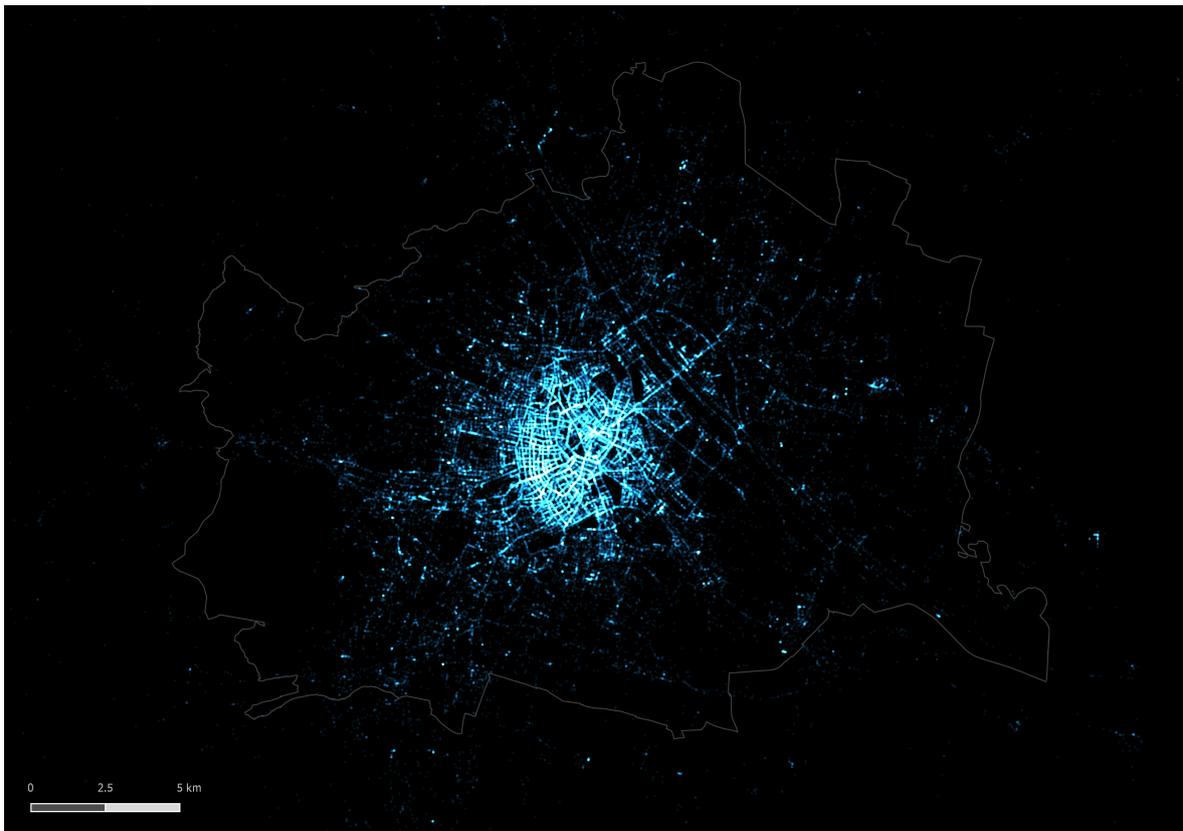


Abbildung 5.22: Visualisierung der Ziele der Bike Citizens Tracks - dargestellt als Punktwolke.
[Eigene Darstellung / Bike Citizens Tracks]

Der Darstellung 5.7. liegt als Geometrie eine einfache Punktwolke zugrunde - und zwar alle Zielpunkte sämtlicher zur Verfügung stehenden *Bike Citizens* Tracks. Auch wenn diese aufgrund des Datenschutzes um 100 Meter vor dem eigentlichen Ziel liegen, bildet die Darstellung trotzdem die Morphologie der Stadt Wien ab. Linienhafte Strukturen, wie markante Straßenzüge oder Flüsse, werden nur über die Darstellung von Zielpunkten sichtbar.

Kurzfasit

Visualisierungen von räumlichen Phänomenen sind in vielen Phasen von Forschungs- und Planungsprojekten ein wichtiger Bestandteil. Mithilfe der Fahrrad-Tracking-Daten können hier maßgeschneiderte, spannende und innovative Darstellungen erzeugt werden. Das Konzept, dass die Daten dabei direkt Nutzerverhalten abbilden, bietet vor allem in der Kommunikation einzigartige Möglichkeiten. Aber auch über die visuelle Analyse auf der Anwenderseite können die Daten schnell räumlich-analytisch begutachtet werden. Auch die Möglichkeit kurzzeitig auftretende, räumliche Phänomene zu begutachten und darzustellen, bietet spannende Möglichkeiten.

5.6.2 Radverkehr und Wetter

Es darf davon ausgegangen werden, dass das Wetter Einfluss auf die Verkehrsmittelwahl hat, vor allem wenn das Fahrrad zur Auswahl steht. Der Einfluss wird in der Literatur sehr unterschiedlich bewertet. Folgende Literatur zur Thematik kann empfohlen werden: [Miranda-Moreno u. Nosal, 2011] oder [Brandenburg u. a., 2007] oder [Thomas u. a., 2012])

Mit der großen Stichprobe der *Bike Citizens* Tracks und den detaillierten Wetterinformationen ist die Frage nach der Wettersensibilität natürlich naheliegend. Wie schon erwähnt, können für alle Aussagen in Zusammenhang mit Wetterdaten nur Tracks aus dem Jahr 2016 als Datenbasis verwendet werden. Ziel ist es hier, nur eine kurze Anwendung zu zeigen, da das Thema aufgrund seiner umfassenden Reichweite hier in keiner Weise komplett gezeigt werden kann. Es sollen lediglich die Potentiale der Tracking-Daten für Analysen zum Zusammenhang Radverkehr-Wetter aufgezeigt werden.

Zunächst soll eine lineare Regressionsanalyse die Einflussfaktoren Temperatur, Luftfeuchtigkeit sowie Windstärke auf ihren konkreten Einfluss auf die Zahl der Tracks je Tag untersuchen. Anzumerken ist bei dieser Analyse, dass ein statistisches Mittel der Wetterattribute über den ganzen Tag (24 Stunden) ermittelt und verwendet wurde.

Tabelle 5.5: Korrelationen Anzahl Tracks pro Tag / Durchschnittliche Wetterausprägung

Korrelation nach Pearson	tracks_count
temperature_mean	,785
humidity_mean	-,627
windspeed_mean	-,244

Mit einem korrigierten R-Quadrat von 0,672 und der vorliegenden Hochsignifikanz kann das Modell als durchaus brauchbar gesehen werden (Die detaillierten Regressionsergebnisse befinden sich im Anhang). Allerdings bildet das Modell durch die tageweise Zusammenfassung der Wetterattribute eher klimatische Unterschiede im Jahr ab als sporadisch auftretende Wetterphänomene. Dieser Jahreszeitenvergleich wurde auch schon in der Vorstellung des Datensatzes deutlich. Den kategorialen Wetterbeschreibungen konnte leider in keiner Analyse ein Einfluss nachgewiesen werden, wobei die Vorzeichen der Korrelationskoeffizienten zumeist die richtige Wirkung darstellten. Vermutet wird hier eine Destabilisierung aufgrund von zu vielen Variablen.

Um das aktuelle Wetter in den Vordergrund zu rücken soll die Analyse nochmal auf Basis der Tracks je Stunde gerechnet werden. Dabei werden nur jene Tracks zur Tageszeit (zwischen 6 und 23 Uhr) berücksichtigt, was Ausreißer durch einzelne Fahrten in der Nacht vermindern soll.

Tabelle 5.6: Korrelationskoeffizienten Anzahl Tracks je Stunde / Aktuelle Wetterausprägung

Korrelation nach Pearson	tracks_count
temperature	,496
humidity	-,491
windspeed	0,37

Das korrigierte R-Quadrat liegt bei 0,315, das Modell ist aber noch immer hochsignifikant. Es zeigt sich nun, dass die Variable „Windgeschwindigkeit“ ihr Vorzeichen geändert hat und somit bei weiteren Berechnungen nicht mehr miteinbezogen werden sollte. Die Korrelationskoeffizienten von Temperatur und Luftfeuchtigkeit sind nun deutlich kleiner und in keinem akzeptablen Bereich mehr.

Kurzfazit

Die Kombination der Wetterdaten mit den Tracks bietet eine Ausgangsbasis für detaillierte und umfassende Hypothesentests. Die große Anzahl an Attributen - sowohl auf Seite der Tracks als auch bei den Wetterdaten bieten unzählige Kombinationsmöglichkeiten.

5.6.3 Radrouten

Der Ausbau von Rad-Langstrecken ist ein im Bereich der Radverkehrsplanung der Stadt Wien beschlossenes Projekt. Ziel ist es, optimale Verbindungen in die Stadtregion und das Stadtumland für den Radverkehr zur Verfügung zu stellen, welche natürlich auch für den Durchzugsverkehr von großer Bedeutung sind. Mittelfristig soll durch die Schaffung attraktiver Radinfrastruktur aus dem Stadtumland in das Zentrum eine Verlagerung des Modal Split auf klassischen Tagespendelrelationen entstehen. Es geht also um bessere Verbindungen die der gesamten Stadtregion zugute kommen.

Erwähnenswert für die Schaffung von Radrouten ist die Festlegung von Qualitätskriterien der Stadt Wien, welche wie folgt aussehen:⁹³

- Weniger enge Kurven, unübersichtliche Stellen und Engstellen
- Möglichst kurze Wartezeit an Ampeln
- Vorrang für die Rad-Langstrecke, wo es möglich ist
- Breite Radfahranlagen, die Überholen ermöglichen

Neben den Qualitätskriterien bleibt die entscheidende Frage nach dem räumlichen Verlauf der Routen.

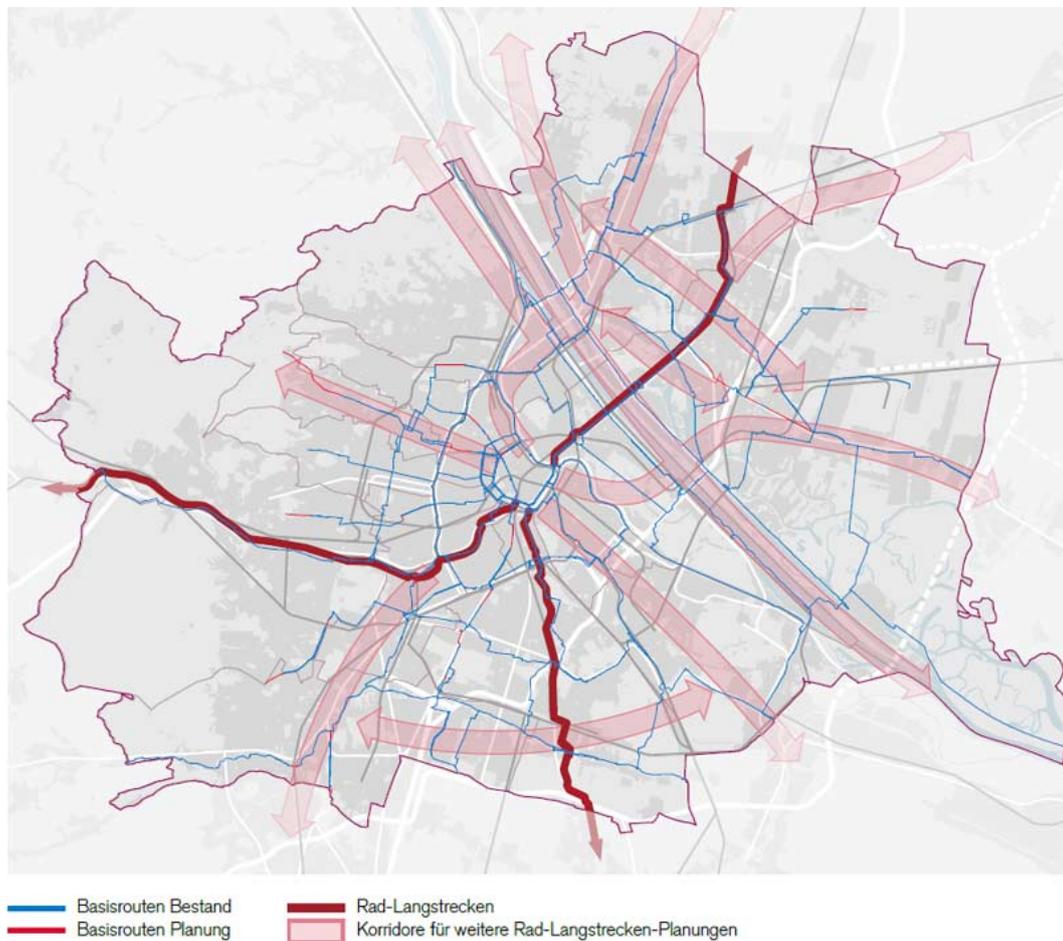


Abbildung 5.23: Verortung der Radrouten-Korridore [StadtWien, 2016b]

Vorrang haben dabei die Routen in Richtung Norden, Süden und Westen, die in der Karte dunkelrot dargestellt sind. An der Stadtgrenze sollen alle Routen in das niederösterreichische RADLGrundnetz übergehen. „Das Fachkonzept Mobilität, das im Dezember 2014 beschlossen wurde, sieht vor, dass die Route Süd bis 2018 als erste in der angestrebten Qualität nutzbar sein soll. Die Route Süd führt vom Karlsplatz über den Hauptbahnhof und die Favoritenstraße bis

⁹³vgl. [StadtWien, 2016b]

zum Anschluss Leopoldsdorf. Bis 2025 sollen weitere Routen umgesetzt werden, insbesondere die Routen Nord und West.“, so die [StadtWien, 2016b].

Für die weitere Analyse der Radrouten in Zusammenhang mit den *Bike Citizens* Tracks bedarf es der Definition einer „Rad-Langstrecke“. Dazu folgende Überlegungen: Die Dimension einer Langstrecke sollte so lang gewählt werden, dass möglichst nur Tracks selektiert werden können, die über die Stadtgrenze hinaus in die Stadtregion führen (sowie vice versa) oder durch das Stadtgebiet nahezu komplett durchführen.

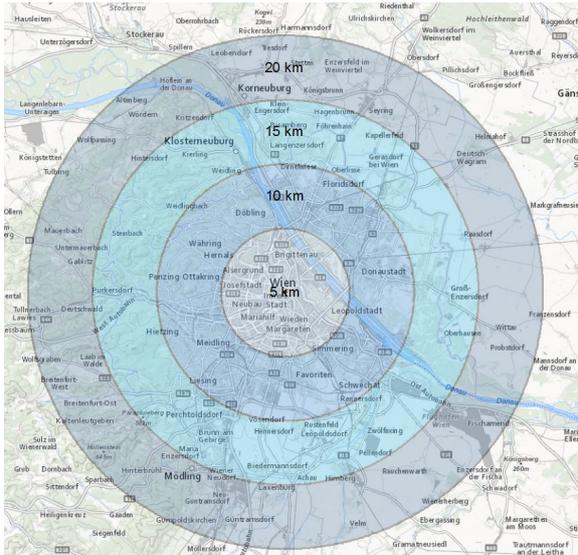


Abbildung 5.24: Luftlinienentfernungsklassen in 5-Kilometer Einteilung rund um den Schwerpunkt von Wien. [Eigene Darstellung, Kartengrundlage: www.basemap.at]

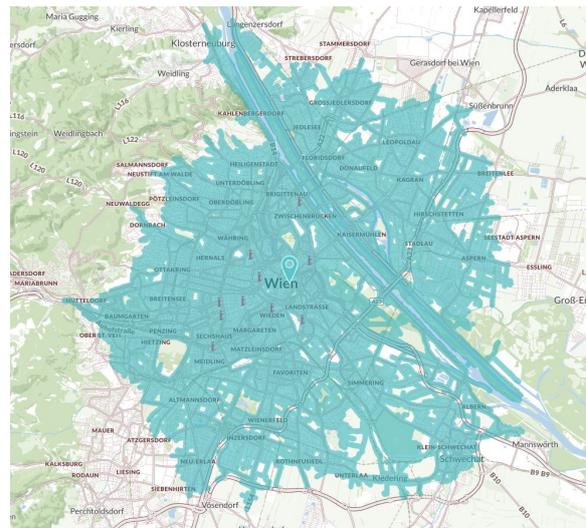


Abbildung 5.25: Bike Citizens Map, dargestellt wird die Erreichbarkeit eines Punktes innerhalb von 30 Minuten ($\varnothing 12,2$ km) mit dem Fahrrad. [<http://map.bikecitizens.net/>]

Zur Ermittlung des Grenzwertes für die Langstreckendefinition dienen die in den beiden Abbildungen 5.13. und 5.13 dargestellten Sachverhalte. Der Screenshot der *Bike Citizens* Map zeigt, dass mit einer Fahrzeit von 30 Minuten, ausgehend vom Mittelpunkt, nahezu jedes Ziel im Stadtgebiet erreicht werden kann. Dies basiert auf einer durchschnittlichen Fahrstrecke von 12,2 km, wobei auf gut ausgebauten Strecken die Erreichbarkeit innerhalb von 30 Minuten teilweise schon über das Stadtgebiet hinaus führt. Zusammen mit der Analyse der Luftlinienentfernung erscheint ein Wert bei 15 sinnvoll.

In der weiteren Vorgehensweise wurden über eine kurze SQL-Syntax jene Tracks extrahiert, welche eine Gesamtlänge über 15 Kilometern aufzeigen. Diese wurden im Anschluss als Heatmap visualisiert. Zusätzlich sollen die drei fix geplanten Radrouten (Nord, Süd, West) mit den Tracks verschnitten werden. Das Ergebnis zeigt die folgende Karte:

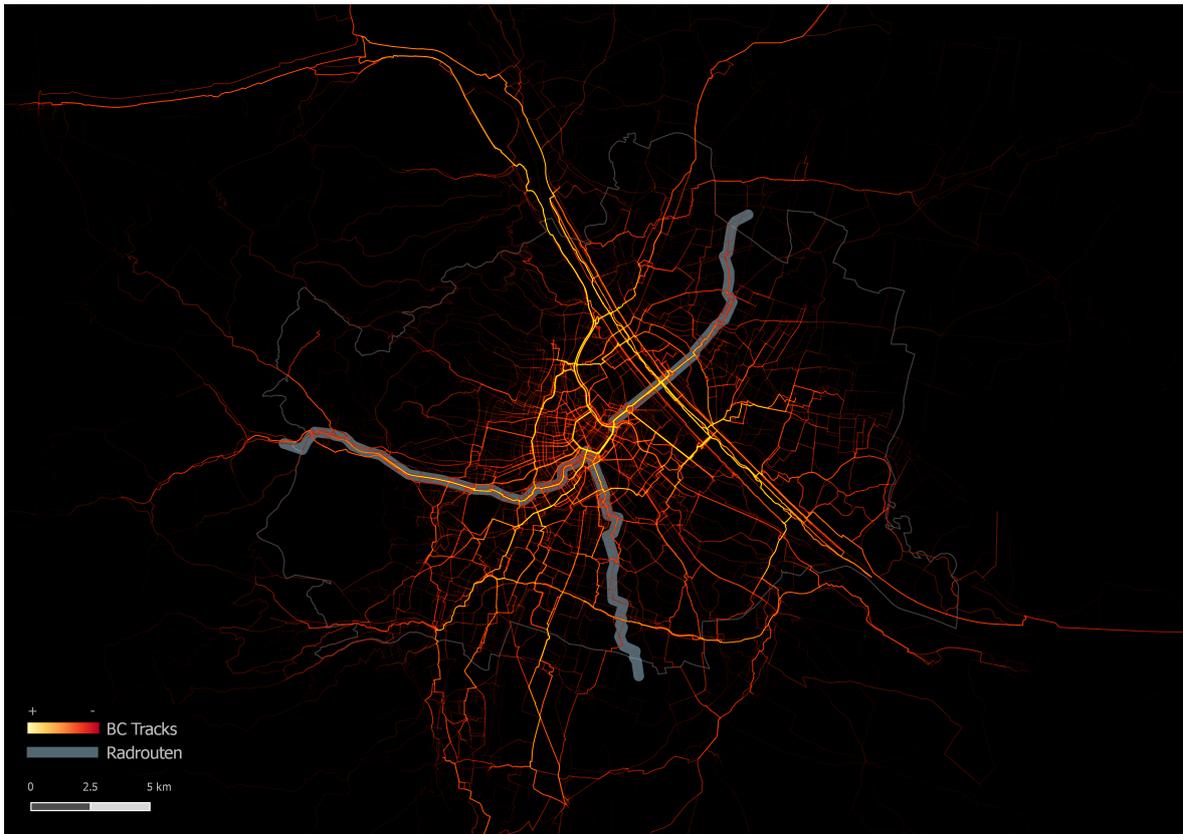


Abbildung 5.26: Langstrecken-Tracks im Vergleich mit den offiziellen Radrouten
[Eigene Darstellung / Bike Citizens Tracks]

Quantitative Aussagen sind aufgrund der aktuell im Datenbestand noch nicht implementierten Zuordnung einzelner Tracks zu Straßenabschnitten (siehe Map-Matching-Problem Kapitel 5.7) nicht möglich. Dennoch bietet die Karte eine solide Gesprächsbasis für weitere Routenentscheidungen. So fällt als erstes auf, dass die Routen Nord und West schon jetzt entlang häufig genutzten Relationen liegen, während die Süd-Route abseits der meisten aktuell gefahrenen Tracks liegt. Dies könnte entweder mit fehlenden Lückenschlüssen (z.B. im Bereich Hauptbahnhof-Sonnwendviertel) zu tun haben oder auch ganz grundsätzlich damit, dass die Relation in ihrer Priorität überschätzt wurde. Stattdessen wäre laut dem Ergebnis der Analyse eine Südroute Richtung Mödling und Baden die Route mit deutlich höherer Frequenz gewesen. Spekulieren lässt sich aber nur über die Grundphilosophie der Planung: Sollen Routen dort ausgebaut werden, wo bereits jetzt rege Nutzung stattfindet oder sollen Routen dort ausgebaut werden, wo zum jetzigen Zeitpunkt kaum Radverkehr herrscht, um eventuelle Umlenkungsprozesse umzusetzen? Ganz klar ersichtlich wird auch der Radverkehr entlang den Donauradwegen zwischen Tulln und Klosterneuburg nach Wien und weiter Richtung Donauauen und Bratislava. Diese Relationen befinden sich aber zum Großteil jetzt schon auf internationalen Radrouten. Dies macht wiederum deutlich, warum im Check-Up Teil der Versuch unternommen wurde, die Nutzer in Freizeit- und Alltagsradverkehr zu teilen. Folgt man der Argumentation der Stadt Wien, so werden die angesprochenen Radrouten vorrangig für den Alltagsradverkehr geplant. Über die Selektion von Fahrten über 15 Kilometern ist aber davon auszugehen, dass oben gezeigte Heatmap einen Großteil an Freizeitradverkehr beinhaltet.

Kurzfasit

Das Kapitel der Radrouten macht deutlich, dass durch die Vorauswahl bestimmter Streckenlängen der Datenbestand auf eine bestimmte Fragestellung hin optimiert werden kann. Das Herauslesen von Quelle-Ziel-Verbindungen ist aber nicht möglich, weshalb auch kein Bezug zwischen quelseitiger Verkehrsnachfrage und zieleitigen Standortfaktoren getroffen werden kann. Durch die Problematik der Abgrenzung des Alltagsradverkehrs vom Freizeitradverkehr können zudem Verzerrungen nicht ausgeschlossen werden.

5.7 Meso-Ebene

Die Meso-Ebene stellt die mittlere Maßstabsebene dar. Ihr Analyseausschnitt wird der Fragestellung angepasst und ist flexibel gewählt. Inhalte der Analysen dieser Ebenen sind Fragen zur Routenwahl und Routencharakteristik sowie Erreichbarkeitsanalysen.

5.7.1 Routensplit

Der Routensplit beschreibt die Aufteilung von Verkehrsströmen auf verschiedene Fahrtrwege⁹⁴. Ziel dieses Kapitels ist es, beispielhaft einen Routensplit zu analysieren und die Tauglichkeit des Datensatzes als Basis einzuschätzen.

Essentiell für diese Analyse ist eine geeignete städteräumliche Relation, auf der es mehrere - im Idealfall auch ähnlich attraktive - Fahrtmöglichkeiten gibt. Die erforderlichen Kriterien erfüllte schlussendlich die Relation Alsergrund (9. Bezirk) Nähe Volksoper (Wohnort des Autors) - Margareten (5. Bezirk). Zwischen diesen beiden Bereichen gibt es - neben etlichen Varianten - drei verschiedene Haupttrouten:

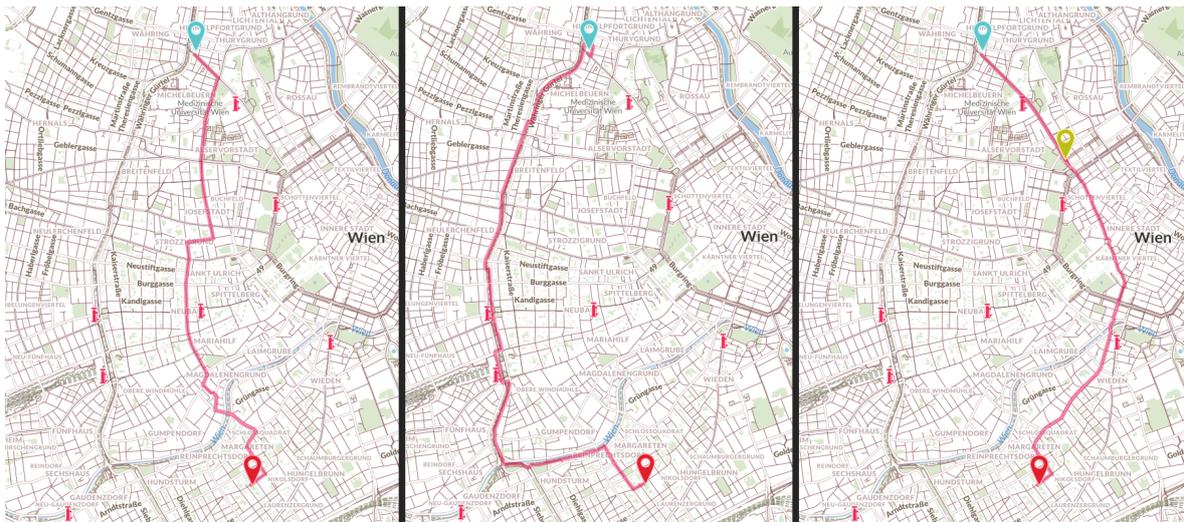


Abbildung 5.27: v.l.n.r.: Route „Direkt“, Route „Gürtel“ und Route „Innenstadt“. [Screenshots Bike Citizens Routing]

Die **Route „Direkt“** führt durch die Bezirke 6, 7 und 8 und folgt dabei annähernd der Luftlinie der Relation. Wichtige Straßenzüge entlang der Route sind die Lange Gasse und die Neubaugasse - dies ist aber von der Fahrtrichtung abhängig: Die Route führt durch über vergleichsweise viele Einbahnen, die nicht alle für den Radverkehr geöffnet sind oder aufgrund des starken Verkehrs auch keine Fahrt gegen die Einbahn empfohlen werden kann.

Die **Route „Gürtel“** führt - wie der Name schon verrät - großteils über den Gürtelradweg, und zwar direkt von der Volksoper bis zur Linken Wienzeile. Von dort folgt sie der U-Bahnlinie 4 stadteinwärts bis - je nach genauem Ziel - in den 5. Bezirk abgelenkt wird.

⁹⁴vgl. [Kirchhoff, 2013]

Die **Route „Innenstadt“** führt, von der Volksoper ausgehend, die Währinger Straße Richtung Ring entlang. Hier bestehen grundsätzlich dann zwei Möglichkeiten: Entweder der Radfahrer folgt dem Ring entlang bis zur Oper oder entscheidet sich für den Weg durch die Herrengasse und die Innenstadt. Über die Operngasse und Margaretenstraße führt die Route weiter in den 5. Bezirk.

Nun der Blick auf die *Bike Citizens* Tracks: Selektiert wurden hier alle Fahrten, die sowohl den grau hinterlegten Zielbereich im 9. Bezirk (Norden) als auch den grau hinterlegten Zielbereich im 5. Bezirk (Süden) berühren. Dabei ist es nicht von Relevanz, ob die Tracks in einem der beiden Gebiete starten oder enden. Die Stichprobengröße liegt bei dieser Analyse bei 333 Tracks.

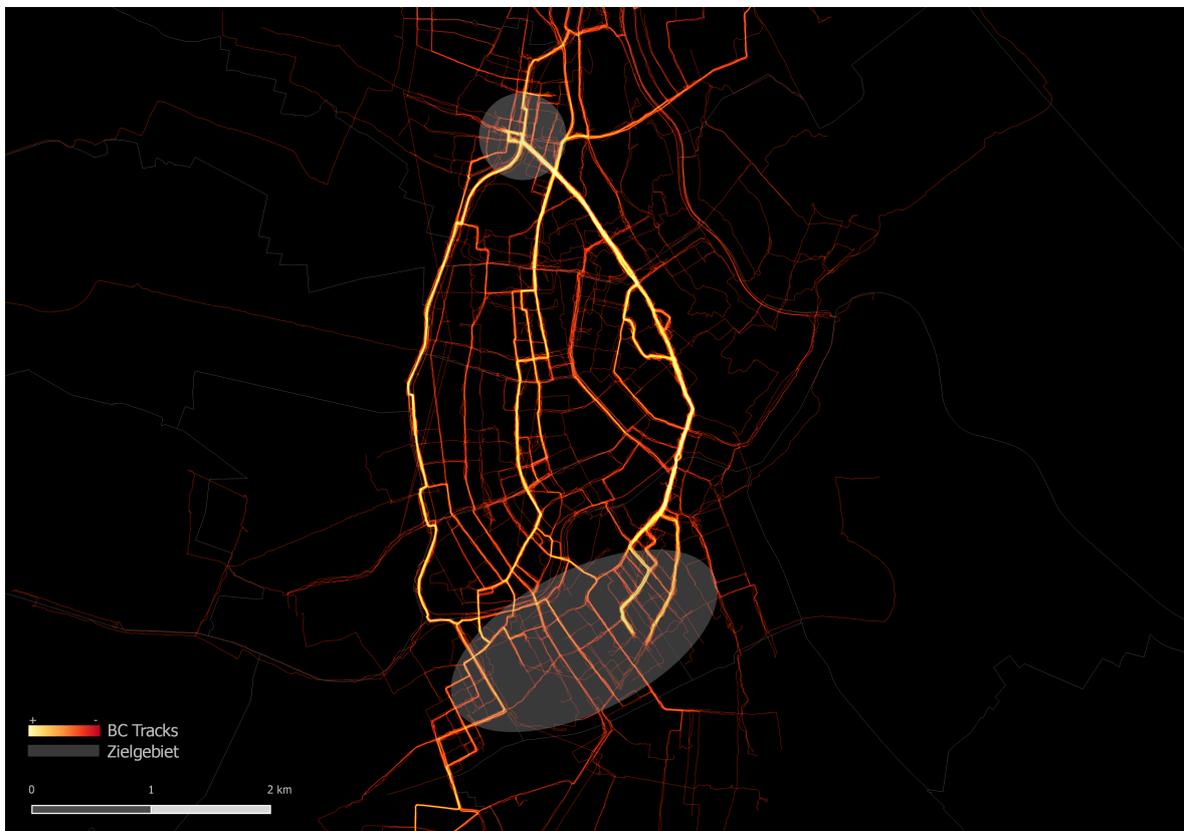


Abbildung 5.28: Bike Citizens Tracks zwischen Gebiet Volksoper im 9. Bezirk und dem 5. Bezirk. Erkennbar sind die in Abbildung 5.14 dargestellten Routenoptionen. (n=133)
[Eigene Darstellung / Bike Citizens Tracks]

Die Routen lassen sich wie folgt im Detail charakterisieren:

Tabelle 5.7: Routensplit im Detail

	Route „Direkt“	Route „Gürtel“	Route „Innenstadt“
Distanz	5,1 km	6,5 km	5,3 km
Fahrzeit BikeCitizens App	19 min	24 min	21 min
Fahrzeit Google Maps	23 min	25 min	20 min
Gemessene Fahrzeit	23,6 min	28,2 min	19 min
Gemessene Strecke	5,8 km	6,1 km	5,3 km
Typ Radinfrastruktur	40% Markierte Anlagen 55% Mischverkehr 5% Getrennte Führung	Großteils getrennte Führung	60% Markierte Anlagen 30% Getrennte Führung 10% Mischverkehr
Anteil BC-Tracks	24%	22%	45%

Anmerkungen zur Tabelle: Für die Ermittlung der Distanz wurden in den Navigationstools von *Bike Citizens* und Google die Adressen Währinger Straße 78, 1090 Wien sowie Siebenbrunnengasse 26, 1050 Wien verwendet. Die Ergebnisse zur gemessenen Fahrzeit und Strecke stammen direkt aus den *Bike Citizens* Tracks. Dazu wurden alle Fahrten selektiert, bei denen sich sowohl Start- als auch Endpunkt innerhalb des definierten Bereichs befinden (Ausreißer durch Rundfahrten wurden entfernt. Ausschlaggebend war hierbei eine Abweichung der Distanz von >25%). Die Werte wurden nach Route gemittelt, wobei die Stichprobengrößen sehr unterschiedlich waren: Die kleinste Menge liegt bei der Route „Direkt“ mit n=6 vor, gefolgt von „Gürtel“ mit n=11 und „Innenstadt“ mit n=77. Der Typ Radinfrastruktur wurde über die zur Verfügung gestellten OpenGeo-Daten auf <http://data.gv.at> ermittelt. Entscheidend war hier das Attribut „Merkmal“. Der Anteil der BC-Tracks wurde durch je vier gemittelte Querschnittszählungen entlang der Routen ermittelt. Die restlichen 9% bewegen sich über andere Wege.

Grundsätzlich lässt sich der Routensplit - bei einer geeigneten räumlichen Umgebung - mit den *Bike Citizens* Tracks gut darstellen und analysieren. Auch die Quantifizierbarkeit der Split-Größen ist durch wiederholte Querschnittszählungen ohne weiteres möglich. Die genaue Abfahrtsadresse innerhalb der Zielrelationen hat zudem auch eine entscheidende Rolle in der Routenwahl, wurde aber aufgrund der sonst zu geringen Datenmengen in diesem Beispiel nicht weiter berücksichtigt (das Phänomen lässt sich aber anhand der gezeigten Heatmap durchaus erkennen). Zudem bleibt die generelle Problematik der Verzerrung durch individuelles Nutzerverhalten, welche sich schon durch die unterschiedlichen Stichprobengrößen bei der gemessenen Fahrzeit und Route zeigt. Die Grundproblematik ist, dass in den Daten keine Zuordnung zu Individuen möglich ist - deshalb kann in der Analyse auch nicht unterschieden werden, ob die Fahrten von einem Nutzer der täglich unterwegs ist oder mehreren die nur wenige Male die Route gewählt haben stammen.

Das größte Problem der Analyse ist aber, dass die versuchte Rückkoppelung auf die Nutzentscheidung, welche Route gewählt wird, mit den *Bike Citizens* Daten aus folgendem Grund nicht möglich ist: Es ist davon auszugehen, dass ein Großteil der Nutzer die App zum Zwecke der Navigation benutzt und der vorgeschlagenen Route des Algorithmus auch folgt. Ergo kann über diese Form der Analyse nahezu ausschließlich auf Entscheidungsfaktoren des *Bike Citizens* Routing-Algorithmus rückgekoppelt werden - nicht jedoch auf die eigentliche Entscheidung des Nutzers. Die Entscheidung der Routenwahl wurde in diesen Fällen vermutlich

fast nie vom Nutzer selbst getroffen - die einzige nutzerbasierte Entscheidung war, die Routenwahl dem *Bike Citizens* Algorithmus zu überlassen.

Laut *Bike Citizens*⁹⁵ ist zukünftig geplant, den Typ des Tracks zusätzlich zu unterscheiden: Zum einen in Tracks, die dem Routing folgen und das Routing verwenden und zum anderen in Tracks, die ohne Routing aufgezeichnet werden. Mit diesen Informationen könnten dann Analysen Richtung Nutzerverhalten realisiert werden.

Kurzfasit

Die Ergebnisse des Routensplits zeigen, dass das Vorhandensein einer großen Menge an Daten für Analysen der Meso-Ebene von entscheidender Bedeutung ist. Diese Menge sollte im Idealfall so groß sein, dass die Verzerrung durch individuelles Verhalten möglichst gering gehalten werden kann. Derzeit sind nicht auf allen dargestellten Alternativstrecken genügend Tracks für aussagekräftige Ergebnisse vorhanden. Zudem sollte aus Anwendersicht im Datenbestand unterschieden werden können, ob der jeweilige Track vom Typ „Tracking mit Routing“ oder „Tracking ohne Routing“ ist. Ohne diese Unterscheidung besteht die Unsicherheit, ob die Rückkopplung auf das Nutzerverhalten oder auf den Routingalgorithmus erfolgt.

5.7.2 Einzugsbereichsanalyse

Erreichbarkeitsanalysen haben in der Stadt- und Regionalforschung einen hohen Stellenwert. Grundlage dafür ist allgemein ein Graph bzw. ein Netz, anhand dessen mithilfe graphentheoretischer Funktionen, Berechnungen durchgeführt werden können. Dazu zählen beispielsweise der kürzeste Weg oder auch die Generierung von Isochronen, also die Menge aller Angebote, die innerhalb einer bestimmten Reisezeit von einem Punkt aus erreicht werden kann. Das Feld der Erreichbarkeitsanalysen ist vielseitig und kann auch projektbezogen adaptiert werden. Über Aussagen zur Erreichbarkeit können auch Aussagen zur Standortqualität untermauert werden - ist diese doch ein wichtiger Standortfaktor.

In der Regionalwissenschaft sind derzeit Isochronenmodelle verbreitet. Es besteht großes Interesse an der Frage, woher Personen kommen und wie weit sie bereit sind zu fahren. Nachdem dafür keine Daten vorhanden sind wird über graphenbasierte Ansätze versucht, möglichst plausible Schätzungen zu berechnen. Die Problematik der Berechnung ist, dass die Kantenimpedanzen bloß eine sehr grobe Annäherung an die Realität darstellen und die Anzahl der Fahrten aus den Modellen nicht hervorgeht. Mithilfe verschiedener Modellansätze - von Interaktions- bis Gravitationsmodell - versucht man der Problematik der Ungewissheit der Größe der Ströme zwischen Quelle und Ziel entgegen zu wirken.

Anders als diese modellbasierte Herangehensweise kann auf Grundlage der Fahrrad-Tracking-Daten beobachtete Interaktion beurteilt und analysiert werden. Die realitätsnähe ist hier jedenfalls größer - werden die Daten doch von der Realität - also den Nutzern - generiert.

⁹⁵vgl. [BikeCitizens, 2017]

Die Erreichbarkeitsanalyse am Beispiel der *Bike Citizens* Daten kann durch die Map Matching Problematik nicht graphenbasiert berechnet werden. Wären die Tracks bereits Kanten zugeordnet und auch eine Lösung für das Übertragen der Fahrzeiten vorhanden, könnten modellbasierte Zugänge mit realitätsnahen Daten weiterentwickelt werden. Vielmehr wurde hier der Ansatz der Einzugsbereichsanalyse gewählt, der analysieren soll, woher die Nutzer einer bestimmten Einrichtung oder eines Point of Interest kommen, wie weit und wie lange sie gefahren sind - also ein etwas anderes Verständnis von Erreichbarkeit als in der Regionalanalyse verbreitete. Nichtsdestotrotz kann mit diesem Ansatz - vor allem mithilfe eines Vergleichs mehrerer Standorte - beispielsweise über die durchschnittliche Distanz, die fahrradbasierte Erreichbarkeit beurteilt werden.

Für das Beispiel wurde der Standort Karlsplatz der TU-Wien ausgewählt. Innerhalb eines Einzugsgebietes von 300 Metern rund um das historische Hauptgebäude befinden sich noch viele weitere wichtige Standorte der Universität. Alle Tracks, die in diesem Zielgebiet enden, werden für die folgenden Analysen berücksichtigt. Dies ergibt eine Stichprobengröße von 2020 Fahrten. Die Abbildung zeigt kein Interaktionsmodell sondern ein Abbild der Interaktionsrealität.

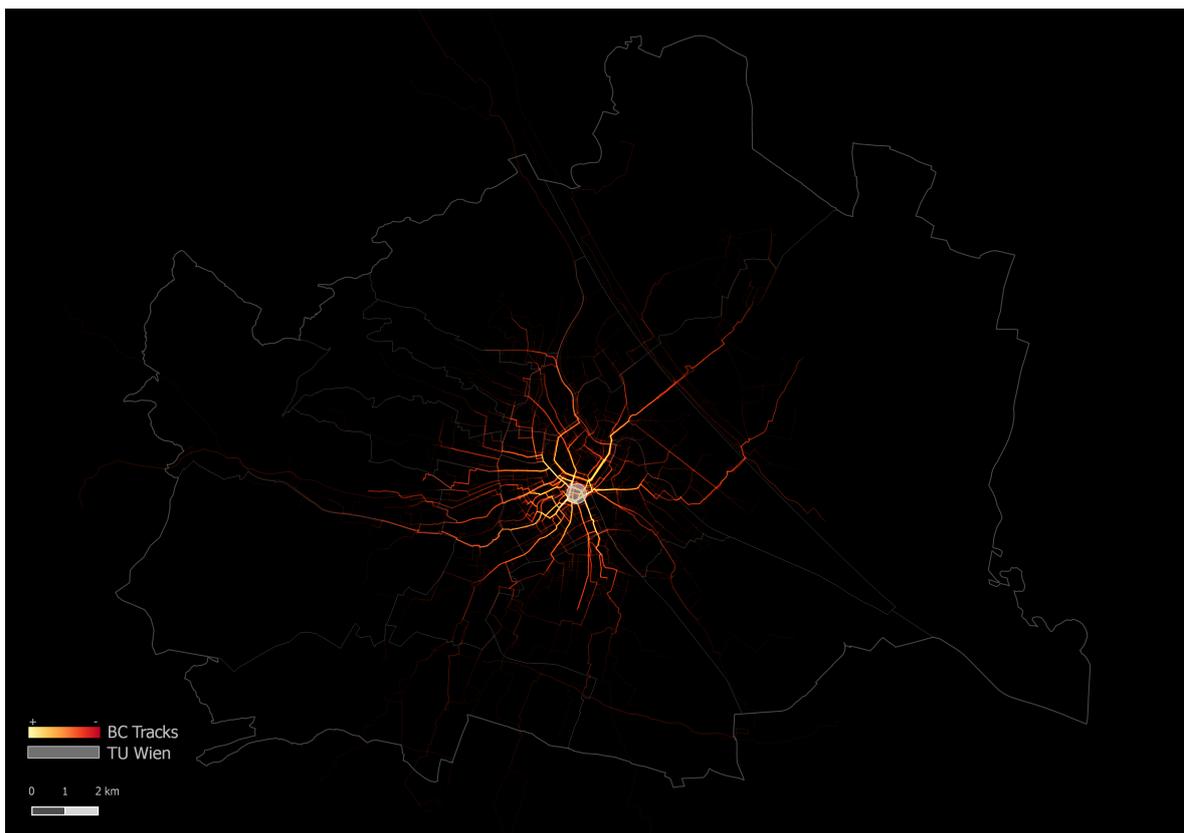


Abbildung 5.29: Alle Bike Citizens Tracks, die im 300m Radius des Hauptgebäudes der TU Wien enden. [Eigene Darstellung / Bike Citizens Tracks]

Zusätzlich zur grafischen Aufbereitung lässt sich die Datenanalyse der Tracks im Bereich der Erreichbarkeitsfrage noch weiter führen. Dies geschieht mithilfe eines Imports der Ergebnisdaten in ein gängiges Statistikprogramm - wie beispielsweise SPSS. Die Analyse der Distanz

zeigt das gewohnte breite Spektrum - das Minimum liegt bei 200 Metern während die größte zurückgelegte Distanz knapp über 60 Kilometer beträgt. Der Mittelwert liegt bei 3271m und der Median bei 2390m, bei einer Standardabweichung von 3235m.

Die Verteilung im Histogramm zeigt aber deutlich, dass auch hier wieder einige verhältnismäßig große Ausreißer im Datensatz vorhanden sind. Um die eigentliche Aussage - die durchschnittlich zurückgelegte Distanz - argumentieren zu können, wird ein Blick auf das Histogramm ohne Ausreißer helfen. Der Grenzwert liegt dabei bei 15 Kilometern.

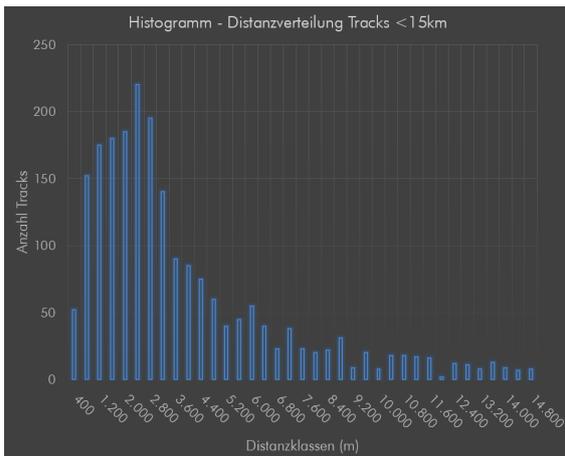


Abbildung 5.30:
Häufigkeitsverteilung der Distanz <15km.
[Eigene Darstellung]

Das nun erstellte Histogramm zeigt die Verteilung der großen Mengen genauer, da diese alle im Bereich um 2500m rangieren. Der Mittelwert liegt nun bei 3071, der Median nahezu unverändert bei 2372.

Dies zeigt, dass durch die Einbeziehung des Grenzwerts nur sehr wenige Fahrten wegfallen (exakt 22, also in etwa 1% - $n_{neu}=1998$). Die durchschnittliche Distanz dürfte also bei ungefähr 2300m liegen, was etwas unter dem allgemeinen Wiener Durchschnitt liegt. Die Kurve macht auch deutlich, dass nur ein sehr geringer Anteil der Personen mehr als 6km zurücklegt.

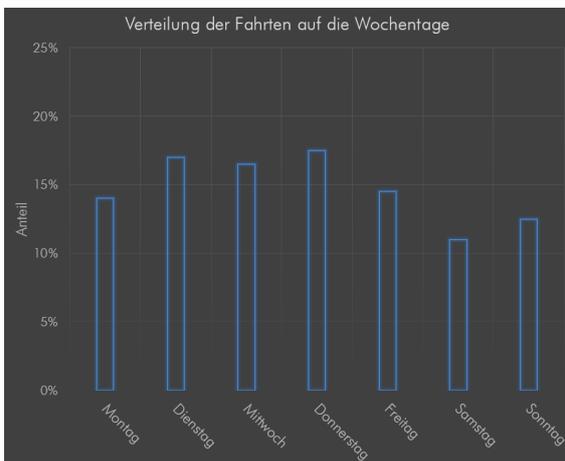


Abbildung 5.31:
Anteilige Verteilung der Tracks auf die Wochentage.
[Eigene Darstellung]

Um Aussagen auf der inhaltlichen Ebene beispielsweise zur Nutzergruppe zu argumentieren, ist es hilfreich, sich die weiteren mitgelieferten Daten und Verteilungen im Detail anzusehen. Dabei gilt ein erster Blick der Häufigkeitsverteilung der Fahrten auf die Wochentage. Folgt man nun dem Ziel, möglichst nur universitäre Nutzer in die Analyse miteinzubeziehen, wäre spätestens hier ein Umdenken erforderlich. Die Werte begründen sich wohl mit dem stark diversifizierten Angebot der Gegend, was diese auch für Nicht-Studierende zum Ziel machen könnte. Der erhöhte Anteil an Werktagen lässt aber trotzdem die Vermutung zu, dass ein großer Teil der Fahrten hier sein Ziel aufgrund der Universität (Mitarbeiter / Studierende) oder einer anderen beruflichen Beschäftigung hat.

Um diese Vermutung weiter zu überprüfen, folgt der nächste Blick auf die Verteilung der Ankunftszeiten im Zielgebiet. Diese ergeben sich aus der Fahrtzeit und dem Startzeitpunkt, wobei alle Fahrten über 75 Minuten aufgrund von großen Ausreißern (Maximum bei 15 Stun-

den Fahrzeit, Datenfehler da Strecke nur 6km) ignoriert wurden (=14 Tracks).

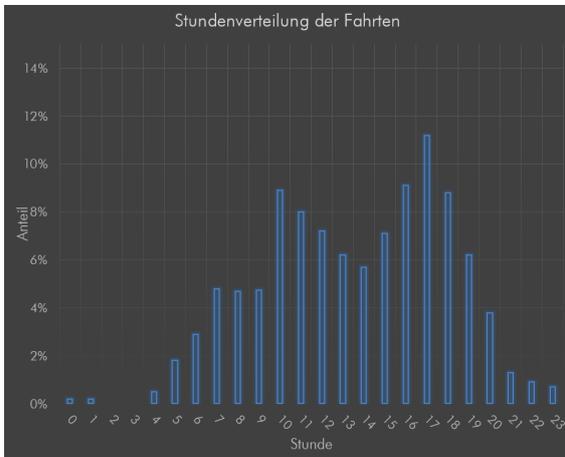


Abbildung 5.32: Anteilige Verteilung der Ankunft der Tracks auf die Stunden an Werktagen. [Eigene Darstellung]



Abbildung 5.33: Anteilige Verteilung der Ankunft der Tracks auf die Monate. [Eigene Darstellung]

Der Verteilung kann ein erster Anstieg der Ankünfte ab 7:00 Uhr früh entnommen werden. Dabei handelt es sich wohl eher um Angestellte der Universität oder lokal ansässigen Unternehmen als um Studierende - die Ankunft letzterer bildet wahrscheinlich eher der Peak um 10 Uhr ab. Auffallend ist auch der Anstieg am Abend gegen 18 Uhr. Eine Erklärung hierfür könnten die vielen Lokale der Gegend sein oder auch - vor allem im Sommer - der Resselpark und der Brunnen vor der Karlskirche als attraktive Freifläche. Zudem darf der Umkreis der Universität auch als Wohnraum nicht unterschätzt werden - es könnten also auch Fahrten in das jeweilige Zuhause enthalten sein.

Ein letzter Blick führt zur anteiligen Verteilung der Tracks auf die Monate. Mit knapp über 20% finden die meisten Fahrten im Juni statt, dem letzten Monat vor den Sommerferien der Universität. Danach fallen die Anteile stark ab. Die gleichmäßige Verteilung in den kühleren Monaten Oktober bis Jänner zeigt wohl die Nutzer, die klimaunabhängig das Fahrrad für ihren Weg benutzen.

Abschließend soll noch auf die datenschutz-basierte Kappung der Tracks am Beginn und Ende der Fahrt um jeweils 100m erinnert werden. Dies wurde in dieser Analyse weitgehend ignoriert, eine genaue Verortung des Ziels würde aber wesentlich exaktere Aussagen zum Zweck der Fahrt ermöglichen.

Die Möglichkeiten im Bereich der Erreichbarkeitsanalyse wurden an einem sehr konkreten Beispiel gezeigt. Naheliegend ist aber, diese Analyse auch für weitere Einrichtungen oder Points of Interest zu wiederholen. Um diesen Prozess zu vereinfachen wurde auf Basis des ArcMap ModelBuilders ein Werkzeug entwickelt, welches die erforderlichen Berechnungen automatisch durchführt. Über die Wahl des passenden Parameters kann diese Analyse auch für alle ausgehenden Fahrten (=Startpunkt) aus dem gewählten Bereich durchgeführt werden. Dieses Modell ist als technische Dokumentation im Anhang der Arbeit zu betrachten.

Kurzfasit

Die Einzugsbereichsanalyse hat gezeigt, dass durch Fahrrad-Tracking-Daten wesentlich realitätsnähere Ergebnisse im Vergleich zu modellbasierten Ansätzen zu erwarten sind. Die Abbildung von realen Interaktionen im Raum, sowie konkreten Menge,n sind hier ebenso als große Potentiale zu nennen wie mögliche Detailanalysen zur zeitlichen Verteilung und die Tatsache, dass nahezu exakte Quelle-Ziel Verbindungen vorhanden sind. Problematisch bleibt aber die Repräsentativität und die relativ kleine Stichprobe sowie die bereits im vorigen Kapitel erklärte Verzerrung durch individuelles Verhalten. Durch die Dimension des Wetters können die Daten auch für spezielle Fragestellungen als Basis herangezogen werden.

5.7.3 Verkehrsspinne

Schon die Analyse jener Tracks, die ihr Ziel im Umkreis der TU Wien haben, hat gezeigt, dass sich Verkehrsströme und die Verteilung im Netz mit den *Bike Citizens* Daten gut visualisieren lassen. Eine ähnliche Vorgehensweise gibt es bei der Darstellung von Verkehrsspinnen.



Abbildung 5.34:
Die 4 Eckpfeiler der Neugestaltung der Hasnerstraße.
[StadtWien, 2016a]

ortelinie sowie U3 Endstation Ottakring im 16. Bezirk auf insgesamt 2,5 Kilometern. Ziel ist es auch, eine für den Radverkehr attraktive Alternative zu den parallel stark vom MIV belasteten Routen Thaliastraße, Koppstraße und Gablenzgasse zu etablieren⁹⁷.

Als konkretes Beispiel wurde die Hasnerstraße im 16. Wiener Gemeindebezirk ausgewählt. „Die erste fahrradfreundliche Straße Österreichs ist in Wien-Ottakring eröffnet worden. Auf zweieinhalb Kilometern Länge haben die Radfahrer auf der Hasnerstraße künftig bei fast jeder Kreuzung Vorrang.“⁹⁶ Zwar wurde die Hasnerstraße schon vor dem Umbau von Radfahrern gerne genutzt, so sollte durch den Umbau aber vor allem eine Erhöhung der Verkehrssicherheit erzielt werden. Die Hasnerstraße verbindet den Gürtel im Bereich U6 Thaliastraße und der Vor-

Die Überprüfung der Attraktivität und die dadurch ableitbare Umwegbereitschaft bietet sich mit den *Bike Citizens* Tracks an. Dazu wurde in der Nähe des Gürtels und in der Nähe des Schuhmeierplatzes ein Querschnitt gezogen und alle Tracks selektiert, die beide Querschnitte passieren. Das Ergebnis - wie gewohnt visualisiert - zeigt folgende Verkehrsspinne:

⁹⁶vgl. <http://wien.orf.at/news/stories/2550826/>, zuletzt aufgerufen am 20.1.2017

⁹⁷vgl. [StadtWien, 2016a]



Abbildung 5.35: Verkehrsspinne Hasnerstraße. [Eigene Darstellung / Bike Citizens Tracks]

Die Darstellung zeigt einen gut erkennbaren Trichtereffekt an beiden Enden der Hasnerstraße. Zum einen ist die Verteilung in die weiteren Außengegenden Wiens ab der Vorortelinie gut erkennbar, zum anderen wird auch die Verteilerwirkung des Gürtels sichtbar. Auch bei der Analyse der Verkehrsspinne bleibt die Problematik der Rückkopplung auf den *Bike Citizens* Algorithmus. Sobald innerhalb der Datenstruktur unterschieden werden kann, ob der Nutzer dem Routing folgt oder nur seinen Weg trackt, können über die Methode Aussagen zur Umwegrentabilität von Fahrradfahrten erhoben werden.

Kurzfasit

Die Verkehrsspinne kann als Spezialform der Einzugsbereichsanalyse gesehen werden, weshalb auch die Herausforderungen und Potentiale ähnlich sind. Über die Systematik der visuellen Analyse kann mithilfe der gezeigten Vorgehensweise die Bedeutung einzelner Abschnitt für das Gesamtnetz aufgezeigt werden.

5.8 Mikro-Ebene

Die Mikroebene stellt die kleinste Maßstabebene dieser Analyse dar. Der Fokus liegt dabei tief im Stadtgefüge, konkret auf einzelnen Straßenabschnitten. Im Gegensatz zu den bisherigen Analysen auf der Makro- und Mesoebene werden im Folgenden nicht die linienhaften Tracks, sondern deren zugrundeliegende Track-Points verwendet. Für diese Vorgehensweise gibt es zwei verschiedene Erklärungen: Zum einen bieten die Tracks durch ihre fehlende Teilbarkeit in Einzelabschnitte auf der Mikroebene kaum Aussagemöglichkeiten. Auf der anderen Seite können die Track-Points aufgrund ihrer überproportional großen Datenmenge nur mit großem Ressourcenaufwand in Analysen auf der Makro- und Mesoebene miteinbezogen werden. Um die wertvollen Informationen dieser Datenschicht aber dennoch ausreichend zu beleuchten, drängt sich der Einsatz in der Mikro-Ebene natürlich nahezu auf.

5.8.1 Geschwindigkeitsanalyse

Die Grundsystematik des *Bike Citizens* Algorithmus zur Datenaufzeichnung ist das Abspeichern der GPS-Position im Intervall von einer Sekunde inklusive Angaben zur Höhe (laut GPS-Signal) und zur Genauigkeit der GPS-Koordinaten. Durch dieses Vorgehen werden pro Fahrt eine große Menge an Daten kreiert. Über eine `track_id`⁹⁸ kann die vorliegende Punktwolke nach einzelnen Fahrten geordnet werden. Wichtig für die korrekte Darstellung ist auch eine durchgehende Nummerierung der Punkte je Track. Das Ergebnis daraus wurde schon behandelt - nämlich die einzelnen Tracks - jedoch liegt nun der gravierende Unterschied vor, dass eine Teilung jederzeit und überall möglich ist.

Für die weiteren Analysen auf der Mikroebene ist die Möglichkeit der Aufspaltung von großer Bedeutung. In einer ersten Analyse soll die Ermittlung der Geschwindigkeit in den Mittelpunkt des Geschehens gerückt werden. Um die Berechnung mit den vorhandenen technischen Ressourcen durchführen zu können, wurde vorab ein Teilausschnitt der Punktwolke definiert. Dieser befindet sich auf der Mariahilfer Straße in Wien innerhalb der Fußgängerzone auf Höhe der Zollerstraße und erstreckt sich über eine Länge von 250 Metern. Es stellt sich die Frage, wie hoch die Durchschnittsgeschwindigkeit innerhalb dieses Abschnittes ist. Angemerkt sei an dieser Stelle noch, dass in Fußgängerzonen für den Radverkehr - sofern dieser erlaubt ist, was im Fall der Mariahilfer Straße zutrifft - maximal Schrittgeschwindigkeit gefahren werden darf.⁹⁹

Anhand dieses Ausschnitts sollen beispielhaft die Eignung der Daten sowie auch die Vorgehensweise vorgestellt werden. Um die Analyse auch mit jedem anderen räumlichen Ausschnitt möglichst einfach durchführbar zu machen, wurde ein Modell entwickelt. Dieses Modell gliedert sich in ein Hauptmodell und ein untergeordnetes Submodell, welches im Anhang als technische Dokumentation detaillierter erklärt wird.

⁹⁸Eindeutige Identifikationsnummer aller Fahrten

⁹⁹vgl. <https://www.help.gv.at/Portal.Node/hlpd/public/content/61/Seite.610300.html> - zuletzt aufgerufen am 20.1.2017

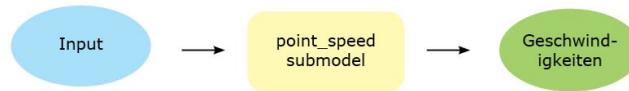


Abbildung 5.36: Grundsystematik des Modells zur Geschwindigkeitsberechnung. [Eigene Darstellung]

Das Ergebnis enthält nun alle Tracks des definierten Ausschnitts sowie die berechneten Informationen. Mithilfe eines Attributes für die Kursrichtung kann nun auch nach Fahrtrichtung selektiert werden. Dabei ist die Logik der Kompassrose für die Weiterverarbeitung von Bedeutung.

Über eine einfache SQL-Abfrage können die Fahrten nach Richtung (=Grad) selektiert werden. Wichtig ist dabei, einen ausreichend großen Toleranzbereich zu wählen, da nur wenige Tracks exakt der gewünschten Richtung entsprechen werden. Bei geraden Streckenabschnitten kann dieser Toleranzbereich ohne Gefahr bei +/- 25 Grad liegen, da die Zielwerte für die Wunschrouten bei einer geraden Strecke 180 Grad voneinander entfernt sind. Bei kurvigen Streckenabschnitten muss berücksichtigt werden, dass die Peilung jeweils die Hauptrichtung jeder Linie - betrachtet vom Start- zum Endpunkt - ermittelt. Die gewünschte Richtung (=Grad) muss dementsprechend adaptiert werden.

Fehlerhafte Ergebnisse wurden an ihrer Anzahl an Punkten innerhalb des Analysegebietes erkannt. Alle Tracks mit weniger als vier Punkten wurden daraufhin aus der Analyse gelöscht. Dies trifft vor allem auf solche Tracks zu, die das Gebiet nur anschnitten oder queren, aber nicht dem Straßenverlauf der Mariahilfer Straße folgen.

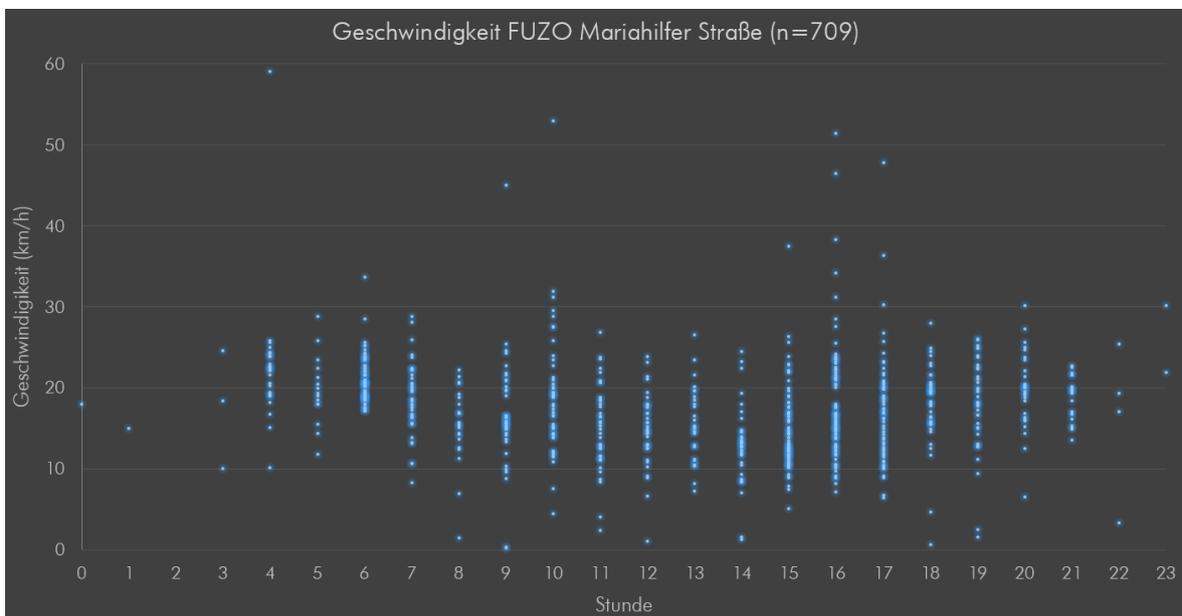


Abbildung 5.37: Geschwindigkeiten beider Fahrtrichtungen im Stundenverlauf [Eigene Darstellung / Bike Citizens Daten]

Die Geschwindigkeitsanalyse zeigt, dass ein Großteil der Fahrten das Tempolimit der Schrittgeschwindigkeit überschreitet. Vor allem außerhalb der Geschäftszeiten - beispielsweise um 6 Uhr früh - werden oft Geschwindigkeiten über 20 km/h erreicht.

Tabelle 5.8: Geschwindigkeitsvergleich je Richtung [Eigene Darstellung]

	Richtung Ring	Richtung Westbahnhof
Anzahl Tracks (n)	342	367
Mittelwert	17,3 km/h	18,1 km/h
Median	17,5 km/h	17,5 km/h
Standardabweichung	5,8	7,4

Der Blick auf die Statistik zeigt je Fahrtrichtung nur geringe Unterschiede. Dies liegt auch daran, dass der gewählte Referenzabschnitt als ausgesprochen flach bezeichnet werden kann. Das Modell ließe sich in weiterer Folge auch dahingehend adaptieren, um Analysen zum Zusammenhang zwischen Steigung und Geschwindigkeit zu erhalten.

Kurzfasit

Durch die Berücksichtigung von realen Fahrtgeschwindigkeiten aus einer großen Stichprobe können Navigations- und Routingapplikationen nachhaltig hinsichtlich ihrer Treffsicherheit verbessert werden. Dabei ist es auch möglich, Geschwindigkeiten nach Tageszeit, Wochentag oder Jahreszeit zu unterscheiden. Mithilfe der vorgestellten Methode können Netzabschnitte identifiziert werden, die aufgrund von besonders niedrigen Durchschnittsgeschwindigkeiten Aufmerksamkeit als potentielle Problemstellen benötigen. Dazu gehören Engstellen, Konflikte mit touristischem Radverkehr oder Baustellen. Dabei können die Daten sowohl bei kurzfristigen Maßnahmen als auch bei längerfristigen Planungen wertvolle Informationen liefern. Voraussetzung für Analysen auf der Punktebene ist eine möglichst gut ausgestattete PC-Infrastruktur, um die Menge an Daten mit akzeptablem Ressourcenaufwand verarbeiten zu können.

5.8.2 Analyse von Kreuzungsbereichen

Kreuzungsbereiche stellen sowohl in der Planung der Infrastruktur für Radfahrer als auch in der Routenplanung Herausforderungen dar. Zum einen soll der Bereich möglichst gefahrenlos und sicher vom Verkehrsteilnehmer überquert werden können, zum anderen können durch Ampelschaltungen und Wegeführungen über mehrere Teiletappen nicht zu unterschätzende Verzögerungen im Routenverlauf entstehen. Verkehrsmodelle berechnen diese Verzögerungen natürlich mit ein. Nichtsdestotrotz kann aus eigener Erfahrung berichtet werden, dass die erwartete Ankunftszeit schon einige Male, aufgrund von unglücklichen Ampelphasen, nicht planmäßig war.

So ist es natürlich naheliegend, ähnlich dem Ansatz zur Geschwindigkeitserhebung, auch die Zeit zum Überwinden eines Kreuzungsbereichs anhand der *Bike Citizens* Daten zu erheben. Auch hier hat die Analyse nicht den Anspruch, etwaige Modellberechnungen zu validieren

und neue Richtwerte für die Verkehrsplanung zu erheben, sondern es soll ein weiterer Anwendungsbereich der Daten aufgezeigt werden. Als Experimentierraum dient eine Kreuzung im 9. Wiener Gemeindebezirk in der Nähe des Wohnorts des Autors. Es ist der Knotenpunkt zwischen Währinger Straße (Südosten nach Nordwesten), der Nußdorfer Straße (Richtung Norden) und der Spitalgasse (Richtung Süden).



Abbildung 5.38: Kreuzungsübersicht Währinger Straße / Nußdorfer Straße / Spitalgasse

Wie in der Übersichtskarte ersichtlich, wurden die vier zu erhebenden Quellen und Ziele mit Punkten markiert, wobei der Punkt jeweils die Straßenmitte repräsentiert. Daraus ergeben sich insgesamt zwölf Relationen: A-B, A-C, A-D, B-C, B-D und C-D sowie jeweils viceversa.

An dieser Stelle soll noch kurz auf die Regeln der Straßenverkehrsordnung dieser Kreuzung eingegangen werden. Erwähnenswert ist hierbei, dass das links Abbiegen nur in der Relation A nach B gestattet ist. Alle anderen Linksabbiegevorgänge (C nach D, B nach C und D nach A) wären demnach ein Verstoß gegen die StVO¹⁰⁰.

Analog dem Modell zur Geschwindigkeitsberechnung sollte auch hier vorab der betreffende Teilbereich der track-points extrahiert werden. Das detaillierte Modell wird wiederum im Anhang erklärt.

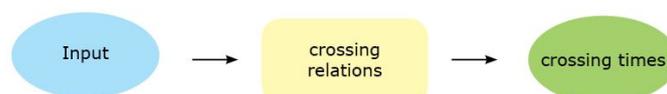


Abbildung 5.39: Grundsystematik des Modells zur Kreuzungsverzögerung. [Eigene Darstellung]

¹⁰⁰ Straßenverkehrsordnung

Die Berechnung ist aufgrund der notwendigen räumlichen Verschneidung der Einzugsgebiete der Referenzpunkte mit den track-points zeitaufwändig und ressourcenintensiv. Anhand des Ergebnisses lassen sich alle Relationen im Detail ansehen:

Tabelle 5.9: Vergleich der Kreuzungsrelationen. Kursiv dargestellt sind jeweils die „illegalen“ Verbindungen. (n=910) [Eigene Berechnung]

	AB	AC	AD	<i>BC</i>	BD	<i>CD</i>	BA	CA	<i>DA</i>	CB	DB	DC
Verteilung	8,0%	29,1%	0,8%	<i>1,1%</i>	10,4%	<i>0,5%</i>	5,8%	21,4%	<i>0,5%</i>	2,2%	16,1%	2,8%
Minimum	8	9	6	<i>7</i>	9	<i>11</i>	9	9	<i>9</i>	4	4	11
Maximum	90	75	35	<i>74</i>	111	<i>68</i>	154	169	<i>55</i>	101	74	49
Mittelwert	27,1	20,5	18,3	<i>38,2</i>	29,7	<i>37,8</i>	28,5	36,4	<i>31,8</i>	23,3	26,2	23,4
Median	16	17	18,5	<i>36</i>	24	<i>28</i>	22	28	<i>33</i>	15	17	19
t bei 15km/h	19,9	21,6	18,48	<i>18,01</i>	21,6	<i>20,4</i>	19,92	21,6	<i>18,48</i>	18,01	21,6	20,4

Ein erster Blick auf die Häufigkeitsverteilung zeigt, dass auch auf den vorher erwähnten, nicht gestatteten Linksabbiege-Relationen Fahrten gefunden wurden. Bei den Werten Minimum und Maximum ist die GPS-Fehleranfälligkeit zu berücksichtigen, daher sind diese Werte eher zu vernachlässigen. Wichtiger ist der Vergleich von Mittelwert und Median mit der geschätzten Fahrtdauer (jeweils in Sekunden) bei 15km/h. Hier zeigen sich im Vergleich große Unterschiede, wobei bei einem sehr geringen Anteil von keiner großen Validität ausgegangen werden kann.

Beispielhaft aufgezeigt wird hier die Häufigkeitsverteilung der Relation AC, die mit knapp 30% zu den am stärksten befahrenen gehört. Gemeinsam mit der Gegenrichtung CA kommt diese Verbindung auf mehr als die Hälfte aller Fahrten. Die berechnete Fahrzeit beträgt 21,6 Sekunden. Die Verteilung in absoluten Werten zeigt aber, dass ein Großteil der Fahrten deutlich schneller die Kreuzung überquert. Das Minimum liegt bei 8 Sekunden, das Maximum bei 90 Sekunden. Aufgrund der großen Streuung ist wiederum der Median an dem Mittelwert vorzuziehen. Die berechnete Zeit bei einer Geschwindigkeit von 15km/h zum Überqueren der Kreuzung beträgt 21,6 Sekunden.

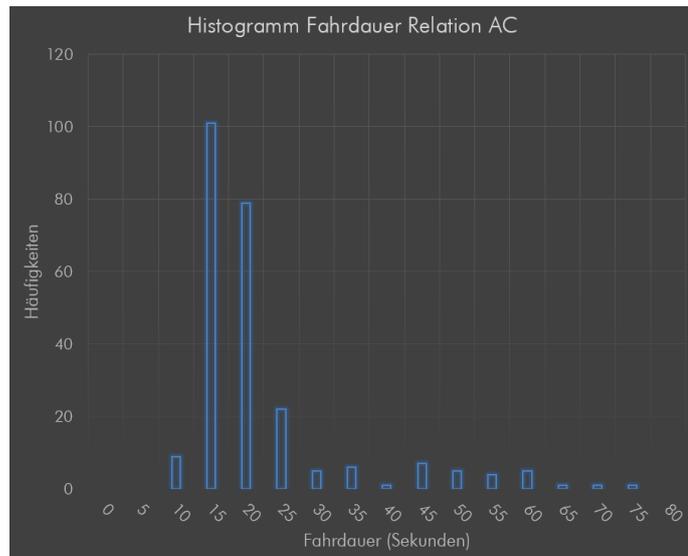


Abbildung 5.40: Häufigkeitsverteilung Relation AC [Eigene Darstellung]

Ein anderer Ansatz im Rahmen der Mikroanalyse versucht, die Punktdichte für die Berechnung von Widerstandsebenen hinsichtlich Routenberechnungen zu verwenden.

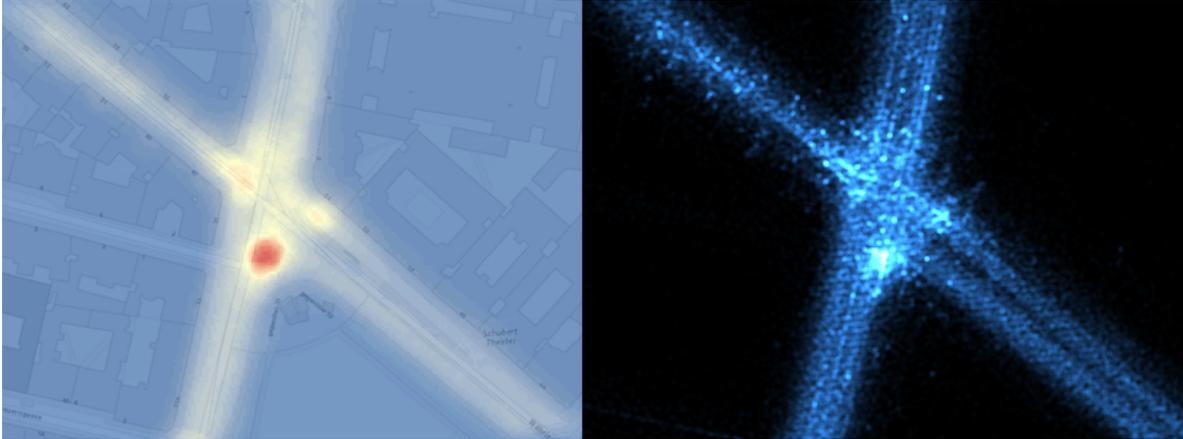


Abbildung 5.41: Darstellung der Punktdichte. Links berechnet mithilfe der Funktion Point Density in ArcGis mit einer Rastergröße von 1m, rechts die rein optische Überlagerung aller track-points. [Eigene Darstellung]

Die Idee dahinter folgt der Annahme, dass es dort wo kleinräumlich eine große Punktdichte auftritt (= sehr viele GPS-Points sehr nah beieinander sind) zu Verzögerungen kommt. Dies passiert durch die sekundliche Lokalisierung - bei langsamer Geschwindigkeit oder gar Stillstand (wie beim Warten an einer Kreuzung) werden folglich mehr Punkte in einem vordefinierten Bereich generiert. Über die Mengen der Punkte kann somit eine Widerstandsebene generiert werden, die bei großer Auflösung jede Relation individuell berücksichtigt.

Kurzfasit

Ähnlich der Geschwindigkeitsanalyse können auch mit der vorgestellten Methode zur Analyse von Kreuzungsbereichen Hindernisse identifiziert werden. Mit der Implementierung als Widerstandsebene können auch diese Ergebnisse Navigations- und Routingapplikationen verbessern.

5.9 Fazit und Diskussion

Kapitel 5 hat bisher ein großes Spektrum an Themenbereichen und Fragestellungen mit den *Bike Citizens* Daten in Relation gesetzt. Die beispielhaften Anwendungen haben Potentiale aber auch Hindernisse aufgezeigt. Im Folgenden sollen Aspekte zusammengefasst werden, die bei der Hands-On Analyse der Beispieldaten aufgetaucht sind, differenziert nach Defiziten und Potentialen.

5.9.1 Defizite

Potentielle Fehlinterpretation der Trackanzahl je Route

Die Daten bieten in der vorliegenden Form keine Möglichkeit zu erkennen, durch wieviele verschiedene Nutzer die einzelnen Tracks generiert werden, was eine mögliche Fehlinterpretation zur Folge hat. Es ist somit keine Zuordnung zu Individuen möglich. Aus datenschutzrechtlicher Sicht abgesichert - stellt diese Tatsache jedoch aus Anwendersicht ein großes Manko dar. Auch eine Unterscheidung zwischen mehreren Fahrten ein und derselben - ohnehin anonymisierten - Person ist nicht möglich. Das Phänomen wird vor allem auf der Meso-Ebene anhand der Analysen zum Routensplit und zur Erreichbarkeit deutlich. Individuelles Verhalten, beispielsweise durch überdurchschnittliche Nutzung der App, führt vor allem in Analysen über einen längeren Zeitraum potentiell zu erheblichen Verzerrungen. So könnte am Beispiel des Routensplits eine Route vom Auge des Betrachters der Analyse als besonders attraktiv bewertet werden, weil dort vergleichsweise viele Fahrten stattfinden. Die Daten geben aber keine Auskunft darüber, ob diese erhöhte Frequenz wirklich in der Qualität der Route oder aber doch durch das Abbild von individuellem Verhalten verursacht wurde. Die abgebildete Frequenz gibt keine Auskunft über die Zahl der Personen sondern nur über die Zahl der Fahrten. Einige wenige Personen, die unzählige Male am Tag aus individuellen Motiven eine Route wählen, hinterlassen beispielsweise den selben Eindruck, wie eine große Zahl an Nutzern, die jeweils bloß einmal pro Tag diese Route fährt. Daraus abgeleitete Aussagen, etwa zur Attraktivität einer Route, müssten in diesem Fall differenziert werden. Eine mögliche Unterscheidung wäre beispielsweise nach Expertenroute, die von wenigen Personen sehr häufig genutzt oder Allgemein angenommene Route, die von sehr vielen Personen befahren wird. Eine derartige Unterscheidung ist aber mit dem aktuellen Datensatz noch nicht möglich. In größerem Kontext, beispielsweise bei allgemeinen Auswertungen zum Radverkehr auf der Makroebene, zeigen sich derartige Verzerrungen nicht auf den ersten Blick. Da aber alle Ebenen zueinander in Wechselwirkungen stehen, ist davon auszugehen, dass auch hier eine Verzerrung - wenngleich diese nicht so leicht deutlich gemacht werden kann - vorhanden ist. Für die Verknüpfung von Wegetappen ist die Zuordnung einzelner Tracks zu anonymisierten Nutzern unabdingbar.

Nutzerverhalten oder Algorithmus

Das Alleinstellungsmerkmal der *Bike Citizens* App ist ein Routingsystem, welches auf dem gesammelten KnowHow von Fahrradkurieren basiert. Laut Auskunft seitens *Bike Citizens*¹⁰¹ werden rund zwei Drittel aller Fahrten mit aktiviertem Routing durchgeführt - ein Drittel trackt die Fahrt ohne Routing-Unterstützung. Die Fahrt verläuft bei ersterem nach der Berechnung des Routing-Algorithmus und nicht nach den räumlichen Kenntnissen des Nutzers. Eine klare Unterscheidung zwischen Fahrten, bei denen das Routing Service nicht genutzt oder ignoriert wird und Fahrten, die den Routing Anweisungen folgen, wäre wünschenswert - zumal diese Unterscheidung bei *Bike Citizens* intern bereits seit November 2015 aufgezeichnet wird. Ohne diese Unterscheidung, ob der Nutzer geleitet wird oder auf Basis eigener Entscheidungen handelt, sind alle Aussagen zum Nutzerverhalten wie Umwegebereitschaft oder Routenattraktivität nur schwer bis kaum möglich.

Spuren von Unbekannten

Die Daten verfügen über keinerlei ergänzende personenbezogene Informationen. Auch Rückschlüsse auf die Anzahl der (aktiven) Nutzer sind nicht möglich. Dies hat einerseits die im Absatz der Fehlinterpretation der Frequenzen dargestellten Folgen und zusätzlich können keine Aussagen zur sozio-demographischen Evidenz getroffen werden. Der Nutzer ist unbekannt - nur seine räumliche Spur bleibt.

Map Matching

Für tiefergehende Analysen und Entwicklung von netzwerkbasierten Anwendungen auf Basis der Daten ist ein Map Matching Verfahren unbedingt notwendig. Der im Rahmen dieser Diplomarbeit umgesetzte Versuch, das Map Matching selbst zu implementieren, hat gezeigt, dass dies in keiner Relation zum benötigten Aufwand steht. Daher muss hier die Empfehlung wiederholt werden, dass die Daten schon vom jeweiligen Anbieter und Erheber automatisiert auf ein Netz übertragen werden und in dieser Form für dem Anwender zur Verfügung gestellt werden. Die kantenbasierten Daten sollen dabei aber keinesfalls die Tracking-Rohdaten ersetzen sondern diese nur ergänzen.

¹⁰¹[BikeCitizens, 2017]

5.9.2 Potentiale

Datenvielfalt

Die Digitalisierung und alle Entwicklungen rund um *Big Data* führen zu einer immer größeren Datenvielfalt. Es ist davon auszugehen, dass dieser Trend durch weitere *Open Data* Bemühungen noch länger bestehen wird. Diese neue Welt der Daten lässt sich jederzeit mit den *Bike Citizens* Tracks in Relation setzen. Alleine am Beispiel der Wetterdaten hätten alle vorgestellten Beispiele noch um etliche Attribute erweitert werden können. Wie sieht der Routensplit an Regentagen aus? Verändert sich die Erreichbarkeit der TU Wien mit dem Fahrrad bei einer Windgeschwindigkeit von über 10 km/h wesentlich? Wirkt sich Schneefall auf die Verzögerung in Kreuzungsbereichen aus? Die möglichen Fragestellungen sind nicht enden wollend. Durch die Erweiterung um inhaltliche Ebenen können noch mehr Kombinationen erforscht werden. Zudem können die Tracks damit auf die jeweilige Forschungsfrage hin inhaltlich erweitert werden - auch im Nachhinein. Diese Tatsache stellt eine der größten Stärken der Daten dar.

Betrachtungsdauer

Ähnlich wie die Datenvielfalt bietet auch der lange Betrachtungszeitraum große Möglichkeiten in der Adaption von Analysen auf eine konkrete Fragestellung. Durch den Zeitstempel sind auch Analysen für bestimmte Tageszeiten - beispielsweise auch im Bereich des nächtlichen Radverkehrs, ohne Probleme möglich. Auch Jahreszeiten, Wochenenden oder Feiertage sind selektierbar. Über eine entsprechende Schnittstelle wäre technisch auch eine (beinahe) Echtzeitauswertung möglich. Dies ist natürlich besonders spannend um (kurzfristige) Veränderungen im Raum in materieller Hinsicht auf ihre räumlich-relationalen Auswirkungen zu untersuchen.

Messung vs. Schätzung

Die Systematik des Vergleichs zwischen gemessenen und geschätzten Werten ist aus Sicht der Regionalwissenschaften eine spannende Materie. Die *Bike Citizens* Daten bieten dabei einen riesigen Datensatz an gemessenen Daten, welche - Map Matching Verfahren vorausgesetzt - geschätzten und berechneten Werten gegenübergestellt werden können. Dadurch könnten die Ergebnisse aus Forschung und Planung ein Stück realitätsnäher werden. Zudem bietet sich die Erweiterung bestehender Simulationsmodelle mit gemessenen/beobachteten Daten an.

Weiterentwicklung der Bike Citizens Daten

Durch die vorgestellten Inhalte beim erfolgreichen Abschlussmeeting im Rahmen dieser Diplomarbeit bei *Bike Citizens* in Graz konnte ein Einblick in bereits bestehende Pläne zur Weiterentwicklung der Daten seitens des Anbieters gewonnen werden. Die genannten Punkte im Fazit treffen also alle auf den aktuell für die Analyse vorliegenden Datensatz zu, zukünftige Analysen können aber eventuell schon von den geplanten Verbesserungen profitieren.

So ist *Bike Citizens* selbst interessiert, den Datensatz in möglichst hoher Qualität und möglichst vielfältigen Anwendungsmöglichkeiten an die Stadtverwaltungen zu vermarkten. Dazu werden in nächster Zeit einige Änderungen und neue Features realisiert.

Zum einen wird versucht, eine Zielgruppenunterscheidung und Kategorisierung der Nutzer in Profile voranzutreiben. Dies beruht einerseits durch automatische Detektion von Verhaltensweisen - beispielsweise durch tägliches sehr flottes Fahren wird ein Nutzer zum Expertenradler eingestuft - und andererseits durch kurze Befragungen innerhalb der App abgesichert. Die Daten könnten zukünftig dem Punkt entgegenwirken, dass über die Nutzer keine Informationen vorhanden sind.

Wie schon erwähnt ist die interne Unterscheidung, ob der Nutzer das Routing nutzt oder nur seine Fahrt trackt, bereits möglich. Dieses Attribut soll auch bei zukünftigen Datensätzen für Analysen mitgeliefert werden. Außerdem sollen die Anreize zum Tracken von Fahrten weiter ausgebaut werden, was aus Anwendersicht begrüßt wird. Je größer der Anteil der Tracks ohne Routing, desto mehr Anwendungsmöglichkeiten ergeben sich in der Analyse.

Eine Zuweisung einzelner Fahrten zu einzelnen Nutzern ist zwar bisher im Hintergrund schon möglich, die Informationen werden aufgrund datenschutzrechtlicher Bedenken und Unsicherheiten aber nicht herausgegeben oder verwendet. Nach rechtlicher Prüfung wäre aber laut *Bike Citizens*¹⁰² die Herausgabe einer tageweise anonymisierten Form möglich. Damit könnten Wegeketten analysiert werden - eine weitere spannende Frage in der Stadt- und Regionalforschung.

¹⁰²[BikeCitizens, 2017]

6 Tracking-Daten in der Stadt- und Regionalforschung - Resümee

6.1 Vergleich der Datencharakteristik

Als Weiterentwicklung der im Kapitel 5 gezeigten Inhalte wird nun der Fokus auf einen Vergleich gelegt. Die Hands-On Analyse und vor allem deren Fazit ließ die Vermutung aufkommen, dass die *Bike Citizens* Daten einer anderen Logik folgen, als jene, bei den in Kapitel 4 gezeigten Konzepte zu gezielten Mobilitätserhebungen. Ein Vergleich nach charakteristischen Aspekten soll überprüfen ob diese Vermutung zutrifft und worin die Unterschiede gegebenenfalls bestehen.

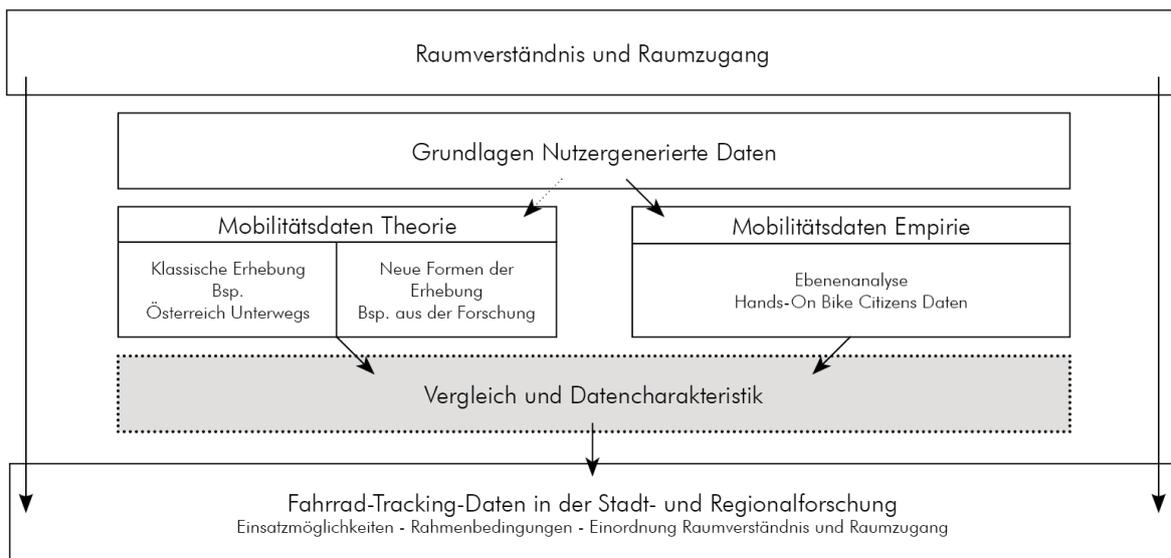


Abbildung 6.1: Aufbau der Arbeit. Der Vergleich ist der letzte Schritt vor der Beantwortung der Forschungsfragen im letzten Block. [Eigene Darstellung]

Dazu wird ein Vergleich gezogen zwischen:

- der im Jahr 2014 durchgeführten Mobilitätserhebung *Österreich Unterwegs* als State of the Art Beispiel,
- *SmartSurvey*, als Projekt einer digitalen Mobilitätserhebung,
- und den Beispieldaten von *Bike Citizens*.

Der Vergleich erfolgt dabei anhand von Merkmalen, die sich im Zuge der Bearbeitung der Kapitel 4 und vor allem 5 als wichtig und wesentlich herausgestellt haben. Es handelt sich dabei um eine persönliche Einschätzung der drei verschiedenen Beispielen und folgt einer dreistufigen Skalenlogik: „+“ für Zutreffend, „o“ für teilweise zutreffend und „-“ für nicht zutreffend.

Tabelle 6.1: Vergleich der Mobilitätsdatenerhebungen nach charakteristischen Merkmalen. [Eigene Darstellung und Beurteilung]

	Österreich U.	SmartSurvey	Bike Citizens
gesicherte Repräsentativität	+	+	-
Demographische Merkmale bekannt	+	+	-
Querschnitt- / Längsschnittdaten	Quer	eher Quer	Längs
Erweiterung hinsichtl. Datenvielfalt	o	+	+
Datengüte	+	+	-
Zweck ist Mobilitätserhebung	+	+	-
Verräumlichung von Verhalten	-	+	+
mögl. Datenfehler durch Nutzer	+	o	-

Anhand dieser Übersicht wird erkannt: Die drei Konzepte weisen in keinem Merkmal eine durchgehende Überschneidung auf, naturgemäß gibt es aber starke Ähnlichkeiten zwischen *Österreich unterwegs* und *SmartSurvey*. Das Erhebungskonzept der *Bike Citizens* Daten hingegen zeigt nur in wenigen Punkten Überschneidungen. Ein grundlegender Unterschied besteht alleine schon im Zweck der Datenerfassung: Bei *Österreich Unterwegs* sowie auch *SmartSurvey* sind Ziel und Zweck der Aktivität, eine Mobilitätsdatenerhebung unter repräsentativen Rahmenbedingungen durchzuführen. Die Teilnehmer werden dazu entweder, wie im Fall von *Österreich Unterwegs*, direkt befragt oder bei *SmartSurvey* ausgewählt. Die App dient bei letzterem Konzept aber nur der Dokumentation von Wegen. Zudem sind bei *SmartSurvey* die demographischen Merkmale der Nutzer bekannt - wenn auch anonymisiert. Die Nutzer der *Bike Citizens* App benutzen diese, um mit dem Fahrrad den besten Weg von A nach B zu finden. Die Aufzeichnung findet passiv und nicht intentional statt. Die Daten von *Bike Citizens* entstehen durch Unbekannte, demographische Merkmale sind - zurzeit - noch keine bekannt. Ein großer Unterschied zu *Österreich Unterwegs* besteht darin, dass bei der Erhebung nur Wegehäufigkeiten und Wegelängen erhoben werden - differenziert nach unterschiedlichen Verkehrsmittel und unterteilt nach vordefinierten Aktivitätskategorien. Bei den Fahrrad-Tracking-Daten liegen die Daten (vorerst) nicht in aggregierter Form da und bieten somit die Möglichkeit eines detaillierten Einblicks.

Datenfehler durch den Nutzer - wie im Theorieteil der Arbeit definiert, beispielsweise durch Überschätzung von Wegen bei *Österreich Unterwegs* oder dem vergessenen Aktivieren der App beim außer Haus gehen bei *SmartSurvey* - können bei *Bike Citizens* nicht passieren, da die App auch keinen Erhebungsanspruch stellt. Es werden jene Wege verwertet, die aufgezeichnet wurden, zu welchem Zweck und wie oft der Nutzer dies macht, ist vorerst zweitrangig. Natürlich verfolgt *Bike Citizens* das Ziel, die Nutzer zur Aufzeichnung möglichst vieler Fahrten zu animieren - aber in einem anderen Verständnis als *Österreich Unterwegs* und *SmartSurvey*. Zudem ist das Konzept der passiven Beobachtung realen Verhaltens bei *Bike Citizens* am klarsten ausgeprägt.

Im Merkmal der Betrachtungsweise zeigt das Konzept von *Bike Citizens* den offensichtlichsten Ansatz der Querschnittsanalyse. Bei *SmartSurvey* hängt die Interpretation von der Projekt-

dauer ab. Aus eigener Erfahrung kann von einem ähnlichen Projekt (Semaphore) berichtet werden, bei dem über einen Zeitraum von 2 Wochen alle Wege über mein Smartphone automatisiert aufgezeichnet wurden. Im Vergleich mit der Datengenerierung bei *Bike Citizens* über mehrere Monate ist dies ein sehr kurzer Zeitraum, wenngleich bei der Frage nach einer Einordnung hinsichtlich Quer- oder Längsschnittanalyse eher zweites zutreffen würde. Es ist aber wichtig herauszustreichen, dass sich die ununterbrochene Datengenerierung bei *Bike Citizens* deutlich von den anderen beiden Konzepten unterscheidet.

Diese gravierenden Unterschiede verlangen nach einer differenzierten Betrachtungsweise der *Bike Citizens* Daten. Dies unterstreicht auch Martin Berger¹⁰³, der nach einer eigenen Sichtweise auf „Tracking Daten“ verlangt und eine klare Differenzierung zu echten Mobilitätserhebungen vorschlägt. Die gemeinsame Bezeichnung und Betrachtung ist mit allen Erkenntnissen dieses Vergleichs nicht mehr zielführend und zu oberflächlich. Passender dafür ist die Bezeichnung (Fahrrad-), „**Tracking Daten**“ - denn es handelt sich ausschließlich um getrackte Wege. Repräsentativität und demographische Merkmale haben dabei keinerlei Bedeutung - die Stärken der Daten liegen in anderen - bereits dargelegten - Aspekten.

6.2 Bewertung der Einsatzmöglichkeiten von Tracking-Daten

1. Welche Einsatzmöglichkeiten von Fahrrad-Tracking-Daten in der Stadt- und Regionalforschung können aus den Hands-On Beispielen abgeleitet werden?

Ausgehend von den Hands-On Ergebnissen der Ebenenbetrachtung werden hier die Einsatzmöglichkeiten der Fahrrad-Tracking-Daten in der Stadt- und Regionalforschung bewertet. Auch auf multiskalare Effekte soll im Zuge der Beantwortung der Forschungsfrage hingewiesen werden.

Makro-Ebene

Die Anwendungen der Makro-Ebene haben durchaus brauchbare Ergebnisse geliefert. Mithilfe der Datenvisualisierung können zur jeweiligen Fragestellung maßgeschneiderte Darstellungen erstellt werden. Diese dienen dabei sowohl dem Anwender zur ersten Vorabanalyse als auch als Kommunikationsmittel für Probleme oder Ergebnisse. Auch die Möglichkeit bestimmte - ebenso kurzzeitig auftretende - räumliche Phänomene abzubilden ist aufgrund des großen Längsschnitts möglich und auf alle anderen Ebenen übertragbar.

Diese lange Betrachtungsdauer ist eine der innovativsten Aspekte hinsichtlich der Einsatzmöglichkeiten in der Stadt- und Regionalforschung. So war es bisher nicht möglich - mit vertretbarem Aufwand - individuelles, tatsächlich (beobachtetes) Verhalten in größerem Umfang und über einen langen Zeitraum zu abbilden. Die Tracks stellen zudem eine sekundengenaue Erhebung dar, was die zeitlich unterschiedliche Nutzung von Infrastrukturen und Räumen visualisierbar sowie auch analysierbar macht. Das Beispiel in Abbildung 6.2 zeigt etwa alle Fahrten in den Sommermonaten Juni bis August zwischen 01:00 Uhr und 05:00 Uhr nachts,

¹⁰³vgl. [Berger, 2017]

wofür keine alternative Datenquelle bisher vorhanden war. Die Datenauswahl kann somit hinsichtlich der jeweilig benötigten Spezifika in der Mobilitätsforschung adaptiert werden - sowohl zeitlich als auch nach Distanz, wie das Beispiel der Radrouten zeigte.

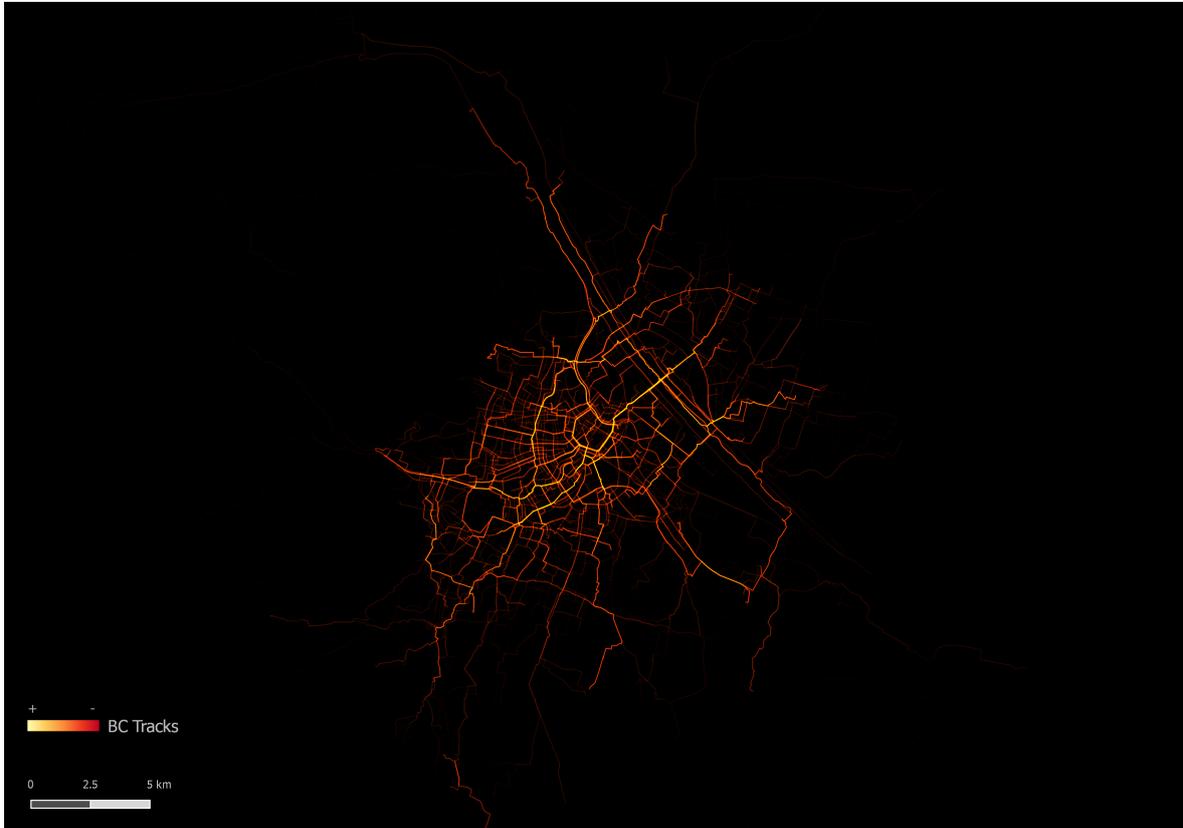


Abbildung 6.2: Alle Tracks in den Sommermonaten Juni bis August 2016 mit einer Startzeit zwischen 01:00 und 05:00 Uhr. Nach dem relationalen Raumzugang wurde auf die Darstellung von Verwaltungsgrenzen verzichtet. [Eigene Darstellung / Bike Citizens]

Die Erweiterungsmöglichkeit mit anderen Inhalten stellt ebenfalls eine spannende Möglichkeit dar, um geographische verortete Gegebenheiten - wie beispielsweise Wetterdaten - mit räumlichen Relationen verschneiden zu können. Die vielfältigen Adaptierungsmöglichkeiten werden dadurch nochmals unterstrichen. Durch die Verknüpfung mit anderen Daten, wie gezeigt am Beispiel mit Wetterdaten, entstehen auch Möglichkeiten für Hypothesentests und weitere statistische Auswertungen.

Mithilfe der Visualisierungen gelingt zwar die Identifikation von Bereichen mit hochfrequenter Nutzung, die Analyse von konkreten Quelle-Ziel-Mustern ist aber nicht möglich da kein Bezug zwischen quellsseitiger Verkehrsnachfrage und zielseitigen Standortfaktoren hergestellt werden kann.

Meso-Ebene

Die Analyse und Ergebnisse zum Routensplit haben gezeigt, dass evidenzbasierte Aussagen zu Nutzerverhalten ohne ausreichend große Stichprobe und der Kennzeichnung von genutzten Routing-Algorithmen nicht möglich sind. Auch die Anzahl unterscheidbarer und unterschiedlicher Nutzer, auf denen die jeweilige Frequenz im Analyseausschnitt basiert, ist für Aussagen zur Akzeptanz von Routen oder der anteiligen Verteilung auf verschiedene Strecken nötig. Für tieferegehende Analysen müssen Subjekte identifiziert werden können, nicht nur einzelne Fahrten.

Die Einzugsbereichsanalyse - so wie auch die Verkehrsspinne als Spezialform - zeigt eine der klassischen Einsatzmöglichkeiten in den Themenfeldern der Erreichbarkeits- und Mobilitätsforschung auf. Vor allem die detailreichen Möglichkeiten in der Datenauswertung lassen sehr spezielle Fragestellungen und Analysen zu. Auch der Aspekt, dass dabei reale (beobachtete) Fahrzeiten herangezogen werden, soll nochmals betont werden. Für evidenzbasierte Ergebnisse müssen aber die im Absatz zuvor beschriebenen Defizite gleichermaßen berücksichtigt werden.

Anwendungsmöglichkeiten bieten - sobald Subjektinformationen verfügbar sind - sich auch im Bereich der Analyse von Mobilitätsmustern. Hier ist allen voran die Erforschung von Wegeketten hervorzuheben - ein seit jeher spannender Aspekt in der Mobilitätsforschung.

Mikro-Ebene

Sowohl die Geschwindigkeitsanalyse als auch die Analyse von Kreuzungssituationen haben spannende Einsatzmöglichkeiten aufgezeigt, die ohne konzeptuelle Einschränkungen durchgeführt werden können. Die Ausweitung dieser Analyse um mehrere Kreuzungssituationen - auch verschiedenen Typs - miteinzubeziehen bedarf aber entsprechender technischer Ressourcen, da die Informationsdichte auf Punktebene enorm ist. Eine Adaption dieser Prozesse hinsichtlich Analysen im Zusammenhang mit der Topographie würde eine weitere Anwendungsmöglichkeit aufzeigen. Damit könnten modellbasierte Berechnungen zum Zusammenhang von Steigungen und Geschwindigkeiten im Radverkehr realisiert werden. Die Ergebnisse der Mikro-Ebene können als Basis von graphenbasierten Berechnungen dienen - Map Matching vorausgesetzt. Dies würde ebenfalls räumliche Simulationen und netzwerkanalytische Prozesse verbessern, da auch dort reales (beobachtetes) Verhalten bisher nicht in dieser Form als Basis angewandt wird.

6.3 Rahmenbedingungen für den Einsatz Tracking-Daten

2. Welche Eigenheiten und Rahmenbedingungen müssen beim Einsatz der Fahrrad-Tracking-Daten in der Stadt- und Regionalforschung beachtet werden?

Der Einsatz von Tracking-Daten als Basis mobilitätsbezogener Fragen ist mit etlichen nachteiligen Aspekten, aber auch mit einzigartigen, innovativen und vorteilhaften Blickpunkten verknüpft. Deswegen ist es von Bedeutung, gewisse Rahmenbedingungen beim Einsatz der Daten zu beachten. Für einige Fragestellungen wird es nötig sein, weitere Daten zur Absicherung der Evidenz zu verwenden.

- Die **datenschutzrechtlichen Rahmenbedingungen** sind zu beachten. Dabei sollte besonders auf die transparente und für die Nutzer klare Generierung und Verwendung der Daten Wert gelegt werden. Durch Bewusstseins-schaffung, dass die Daten für gesellschaftlich relevante Projekte eingesetzt werden, soll die Akzeptanz der Datenverwendung steigen. Auch wenn die Umsetzung dieser Empfehlung auf Anbieterseite erfolgen muss, so sollte die Anwenderseite als Kontrollinstanz die datenschutzrechtlichen Rahmenbedingungen überprüfen.

Datenschutz ist wichtig. Wenn zukünftig innovative Projekte jedoch nur an datenschutzrechtlichen Bedenken scheitern, wird es für die Stadt- und Regionalforschung schwierig, immer am Puls der Zeit zu sein. Deswegen ist eine Abwägung zwischen Daten schützen und Daten nutzen unumgänglich. Mit klaren und transparenten Statuten sollte einer Verwendung nichts im Wege stehen - natürlich innerhalb des rechtlichen Rahmens.

- **Sozio-demographische** Aspekte werden mit Tracking-Daten auch zukünftig schwierig zu berücksichtigen sein. Auch wenn - wie erwähnt - *Bike Citizens* an einer Profilierung der Nutzer arbeitet, so kann diese wahrscheinlich nicht das Niveau von anderen zum Zweck der Erhebung ausgeführten Befragungen und Analysen erreichen. Beim Einsatz der Daten sollte dies im Bewusstsein des Anwenders sein.
- Die **Repräsentativität** der Daten sollte für jede Datenbezugsquelle mit den vorhandenen Möglichkeiten individuell überprüft werden. Für die Aussagekraft ist nicht nur der Stichprobenumfang von Bedeutung, auch eine akzeptable Verteilung im Raum sollte stets angestrebt werden. Zudem kann bei Tracking-Daten - zumindest zur Zeit - von keiner Zufallsstichprobe ausgegangen werden, da die Verwendung einer spezifischen App keines Falls gesellschaftlich gleichverteilt ist.
- Rückschlüsse von einzelner beobachtetem **Nutzerverhalten** auf die Grundgesamtheit sind mit Tracking-Daten sind nur mit weiteren Subjektinformationen möglich. Ohne diese Informationen kann keine Unterscheidung gemacht werden, ob die beobachteten Fahrten von einem Nutzer oder von x-Nutzern durchgeführt wurden. Dadurch entsteht die Gefahr einer Ergebnisverzerrung durch individuelles Verhalten. Eine Differenzierung nach personenbezogener Frequenz ist also für derartige Analysen erforderlich.

6.4 Tracking-Daten aus raumtheoretischer Perspektive

3. Wie können die Fahrrad-Tracking-Daten aus unterschiedlichen raumtheoretischen Perspektiven von Nutzen sein?

Physisch-materieller / geographischer Raum

Während bei *Österreich Unterwegs* durch die Erhebung „Informationen aus dem *Häferl* abgefragt“ werden und somit der Typ des physisch-materiellen / geographischen Raums zum Einsatz kommt, spielt dieser bei den Tracking-Daten eine eher untergeordnete Rolle.

Angesprochen wurden aber bereits am Beginn der Arbeit die Wechselwirkungen der räumlichen Zugänge, so bestimmen genaue Voraussetzungen und Gegebenheiten andere Ebenen und umgekehrt. So sind die aufgezeichneten Wege - also die Relationen - in ihrer räumlichen Ausprägung stark von den morphologischen Gegebenheiten der Stadt - beispielsweise der Gebäudesubstanz - abhängig. Genauso ist eine Rückkopplung von den Tracks auf die Morphologie möglich. Wenn im relationalen Raumverständnis die Relationen entfernt werden, dann bleibt - wie schon erwähnt - nichts übrig. Werden diese Wege nun als befahrbare Trassen im physisch-materiellen Zugang verstanden und anschließend vom „Raum“ subtrahiert, so bleibt die Morphologie übrig.

Ergebnis ist ein ähnliches Bild, das Planer bereits unter dem Begriff „Schwarzplan“ kennen. Zwar stellt dieser in seiner originalen Bedeutung die Baumasse einer Stadt dar, so zeigt die folgende Darstellung auch Freiflächen - wie Gewässer oder weitläufige Parks sowie Areale, die sich für eine Fahrradnutzung nicht eignen (wie Eisenbahnareale oder Betriebsgebiete).

Für die korrekte Deutung der folgenden Grafik ist es wichtig, die dunklen/schwarzen Flächen zu fokussieren:



Abbildung 6.3: Der Schwarzplan offenbart die Morphologie der Stadt - dargestellt anhand der dunklen Bereiche.
[Eigene Darstellung / Bike Citizens]

Der physisch-materielle Raumtyp wurde bisher in der Mobilitätsforschung auch für relationale Beziehungen, wie beispielsweise Pendlerdaten, genutzt. Die Pendlermatrix beschreibt dabei, wie viele Personen von einem Behältnis (=Erhebungs- oder Verwaltungseinheit) zu einem anderen fahren. Das Ergebnis ist die Interaktionshäufigkeit zwischen Behältnissen. Zwar wurde versucht, durch immer kleinere Behältnisse - bis hin zu Adresspunkten - eine Annäherung an den relationalen Raumzugang zu schaffen, aber eine technische Möglichkeit Informationen relational zu erfassen war bisher nicht vorhanden. Die kleinräumigen Behältnisse bieten darüber hinaus aber zukünftig eine interessante Datenbasis, die mit relationalen Tracking-Informationen verschneiden werden kann. Vor allem bei der Analyse quell- und zieleitiger Standortfaktoren - wie beispielsweise die Anzahl der Fahrräder pro Haus oder die Anzahl der Fahrradabstellplätze am Point of Interest - können Informationen aus dem physisch-materiellen Raum fehlende Details der Tracking-Daten ergänzen. Die multiskalare Verschneidung wird auch zukünftig ein spannender Aspekt der Regionalwissenschaften sein. Bei der Verschneidung muss aber auch weiterhin mit Unschärfen gerechnet werden. Datenschutzrechtliche Aspekte oder auch technische Phänomene wie ungenaue GPS-Verbindungen werden die Anknüpfung von relationalen Informationen an Behältnisse erschweren.

Relationaler Raum

Wie schon in der Einleitung der Arbeit angedeutet, kommt dem Zugang des Relationalen Raumes in Zusammenhang mit Tracking-Daten eine große Bedeutung zu. Der *Bike Citizens* Datensatz bildet in seiner inhaltlichen Gestalt grundsätzlich Relationen ab. Wo diese Verbindungen verlaufen, ist dabei in erster Linie abhängig von der räumlichen Anordnung von Quelle und Ziel jeder Fahrt. Mithilfe der Visualisierung der Tracking-Daten ist es möglich, diese Beziehungen zu veranschaulichen. Auch Zeitaufwände oder Erreichbarkeitsanalysen sind Teil dieses Verständnisses.

Die einzelnen Fahrten sind aber nicht nur von der Relation Quelle - Ziel beeinflusst, auch die Anordnung attraktiver Fahrradinfrastruktur spielt eine Rolle, was wiederum eine Wechselwirkung zum physisch-materiellen Raum darstellt.

Eine Verschneidung des relationalen Zugangs mit Elementen aus dem geographischen Raum passiert, wenn klar verortbare Gegebenheiten mit den Daten verknüpft werden. Ein Beispiel für eine derartige Verschneidung wären Wetterphänomene oder topographische Eigenschaften wie Steigung oder Gefälle, die auch wiederum Einfluss auf die relationale Ebene haben.

Grenzen - ob in baulicher oder in Verwaltungshinsicht - existieren in der Darstellung des relationalen Raumes nicht. Vor allem die Unbedeutsamkeit von Verwaltungsgrenzen in der Realität und dem täglichen Leben kann mithilfe der Tracking-Daten visualisiert werden. In baulicher Hinsicht sind die Fahrten zwar von den morphologischen Gegebenheiten abhängig - es kann nur dort gefahren werden wo auch ein Weg ist - in einem relationalen Raumverständnis ist diese Tatsache aber zu vernachlässigen. Schließlich könnten die einzelnen Relationen auch nur über eine einfache, abstrakt-lineare Verbindung von Quelle und Ziel dargestellt werden, die keine Rücksicht auf die Bausubstanz nimmt.

Der spannende Aspekt der Tracking-Daten liegt in der Tatsache, dass im Gegensatz zu vielen modellbasierten Ergebnissen tatsächlich existierende Relationen abgebildet werden können, da die Daten direkt vom Nutzer generiert wurden.

Sozialer Raum

Der Zugang des sozialen Raums, ein durch Handlungspraxis konstituiertes und gesellschafts- bzw. entwicklungsbedingtes soziales Konstrukt, spielt auch in Zusammenhang mit den Tracking-Daten eine große Rolle. Seine stark ausgeprägte Wechselwirkung mit dem relationalen Raumverständnis zeigte sich bereits: sind die visualisierten Daten doch eine Abbildung von sozialer Interaktion im Raum. Das Ergebnis ist ein sozialräumliches Abbild.

Aber auch normativ-politische Regionalisierungen beeinflussen die abgebildeten Relationen. Gesetze können als bestimmte Norm verstanden werden, die von der jeweiligen Politik abhängig ist. Die Inhalte dieser beeinflussen das Verhalten im Raum, was wiederum mit den Tracking-Daten visualisiert werden kann. Einbahnen oder Fahrverbote sind hier plakative Beispiele für eine normativ-politische Beeinflussung des Raums im Kontext der Tracking-Daten, genauso wie weitere Ver- und Gebote aus den Alltagswelten. Aber auch die generelle Einstellung der jeweiligen Politik gegenüber dem Radverkehr beeinflusst die Tracking-Daten indirekt

- durch radverkehrsfördernde Maßnahmen kann beispielsweise die Gesamtzahl an Fahrrad-Relationen gesteigert werden.

Aus Sicht der Mobilitätsforschung sind im Zusammenhang mit dem relationalen Raum auch Fragen nach Motiven der Routenwahl spannend, primär hinsichtlich der Attraktivität und Wohlfühlfaktoren von Räumen. Das individuelle Empfinden und persönliche Wahrnehmungen von verschiedenen räumlichen Ausstattungsmerkmalen stellt in der Forschung ein spannendes Feld dar, zu dem es bisher nur wenige relevante Datengrundlagen gibt. Sowohl der physisch-materielle als auch der relationale Zugang können auf diese Fragen keine Antworten liefern. Die Tracking-Daten müssten aber mit weiteren - eventuell über App-interne Befragungen generierten - Attributen angereichert werden, um auch qualitative Aspekte in der Analyse verschneiden zu können. Gesellschaftlich relevante räumliche Wahrnehmungen - als mögliches Ergebnis einer derartigen Analyse - können zudem eine äußerst hilfreiche Grundlage für zielgruppenspezifische Planungen sein. Die Ebene des sozialen Raums bietet Möglichkeit, die Tracking-Daten mit vielfältigen qualitativen Informationen anzureichern.

6.5 Ausblick

Tracking-Daten stellen für die Regionalwissenschaft eine zukunftsweisende und innovative Materie dar, die vielen Projekten neue und spannende Aspekte bringen kann. Die Arbeit zeigt die vielseitigen Einsatzmöglichkeiten auf und weist auf wichtige Rahmenbedingungen hin. Diese sollten auch bei zukünftigen Projekten berücksichtigt werden. Die angesprochene gesellschaftliche Entwicklung in Zusammenhang mit der Digitalisierung ist schwer zu prognostizieren. Viele Anzeichen sprechen aber dafür, in Zukunft noch weitere Daten von diesem Typus zur Verfügung zu haben. Allgemein kann mit den Tracking-Daten ein beginnender Trend beschrieben werden, der weg von modellbasierten Hochrechnungen hin zu gemessenen und real beobachteten Datenbeständen führt. Derartige Möglichkeiten, Verhalten anhand großer Stichproben zu dokumentieren waren bisher nicht vorhanden. Zusammen mit der starken räumlichen Ausprägung - verursacht durch Smartphones und andere mobile Sensoren - kann diese Entwicklung die Mobilitätsforschung nachhaltig verändern und bei zweckmäßigem Dateneinsatz auch verbessern. Die in der Arbeit angesprochenen Rahmenbedingungen und Defizite, vor allem hinsichtlich Aussagekraft, sollen aber auch zukünftig beachtet werden. Auch der Wunsch, aus *Big Data* sollte *Big Information* werden, wird zukünftig an Bedeutung gewinnen. Vor dem Hintergrund der weiter zunehmenden Menge an Daten ist es wichtig, die notwendigen Informationen zur Zielerreichung auch in den Datensätzen vorzufinden. Vor allem aus Anwendersicht kann hier deshalb der Wunsch geäußert werden, dass einige wichtige Details für Auswertungen bereits beim Anbieter im Zuge der Datenaufbereitung organisiert werden.

Die Analyse der Daten im raumtheoretischen Kontext hat gezeigt, dass es je nach Zugang innerhalb der Raumplanungswelt Anknüpfungspunkte gibt. Die Tracking-Daten stellen dabei ein Art Bindeglied dar, um vorhandenen Informationen eine neue Dimension zu verleihen.

Die vorliegende Diplomarbeit bietet in vielen Bereichen Anknüpfungspunkte für weitergehende vertiefende Analysen, deren Umsetzung zu empfehlen ist. Je größer das Verständnis des Datentyps ist, desto besser werden auch die Ergebnisse in der Stadt- und Regionalforschung sein.

7 Literaturverzeichnis

- [Alexa 2016] ALEXA: *The top 500 sites on the web*. <http://www.alexa.com/topsites>, 2016. – (aufgerufen am 13.12.2016)
- [Berger 2017] BERGER, Martin: *Experteninterview in Wien am 17.1.2017*. 2017
- [Bierlaire u. a. 2013] BIERLAIRE, Michel ; CHEN, Jingmin ; NEWMAN, Jeffrey: A probabilistic map matching method for smartphone GPS data. In: *Transportation Research Part C: Emerging Technologies* 26 (2013), S. 78–98
- [BikeCitizens 2017] BIKECITIZENS: *Persönliches Meeting mit Daniel Kofler, Thomas Rath und Steffen Wirth von Bike Citizens sowie Robert Kalasek und Aggelos Soteropoulos von der TU Wien am 1.2.2017 in Graz*. 2017
- [Bmvit 2016] BMVIT: *Der Faktor Zeit im Radverkehr*. Wien : Bundesministerium für Verkehr, Innovation und Technologie, 2016
- [Brandenburg u. a. 2007] BRANDENBURG, Christiane ; MATZARAKIS, Andreas ; ARNBERGER, Arne: Weather and cycling—a first approach to the effects of weather conditions on cycling. In: *Meteorological applications* 14 (2007), Nr. 1, S. 61–67
- [Cik 2016] CIK, Michael: Floating-Phone-Daten und Mobilitätsdaten im öffentlichen Verkehr. In: *Evolution vs. Revolution der Erhebung und Anwendung von Mobilitätsdaten - FSV Tagungsband* (2016), S. 74–76
- [ESRI 2017] ESRI, Arcmap O.: *Kurzer Überblick über die Verwendung von Iteratoren*. <https://goo.gl/d75CcI>, 2017. – (aufgerufen am 20.01.2017)
- [EuropäischesParlament 2016] EUROPÄISCHESPARLAMENT: *Parlament verabschiedet EU-Datenschutzreform – EU fit fürs digitale Zeitalter*. <https://goo.gl/Am4kgU>, 2016. – (aufgerufen am 13.12.2016)
- [Fellendorf u. a. 2011] FELLENDORF, Martin ; HERRY, Max ; KARMASIN, Helene ; KLEMENTSCHITZ, Roman ; KOHLA, Birgit ; MESCHIK, Michael ; REHRL, Karl ; REITER, Thomas ; SAMMER, Gerd ; SCHNEIDER, Cornelia u. a.: KOMOD-Konzeptstudie Mobilitätsdaten Österreichs: Handbuch für Mobilitätserhebungen. In: *KOMOD-concept study mobility data Austria: manual for mobility surveys, Project report, Vienna* (2011)
- [Fischer 2017] FISCHER, Eric: *Locals and Tourists #35 (GTWA #26): Vienna*. <https://flic.kr/p/87NNrv>, 2017. – (aufgerufen am 20.01.2017)
- [Freier u. Marco 2014] FREIER, Dominik ; MARCO, Leutenegger: *Internet of Things*. Hochschule für Technik Rapperswil, 2014

- [Gicycle 2016] GICYCLE: *A simple map matching approach for bicycle GPS tracks*. <https://gicycle.wordpress.com/2016/05/13/a-simple-map-matching-approach-for-bicycle-gps-tracks/>, 2016. – (aufgerufen am 20.01.2017)
- [Giffinger 2016] GIFFINGER, Rudolf: *Vorlesung zu Forschungs- und Planungsdesign im Curriculum des Masterstudiums Raumplanung und Raumordnung an der TU Wien*. 2016
- [Goodchild 2007] GOODCHILD, Michael: Citizens as sensors: the world of volunteered geography. In: *GeoJournal* 69 (2007), Nr. 4, S. 211–221
- [Greve 2002] GREVE, Klaus: Vom GIS zur geodateninfrastruktur. In: *Standort* 26 (2002), Nr. 3, S. 121–125
- [Haklay 2010] HAKLAY, Mordechai: How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. In: *Environment and Planning B: Planning and Design* 37 (2010), Nr. 4, S. 682–703
- [Herry u. Tomschy 2016] HERRY, Max ; TOMSCHY, Rupert: Zur Methode der Gewichtung und Hochrechnung - Österreich Unterwegs. In: *Evolution vs. Revolution der Erhebung und Anwendung von Mobilitätsdaten - FSV Tagungsband* (2016), S. 17–20
- [ITF u. OECD 2015] ITF ; OECD: *Big Data and Transport - Understanding and assessing options*. International Transport Forum, 2015
- [Kalasek u. Weninger 2015] KALASEK, Robert ; WENINGER, Kurt: Open GeoData. In: *Der öffentliche Sektor - The Public Sector* 1 (2015), S. 17–28
- [Kirchhoff 2013] KIRCHHOFF, Peter: *Städtische Verkehrsplanung: Konzepte, Verfahren, Maßnahmen*. Springer-Verlag, 2013
- [Kirnbauer 2016] KIRNBAUER, Roman: Projekthintergrund zur Erhebung Österreich unterwegs. In: *Evolution vs. Revolution der Erhebung und Anwendung von Mobilitätsdaten - FSV Tagungsband* (2016), S. 9–11
- [Koll-Schretzenmayr 2013] KOLL-SCHRETZENMAYR, Martina: Einsatz von Mobiltelefonaten für die Raumplanung: The Use of Mobile Telephone Data for Spatial Planning. In: *disP - The Planning Review* 49 (2013), Nr. 2, S. 118–120
- [Körnig-Pich u. a. 2010] KÖRNIG-PICH, Rebecca ; KEBBEDIES, Guido ; ZEILE, Peter: *Die Potenziale aktueller WebGIS-und Web-2.0-Entwicklungen als Planungsinstrumente—der Planer als Eichhörnchen?!* 2010
- [Krammer 2014] KRAMMER, Rainer: *Ausgewählte Analyse der Qualität von OSM Daten im alpinen Raum*, TU Wien, Diplomarbeit, 2014
- [Krampe u. Pucher 2016] KRAMPE, Stefan ; PUCHER, Gernot: Floating-Phone-Daten und Mobilitätsdaten im öffentlichen Verkehr. In: *Evolution vs. Revolution der Erhebung und Anwendung von Mobilitätsdaten - FSV Tagungsband* (2016), S. 62–70
- [Lechner 2011] LECHNER, Marco: *Nutzungspotentiale crowdsourcing-erhobener Geodaten auf verschiedenen Skalen*, Albert-Ludwigs-Universität Freiburg, Dissertation, 2011

- [Lessmann u. a. 2008] LESSMANN, Martin ; LESSMANN, Nicole ; STEFAN, Voß: Crowdsourcing: Systematisierung praktischer Ausprägungen und verwandter Konzepte. (2008), S. 273–274
- [von Lucke 2011] LUCKE, Jörn von: *Innovationsschub durch Open Data, Datenportale und Umsetzungswettbewerbe*. 2011
- [Miranda-Moreno u. Nosal 2011] MIRANDA-MORENO, Luis ; NOSAL, Thomas: Weather or not to cycle: Temporal trends and impact of weather on cycling in an urban environment. In: *Transportation Research Record: Journal of the Transportation Research Board* (2011), Nr. 2247, S. 42–52
- [MobileMarketingAssociation 2016] MOBILEMARKETINGASSOCIATION: *Mobile Communications Report 2015*. <http://www.mmaaustria.at/>, 2016. – (aufgerufen am 06.08.2016)
- [nast 2016] NAST, Consulting: *Radverkehrsdaten Wien*. <http://www.nast.at/verkehrsdaten/>, 2016. – (aufgerufen am 12.01.2017)
- [OECD 2016] OECD: *Participative Web*. <http://www.oecd.org/dataoecd/57/14/38393115.pdf>, 2016. – (aufgerufen am 13.12.2016)
- [OpenStreetMapStats 2016] OPENSTREETMAPSTATS: *OpenStreetMap stats report*. <http://www.openstreetmap.org/stats/datastats.html>, 2016. – (aufgerufen am 13.12.2016)
- [Parlament u. Rat 2007] PARLAMENT, Europäisches ; RAT, Europäischer: Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE). In: *Amtsblatt der Europäischen Union* 50 (2007), S. 1–14
- [Quddus u. a. 2007] QUDDUS, Mohammed A. ; OCHIENG, Washington Y. ; NOLAND, Robert B.: Current map-matching algorithms for transport applications: State-of-the art and future research directions. In: *Transportation research part c: Emerging technologies* 15 (2007), Nr. 5, S. 312–328
- [Rehrl u. a. 2016] REHRL, Karl ; BRUNAUER, Richard ; GRÖCHENIG, Simon: Collecting floating car data with smartphones: results from a field trial in Austria. In: *Journal of location Based services* 10 (2016), Nr. 1, S. 16–30
- [Riegler 2016] RIEGLER, Sebastian: Zur Methode der Erhebung - Österreich Unterwegs. In: *Evolution vs. Revolution der Erhebung und Anwendung von Mobilitätsdaten - FSV Tagungsband* (2016), S. 13–16
- [StadtWien 2016a] STADTWIEN: *Hasnerstraße ist erste "fahrradfreundliche Straße" Wiens*. <https://www.wien.gv.at/verkehr/strassen/bauen/radprojekte/hasnerstrasse.html>, 2016. – (aufgerufen am 12.01.2017)
- [StadtWien 2016b] STADTWIEN: *Konzept für Rad-Langstrecken in Wien*. <https://www.wien.gv.at/stadtentwicklung/projekte/verkehrsplanung/radwege/langstrecken/>, 2016. – (aufgerufen am 12.01.2017)
- [Straub u. Graser 2015] STRAUB, Markus ; GRASER, Anita: Learning from experts: Inferring road popularity from gps trajectories. In: *GI_Forum J. Geogr. Inf. Sci* 2015 (2015), S. 41–50

-
- [Strava 2016] STRAVA: *Strava Metro - Better Data for Better Cities*. <http://metro.strava.com/>, 2016. – (aufgerufen am 20.01.2017)
- [Strobl 2005] STROBL, Josef: Hierarchische Aggregation: Detailinformation versus Datenschutz am Beispiel adressbezogen georeferenzierter Datensätze. In: *Salzburger Geographische Arbeiten* 38 (2005), S. 163–171
- [Thomas u. a. 2012] THOMAS, Tom ; JAARSMA, Rinus ; TUTERT, Bas: *Exploring temporal fluctuations of daily cycling demand on Dutch cycle paths: the influence of weather on cycling*. Springer Science + Business Media, 2012
- [Tomschy u. a. 2016] TOMSCHY, Rupert ; HERRY, Max ; SAMMER, Gerd ; KLEMENTSCHITZ, Roman ; RIEGLER, Stefan ; FOLLMER, Robert ; GRUSCHWITZ, Dana ; JOSEF, Felix ; GENSASZ, Stefan ; KIRNBAUER, Romoan ; SPIEGEL, Thomas: *Ergebnisbericht zur österreichweiten Mobilitätserhebung „Österreich unterwegs 2013/2014 Maßnahmen*. Wien : Bundesministerium für Verkehr, Innovation und Technologie, 2016
- [Weichhart 2008] WEICHHART, Peter: *Entwicklungslinien der Sozialgeographie; von Hans Bobek bis Benno Werlen*. Stuttgart : Steiner, 2008 (Sozialgeographie kompakt; 1)
- [Widhalm 2016] WIDHALM, Peter: Smart Survey - Mobilitätserhebungen mit Smartphones. In: *Evolution vs. Revolution der Erhebung und Anwendung von Mobilitätsdaten - FSV Tagungsband* (2016), S. 46–48
- [Wildfiretoday 2017] WILDFIRETODAY: *Air traffic animation*. http://wildfiretoday.com/wp-content/uploads/2009/03/Air_Traffic_3.jpg, 2017. – (aufgerufen am 20.01.2017)
- [Wolf 2016] WOLF, Eike: Die rechtliche Situation zu Dateneigentum, Datenschutz, Big Data und Internet der Dinge. In: *Evolution vs. Revolution der Erhebung und Anwendung von Mobilitätsdaten - FSV Tagungsband* (2016), S. 39–45
- [Wonka 2010] WONKA, E: Regionalstatistik in Österreich auf der räumlichen Bezugsbasis von regionalstatistischen Rastereinheiten. In: *Hrsg.: Josef Strobl, Institut für Geographic Information Science, Österreichische Akademie der Wissenschaften gemeinsam mit STATISTIK AUSTRIA* (2010)
- [Zipf 2009] ZIPF, Alexander: Nutzungspotentiale und Herausforderungen von „Volunteered Geography“ - Zur Kombination von GDI-Technologie und nutzergenerierten Geomassendaten. In: *Wiener Schriften zur Geographie und Kartographie* 19 (2009), S. 121–134

8 Abbildungsverzeichnis

1.1	Übersicht Datenmengen	9
1.2	Aufbau der Arbeit	13
2.1	Pendlersaldo als Beispiel materieller Raum	17
2.2	Flugverkehrskarte als Beispiel für relationalen Raum	18
2.3	Darstellung Sozialer Raum REK Pongau	19
3.1	Flickr Verortete Fotos	23
3.2	BikeCitizens-Tracks auf der Praterbrücke zeigen GPS-Ungenauigkeit	28
4.1	Ablauf des Prozesses rund um Österreich Unterwegs	35
4.2	Schematische Darstellung Befragungsablauf ÖU	36
4.3	Österreich Unterwegs Fragebogen	37
4.4	Ergebnisstruktur Österreich Unterwegs	38
4.5	Geschwindigkeits und Beschleunigungssensor	43
4.6	Verkehrsmitteldetektion AIT.	43
4.7	Vergleich Frequenz Trajektorien Bahn MIV	45
4.8	Quell- und Zielgebiete Mobilfunkdatentool	45
5.1	Distanzverteilung Österreich Unterwegs	48
5.2	Aufbau Kapitel 5	49
5.3	Screenshot Bike Citizens App	51
5.4	Anzahl der Tracks im Monatsvergleich	53
5.5	GPX-Skizze	53
5.6	Beispielausschnitt Tracks	56
5.7	Beispielausschnitt Points	56
5.8	Strava Ausschnitt Paris	57
5.9	Matching Beispiel Gürtel	58
5.10	Rasteranalyse Testausschnitt	59
5.11	Gicycle Map Matching Ansatz	60
5.12	Modelbuilder Grundidee	64
5.13	Verortung Zählstellen	66
5.14	Konfidenzintervalle	68
5.15	Koeffizienten BC-Data / Zählstellen	69
5.16	Koeffizienten BC-Data / Zählstellen	69
5.17	Wochentagsverteilung Tracks	71
5.18	Stundenverteilung Tracks	72
5.19	Distanzverteilung Bike Citizens Tracks	72
5.20	Histogramm Distanz	73
5.21	Visualisierung Bike Citizens Tracks	74

5.22	Visualisierung Bike Citizens Track-Ziele	75
5.23	Verortung Radrouten	78
5.24	Entfernungsklassen Wien Schwerpunkt	79
5.25	Entfernungsklassen Bike Citizens Tool	79
5.26	Langstrecken-Tracks im Vergleich mit den offiziellen Radrouten	80
5.27	Routenvergleich Screenshot	82
5.28	Routenvergleich Heatmap	83
5.29	Bike Citizens Tracks zur TU Wien	86
5.30	Häufigkeitsverteilung der Distanz <15km.	87
5.31	Anteilige Verteilung der Tracks auf die Wochentage	87
5.32	Anteilige Verteilung der Ankunft der Tracks auf die Stunden an Werktagen	88
5.33	Anteilige Verteilung der Tracks auf die Monate	88
5.34	Die 4 Eckpfeiler der Neugestaltung der Hasnerstrasse	89
5.35	Verkehrsspinne Hasnerstraße	90
5.36	Grundsystematik Geschwindigkeitsmodell	92
5.37	Geschwindigkeitsverteilung über die Stunden	92
5.38	Kreuzungsübersicht Währinger Straße / Nußdorfer Straße	94
5.39	Grundsystematik Kreuzungsmodell	94
5.40	Häufigkeitsverteilung Relation AC	95
5.41	Dichteberechnung Kreuzung	96
6.1	Aufbau der Arbeit	101
6.2	BC Heatmap Nacht	104
6.3	Schwarzplan Bike Citizens	108

9 Tabellenverzeichnis

4.1	Einteilung der Tracking-Systematik	41
5.1	Verwendete Datengrundlagen und deren Quelle	63
5.2	Deskriptive Statistik des Attributs Distanz (in Metern)	65
5.3	Anteile der Bike Citizens Tracks an den Zählstellenwerten an Werktagen	67
5.4	Vergleich der absoluten Werte an den Zählstellen, differenziert jeweils nach ZS (Zählstellenwert) und BC (Bike Citizens Tracks). [Eigene Darstellung]	67
5.5	Korrelationen Anzahl Tracks pro Tag / Durchschnittliche Wetterausprägung	76
5.6	Korrelationskoeffizienten Anzahl Tracks je Stunde / Aktuelle Wetterausprägung	77
5.7	Routensplit im Detail	84
5.8	Geschwindigkeitsvergleich je Richtung	93
5.9	Vergleich der Kreuzungsrelationen	95
6.1	Vergleich der Mobilitätsdatenerhebungen nach charakteristischen Merkmalen	102

10 Anhang

Zu Kapitel 5.5.2: Zählstellen-Vergleichs-Tabellen

Anteile der Bike Citizens Tracks an den Zählstellenwerten an Samstagen. Fett markiert jeweils der maximale Wert. [Eigene Darstellung]

ID / Ort	Jan16	Feb16	Mar16	Apr16	May16	Jun16	Jul16	Aug16	Mittelwert
1 Lassallestraße	0,08%	0,26%	0,30%	0,42%	0,44%	0,45%	0,37%	0,42%	0,34%
2 Argentinierstraße	0,58%	0,36%	0,47%	0,61%	0,68%	0,82%	0,64%	0,63%	0,60%
3 Donaukanal	0,00%	0,14%	0,04%	0,16%	0,26%	0,26%	0,24%	0,23%	0,17%
4 Langobardenstraße	0,00%	0,15%	0,00%	0,25%	0,49%	0,49%	0,72%	0,61%	0,34%
5 Liesingbach	0,30%	0,14%	0,11%	0,38%	0,50%	0,19%	0,40%	0,44%	0,31%
6 Wienzeile	0,25%	0,23%	0,11%	0,42%	0,46%	0,48%	0,57%	0,38%	0,36%
7 Neubaugürtel	0,20%	0,29%	0,24%	0,40%	0,36%	0,34%	0,25%	0,33%	0,30%
8 Margaritensteg	0,37%	0,28%	0,42%	0,35%	0,34%	0,21%	0,40%	0,39%	0,35%
9 Praterstern	0,14%	0,22%	0,36%	0,43%	0,37%	0,41%	0,42%	0,50%	0,36%
10 Operngasse	0,20%	0,53%	0,55%	0,55%	0,28%	0,39%	0,53%	0,51%	0,44%
11 Opernring	0,66%	0,36%	0,65%	0,43%	0,37%	0,39%	0,41%	0,28%	0,44%
Mittelwert	0,25%	0,27%	0,30%	0,40%	0,41%	0,40%	0,45%	0,43%	0,36%

Anteile der Bike Citizens Tracks an den Zählstellenwerten an Sonn- und Feiertagen. Fett markiert jeweils der maximale Wert. [Eigene Darstellung]

ID / Ort	Jan16	Feb16	Mar16	Apr16	May16	Jun16	Jul16	Aug16	Mittelwert
1 Lassallestraße	0,10%	0,15%	0,57%	0,35%	0,23%	0,41%	0,45%	0,30%	0,32%
2 Argentinierstraße	0,14%	0,29%	0,69%	0,49%	0,26%	0,42%	0,52%	0,21%	0,38%
3 Donaukanal	0,00%	0,12%	0,27%	0,31%	0,23%	0,19%	0,22%	0,27%	0,20%
4 Langobardenstraße	0,00%	0,12%	0,26%	0,24%	0,38%	0,56%	0,19%	0,21%	0,24%
5 Liesingbach	0,00%	0,00%	0,49%	0,54%	0,42%	0,42%	0,55%	0,51%	0,37%
6 Wienzeile	0,00%	0,06%	0,48%	0,29%	0,26%	0,38%	0,27%	0,29%	0,25%
7 Neubaugürtel	0,12%	0,28%	0,18%	0,22%	0,16%	0,28%	0,26%	0,17%	0,21%
8 Margaritensteg	0,22%	0,27%	0,15%	0,16%	0,18%	0,26%	0,33%	0,23%	0,22%
9 Praterstern	0,06%	0,11%	0,34%	0,35%	0,26%	0,26%	0,38%	0,29%	0,26%
10 Operngasse	0,20%	0,36%	0,36%	0,31%	0,16%	0,30%	0,35%	0,26%	0,29%
11 Opernring	0,45%	0,45%	0,54%	0,40%	0,21%	0,31%	0,27%	0,20%	0,35%
Mittelwert	0,12%	0,20%	0,39%	0,33%	0,25%	0,34%	0,34%	0,27%	0,28%

Konfidenzintervallgrenzen der Zählstellen in den Monaten Jänner bis April 2016. (unt. = untere Intervallgrenze, obe. = obere Intervallgrenze) [Eigene Darstellung]

	Jan16		Feb16		Mar16		Apr16	
	unt.	obe.	unt.	obe.	unt.	obe.	unt.	obe.
Lassallestraße	-0,14%	0,34%	-0,09%	0,56%	0,08%	0,86%	0,26%	0,90%
Argentinerstraße	-0,06%	0,54%	-0,03%	0,55%	0,10%	0,78%	0,18%	0,71%
Donaukanal	-0,20%	0,30%	-0,21%	0,42%	-0,14%	0,30%	-0,07%	0,23%
Langobardenstraße	0,00%	0,00%	-0,45%	1,25%	-0,29%	1,52%	-0,14%	0,98%
Liesingbach	-0,46%	0,64%	-0,24%	0,30%	-0,32%	0,63%	-0,17%	0,86%
Wienzeile	-0,27%	0,59%	-0,21%	0,83%	-0,09%	0,94%	0,03%	0,72%
Neubaugürtel	-0,13%	0,47%	-0,08%	0,58%	-0,04%	0,56%	0,10%	0,64%
Margaritensteg	-0,23%	0,81%	-0,17%	0,99%	-0,07%	0,90%	0,05%	0,75%
Praterstern	-0,09%	0,24%	-0,06%	0,36%	0,10%	0,68%	0,24%	0,74%
Operngasse	0,09%	0,63%	0,12%	0,67%	0,19%	0,70%	0,27%	0,65%
Opernring	0,18%	0,80%	0,19%	0,75%	0,28%	0,79%	0,30%	0,67%

Konfidenzintervallgrenzen der Zählstellen in den Monaten Mai bis August 2016. (unt. = untere Intervallgrenze, obe. = obere Intervallgrenze) [Eigene Darstellung]

	May16		Jun16		Jul16		Aug16	
	unt.	obe.	unt.	obe.	unt.	obe.	unt.	obe.
Lassallestraße	0,26%	0,87%	0,36%	0,86%	0,32%	0,77%	0,28%	0,73%
Argentinerstraße	0,34%	0,98%	0,47%	1,09%	0,43%	1,07%	0,28%	0,85%
Donaukanal	-0,06%	0,27%	-0,01%	0,35%	0,00%	0,36%	0,04%	0,46%
Langobardenstraße	0,00%	1,43%	0,02%	1,08%	0,10%	1,29%	0,00%	1,05%
Liesingbach	-0,13%	0,94%	0,05%	1,20%	-0,02%	1,00%	-0,05%	0,95%
Wienzeile	0,14%	0,89%	0,22%	0,92%	0,16%	0,73%	0,19%	0,80%
Neubaugürtel	0,10%	0,64%	0,22%	0,77%	0,10%	0,56%	0,16%	0,72%
Margaritensteg	0,00%	0,60%	0,10%	0,70%	0,13%	0,78%	0,09%	0,72%
Praterstern	0,21%	0,68%	0,31%	0,72%	0,32%	0,74%	0,28%	0,69%
Operngasse	0,27%	0,61%	0,35%	0,70%	0,36%	0,75%	0,31%	0,70%
Opernring	0,29%	0,64%	0,34%	0,65%	0,28%	0,59%	0,24%	0,53%

Zu Kapitel 5.6.2: Regressionsergebnisse Radverkehr und Wetter

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.	95,0% Konfidenzintervalle für B	
		Regressionskoeffizient B	Standardfehler	Beta			Untergrenze	Obergrenze
1	(Konstante)	575,359	60,603		9,494	,000	455,978	694,740
	temperature_mean	13,051	1,061	,588	12,305	,000	10,962	15,141
	humidity_mean	-3,797	,636	-,281	-5,968	,000	-5,051	-2,544
	windspeed_mean	-15,639	3,810	-,157	-4,105	,000	-23,144	-8,134

a. Abhängige Variable: tracks_date_count

Koeffizienten zur Tabelle 5.6 [Eigene Darstellung / SPSS]

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Änderung in R-Quadrat	Statistikwerte ändern			Sig. Änderung in F
						Änderung in F	df1	df2	
1	,822 ^a	,676	,672	107,589	,676	166,724	3	240	,000

a. Einflussvariablen : (Konstante), windspeed_mean, humidity_mean, temperature_mean

Modellergebnisse zur Tabelle 5.6 [Eigene Darstellung / SPSS]

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Änderung in R-Quadrat	Statistikwerte ändern			Sig. Änderung in F
						Änderung in F	df1	df2	
1	,562 ^a	,315	,315	14,754	,315	767,564	3	4999	,000

a. Einflussvariablen : (Konstante), windspeed_mean, temperature_mean, humidity_mean

Modellergebnisse zur Tabelle 5.7 [Eigene Darstellung / SPSS]

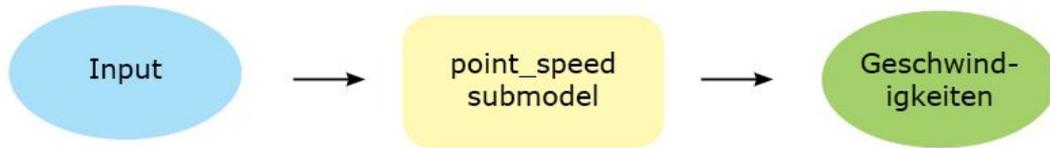
Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.
		Regressionskoeffizient B	Standardfehler	Beta		
1	(Konstante)	35,359	1,509		23,427	,000
	temperature_mean	,681	,030	,323	22,546	,000
	humidity_mean	-,318	,015	-,316	-21,647	,000
	windspeed_mean	-,048	,096	-,006	-,499	,618

a. Abhängige Variable: track_count

Koeffizienten zur Tabelle 5.7 [Eigene Darstellung / SPSS]

Zu Kapitel 5.8.1: ArcMap-Modell zur Berechnung der Geschwindigkeit



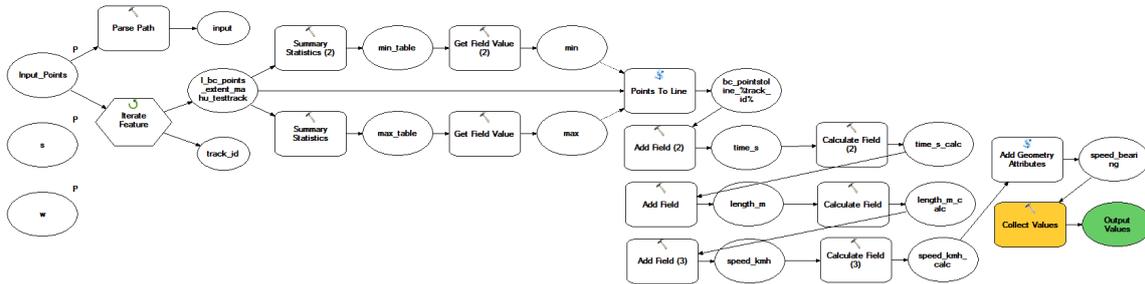
Grundsystematik des Modells zur Geschwindigkeitsberechnung. [Eigene Darstellung]

In dieser Berechnung findet der Großteil der Berechnung in einem Submodell statt. Dies ist ein dem Hauptmodell untergeordnetes Modell und hat den Vorteil, dass die Zusammenführung der Daten (Merge) erst nach Finalisierung der eigentlichen Geschwindigkeitsberechnungen je Track durchgeführt wird. Dies spart Rechenaufwand und somit Zeit im Durchlauf.

Um das Modell starten zu können, ist die Definition eines work-workspaces (für die Endergebnisse) und eines scratch-workspaces (für die Zwischenergebnisse) erforderlich. Durch die große Datenmenge, die je nach Größe des Ausschnitts generiert wird, empfiehlt sich zumindest für die Zwischenergebnisse eine neue Geodatenbank zu erstellen.

Das Modell gliedert sich in folgende Elemente:

- Input: Pointset des betreffenden räumlichen Ausschnitts
- Geoverarbeitung
 - *Iterate Feature* wählt systematisch jede track-ID aus, um zusammenhängende Punkte zu bekommen
 - *Summary Statistics* berechnet das Minimum und das Maximum der Point-Nummerierung
 - *Points To Line* verbindet die Punkte unter Berücksichtigung der fortlaufenden Nummerierung zu einer Linie
 - Über die Kombination *Add Field* und *Calculate Field* werden die zurückgelegte Strecke (Meter), die Zeit (Differenz aus Minimum und Maximum von *Summary Statistics*) und die durchschnittliche Geschwindigkeit (km/h) berechnet
 - *Add Geometry Attributes* gibt jeder Linie eine Peilung in Grad (0 bis 360)
 - *Collect Values* speichert das Ergebnis ab. Die Iteration ist hier abgeschlossen und beginnt erneut von vorne, bis alle track-IDs abgearbeitet sind
 - *Merge* fügt im Hauptmodell alle gesammelten Teilergebnisse aus *Collect Values* zusammen
- Output: Ist ein Shape-File mit allen Tracks und den neuen Attributen Zeit, Strecke, Durchschnittsgeschwindigkeit und Kurs



Sub-Modell Geschwindigkeitsberechnung [ModelBuilder Screenshot / Eigene Darstellung]

Zu Kapitel 5.8.2: ArcMap-Modell zur Kreuzungsanalyse

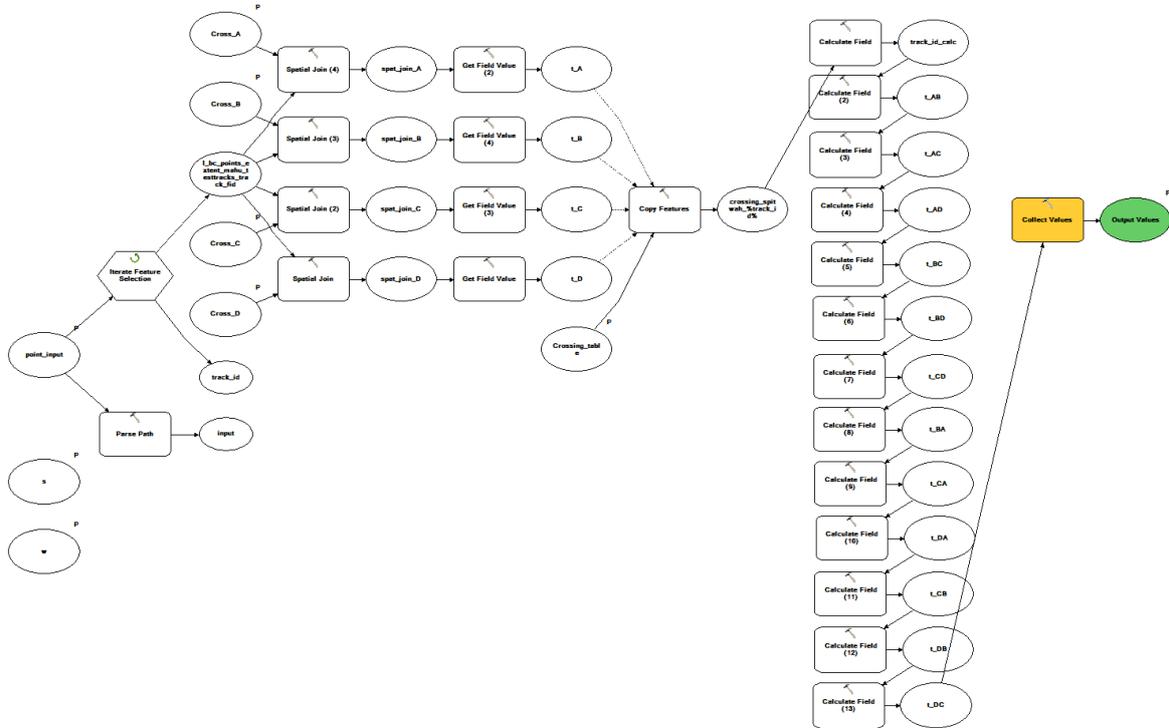


Grundsystematik des Modells zur Kreuzungsverzögerung. [Eigene Darstellung]

Das Modell gliedert sich in folgende Elemente:

- Input
 - Pointset des betreffenden räumlichen Ausschnitts
 - Relationspunkte A, B, C und D als point-shape
- Geoverarbeitung
 - *Iterator* wählt über die track-ID eine zusammengehörende Punktmenge aus
 - *Spatial Join* mit der Einstellung *closest to* findet von den vier Bezugspunkten A bis D jeweils den am nächsten liegenden *track-point*.
 - Mithilfe von *Get Field Value* wird von diesem gefundenen *track-point* jeweils die fortlaufende Nummerierung gespeichert
 - Mithilfe einer Abfolge von *Calculate Field* wird für jede Relation die Zeitdauer in Sekunden ermittelt. Die Fahrtrichtung wird über einen einfachen größer/kleiner Vergleich überprüft
 - *collect values* sammelt die Ergebnisse je track-ID und der *Iterator* beginnt mit der Auswahl der nächsten

- Output: Ist ein Shape-File mit allen Tracks und den neuen Attributen Zeit, Strecke, Durchschnittsgeschwindigkeit und Kurs



Sub-Modell Kreuzungsanalyse [ModelBuilder Screenshot / Eigene Darstellung]