



TECHNISCHE
UNIVERSITÄT
DRESDEN

Angela Francke • Sven Lißner

Big Data im Radverkehr

Ein anwendungsorientierter Leitfaden zur
Nutzung von smartphone-generierten
Radverkehrsdaten

Impressum

Big Data im Radverkehr

Ein anwendungsorientierter Leitfaden zur
Nutzung von smartphone-generierten
Radverkehrsdaten

Dresden
November 2017

Herausgeber: Technische Universität Dresden, Professur für
Verkehrsökologie und Professur für Verkehrspsychologie

Autoren: Dipl.-Ing. Sven Lißner,
Dipl.-Verk.wirtsch. Angela Francke

Layout: Lisa-Marie Schaefer

Kontakt: verkehrsoekologie@tu-dresden.de

Gefördert durch:



Bundesministerium
für Verkehr und
digitale Infrastruktur

aufgrund eines Beschlusses
des Deutschen Bundestages

Geleitwort

des Bundesministeriums für Verkehr und digitale Infrastruktur

Der vorliegende Leitfaden baut auf zwei äußerst positiven Entwicklungen auf:

1. Radfahren macht Spaß, es ist gesund, es schont den Geldbeutel und die Umwelt. Zudem ist das Fahrrad in der Stadt ein äußerst flexibles und effizientes Verkehrsmittel; der Radverkehr steigert also nicht nur die Lebensqualität der Radfahrerinnen und Radfahrer selbst, sondern auch die Lebensqualität einer Stadt insgesamt.
2. Die rasant fortschreitende Digitalisierung erschließt neue Möglichkeiten der Datenerfassung und -auswertung, partizipative neue Dateninhalte und die Reorganisation vorhandener Prozesse mittels digitaler Daten.

Warum nicht diese beiden positiven Entwicklungen kombinieren und die offenkundigen Vorteile der Digitalisierung für die kommunale Radverkehrsplanung nutzen? Eine gute Radverkehrsinfrastruktur – Voraussetzung für einen hohen Anteil des Radverkehrs am Modal Split – muss genau dort sein, wo die Radfahrerinnen und Radfahrer sie brauchen. Fast erstaunlich, aber doch Realität: Bislang liegen vielen kommunalen Verkehrsplanungen keine belastbaren Daten über die genauen Wege der Radfahrerinnen und Radfahrer vor. Hier kann die Digitalisierung helfen.

Im Forschungsprojekt „Mit Smartphones generierte Verhaltensdaten im Radverkehr“ wurden daher modellhaft für die Radverkehrsplanung in Dresden Bewegungsdaten des Radverkehrs auf „Herz und Nieren“ geprüft. Im Ergebnis bestätigte sich, dass die Bewegungsdaten nicht nur interessante Einblicke in die Radverkehrsströme vor Ort geben, sondern auch helfen können, die Bedürfnisse und Erwartungen der Radfahrerinnen und Radfahrer besser wahrzunehmen und in die Verkehrsplanung einzubeziehen.

Der vorliegende Leitfaden bietet den kommunalen Planerinnen und Planern sowie der interessierten Öffentlichkeit einen handlichen, praxisorientierten Einstieg in das Thema GPS-Daten und zeigt dabei Potenziale für die Radverkehrsplanung sowie mögliche Stolpersteine bei der Nutzung dieser Daten auf. Er ermutigt, die Digitalisierung auch in diesem Bereich zu nutzen und spricht dabei Schlüsselthemen, wie den Datenschutz oder Herausforderungen bei der Interpretation der Daten, an.

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

Zusammenfassung

Für einen attraktiven Radverkehr bedarf es einer qualitativ hochwertigen Infrastruktur. Bisher liegen durch den hohen Aufwand von Vor-Ort-Erfassungen nur punktuelle Radverkehrsstärken vor. Die aktuell wohl zuverlässigsten und tauglichsten Werte liefern bisher fest installierte automatische Radverkehrszählstellen, wie sie bereits viele Kommunen eingerichtet haben. Ein Nachteil ist hierbei, dass für eine flächige Abdeckung mit einer besseren Aussagekraft für die gesamte Stadt oder Kommune die Anzahl der Erhebungspunkte meist deutlich zu gering ist. Die Bedeutung des Nebennetzes für den Radverkehr wird somit nur unvollständig erfasst. Für weitere Parameter, wie Wartezeiten, Routenwahl oder Geschwindigkeiten der Radfahrenden, fehlen dagegen meist die Daten. Perspektivisch kann diese Lücke unter anderem durch GPS-Routendaten gefüllt werden, was durch die mittlerweile sehr hohe Verbreitung von Smartphones und den entsprechenden Tracking-Apps ermöglicht wird. Die im Leitfaden vorgestellten Erkenntnisse sind im Rahmen eines umfangreichen Forschungsprojektes zu der Nutzbarkeit von mit Smartphones generierten Nutzerdaten einer App für die kommunale Radverkehrsplanung entstanden. Das Projekt wurde durch das Bundesministerium für Verkehr und digitale Infrastruktur (BMVI) im Rahmen des Nationalen Radverkehrsplans 2020 gefördert.

Zusammenfassend lässt sich sagen, dass unter Beachtung der im folgenden Leitfaden beschriebenen Faktoren GPS-Daten, im vorliegenden Fall exemplarisch untersucht anhand von Daten der Firma Strava Inc., mit einigen Einschränkungen für die Radverkehrsplanung nutzbar sind. Bereits heute sind damit Auswertungen möglich, die zeigen, wo, wann und wie sich Radfahrende im gesamten Netz bewegen. Die mittels Smartphone-App generierten Daten können sehr sinnvoll als Ergänzung zu bestehenden Dauerzählstellen von Kommunen genutzt werden. Berücksichtigt werden sollten bei der Auswertung und Interpretation der Daten jedoch einige Aspekte, wie der eher sportlich orientierte Kontext der erfassten Routen in den untersuchten Beispielen. Des Weiteren werden aktuell die Daten zum Teil noch als Datenbank- oder GIS-Dateien zur Verfügung gestellt bzw. befinden sich online Masken zur einfacheren Nutzung im Aufbau oder einem ersten Nutzungsstadium. Die Auswertung und Interpretation erfordert also weiterhin Fachkompetenz und auch personelle Ressourcen. Der Einsatz dieser wird jedoch voraussichtlich durch die Weiterentwicklung von Web-Oberflächen und unterstützenden Auswertemaschinen abnehmen. Hier gilt es zukünftig, in Zusammenarbeit mit den Kommunen, die benötigten Parameter sowie die geeignetsten Aufbereitungsformen zu erarbeiten. Im Forschungsprojekt erfolgte ein Ansatz der Hochrechnung von Radverkehrsstärken aus Stichproben von GPS-Daten auf das gesamte Netz. Dieser konnte auch erfolgreich in einer weiteren Kommune verifiziert werden. Jedoch ist auch hier in Zukunft noch Forschungsbedarf vorhanden bzw. die Anpassung auf lokale Gegebenheiten notwendig.

In naher Zukunft ist es notwendig, den Praxisnachweis für die Nutzbarkeit von GPS-Daten zu erbringen. Vorbilder hierfür können die Städte Dresden, Leipzig oder Mainz sein, die jeweils bereits erste Schritte zur Nutzung von GPS-Daten in der Radverkehrsplanung und -förderung unternommen. Diese Schritte sind vor dem Hintergrund der weiteren Digitalisierung von Mobilität und Verkehrsmitteln und dem damit wachsenden Datenangebot – auch trotz der bisherigen Einschränkungen der Daten – sinnvoll, um in den Verwaltungen frühzeitig entsprechende Kompetenzen aufzubauen. Langfristig bietet die Nutzung von GPS-Daten einen Mehrwert für die Radverkehrsplanung. Der aktive Einbezug von Radfahrenden eröffnet zudem neue Möglichkeiten in der Kommunikation und der Bürgerbeteiligung – auch ohne Fachwissen vorauszusetzen. Der vorliegende Leitfaden liefert dafür einen praxisorientierten Einstieg in das Thema und weist umfassend auf Angebote, Hindernisse und Potenziale von GPS-Daten hin.

Inhaltsverzeichnis

1	Einleitung	8
	1.1 An wen richtet sich der Leitfaden?	
	1.2 Überblick: GPS-Daten im Radverkehr	
2	Was sind GPS-Daten?	11
	2.1 Allgemeines	
	2.2 Welche Daten werden angeboten?	
	2.3 Welche Parameter können untersucht werden?	
3	Welche Möglichkeiten bieten GPS-Daten?	19
	3.1 Überblick	
	3.2 Netzplanung	
	3.3 Anlagenplanung	
	3.4 Priorisierung von Maßnahmen	
	3.5 Monitoring und Evaluation	
	3.6 Wie können Begleitmaßnahmen aussehen?	
	3.7 Mögliche Schwächen	
4	Anwendungsbeispiel: Pilotkommune Dresden	24
	4.1 Projektrahmen	
	4.2 Anwendungsbeispiel – Vorgehen	
5	Ergebnisprüfung	37
	5.1 Wann ist der Datensatz nutzbar?	
	5.2 Was kann schiefgehen?	
6	Ausblick – Zukünftige Entwicklungen des Marktes	39
	Anhang	40
	Detailwissen I-VII	
	Glossar	
	Häufige Fragen	

1 Einleitung

Die zukünftige Verkehrsplanung in einer immer stärker digitalisierten Welt kann sich neuen, smarten Technologien nicht verschließen. Die Mobilitätsforschung steht an der Schwelle zu einem Wandel hin zu mehr Bürgerbeteiligung durch die (automatische) Speicherung von Mobilitätsgewohnheiten. Der vorliegende Leitfaden ermöglicht einen Einstieg in das Thema GPS-Daten, das als eine Möglichkeit einer neuen, digitalen Datenerhebung in der Radverkehrsplanung dient.

1.1 An wen richtet sich der Leitfaden?

Zielgruppe

Im Projekt „Smartphonegenerierte Verhaltensdaten im Radverkehr“ hat sich die Forschungsgruppe der Verkehrsökologie und -psychologie der TU Dresden mit der praktischen Verwendbarkeit von freiwillig generierten Radverkehrs-Bewegungsdaten (GPS-Daten) in der allgemeinen Radverkehrsplanung befasst.

Mit diesem Leitfaden liegt eine praxisnahe Anleitung für den Einsatz von GPS-Daten in der Radverkehrsplanung vor. Egal, ob bereits mit GPS-Daten gearbeitet wurde oder noch keine Erfahrungen auf diesem Gebiet vorliegen, kann der Leitfaden sowohl als Einstieg in das Thema als auch als praktische Strukturierungshilfe zur Qualitätsbeurteilung eingesetzt werden. Vertiefende Projektergebnisse finden sich im Projektbericht, der auf der NRVP-Website zum Download zur Verfügung steht.

Leitfadenaufbau

Die Kapitel sind unabhängig voneinander aufgebaut und können nach Bedarf spezifisch genutzt werden. Es wird durch den kompletten Prozess der GPS-Datenbasierten Radverkehrsplanung hindurchgeführt, angefangen mit einem theoretischen Einstieg in das Thema und Zugangsmöglichkeiten zu den GPS-Daten. In Kapitel 2 werden die Eigenschaften, Möglichkeiten und Grenzen von GPS-Daten aufgezeigt und in Kapitel 3 wird auf sinnvolle Anwendungsfälle hingewiesen. In einem umfangreichen Anwendungsbeispiel in Kapitel 4 kann ein Planungsvorgang modellhaft nachvollzogen, auf die eigene Kommune übertragen und auf dieser Basis individuelle Einsatzideen entwickelt werden. Schlussendlich helfen Kriterien für die Nutzbarkeit und Interpretation der Daten, Fehlschlüsse zu vermeiden und die Potentiale vollständig auszuschöpfen (Kapitel 5). Ein umfangreiches Detailwissen, ein Glossar (Schlagwörter im Text sind in **dunkelblau** markiert) sowie eine Liste mit besonders häufigen Fragen und Antworten (FAQ) zum Thema ermöglichen eine effiziente Arbeit mit dem Leitfaden.

1.2 Überblick: GPS-Daten im Radverkehr

Datenlage im Radverkehr

Für den **motorisierten Individualverkehr (MIV)** liegen bereits aussagekräftige Daten aus unterschiedlichen Quellen vor. Die klassische Datenlage in der Radverkehrsplanung hingegen basiert auf der aufwändigen und aktiven Datensammlung durch Zählstellen, Verkehrsbeobachtungen und -befragungen. Die Bewegungsmuster von Verkehrsteilnehmerinnen und Verkehrsteilnehmern sind dementsprechend Momentaufnahmen, die selten miteinander in Bezug gesetzt werden können. Die zunehmende digitale Vernetzung und Verbreitung von GPS-fähigen Endgeräten bietet nun die Möglichkeit, Verkehrsteilnehmende in ihren Bewegungsmustern zusammengefasst zu betrachten und daraus eine bedarfsorientierte Radverkehrsplanung zu etablieren.

In der Arbeit mit GPS-Daten für die Radverkehrsplanung haben sich im Vorfeld und im Rahmen dieses Projektes einige positive Ergebnisse gezeigt:

- ✓ *Trend zur smartphonegestützten Mobilitätskontrolle der Nutzerinnen und Nutzer vorhanden,*
- ✓ *Verfügbarkeit der Daten gegeben,*
- ✓ *als Grundlage für verschiedene Planungszwecke einsetzbar,*
- ✓ *Wissenschaftliche Evaluation der Nutzbarkeit ergibt ein positives Bild.*

Die Nutzerzahl von Smartphones wächst seit Jahren konstant. Im Jahr 2016 besaßen nach Angaben des Statistischen Bundesamtes bereits 49 Mio. Deutsche ein Smartphone. Diese Geräte sind fast alle mit GPS-Funktionen ausgestattet, die vor allem zur Navigation in Verbindung mit Karten-Applikationen eingesetzt werden. Seit einigen Jahren kommen darüber hinaus **Smartphone-Applikationen (Apps)**, die diese Funktionen nutzen und auswerten, auf den Markt. Sie bieten Sportbegeisterten die Möglichkeit, Trainingsstrecken, Bewegungsmuster und -geschwindigkeiten aufzuzeichnen und zu speichern. Nutzerinnen und Nutzer erhalten so eine umfangreiche Übersicht über ihre Leistung und können sie mit anderen vergleichen. Die Nutzerzahlen dieser Apps sind ebenfalls steigend. Die meisten App-Betreiber lassen sich von den Nutzerinnen und Nutzern Verwendungsrechte übertragen. Sie können daher die gesammelten Daten in anonymisierter Form auf dem freien Markt zum Kauf anbieten. Diese Daten können unter anderem für Kommunen interessante Einblicke bieten. Eine Übersicht über den Bezug der Daten bietet Unterkapitel 2.3.

Datenanbieter und Drittanbieter bauen ihre Dienste hinsichtlich der Datenaufbereitung und -auswertung zunehmend kundenorientiert aus. Erfahrungsgemäß sind ein detailliertes Anforderungsprofil und das entsprechende GIS-Straßenknotennetz der Stadt ausreichend, um den Anbieter die entsprechende Datenauswertung nach individuellen Wünschen vornehmen zu lassen. Folglich liegt der Schwerpunkt des Leitfadens vielmehr auf der Bestellung und Verwendung von GPS-Daten in der Radverkehrsplanung als auf der technischen Auswertungspraxis der Rohdaten. Wie die Daten erworben werden können, wird in Kapitel 2 ausgeführt.

Die Grundeigenschaften der nutzergenerierten GPS-Daten bieten einige entscheidende Vorteile gegenüber konventionellen Datenquellen wie Verkehrszählungen und Befragungen. Sie bieten besonders in den Parametern Verkehrsmengen, Quelle-Ziel-Matrizen und Geschwindigkeiten gute Eigenschaften und sind damit in vielen zentralen Bereichen der bedarfsorientierten Radverkehrsplanung einsetzbar. In Kapitel 3 werden die Anwendungsbereiche Netzplanung, Anlagenplanung, Priorisierung von Maßnahmen sowie Monitoring und Evaluation vorgestellt und die entsprechenden Parameter in Hinblick auf GPS-Daten erläutert. Abschließend werden Ideen und Beispiele zur öffentlichkeitswirksamen Nutzung der GPS-Datenbasierten Planung im Rahmen von Bürgerpartizipation und Öffentlichkeitsarbeit vorgestellt.

Im Rahmen des Forschungsprojektes waren die anonymisierten, käuflichen GPS-Daten einer (sport-orientierten) App für Smartphones der Firma Strava Inc. Basis für die **Validierung**. Verglichen wurden diese mit empirisch im Straßenverkehr erhobenen Daten (Zählstellen, Geschwindigkeitsmessungen) in der Pilotkommune Dresden. Im Projektverlauf wurden nicht nur die zur Verfügung gestellten Daten auf ihre Verallgemeinerbarkeit in Bezug auf Nutzerstruktur, Bewegungsmuster und Netzabdeckung überprüft, sondern darüber hinaus für die Radverkehrsplanung praxisrelevante Anwendungsvorschläge erarbeitet und in diesem Leitfaden festgehalten (Kapitel 4). Nach der umfangreichen wissenschaftlichen Evaluation kann eine Empfehlung für eine GPS-Datenbasierte Radverkehrsplanung ausgesprochen werden, sofern eine reflektierte und sorgfältige Interpretation der Daten erfolgt. In Kapitel 5 werden hierfür einige Empfehlungen gegeben.

Smartphone-Nutzung

Datenverfügbarkeit

Vorteile von GPS-Daten

Potential der Datennutzung



Planungspraxis mit GPS-Daten

Diese Aufstellung bietet einen Überblick über Chancen und Herausforderungen bei der Arbeit mit GPS-Daten, die in Kapitel 2 näher erörtert werden. Eine detaillierte Erörterung der einzelnen Punkte und Anwendungsbeispiele finden sich in Kapitel 4.

Chancen

Datengewinnung

- Generierung der Daten in Echtzeit
- Filterbarkeit nach verschiedenen Kriterien
- Überschaubarkeit des Kostenaufwandes
- Ersetzung von wartungsintensiven Zählstellen
- Gewinnung von sozio-demografischen Angaben

Datenstruktur & Anwendungsfelder

- Schneller Überblick über das Radverkehrsaufkommen in Haupt- /Nebennetzen
- Vorher-Nachher-Evaluation von Maßnahmen
- Flächendeckende Verkehrsmengenkarten
- Quelle-Ziel-Matrizen zwischen Polygonen
- Wartezeiten an Knotenpunkten berechenbar
- Abschnittsfeine Durchschnittsgeschwindigkeiten
- Routen von Einzelwegen innerhalb grober Korridore

Herausforderungen

- Systematische Datenfehler durch Ausrichtung und Nutzerschaft der Applikationen
- Personen ohne geeignetes Endgerät werden systematisch aus der Stichprobe ausgeschlossen
- Skepsis gegenüber Datenerwerb durch Bedenken zum Datenschutz
- Gesetzliche Vorgaben besonders zum Datenschutz und damit verbundene Einschränkungen bei der Datennutzung

Was sind GPS-Daten? 2

GPS-Daten stehen stellvertretend für globale, satellitengestützte Positionsdaten. Durch deren Freigabe im Jahr 2000 erlangten vor allem Navigationsgeräte eine starke Bedeutung. Mit der ansteigenden Zahl von Empfängern und einfacher Software zur Aufzeichnung können diese Daten ein genaues Abbild der täglichen Mobilität liefern.

2.1 Allgemeines

Ein wachsender Anteil an Menschen besitzt ein Smartphone und nutzt dieses für die Realisierung von täglichen Mobilitätsbedürfnissen. Die Möglichkeit mit GPS-Daten Bewegungen sowohl individuell als auch über große Personenzahlen zusammengefasst abzubilden wird in der zukünftigen Verkehrsplanung einen immer größeren Stellenwert einnehmen. Erkenntnisse aus Verkehrsbefragungen, wie **SrV** oder **MiD**, können um solche freiwillig erhobenen GPS-Daten diverser App-Anbieter ergänzt und kostengünstig im Jahresrhythmus fortgeschrieben werden. Der besseren Lesbarkeit halber werden im Folgenden mit einem Smartphone generierte GPS-Aufzeichnungen kurz als ‚GPS-Daten‘ bezeichnet.

Erste Anbieter sehen in der Bereitschaft der Nutzerinnen und Nutzer zur Datenpreisgabe marktfähige Modelle für Datenverkauf und -analyse. Diese Entwicklung erfordert eine kritische wissenschaftliche Begleitung und eine **Validierung** der entstehenden Daten, um den Weg hin zu einer qualitativ hochwertigen Planungsgrundlage zu gestalten. Im Rahmen eines umfangreichen Forschungsprojektes wurde daher modellhaft das Datenangebot der Firma Strava Inc. (zum Zeitpunkt des Projektstarts der größte Datenanbieter am Markt, aktuell nutzen weltweit 180 Städte und Kommunen die Daten, Stand: 08/2017) auf seine Nutzungsmöglichkeiten untersucht.

Welche Form haben die GPS-Daten?

GPS-Daten werden zumeist im sportiven Kontext zu Trainingszwecken aufgezeichnet oder fallen als Nebenprodukt eines Navigationsvorgangs an. Es gibt aber auch eine Vielzahl weiterer Gründe, seine Wege aufzuzeichnen, was in einer großen Bandbreite von Nutzerinnen und Nutzern resultiert. Die aufgezeichneten Routen werden durch Datenprovider gesammelt, anonymisiert, zusammengefasst und für den Kunden – d.h. beispielsweise mit einer Gemeinde oder Kommune – auf eine abgestimmte Karte projiziert. Es handelt sich beim fertigen Produkt also zumeist nicht um Routendaten, sondern um eine sehr feine Kartierung von Verkehrsmengen. Sie gleichen den Verkehrsmengenkarten im MIV in ihrer Darstellung von Kantenbelastungen der mit dem Rad zurückgelegten Wege (siehe Abbildung 1).

Die angebotenen Daten bestehen aus der Verortung von Koordinaten in Längen- und Breitengraden zu einem jeweiligen Zeitstempel mit einer Aufzeichnungsrate von im Regelfall 1 Hz. Geringere Raten sind möglich und werden durch einige Endgeräte bestimmt, sind aber für den Planungsfall nicht unbedingt sinnvoll. Aus diesen verorteten Zeitstempeln lassen sich anschließend mittels einfacher Berechnungen Entfernungen oder Geschwindigkeiten ableiten. Strava unterscheidet beispielsweise zwischen:

- » **Kantendaten:** Referenzierung von Verkehrsmengen auf ein **GIS-Element** (Kante),
- » **Knotenpunktdaten:** Referenzierung von Verkehrsmengen und Wartezeiten) auf ein **GIS-Element** (Knoten).

*Nutzungs-
möglichkeiten*

Hintergrund

Datenform

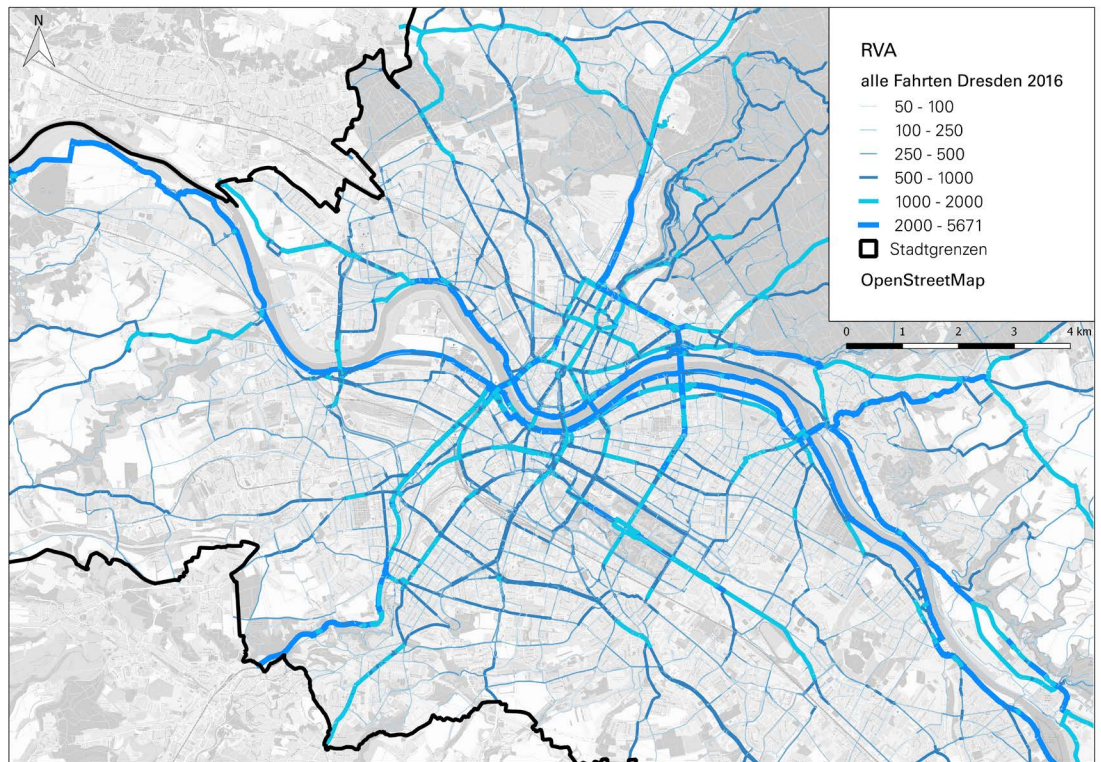


Abbildung 1: Radverkehrsaufkommen (RVA) aller Strava-Nutzerinnen und -Nutzer für Dresden im Zeitraum 06/2015 - 06/2016

Datenschutzgesetz

Der Datenschutz spielt bei der smartphonegestützten GPS-Datensammlung eine wichtige Rolle. In Deutschland wird dieser in den einzelnen Bundesländern meist durch das Landesdatenschutzgesetz geregelt. Hierbei ist der Anwendungsbereich durch die Verarbeitung personenbezogener Daten durch Behörden und sonstige der Aufsicht des Landes unterstehenden juristischen „Personen des öffentlichen Rechts“ gekennzeichnet. Wichtig ist hierbei der Begriff „personenbezogene Daten“. Das Sächsische Datenschutzgesetz beispielsweise definiert diese wie folgt: „personenbezogene Daten sind Einzelangaben über persönliche oder sächliche Verhältnisse einer bestimmten oder bestimmbarer natürlichen Person“*. Also findet das Datenschutzgesetz immer dann Anwendung, wenn aus einem Datensatz eine natürliche Person bestimmbar ist und dieser Person Daten zugeordnet werden können. Folglich bieten die Datenprovider in der Regel anonymisierte Datensätze an, bei denen eine Zuordnung von Nutzungsmustern zu bestimmten Personen für die Kunden nicht möglich ist.

Wie funktioniert die Anonymisierung?

Daten-anonymisierung

Wichtig zu wissen ist, dass keine Routendaten an den Kunden, in dem Fall die Kommune, übermittelt werden, sondern ausschließlich aggregierte (zusammengefasste) Daten auf den Kanten des genutzten GIS-Netzes. Diese Routeninformationen können auch einzeln, auf die GIS-Kanten referenziert, als Basisdaten übergeben werden. Hier ist es zumindest theoretisch möglich, dass anhand der Zeitstempel in wenig belasteten Netzen ein Routenverlauf rekonstruierbar wäre. Um hier einen drohenden Personenbezug zu umgehen, werden jeweils die ersten und letzten 100 Meter der Route vor der Übermittlung durch den Datenprovider entfernt. In den überwiegenden Fällen werden Kommunen jedoch fertige Kartenprodukte erwerben, welche sich nicht im Geltungsbereich des Datenschutzgesetzes befinden. Für nähere Informationen empfiehlt sich der Blick in das jeweilige Datenschutzgesetz des entsprechenden Bundeslandes. In den meisten Fällen sind diese online zu finden.

* Sächsisches Datenschutzgesetz vom 25. August 2003 (SächsGVBl. S. 330), das zuletzt durch Artikel 17 des Gesetzes vom 29. April 2015 (SächsGVBl. S. 349) geändert worden ist.

Welche Datenschutzbedenken haben Nutzerinnen und Nutzern?

Nicht nur die gesetzlichen Rahmenbedingungen geben Anlass zu weiterführenden Datenschutzüberlegungen. Seit Beginn der umfangreichen Datensammlung und Vernetzung ist die Datensicherheit ein wichtiges und viel umstrittenes Thema in Politik, Wirtschaft und Gesellschaft. Durch Smartphones, die nicht nur ständig nah am Körper getragen werden, sondern auch Bewegungs- und Nutzungsdaten aufzeichnen, speichern und übertragen können, erreicht die Debatte eine neue Dimension. Wenig verwunderlich ist, dass eine Nutzung von GPS-Daten, die für einen sachfremden Zweck gesammelt und weiterveräußert wurden, für Kritik sorgen kann. Noch fehlt die nötige Routine im Umgang mit digital generierten persönlichen Daten, so dass die Auswirkungen und Verwendungszwecke vielen Menschen bisher nicht vollständig bewusst sind. Auch zeigt aktuelle Forschung, dass Menschen ganz unterschiedlich mit Datenschutz und neuen Medien umgehen. Hierbei spielen Vertrauen, persönlicher Mehrwert und generationsspezifische Faktoren eine wichtige Rolle. Einen vielversprechenden Ansatz zur Partizipation der Bürgerinnen und Bürger bietet beispielsweise die Initiative Radwende, auf welche in Kapitel 3 eingegangen wird.

Zwar nehmen die Nutzerinnen und Nutzer die Nutzungsbedingungen der Applikationen bei der Installation an, dennoch ist vielen nicht bewusst, dass die von ihnen generierten GPS-Daten für sachfremde Zwecke an Dritte weiterveräußert werden können. Dies gilt es bei der Arbeit mit entsprechend sensiblen Daten nicht außer Acht zu lassen. Nutzerbefragungen haben darüber hinaus ergeben, dass besonders zu Forschungs- oder Planungszwecken Daten bereitwilliger weitergegeben werden, als zur kommerziellen Verwendung. So kann ein transparenter Umgang mit Erwerb und Verwendung vertrauensfördernd wirken.

2.2 Welche Daten werden angeboten?

Das Feld möglicher GPS-Daten ist durchaus vielfältig. Aktuell sind die gebräuchlichsten Lösungen die Daten von App-Anbietern mit ganz unterschiedlicher Nutzerstruktur und Datenlage, wie z.B. BikeCitizens oder Strava, Daten aus Fahrradverleihsystemen oder Daten lokaler Initiativen, wie beispielsweise des ADFC.

Smartphone-Applikationen, wie Strava oder BikeCitizens, generieren in der Regel **Routendaten**, die mit demographischen Angaben des Nutzers aus der Applikation gemeinsam in Datenbanken gespeichert werden. Diese Routendaten enthalten allerdings datenschutz sensible Informationen, wie Wohnort oder Arbeitsplatz des Anwenders, die außerdem mit Profilinformatoren, wie Name, Alter, Geschlecht und anderen freiwillig übermittelten Informationen, verknüpfbar sind. Die Anbieter verpflichten sich bei der Datenweitergabe an Dritte, diese entsprechend der Datenschutzgesetze und allgemeinen Geschäftsbedingungen zu anonymisieren. Folglich erwirbt der Käufer bereits fertig aggregierte Daten, die keine Rückschlüsse auf die Datenerzeuger zulassen. Anonymisierte demographische Informationen, wie Geschlecht, Alter, können hierbei im Datensatz bestehen bleiben. Die Daten globaler Anbieter von Smartphone-Applikationen bieten dabei die größte Spannweite und Anzahl möglicher Nutzerinnen und Nutzer. Hier können sich größere Unterschiede innerhalb der Nutzerstruktur offenbaren. Die Daten werden sekundlich erhoben, nach Fahrtende gespeichert und auf einen Server übermittelt. Diese können durch Nutzerinnen und Nutzer auf dem Smartphone eingesehen werden und mit anderen geteilt werden. Aus diesem sozialen Faktor speist sich z.T. auch die Nutzer motivation von sportlich motivierten Applikationen, wie Strava, beispielsweise die zurückgelegte Route mit anderen zu teilen oder ein Trainingstagebuch zu führen. Grundsätzlich kann zwischen dem Hintergrund der App in jene mit sportivem Hintergrund und jene

*Umgang mit
Datenschutz-
bedenken*

*App-Nutzungs-
bedingungen*

*GPS-Radverkehrs-
daten*

*Smartphone-
generierte
Nutzerdaten*

Datenlieferanten

mit einem anderen Motivationsmodell unterschieden werden. Zu letzterer Gruppe gehören routingfähige Apps, wie Naviki oder BikeCitizens, die ihren Mehrwert für Nutzerinnen und Nutzer durch Navigationsfunktionen aber auch andere Motivatoren, wie soziale Verknüpfung oder Belohnungssysteme, haben. Bekannte Angebote gibt es nach heutigem Stand (08/2017) von Strava, BikeCitizens und Scholz&Volkmer.

Von wem kommen die Daten?

Bei der Auswahl des Datenlieferanten sollte sich also ausreichend Zeit für grundlegende Gedanken über die zu erwartenden Nutzerstrukturen, Datenquellen und Motivationsmodelle genommen werden, da diese einen direkten Einfluss auf die Datenlage, die Stichprobe und die Interpretationskriterien haben.

Durch personalisierte Benutzerprofile in Smartphone-Applikationen können die Datenanbieter in der Regel sehr präzise Angaben über die Nutzerstruktur ihrer Dienste machen. Entsprechend dem Einsatzgebiet der Daten in der Radverkehrsplanung sollte die Stichprobe auf Basis von Erfahrung und vorliegenden Daten aus anderen Quellen analysiert und in Hinblick auf ihre **Repräsentativität** bewertet werden. Dies bedeutet nicht unbedingt, dass die Nutzergruppe der breiten Radfahrerstruktur der Kommune entsprechen muss, wie auch das Anwendungsbeispiel in Kapitel 4 verdeutlichen soll. Bei einer starken Diskrepanz der Nutzergruppe können allerdings entsprechende Anpassungen der Daten notwendig sein, um eine belastbare Interpretation zuzulassen.

App-Anbieter

Im durchgeführten Forschungsprojekt wurde hierfür beispielhaft die **Reliabilität** der erhältlichen GPS-Datensätze von der Firma Strava qualitativ und quantitativ untersucht. Die sportlich motivierten Nutzerinnen und Nutzer bei Strava unterscheiden sich beispielsweise vom Durchschnittsradfahrenden, da sie eher dem Spektrum der ambitionierten „Poweruser“ zuzuordnen sind. Dies spiegelt sich vor allem in einer größeren durchschnittlichen Wegelänge, Radnutzungshäufigkeit und einer höheren Geschwindigkeit wieder. Wegen der starken Motivation durch die Gemeinschaft und der eigenen Leistungskontrolle ist hier ein sorgfältiges und umfangreiches Fahrtracking zu erwarten. Im Verlauf des Leitfadens werden weitere Hinweise zu diesem wichtigen Aspekt ausgeführt. Der navigationsfokussierte App-Anbieter BikeCitizens hingegen spricht eher Alltagsfahrer an. Diese entsprechen möglicherweise in Fahrweise und Gewohnheiten stärker dem Durchschnitt, sind aber unter Umständen weniger stark intrinsisch motiviert, sämtliche Strecken detailliert von der Applikation aufzeichnen zu lassen. Andere App-Anbieter sind beispielsweise komoot, Naviki und BikeMap, die allerdings zum jetzigen Zeitpunkt keine, den Autoren bekannte, Daten für die Planung anbieten.

Lokale Apps

Neben den global angebotenen Smartphone-Applikationen gibt es noch eine Anzahl lokaler Varianten. Einige Bundesländer bieten Radroutenplaner-Apps an und auch Apps von Verkehrsverbänden berücksichtigen Fahrräder (z.B. MVV oder VVS). Daneben existieren eine Reihe von lokalen Initiativen, wie die Initiative Radwende in u.a. Mainz und Wiesbaden, RADschlag in Düsseldorf, die BikeApp des ADFC Verbandes Speyer oder BBBike, die eine Datengrundlage bieten können. Eine eigene Datenerhebung, wie beispielsweise bei der Initiative Radwende, ist jedoch mit großem personellem Aufwand und Know-How verbunden. Hier ist nicht allein die Programmierung einer App ein Problem, sondern vielmehr die dauerhafte Akquise einer ausreichend großen Stichprobe. Auch die Datenverarbeitung und das aggregierte Projizieren der Routendaten auf ein Wegenetz ist als aufwändig zu bezeichnen. Nähere Angaben dazu werden in Kapitel 4.2 gemacht. Für die zusätzliche Differenzierung der Ergebnisdaten sowie deren Plausibilisierung und Aufbereitung für eine Visualisierung sind zudem erweiterte Kenntnisse für die Anwendungen von **Datenbanken**, wie SQL oder PostgreSQL, nötig.

Wichtig ist es zu beachten, dass bei Daten, die auf Smartphone-Nutzung basieren, Nutzerinnen und Nutzer ohne ein Smartphone systematisch ausgeschlossen werden. Ein weiterer möglicher Datenlieferant mit einer völlig anderen Nutzerstruktur können beispielsweise Fahrradverleihsysteme sein. In den Fahrrädern verbaute Chips ermöglichen für einige Systeme die Ortung der Räder. Hierbei ist insbesondere auf die Aufzeichnungsintervalle zu achten. Viele Verleihsysteme zeichnen nur den Startort und das Ende der Fahrten auf. Dies liefert vor allem für den Fall Start=Ziel wenig bis keinen Aufschluss über die gewählten Routen. Auch ist zu vermuten, dass Leihfahrräder vorwiegend in Innenstädten oder für fest definierte Kurzstrecken im Rahmen einer Wegekette genutzt werden. Folglich können die regulären täglichen Wege eines monomodal Radfahrenden nicht mit abgebildet werden. Jedoch können Daten aus Verleihsystemen andere Datensätze sinnvoll ergänzen. Auf diesem Gebiet findet aktuell eine sehr dynamische Entwicklung statt. Anbieter, wie Mobike, YoBike, oBike und Gobeebike, drängen in deutsche Städte und haben neben der Nahmobilität auch die Auswertung von Nutzerdaten im Fokus. Wie genau diese Angebote aussehen werden, ist aktuell noch nicht absehbar – jedoch ist in naher Zukunft mit weiteren Datenquellen zu rechnen.

Zukünftig können auch Apps mit Belohnungssystemen für zurückgelegte Kilometer, wie beispielsweise Radbonus, eine Rolle in der Radverkehrsplanung spielen. Entweder direkt über eine eigene Datenaufzeichnung oder mittelbar als motivatorisches Element im Zuge der Nutzung anderer Tracking-Apps. Hier ist eine Vernetzung denkbar und sinnvoll. Weitere Akteure sind unter anderem Krankenkassen. Hier gibt es erste Ansätze durch die AOK aktive Mobilität zu belohnen. Weitere potenzielle Akteure sind Hersteller von Fahrradnavigationssystemen, wie Garmin, Teasi oder Falk. Diese liefern zwar Daten mit einem starken Freizeitfokus, können aber im Bereich Stadtmarketing oder Tourismus sowie im ländlichen Raum interessante Einblicke geben. Durch die sehr schnelle Entwicklung des Marktes, der Technologie und der Preisgestaltung in diesem Bereich kann an dieser Stelle keine vollständige Übersicht über alle Anbieter am Markt gegeben werden. Eine eigene Anbieterrecherche kann interessante neue Lösungen ausfindig machen und auf Marktentwicklungen reagieren.

Wie kann man die Daten beziehen? Wie muss man sie ausschreiben?

Es gibt verschiedene Wege, die GPS-Daten bei den Anbietern zu erwerben. Besonders durch den noch wenig erschlossenen Markt befinden sich auch die Anbieter noch

in einer Testphase und gehen sehr individuell auf Kundenwünsche ein. Das unter Berücksichtigung der Strukturen in vielen Kommunen und in Zusammenarbeit mit Radverkehrsbeauftragten entwickelte, empfohlene System ist in Abbildung 2 illustriert. Entsprechend dieses Ablaufs bleibt der personelle Aufwand auf Seiten des Auftragsgebers gering und die fachlich spezifischen Anforderungen für die Aufbereitung der GPS-Daten liegen beim Anbieter.

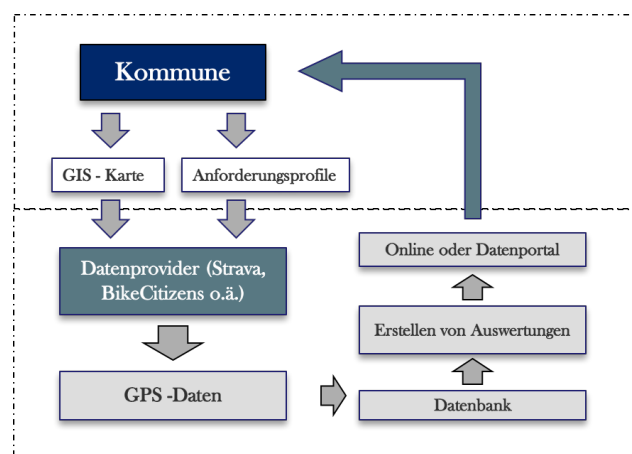


Abbildung 2: Empfohlene kommunale Herangehensweise zum Datenbezug

Einige Anbieter sind bereits darauf eingerichtet, die Daten intern aufzubereiten und niederschwellig zu Verfügung zu stellen. Die Kommune übermittelt dem Anbieter



Abbildung 3: Beispieldarstellung Strava-Metro (Quelle und Copyright: Strava Inc.)

eine geeignete GIS-Karte der Wegenetze (siehe Kapitel 4) und die individuellen Anforderungsprofile entsprechend der Fragestellung. Der Anbieter bereitet dann im nächsten Schritt die Daten vollständig auf und stellt diese online oder in einem Datenportal zur weiteren Nutzung zu Verfügung (siehe Abbildung 3). Im Planungsprozess kann nun auf die Datensätze zugegriffen und je nach Wunsch ausgewertet werden. Wichtig ist es, im Verlauf der Datenakquise klare Anforderungen zu definieren. Sollte zu Beginn des Prozesses die Entscheidung getroffen werden, sich auf vorgefertigte Datenauszüge und -ansichten zu beschränken, ist hier bereits darauf zu achten, dass diese alle gewünschten Informationen

enthalten. Auch dem Wegenetz kommt dabei eine hohe Bedeutung zu. Dieses sollte vorab unbedingt auf etwaige Netzlücken geprüft werden. Mehr Informationen dazu sind in Kapitel 4 zu finden.

Bei der Nutzung von Datenportalen sind unterschiedliche Herangehensweisen möglich. Grundsätzlich wird als Darstellungsform aber immer eine Kartenform gewählt, bei der über Schalter Informationen ein- oder ausgeblendet werden können. Diese sind dann on-click oder als Mouseover für den Fachnutzer ersichtlich. Die Firma Strava bietet dabei neben der Gesamtfahrtanzahl auch Angaben zu Alltagsfahrten, Nutzerzahlen oder Wartezeiten an. Außerdem können auch Quelle-Ziel-Beziehungen analysiert werden.

Die Ausgabe von statistischen Daten für einen bestimmten Zeitraum ist hier in Absprache mit dem Datenanbieter möglich. Eigene Aggregationen oder Auswertungen sind jedoch nur bis zu einer bestimmten Detailtiefe möglich. Die Ausgabe der Auswertung erfolgt über einen Internet-Browser, so dass keine spezielle Software auf Seiten der Kommunen nötig ist. Eine Einbindung von Dauerzählstellen wird in naher Zukunft nicht möglich sein. Jedoch können für Quellverkehrszellen zukünftig die darin beginnenden und endenden Routen aggregiert dargestellt werden.

BikeCitizens gehen mit ihrem Bike Analytics Tool (siehe Abbildung 4) dabei einen etwas anderen Weg. Hier steht nicht allein die makroskopische Analyse im Vordergrund, sondern werden on-click-Befehle sinnvoll mit mikroskopischen



Abbildung 4: Beispieldarstellung Bike Citizens Analytics (Quelle und Copyright: BikeCitizens)

Informationen verknüpft. Dabei sind unterschiedliche Auswertungen möglich: Neben Radverkehrsaufkommen können so Routen, bei denen ein markiertes Netzelement enthalten ist, sowie einzelne Tagesgänge visualisiert werden. Diese Form der Analyse bietet deutlich tiefere Einblicke in Verhaltensmuster der Radfahrenden und erlaubt detailliertere Auswertungen entlang einzelner Kanten im Wegenetz.

2.3 Welche Parameter können untersucht werden?

Verkehrsmengen geben die Anzahl an Radverkehrsteilnehmern pro Zeiteinheit an. Da bei smartphone-basierten Erhebungen nur ein Teil der Radfahrenden an der Datenerfassung teilnimmt, ist eine anschließende Hochrechnung auf die Gesamtheit der Radfahrerinnen und Radfahrer notwendig, um Schlüsse für alle Radfahrenden in einem Gebiet ziehen zu können. Ein beispielhaftes Hochrechnungsverfahren, validiert anhand von Daten aus Zählstellen in der Pilotkommune Dresden, ist in Kapitel 4.2, Schritt 5 aufgezeigt.

Vor dem Hintergrund der Nutzbarkeit der Daten für die Radverkehrsplanung ist es sinnvoll, die genaue Definition der Pendler-/Alltagsfahrten durch den jeweiligen Anbieter zu betrachten, da diese unter Umständen auch regelmäßig wiederkehrende Freizeit- oder Sportfahrten mit einschließt. Der Anteil der reinen Sportfahrten sollte innerhalb eines Datensatzes dabei immer möglichst gering sein, beziehungsweise sollten diese wirksam markiert und gefiltert werden können.

Da aus Datenschutzgründen nicht die vollständigen Fahrten der Nutzerinnen und Nutzer sondern nur aggregierte Daten bezogen werden können, sind auf der Ebene der Netzkanten nur Verkehrsmengen nicht aber Routenwahlentscheidungen oder Quellen und Ziele der Radfahrenden vorhanden. Diese Einschränkung lässt sich umgehen, indem die Daten vom Anbieter in Blöcken zusammengefasst werden. Diese Blöcke können beispielsweise Postleitzahlenbereiche oder Ortsamtsbereiche, Verkehrszellen oder Blockkarten sein. Während Postleitzahlenbereiche und Ortsamtsbereiche nur eine sehr grobe Zuordnung zulassen, sind Blockkarten so fein aufgelöst (50m und weniger), dass eine Anonymisierung nicht mehr unbedingt gegeben ist. In Städten mit Verkehrsmodellen liegen meist Verkehrszellen vor, die sich je nach Qualität in der Auflösung der Ortsamtsbereiche oder etwas feiner bewegen. Diese bieten eine Möglichkeit zum Aufbau einer Quelle-Ziel-Matrix der Radfahrenden, liegen jedoch bundesweit unterschiedlich aufgelöst oder gar nicht vor.

Für die Bildung eines Quelle-Ziel-Netzes wurde im Rahmen des Forschungsprojektes das einheitliche europäische Gittersystem für amtliche Statistik verwendet. Dazu wurden Polygone als regelmäßige Quadrate mit einer Kantenlänge von 1.000m gebildet. Dieses Verfahren genügt zum Einen den Bestimmungen des Datenschutzes und bietet zum Anderen eine gute überregionale Vergleichbarkeit der Ergebnisse. Die genaue Methodik wird im Anwendungsbeispiel in Kapitel 4.2, Schritt 4 erläutert.

Für jede Kante im Wegenetz liegen richtungsgetrennt zwei unterschiedliche Geschwindigkeiten vor. Die berechnete Gesamtgeschwindigkeit sollte dabei, wenn sie separat ausgewiesen wird, dem über die Verkehrsmenge gewichteten Mittelwert beider Richtungen entsprechen. Bei Steigungsstrecken sind große Differenzen der Richtungsgeschwindigkeiten zu beachten. Da innerhalb der Basiskarten meist kein Muster für die Festlegung der Digitalisierungsrichtung erkennbar ist, ist eine Darstellung aller Bergab- bzw. aller Bergan-Fahrten mitunter nicht einfach lösbar. Hier sind dennoch getrennte graphische Darstellungen per Mouseover denkbar.

GPS-basierte Wartezeiten sind geeignete Indikatoren für die Knotenpunktbewertung aus Radfahrersicht. Die Daten der Firma Strava waren ohne starke, korrigierende Aufbereitung allerdings nicht nutzbar. Mit einigen Eingriffen erlauben sie trotzdem zumindest eine oberflächliche Bewertung von Knotenpunkten anhand des Anteils an ungehinderten Durchfahrten (siehe Projektbericht). Vielversprechende Ansätze diesbezüglich verwendet der Anbieter BikeCitizens. An dieser Stelle ist es wichtig, bei der Datenübergabe darauf zu achten, ob tatsächlich Wartezeiten oder eher Durchfahrtszeiten ausgegeben werden. Ausschließlich georeferenzierte Modelle

Verkehrsmengen

*Pendler-/
Alltagsfahrten*

*Quelle-Ziel-
Matrix*

Geschwindigkeit

Wartezeiten

zur Berechnung der Wartezeit geben meist nur Summen aus Wartezeit und Durchfahrtszeit aus. Dies resultiert, wie im Fall der Strava-Daten, in deutlich zu wenigen ungehinderten Durchfahrten. Im Idealfall sollte der Datenprovider für die Berechnung von Wartezeiten neben der Georeferenzierung um den Knotenpunkt und den entsprechenden Zeitstempeln auch auf Geschwindigkeiten und Beschleunigungswerte der Radfahrenden zurückgreifen (siehe Abbildung 5).

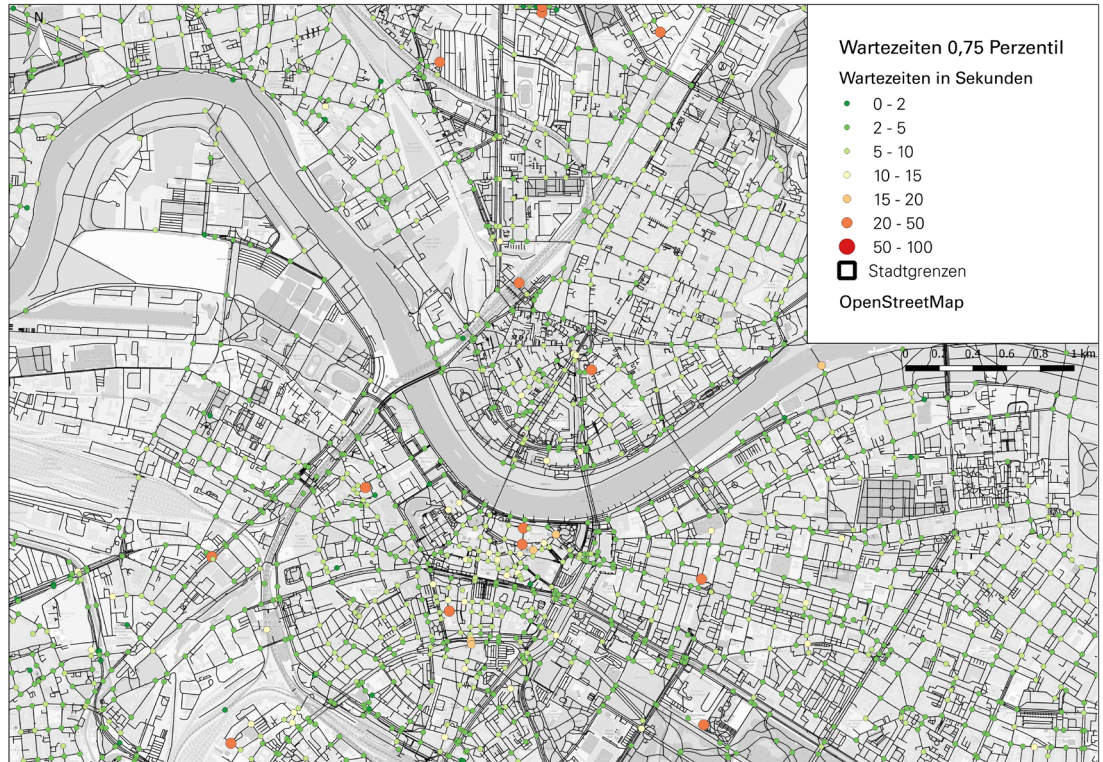


Abbildung 5: Wartezeiten aller Strava-Nutzerinnen und -Nutzer im Zeitraum 06/2015 - 06/2016

Welche Möglichkeiten bieten 3 GPS-Daten?

GPS-Daten bieten die Möglichkeit zu analysieren wo, wann und wie sich Radfahrende im gesamten Netz bewegen. Dabei sind verschiedene Anwendungsfelder der in Kapitel 2 vorgestellten GPS-Daten, wie beispielsweise Netzplanung, Anlagenplanung, Prioritätensetzung, sowie Monitoring und Evaluation denkbar. Dieses Kapitel gibt eine Übersicht über diese Anwendungsfelder und schlägt einige Begleitmaßnahmen vor (Kapitel 3.6), die darauf abzielen, die Bürgerinnen und Bürger in den Prozess der Datensammlung und -verwendung mit einzubeziehen.

3.1 Überblick

GPS-Daten bieten eine Vielfalt an Anwendungsbereichen, die durch herkömmliche Datenquellen nur unzureichend oder aufwändig abgedeckt werden können. Im Folgenden werden die verschiedenen Anwendungsbereiche und die entsprechenden Parameter im GPS-Datensatz vorgestellt.



Anwendungsmöglichkeiten

Im Folgenden dargestellt ist eine Übersicht zu den Anwendungsmöglichkeiten der drei Parameter Verkehrsmenge, Quelle-Ziel-Matrix und Geschwindigkeit (vgl. Kapitel 2.4) für die Radverkehrsplanung. Eine detaillierte Beschreibung findet sich in den Unterkapiteln 3.2-3.5.

Tabelle 1: Anwendungsmöglichkeiten von GPS-Daten nach den Parametern Verkehrsmenge, Quelle-Ziel-Matrix und Geschwindigkeit

	Verkehrsmengen	Quelle-Ziel Matrix	Geschwindigkeiten
Netzplanung	✓	✓	
Anlagenplanung		✓	✓
Priorisierung von Maßnahmen	✓	✓	
Monitoring & Evaluation	✓		✓

3.2 Netzplanung

Für die Planung von Zielnetzen oder die Erarbeitung einer Radverkehrskonzeption können GPS-Daten einen wertvollen Input liefern. Sie können zeigen, welche Wege Radfahrende wählen, wie groß die Nachfrage auf bestimmten Strecken ist und wo Bedarf für neue Infrastrukturen besteht.

Vor allem über Radverkehrsströme in **Nebennetzen** herrscht oft eine schlechte Datenlage, da Dauerzählstellen meist an Haupttrouten liegen und Kurzzeitzählungen häufig nicht im erforderlichen Umfang vorliegen. Innerhalb bereits konzipierter Radverkehrsnetze ist es möglich, Netzlücken zu erkennen und zu schließen bzw. auf Basis der Daten aus **städtischen Geoinformationssystemen** ungenügende Anlagenqualitäten zu identifizieren und zu beheben. Für diese Schritte sind vor allem die Radverkehrsmengen auf den Netzkanten von Belang. Zukünftig wird voraussichtlich bei steigenden Nutzerzahlen auch die Identifikation von Kapazitätsengpässen nach erfolgter Hochrechnung möglich sein (beispielhafte Hochrechnungen siehe Kapitel

*Zielnetze und
Radverkehrskonzeption*

*Netzlücken und
Nebennetz*

Quelle-Ziel-Relationen

4). Routendaten können hier auch dezidiert Auskunft über gefahrene Umwege und fehlende Netzelemente geben. Bei vielen verfügbaren Datensätzen sind diese jedoch nicht enthalten, da im Zug der aus Datenschutzgründen notwendigen Aggregation die Informationen zu Einzelfahrten verloren gehen. Dass derartige Auswertungen jedoch bereits möglich sind, zeigen beispielsweise Angebote von Bikeprint.nl oder BikeCitizens.

Angebotsplanung

Abhilfe kann zunächst die Nutzung der **Quelle-Ziel-Matrizen** liefern. Häufig nachgefragte **Quelle-Ziel-Relationen** bieten Informationen zu einem notwendigen Ausbau von Verbindungen und weisen auf Orte hin, an denen die Anlagenqualität überprüft werden sollte. Infrastrukturmängel zeigen sich vor allem für die Querung linienhafter Elemente mit großer Zerschneidungswirkung, wie beispielsweise Flüsse, oder Bahndämme. Mitunter schon häufig genutzte Quelle-Ziel-Relationen können dabei Auskunft über die Notwendigkeit weiterer Querungselemente geben.

Bei der Netzplanung sollte jedoch trotz allem auch die Angebotsplanung für Orte ohne erkennbare Nachfrage als Mittel der Wahl gelten, denn Nichtnutzer werden, wie bei fast allen anderen Erhebungsmethoden auch, wenig bis gar nicht abgebildet.

3.3 Anlagenplanung

Geschwindigkeitsniveau als Indikator

Anhand von niedrigeren Geschwindigkeitsniveaus bei vergleichbaren ebenen Netzelementen lassen sich Qualitätslücken, wie beispielsweise mangelhafte Oberflächen oder überlastete Infrastruktur, erkennen. Der Parameter Geschwindigkeit kann damit bereits erste Erkenntnisse über eventuelle Verbesserungspotentiale an konkreten Orten geben. Dadurch können Verbesserungspotentiale für bestimmte Abschnitte mit dem Gesamtnetz in Bezug gesetzt werden.

Bei der Nutzung des Parameters Geschwindigkeit ist die Nutzergruppe der verwendeten **App** zu beachten. Hierfür ist zumindest bei Apps mit sportivem Hintergrund die Durchschnittsgeschwindigkeit zu senken. Um das Niveau der Alltagsradfahrenden zu erreichen, sollte man die Geschwindigkeit beispielsweise in Dresden um 5,5km/h korrigieren (siehe Kapitel 4).

Radverkehrsaufkommen

Anhand besonders hoher und niedriger Werte des Radverkehrsaufkommens, welche aus der Quelle-Ziel-Matrix ersichtlich sind, kann auch ein Bedarf an Abstellanlagen geschätzt bzw. überprüft werden. Gleiches gilt für eine Überprüfung der Anlagenqualität entlang besonders häufig genutzter Relationen. Quelle-Ziel-Matrizen können außerdem genutzt werden, um Defizite bei der Anzahl von Abstellanlagen zu erkennen.

Ergebnisprüfung

Weitere Möglichkeiten innerhalb der Anlagenplanung betreffen die Überprüfung von erfolgten Projekten. Hier kann beispielsweise anhand der realisierten Verkehrsmengen visualisiert werden, ob eine Maßnahme das gewünschte Ergebnis erreicht, oder ob beispielsweise weitere begleitende Informations- und Kommunikationsmaßnahmen oder Ausschilderungen nötig sind.

3.4 Priorisierung von Maßnahmen

Verkehrsmengen als ergänzender Indikator

Bei der Prioritätensetzung für eine Umsetzung von Radverkehrsmaßnahmen werden üblicherweise Parameter, wie die Verkehrssicherheit und geplante grundhafte Sanierungen, berücksichtigt. Diese Parameter können nun durch Informationen über die vorhandene Verkehrsmenge ergänzt werden. Gerade für das Nebennetz kann so der Einsatz kostengünstiger Maßnahmen priorisiert werden. Dazu gehören die

Markierung von Schutzstreifen, die Freigabe von Einbahnstraßen in Gegenrichtung sowie das Setzen von Geschwindigkeitsbegrenzungen für den MIV. Vor allem bei Entscheidungen über die Umsetzung vorhandener Radverkehrskonzepte kann auf Verkehrsmengendaten zurückgegriffen werden. Sie können bei der Priorisierung bestimmter Maßnahmen ebenso mit einbezogen werden, wie weitere Parameter, beispielsweise die Entschärfung von Unfallhäufungsstellen. Für Bereiche mit hohem Quell- und Zielverkehr ist vor allem für die Umsetzung von Maßnahmen, die den ruhenden Radverkehr betreffen, eine Priorisierung möglich.

Der Einsatz von GPS-Daten kann neben den üblichen Instrumenten der Radverkehrsplanung weitere Klarheit über einen sinnvollen zeitlichen Verlauf der Maßnahmen bringen, sollte jedoch keinesfalls als ausschließlicher Indikator dienen. So sollte die Verkehrssicherheit selbstverständlich weiterhin als wichtigster Parameter für die Realisierung von Bauprojekten dienen.

3.5 Monitoring & Evaluation

Nach der Umsetzung einzelner Maßnahmen ist eine Evaluation sinnvoll. Größere Änderungen im Verkehrsangebot ziehen dabei meist abbildbare Reaktionen nach sich. So lassen sich anhand der GPS-Daten beispielsweise Verdrängungsverkehre aufgrund von Baustellen ebenso abbilden wie die Zunahme der Verkehrsstärke aufgrund neuer oder veränderter Infrastruktur.

Zur Darstellung lassen sich u.a. Differenznetze verwenden. Ein Beispiel zu einer Differenzkarte im Stadtgebiet Dresden ist in Abbildung 6 dargestellt. Während auf den meisten Strecken eine deutliche Zunahme der Radverkehrsmenge zu erkennen ist, sind einzelne Abschnitte mit einer Abnahme zu identifizieren. Diese Abschnitte weisen auf verändertes Verhalten der Radfahrerinnen und Radfahrer hin. Mögliche Defizite in der Infrastruktur oder Auswirkungen gezielt vorgenommener Veränderungen lassen sich so im Gesamtnetz erkennen. Dabei muss stets auch eine mögliche Zu- oder Abnahme der App-Nutzerzahlen in Betracht gezogen werden. Es bietet sich an, die

*Zeitlicher
Maßnahmen-
ablauf*

*Maßnahmen-
evaluation*

Differenznetze

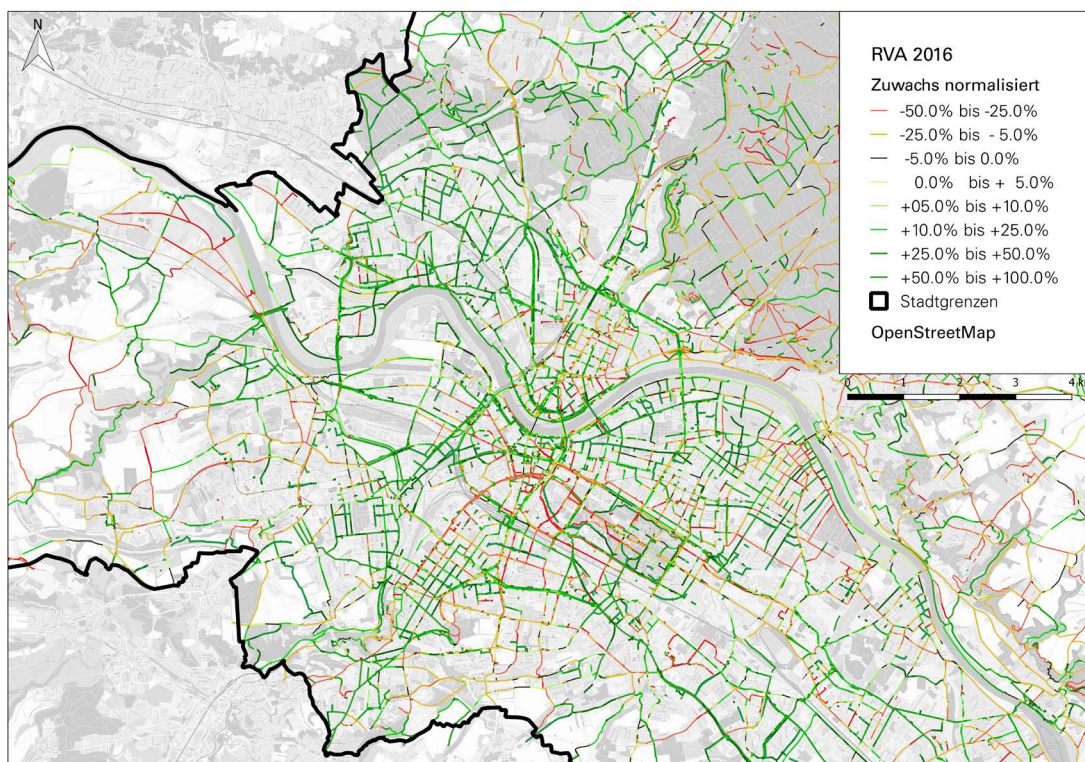


Abbildung 6: Differenznetz Radverkehrsaufkommen (RVA), prozentuale Entwicklung, normiert 01/06/2015 zu 01/06/2016, alle Fahrten

Änderungen des Geschwindigkeitsniveaus

Daten auf ein Basisjahr zu normieren oder prozentuale Werte zu vergleichen. Wichtig ist hierbei vor allem ein konsistentes Wegenetz, denn Vergleiche sind grundsätzlich am besten über einen eindeutigen, konsistenten Index der Netzelemente möglich.

Ein weiterer Indikator für das Monitoring von Maßnahmen kann die Geschwindigkeit der Radfahrenden darstellen, insofern es sich bei der durchgeführten Maßnahme um eine Verbesserung der Oberflächenqualität oder eine Veränderung des Infrastrukturangebotes handelt. So wird zum Beispiel im Idealfall ein separater Radweg einen deutlich schnelleren Verkehrsfluss und somit höhere Geschwindigkeiten aufweisen als eine Führung als gemeinsamer Geh- und Radweg. Aber auch gemeinsame Führungsformen mit dem Fußverkehr können so auf ihre Tauglichkeit geprüft werden. Hohe Geschwindigkeitsniveaus im Bereich gemeinsamer Führung deuten auf eine Untauglichkeit der gewählten Führungsform oder zumindest auf die Notwendigkeit weiterer regulierender Maßnahmen hin.

3.6 Wie können Begleitmaßnahmen aussehen?

Partizipation und Öffentlichkeitsarbeit

GPS-Daten haben durch ihren interaktiven und realitätsnahen Charakter ein hohes Potential für Öffentlichkeitsarbeit und internes Marketing. Herkömmliche Planungswerkzeuge sind für Bürgerinnen und Bürger entweder kaum spürbar (Verkehrszählungen, Geschwindigkeitsmessungen) oder bieten keinen direkten objektiven Mehrwert (Bürgerbefragungen etc.). Da hingegen die GPS-Datengenerierung ausschließlich durch aktive Radfahrerinnen und Radfahrer erfolgt, liegt es nahe, dies auch nach außen zu kommunizieren. Langfristig erhält die Nutzung von GPS-Daten durch den aktiven Einbezug von Radfahrenden einen wichtigen Mehrwert. Eine in der App implementierte direkte Kommunikation zwischen Bürgerinnen bzw. Bürgern und den Planungsstellen reduziert Barrieren, die Ortsbeiratssitzungen oder Planauslagen aufbauen können. Auch weniger fachkundige Bürgerinnen und Bürger können so partizipieren. Dieser Vorteil kommt beim im Forschungsprojekt modellhaft betrachteten Datenanbieter Strava aufgrund der geringen Interaktionsmöglichkeit zwischen App-Betreiber und Nutzerinnen und Nutzern leider kaum zum Tragen. Eine Nutzeransprache durch die Kommunen ist daher schwer realisierbar. Dagegen bieten weitere Mitbewerber am Markt deutlich größere Gestaltungs- und Interaktionsspielräume für die Bürgerbeteiligung an, die zukünftig sicherlich ausgebaut bzw. durch zukünftige Anbieter abgedeckt werden. Einige Beispielstädte zeigen, dass Bürgerinnen und Bürger bereits beim Generieren der GPS-Daten mit eingebunden werden können, indem für die eigene Kommune Apps angeboten werden, die die Wege der Nutzerinnen und Nutzer aufzeichnen und an die Stadt übermitteln. So kann eine persönliche Identifikation mit der Kommune unterstützt werden.

Beispiel-App Radwende

Ein Beispiel dafür ist die App zur *Radwende* der Agentur Scholz&Volkmer in den Städten Mainz und Wiesbaden. Zwar steht diese für eine stark lokal eingefärbte Lösung, die aufgrund ihrer fehlenden Skalierbarkeit zum jetzigen Zeitpunkt weniger geeignet ist, im globalen Stil Daten zu sammeln und aufzubereiten. Die Datenerhebung und Kommunikation ist jedoch aufgrund der ausgezeichneten medialen Begleitung während der Erhebungsphase als ausgesprochenes Positivbeispiel hervorzuheben. Hierbei wurden unter anderem die aufgezeichneten Radrouten zur Kunstform erhoben und visualisiert. Dies geschah vor allem im Zusammenhang mit Akteurinnen und Akteuren der lokalen Kultur, wie beispielsweise dem Staatstheater Mainz. So konnte ganz gezielt eine ansonsten für die App-Nutzung eher weniger affine Zielgruppe angesprochen werden. Zusätzlich wurden durch die Stadt Mainz für Meilensteine angefahrenen Kilometern Maßnahmen für den Radverkehr durchgeführt. Darüber hinaus gab es Anreize in Form von Vergünstigungen beim lokalen Einzelhandel, die ab einer bestimmten Fahrweite in Anspruch genommen werden konnten. Insgesamt

hatte die Kampagne eine sehr starke Außenwirkung, besonders durch die anschauliche Visualisierung der Radverkehrsströme, und erreichte auch sonst für die klassische Kommunikation schwer erreichbare Zielgruppen. Die Auswertung der Daten fand bisher jedoch nicht auf einem höheren Niveau als der Erstellung einer Heatmap statt. Insgesamt wurden in beiden Städten 93.000km und ca. 16.000 Fahrten von den Nutzerinnen und Nutzern aufgezeichnet. Dies kann als klarer Erfolg des Konzeptes gewertet werden.

Ein weiteres Beispiel ist die Bike Track App des ADFC Speyer, die mit der durch die Partizipation der Bürgerinnen und Bürger möglichen Verbesserung der Radverkehrsinfrastruktur wirbt, sowie RADschlag in Düsseldorf. Auch bei RADschlag ist eine Interaktion mit der Stadt Düsseldorf als Betreiberin möglich, um zum Beispiel Infrastrukturdefizite zu melden.

Wie bereits an den Beispielen Mainz und Wiesbaden gezeigt, liegt ein großer Vorteil von GPS-Radverkehrsdaten in der Anschaulichkeit der Datensätze. In Kombination mit GIS-Netzen und Straßenkarten bietet sich die Möglichkeit, verschiedene Fragestellungen zu illustrieren und auch für Laien verständlich zu präsentieren. Neben der Öffentlichkeitsarbeit nach außen können die Datensätze auch für die interne Kommunikation in Verwaltung und Politik genutzt werden. So verdeutlichen Visualisierungen Handlungsbedarfe in der Verkehrsplanung und liefern den beteiligten Akteuren unterstützende und fundierte Argumente für anstehende Entscheidungsprozesse.

3.7 Mögliche Schwächen

Betrachtet man die Anlagen- und Netzplanung, so ist es zum jetzigen Zeitpunkt noch nicht möglich, Nutzerinnen bzw. Nutzer einem befahrenen Infrastrukturelement zuzuordnen. Beispielsweise kann der Nachweis, ob Radfahrende im Mischverkehr oder auf dem freigegebenen Fußweg fahren, bisher nicht zweifelsfrei erbracht werden. Abhilfe schafft hier in den nächsten Jahren möglicherweise das neue Positionssystem Galileo, das eine höhere Genauigkeit verspricht. Mit dem zukünftig zu nutzenden Standard Galileo kann die Genauigkeit von aktuell 6,8m horizontal sogar noch erhöht werden. Dies steigert auch den Nutzen für die Planung. Smartphones der neusten Generation verwenden eine Mischung aus allen empfangbaren Satellitensystemen. So kann auch bei schlechten Empfangsbedingungen eine hohe Genauigkeit erreicht werden.

GPS-Daten sind aktuell noch stark homogene Datensätze. Zu den Nutzerinnen und Nutzern fehlt es bisher an vertiefenden Studien. Hier setzt das Projekt RadVerS der TU Dresden, gefördert im Rahmen des Nationalen Radverkehrsplans 2020, an. Dabei steht die Erforschung der Nutzergruppen im Vordergrund, wodurch diese Wissenslücke in naher Zukunft voraussichtlich kleiner werden wird. Des Weiteren sind auch die Themen Nichtnutzer und systematische Exklusion von Nutzergruppen weitere Forschungsfelder. Hier muss in naher Zukunft nachgewiesen werden, wie eine repräsentative Stichprobe geschaffen werden kann.

Abschließend bleibt noch anzumerken, dass die hohe Motivation durch den Wettbewerbscharakter und den Trainingsmehrwert für die regelmäßige Nutzung von Sport-Apps, wie Strava, bei anderen App-Konzepten nicht unbedingt gegeben ist. Um Nutzerinnen und Nutzer langfristig von der aktiven App-Nutzung zu überzeugen, müssen entsprechende Motivationselemente geschaffen werden. Dies bedeutet einen höheren Aufwand, der von den interessierten Kommunen frühzeitig berücksichtigt und in Informations- und Kommunikationskampagnen eingebunden werden sollte.

*Beispiel-Apps
ADFC Speyer
RADschlag*

*Unterstützung der
internen Kommunikation*

*Genauigkeit der
GPS-Daten*

*Stichproben-
zusammensetzung*

Erhebungsaufwand

4

Anwendungsbeispiel Pilotkommune Dresden

Im folgenden Kapitel werden anhand eines Fallbeispiels Datenerwerb, -auswertung und -interpretation diskutiert, um Planerinnen und Planer auch für die entsprechenden kritischen Aspekte zu sensibilisieren und mögliche Strategien zum Umgang mit den Daten aufzuzeigen. Die Datenüberprüfung am Beispiel der Pilotkommune Dresden wurde quantitativ durch Mapmatching von Radverkehrsdaten aus anderen Messungen und qualitativ durch eine Nutzerbefragung durchgeführt. Verzerrungen durch die sportlich-orientierte Zielgruppe kann mathematisch und qualitativ begegnet werden.

4.1 Projektrahmen

Projekthalt

Für das Projekt „GPS-Daten im Radverkehr“ haben die Professuren für Verkehrsökologie und Verkehrspsychologie an der TU Dresden eine vollständige Analyse des GPS-Datensatzes für die Stadt Dresden exemplarisch durchgeführt. Die Vorgehensweise und einschlägige Ergebnisse werden in diesem Kapitel detailliert dargestellt. Sie können bei der Frage, ob und wie angebotene **GPS-Daten** auch für eine Kommune nutzbar sind, weiterhelfen und vermitteln ein Bild davon, welche Einschränkungen und Besonderheiten in der praktischen Anwendung und Interpretation der Datensätze auftreten können. Detailliertere Ergebnisse und Auswertungsbeispiele aus dem Projekt finden sich im Anhang ‚Detailwissen‘.

Datensatz und Kosten

Die hier vorgestellten Daten wurden nach einer Anbietersauswahl von dem Anbieter Strava bezogen. Dabei wurden zwei verschiedene Datensätze mit einem zeitlichen Umfang von 18 Monaten bezogen. Für die Pilotkommune Dresden (ca. 550.000 Einwohner) besteht der Datensatz dabei aus 70.500 Fahrten von rund 3.200 Nutzerinnen und Nutzern. Die Preisgestaltung ist bei den einzelnen Anbietern unterschiedlich. Strava beispielsweise kalkuliert den Endpreis aktuell anhand der Nutzerzahl je Jahr. Dafür wird momentan ca. 1€ je Nutzerin bzw. Nutzer veranschlagt. BikeCitizens verfolgt ein alternatives Modell, welches mit den beteiligten Kommunen abgestimmt wird und z.B. Kampagnen in Zusammenarbeit mit dem Einzelhandel einbindet. Hier kann jährlich von einem unteren bis mittleren fünfstelligen Betrag ausgegangen werden.

4.2 Anwendungsbeispiel – Vorgehen

Zielsetzung und Richtlinien

Schritt 1: Kann ich GPS-Daten für meine Kommune nutzen?

Für die Pilotkommune Dresden wurde eine umfassende Analyse durchgeführt, welche möglichst viele Fragestellungen der Radverkehrsplanung beantwortet und die Daten in ihrem Umfang bewertbar macht. Dazu wurde vor allem Fragen der **Validierung** nachgegangen. Wie bereits in Kapitel 3 dargestellt, sind die Möglichkeiten in der Radverkehrsplanung darüber hinaus vielfältig. Eine genaue Zielsetzung ist daher zu Beginn des Prozesses notwendig. Gegebenenfalls müssen beim Erwerb der Daten geltende Vergaberichtlinien beachtet werden.

Welche Fragen sollen beantwortet werden?

Definition einer Fragestellung

Bevor Daten erworben werden, ist es notwendig, sich genau über die eigene Fragestellung, die mit diesen Daten beantwortet werden soll, im Klaren zu sein. Diese Fragestellung sollte zu Beginn definiert werden. Dazu sind die Grenzen und Möglichkeiten der Methode, wie sie in diesem Leitfaden beschrieben werden, zu berücksichtigen. Wie bereits in Kapitel 2 erläutert wird, ist der komplette Datensatz

aus Datenschutzgründen nicht zu beziehen und verbleibt stets beim Anbieter, welcher die Daten für den Kunden aufbereitet. Zur Anonymisierung werden dabei die ersten und letzten 100m der Route entfernt und aggregierte Datensätze aus den einzelnen aufgezeichneten Fahrten erzeugt.

Im hier beschriebenen Anwendungsbeispiel in der Pilotkommune Dresden sollten die am Markt angebotenen Daten kritisch untersucht werden, um deren Möglichkeiten für die Radverkehrsplanung aufzuzeigen. In einem ersten Schritt wurden dazu die folgenden Anforderungen an die Daten zur Verwendung in der Radverkehrsplanung definiert:

- » Radverkehrsstärke
- » Geschwindigkeiten der Radfahrenden
- » Sozio-demografische Angaben zu den App-Nutzenden
- » Wartezeiten an **Knotenpunkten**
- » **Quelle-Ziel-Matrizen**
- » Möglichkeit zu einer Aufteilung der Daten in **Zeitscheiben**, um Ganglinien anhand von Zählstellen vergleichen zu können

Neben diesen Minimalanforderungen sollten als zusätzliche Kriterien die flächenhafte Ausdehnung der GPS-Daten, die Nutzerzahl und schlussendlich der Preis in die Auswahlentscheidung einfließen. In Kapitel 3.1 findet sich eine Orientierung, welche Daten für welche Analysen benötigt werden. Obwohl die detaillierte Datenlage umfangreiche Auswertungsmöglichkeiten zulässt, gehen mit der Stichprobe und dem Erhebungsverfahren einige Einschränkungen einher (siehe Kapitel 2.2).

Schritt 2: Welche Anbieter gibt es?

Auf Grundlage der in Schritt 1 definierten Anforderungen, wurden die Angebote der verschiedenen Anbieter bewertet. Während zum Zeitpunkt der Analyse (Stand 02/2016) nur wenige Anbieter die erhobenen Nutzerdaten zum Verkauf stellten, kann eine stetig wachsende Zahl von Akteuren am Markt mit vielfältigen Datenangeboten und Aufbereitungen verzeichnet werden. Lediglich die Daten der Firma Strava entsprachen zum Analysezeitpunkt in Umfang und Güte den Anforderungen an einen für die Radverkehrsplanung geeigneten Datensatz für Dresden. Daher fiel der Entschluss zum Kauf von zwei Datensätzen:

Datensatz 1: Bestehend aus 439.570 Fahrten von 19.615 Nutzerinnen und Nutzern im Zeitraum 10/2014 bis 06/2016 in den Bundesländern Sachsen, Berlin und Brandenburg. Dieser Datensatz dient der überregionalen Vergleichbarkeit der Ergebnisse.

Datensatz 2: Bestehend aus 70.500 Fahrten von 3.200 Nutzerinnen und Nutzern im Zeitraum 06/2015 bis 06/2016 in Dresden. Dieser Datensatz bildet die Grundlage des hier vorgestellten Anwendungsbeispiels.

Um die Eignung der Daten für die vom Kunden gewählte Fragestellung bereits im Vorfeld zu überprüfen, ist es sinnvoll den ausgewählten Anbieter nach einem Beispieldatensatz zu fragen. So kann dieser bereits im Vorfeld des geplanten Datenerwerbs in einer groben Analyse mit Erfahrungswerten abgeglichen und so auf generelle Plausibilität und mögliche Verzerrungen geprüft werden.

Wer sind die App-Nutzenden?

Eine Darstellung der Nutzerinnen und Nutzer in Dresden soll ein erstes Bild der möglichen Nutzerschicht geben. Während einige Apps mit eher sozialen Motivatoren arbeiten und daher eine etwas breiter gefächerte Nutzerschicht aufweisen, arbeitet

*Anforderungs-
profil der Daten*

Anbieterauswahl

Beispieldatensatz

App-Nutzerprofil

Strava ausschließlich mit sportiven und kompetitiven Anreizen. Dementsprechend gibt es noch Vorbehalte bei der Nutzung der Daten durch die Ämter, da die Daten eher von jungen, sportlich fahrenden, männlichen Radfahrern zu stammen scheinen. Die Altersverteilung in Dresden fällt jedoch besser aus als anfangs vermutet und weist Nutzerinnen und Nutzer in allen Altersgruppen auf (siehe Abbildung 7).

Altersverteilung der Strava-Nutzerinnen und -Nutzer in der Beispielstadt Dresden

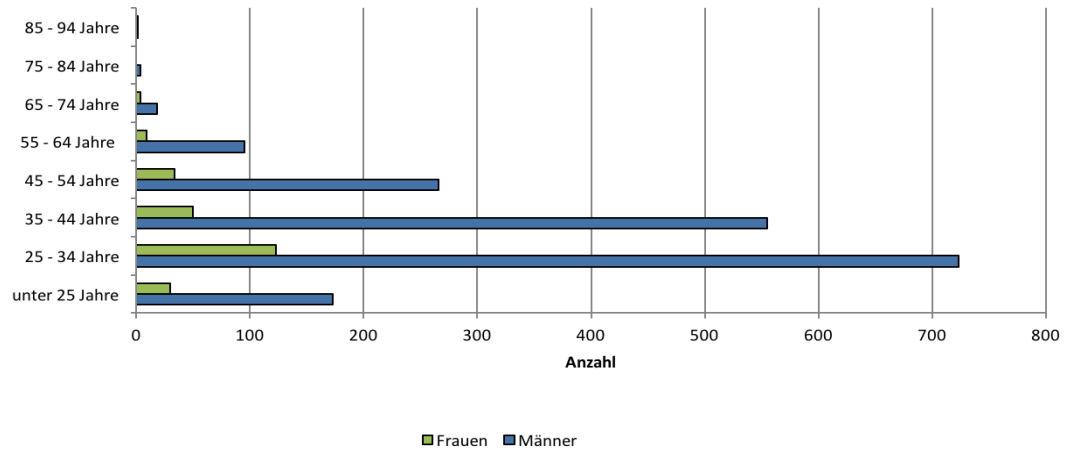


Abbildung 7: Altersverteilung der Strava-Nutzerinnen und -Nutzer in Dresden im Zeitraum 01/2016 - 06/2016 (N =3.016)

Altersverteilung

Die geäußerten Vorbehalte der Kommunen stimmen demnach nur zum Teil, da sportlich fahrende, männliche Nutzer überrepräsentiert sind. Eine Repräsentativitätsprüfung erscheint daher neben der reinen Datenanalyse notwendig, um die Aussagekraft der Daten zu erhöhen (siehe Kapitel 4.2, Schritt 5). Im Projekt wurde dies durch eine Befragung der Strava-Nutzenden geleistet.

Einfluss auf die Geschwindigkeiten

Die Form der Motivation bei Strava hat zu einem Auswirkung auf die Radverkehrsleistung der Nutzerinnen und Nutzer. Es wird intrinsisch motiviert eine hohe Anzahl an Fahrten mit hoher Fahrtweite aufgezeichnet (Ø 22 Fahrten/Jahr mit Ø 20km Länge). Die Kehrseite ist die durch die hohe Anzahl an überdurchschnittlich trainierten Radfahrenden bedingte höhere Geschwindigkeit. Diese liegt durchschnittlich rund 5,5km/h über der der Durchschnittsradfahrenden (siehe Detailwissen VI). Innerhalb des Datensatzes lassen sich bis zu einem gewissen Grad die reinen Sportfahrten von den Alltagsfahrten trennen, jedoch sind auch die Alltagsfahrten durch diese Nutzergruppe beeinflusst.

Beeinflussung des Routenwahlverhaltens

Betrachtet man die Netzabdeckung, scheint dieser Sachverhalt auf die Routenwahl bei Alltagsfahrten dennoch keinen besonders großen Einfluss zu haben – vielmehr handelt es sich bei dieser Nutzergruppe mutmaßlich um besonders gut informierte Radfahrende. Bei der Dateninterpretation ist jedoch beispielsweise auf Radsportevents zu achten. Diese sollten dem Datenprovider zur Sicherheit immer genannt werden, da sie sonst im Nachhinein schwer erklärbare Aufkommensspitzen schaffen können.

Nutzerbefragung

Zusätzlich wurde eine Nutzerbefragung durchgeführt, auf deren Basis das in den Daten abgebildete Nutzungsverhalten mit subjektiven Einschätzungen der Nutzerinnen und Nutzer abgeglichen werden konnte. Insgesamt decken sich die Ergebnisse mit der im Folgenden dargestellten Datenlage. So sind Motivation und Geschwindigkeit der Strava-Nutzerinnen und -Nutzer erhöht, bei Routenwahl und Routenlänge bestehen dagegen kaum Abweichungen auf Alltagsradfahrten (für eine detaillierte Betrachtung der Nutzerbefragung siehe Projektbericht).

Schritt 3: Wie funktioniert die Bereitstellung durch den Anbieter?

Die von dem App-Anbieter Strava erworbenen Daten basieren auf GPS-Punkten, die zu Analysezwecken einem GIS-basierten Wegenetz zugeordnet werden. Bei Strava liegen zunächst alle GPS-Positionen mit einem Nutzer-, Zeit- und Fahrtstempel vor. Diese werden von Strava einem Wegenetz mittels eines Mapmatching-Algorithmus zugeordnet. Der genaue Mapmatching-Algorithmus ist dem Projektteam unbekannt. Da während der Auswertung jedoch eine hohe Anzahl an Duplikaten in Bereichen hoher Netzkantendichte festgestellt wurde, ist davon auszugehen, dass es sich um einen einfachen **Point-to-Point** oder **Point-to-Curve** Ansatz handelt. Eine genaue Prüfung der gelieferten Daten sowie ein statistischer Test auf Duplikate ist daher unabhängig vom Anbieter immer zu empfehlen.

Abbildung 8 zeigt die mehrfache Projektion von Radverkehrsaufkommen auf unterschiedliche parallele Kanten. Davon sind jeweils nur die äußeren für den Radverkehr befahrbar. Mittig wurden Radfahrende auf eine Hauptverkehrsachse sowie einen Tunnel projiziert.

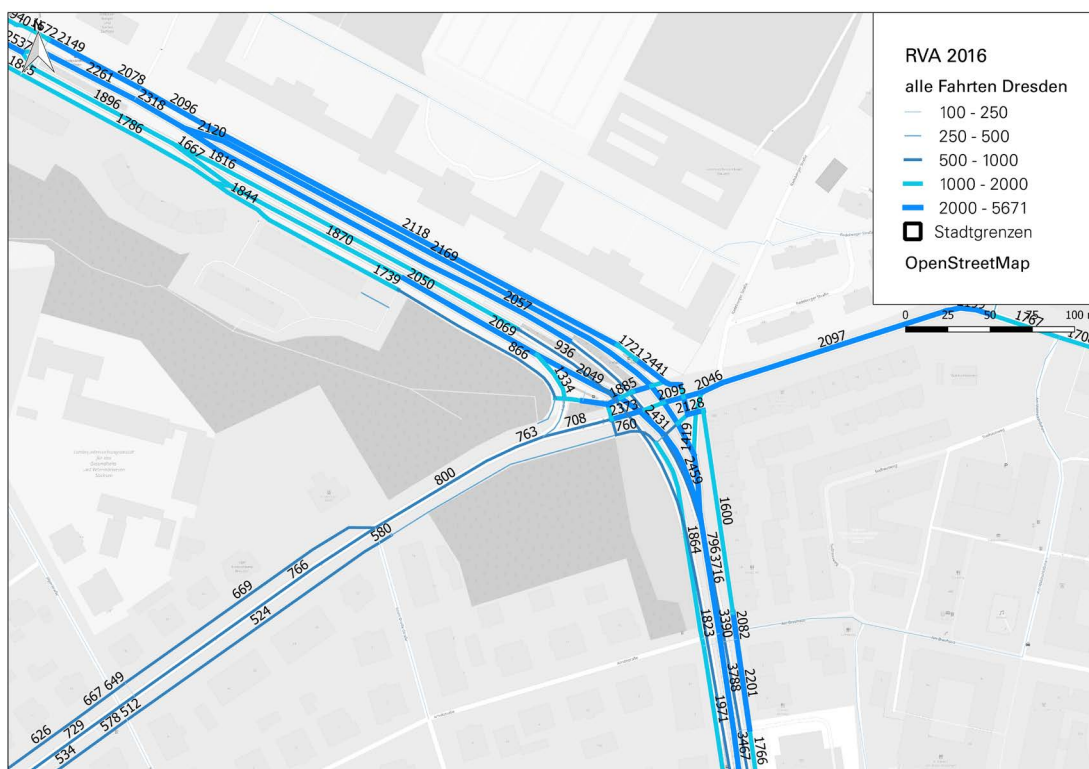


Abbildung 8: Visualisierung von Duplikaten im Radverkehrsaufkommen (RVA), OSM-Wegenetz Stadt Dresden, 2017

Welches Wegenetz soll verwendet werden?

Das dem **Mapmatching** zugrundeliegende Wegenetz kann frei durch den Kunden gewählt werden und wird in der Regel auch von diesem zur Verfügung gestellt. Zunächst musste durch das Projektteam somit geklärt werden, welches Wegenetz der Pilotkommune Dresden für das Mapmatching an den Datenprovider Strava übermittelt und somit auch Grundlage der späteren Analyse sein wird (siehe Tabelle 2).

In der Frühphase des Projektes wurde entschieden, zwei unterschiedliche Wegenetze als Grundlage für das Mapmatching zu verwenden: zum Einen **OpenStreetMap (OSM)**, um auch eine überregionale Vergleichbarkeit innerhalb der drei Bundesländer Sachsen, Brandenburg und Berlin zu ermöglichen. Zum Anderen wurden die von den Stadtverwaltungen zur Verfügung gestellten erweiterten Straßenknotennetze genutzt,

Datenbereitstellung

Prüfung auf Duplikate

Auswahl der Kartengrundlage

die letztendlich an das Amtlich Topographisch-Kartographische Informationssystem (ATKIS) angebunden sind und diverse Zusatzinformationen über die Infrastruktur enthalten können. Zudem bietet es die Möglichkeit, dass die Stadtverwaltungen die Karten zu einem späteren Zeitpunkt besser für ihre eigenen Planungszwecke verwenden können. Die städtischen Netze sind ohne umfangreiche Weiterverarbeitung verwendbar gewesen, da jeder Verkehrsweg aus einer **Kante** mit diversen Attributen, wie Radverkehrsinfrastruktur, Fahrstreifenanzahl oder zulässiger Geschwindigkeit, besteht.

Mitunter schwieriger gestaltet sich die Nutzung der detaillierten OSM-Kartengrundlage. Um eine Datengrundlage für das Mapmatching auf Basis von OSM bereitstellen zu können, musste ein geeignetes Wegenetz zusammengestellt werden. Dafür wurden im ersten Schritt Wegenetze mit möglichst vollständigen Datensätzen von OSM verwendet. Diese wurden in Form von **Metro Extracts** von einem Online-Kartenportal exportiert. Die Kartenexporte sind in ihrer Vollständigkeit sehr nah an den Rohdaten von OSM, bieten also eine umfangreiche Datengrundlage. Probleme entstehen, wenn Kanten übereinander oder räumlich sehr dicht parallel zueinander verlaufen, da die zugeordneten Fahrten zwischen den Kanten springen können, wie in Abbildung 8 visualisiert. Außerdem wird bei der Auswertung der Daten die Analyse von **Querschnitten** erschwert. Um dies zu vermeiden, wurden die Netze vereinfacht, so dass nur noch einzelne Kanten pro Weg dargestellt werden. Das genaue Vorgehen kann im Detailwissen I nachvollzogen werden.

Tabelle 2: Vergleich der Wegenetze, die in Dresden berücksichtigt wurden

Anbieter	Vorteile	Nachteile
openstreetmap.org	Umfangreiche Datensätze Überregionale Vergleichbarkeit	Nur Export von Ausschnitten möglich, xml Format
mapzen.com	Umfangreiche Datensätze	Nur Export von ausgewählten Städten möglich
geofabrik.de	Reduzierte Kantenmenge	Eingeschränkter Datenumfang
ESKN 5*	Wird städtisch verwendet	z.T. wenige Kanten (Brücken)

* Erweitertes Straßenknotenetz, durch die Stadt zur Verfügung gestellt.

Was liefert der Anbieter?

Nach der Übergabe der Kartengrundlage liefert der Anbieter in der Regel die gewünschten Daten aggregiert und auf das Basisnetz referenziert zurück. Die im Projekt betrachteten Strava-Daten lagen dabei in zwei Datenpaketen vor. Es handelt sich um einen Datensatz, der auf die Kanten des initialen GIS-Netzes referenziert (siehe Detailwissen II) und einen Datensatz, der mit den Knoten verknüpft ist (siehe Detailwissen III). Dies geschieht über eine einfache Indizierung der jeweiligen Datensätze. Die Daten liegen ausschließlich bezogen auf die jeweiligen Netzelemente vor. Zusammenhängende Routen sind daraus, um dem Datenschutz gerecht zu werden, nicht direkt ableitbar. Aus diesen Grunddaten können nach Wunsch des Kunden unterschiedliche, ‚Roll-up‘ genannte, Aggregationsformen erzeugt werden. Diese werden meist nach Monaten, Tageszeiten oder Jahreszeiten angeboten. Beispielsweise können anhand der Zeitstempel alle Daten der Monate Januar bis März eines Jahres aggregiert werden. Diese Daten können entweder eigenhändig in den städtischen Onlineauftritt eingebunden werden oder durch den Anbieter interaktiv auf einer Plattform zur Verfügung gestellt werden.

Schritt 4: Was sagen die Daten aus?

Basierend auf den Kartengrundlagen konnte mit der explorativen Datenanalyse begonnen werden. Im Fokus standen dabei zunächst Verkehrsstärken und Geschwindigkeiten. Hierbei zeigt sich im Großen und Ganzen eine recht hohe Aussagekraft des vorhandenen Datensatzes.

Die generelle Netzabdeckung durch die Nutzerinnen und Nutzer der App ist zumindest für das Gebiet der Stadt Dresden sehr zufriedenstellend. In Abbildung 9 ist dies ersichtlich anhand der blauen Linien, die engmaschig über das gesamte Stadtgebiet verteilt sind. Anzumerken ist, dass in dieser Form der Darstellung noch keine Differenzierung nach verschiedenen Wegezwecken, wie Arbeitswegen oder reinen Freizeitfahrten, erfolgte. Klar ersichtlich sind jedoch die Hauptradrouten in der Stadt sowohl in Ost-West-, als auch in Nord-Süd-Richtung.

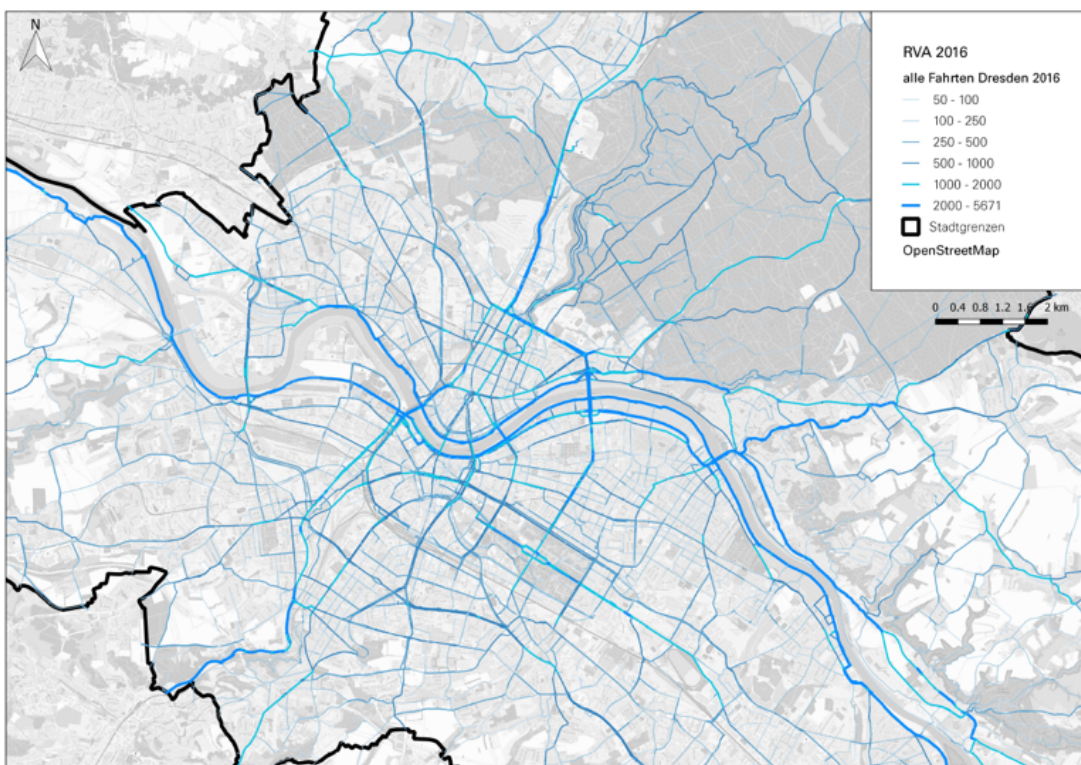


Abbildung 9: Radverkehrsaufkommen (RVA) aller Strava-Nutzerinnen und -Nutzer im Zeitraum 06/2015 – 06/2016

Ein detaillierter Blick offenbart die konkreten Möglichkeiten, die sich den Planenden bieten (siehe Abbildung 10). So ist in der unten stehenden Abbildung mittig ein nahezu knotenpunktfreier, gemeinsamer Geh- und Radweg durch einen Grünzug (Weißeritz-Grünzug) zu sehen, während die äußeren, deutlich stärker belasteten Kanten, jeweils gemeinsame Geh- und Radwege an stark befahrenen Hauptverkehrsstraßen darstellen. Offensichtlich werden diese jedoch deutlich intensiver genutzt als die verkehrsferne Verbindung mit der sehr viel höheren Aufenthaltsqualität. Das Fallbeispiel steht hier nicht ausschließlich für Arbeitswege, bei denen das Argument der kürzesten Entfernung ein sehr starkes ist, sondern auch für Freizeitfahrten. Dementsprechend wurde auf den Filter ‚Alltagsfahrten‘ verzichtet. Für den Stadt- oder Radverkehrsplaner können sich im Zuge dieser Form der Evaluation nun folgende Fragen stellen: Ist das Angebot in seiner jetzigen Form attraktiv für Radfahrende? Hier jedoch ist die Angebotsqualität als sehr hoch einzustufen. Es handelt sich um eine wenig frequentierte und sehr ruhige Route im Grünen mit ausreichend Sitzgelegenheiten und Spielplätzen, die auch zum Verweilen einladen. Darin ist der Grund für die geringe Frequentierung also wahrscheinlich nicht zu suchen. Eine weitere Frage stellt sich jedoch in Bezug auf die

Netzabdeckung

Untersuchung einzelner Netzelemente

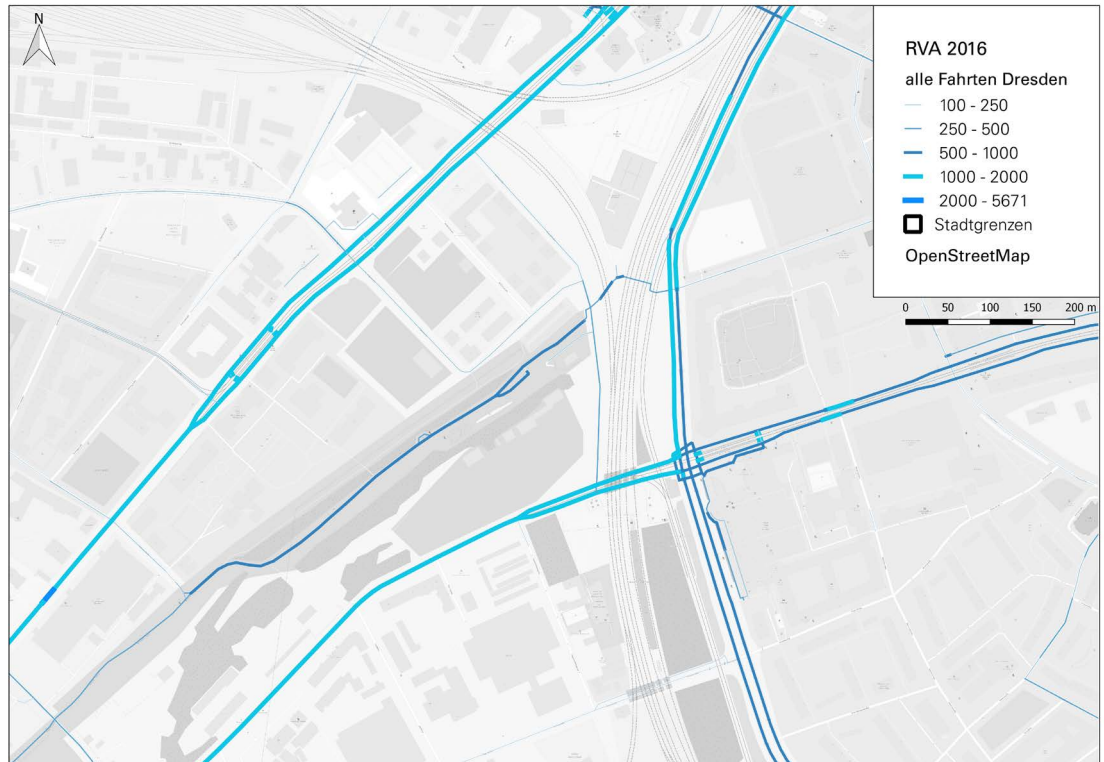


Abbildung 10: Kartenausschnitt des Radverkehrsaufkommens (RVA) aller Strava-Nutzerinnen und -Nutzer im Zeitraum 06/2015 – 06/2016, der die Nutzung des Weißeritz-Grünzuges zeigt, sowie Fotos dieser Stelle

Im betrachteten Fallbeispiel ist die These der mangelnden Sichtbarkeit als wahrscheinlich anzusehen. Der Grünzug als solcher ist für die Radfahrenden nicht überall ausgeschildert und hinter einer Eisenbahnunterführung etwas versteckt zu finden (Abbildung 10, siehe Fotos). Zunächst sollte in einem derartigen Fall also die Möglichkeit einer verbesserten Beschilderung eruiert werden. Anschließend können im Folgejahr die Nutzerzahlen auf eine veränderte Maßnahmenwirkung hin überprüft werden.

Pendlerfahrten bzw. Alltagsfahrten

Abbildung 11 zeigt die durch Strava als Alltags- bzw. Pendlerfahrten gekennzeichneten Wege (siehe Detailwissen IV). Auf den ersten Blick ist eine Verringerung der Verkehrsmengen im Vergleich zu Abbildung 9 über die gesamte Karte zu verzeichnen. So liegt beispielsweise der Anteil der Alltagsfahrten auf Segmenten des Elberadweges bei lediglich 50%. Auch im nordöstlichen Bereich des Stadtwaldes sind Rückgänge der Fahrtenzahlen zu beobachten. Die These der überdurchschnittlich repräsentierten Sportfahrten kann also zumindest für diesen Bereich als bestätigt angesehen werden. Da es sich bei den Routen sowohl um asphaltierte als auch unbefestigte Wege handelt, verwenden offenbar Nutzerinnen und Nutzer von Rennrad und Mountainbike die App gleichermaßen. Es ist also wahrscheinlich, dass sowohl Rennrad- als auch Mountainbikefahrerinnen und -fahrer die App nutzen.

Beachtung der Wegzwecke

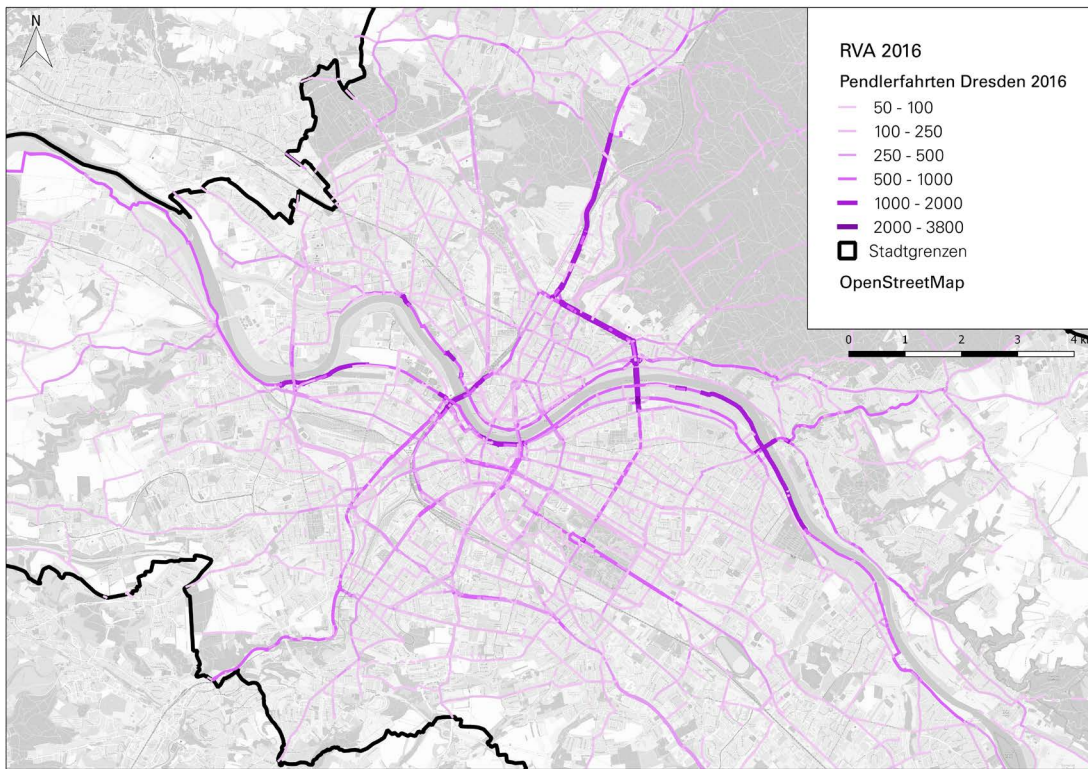


Abbildung 11: Als ‚Commute‘ (pendeln) gekennzeichnete Fahrten der Strava-Nutzerinnen und -Nutzer

Während sich der Geländeradspport aufgrund der genutzten Wege relativ leicht lokalisieren und anschließend auch aus Kartendarstellungen exkludieren lassen, ist dies für den Straßenradspport nicht ohne weiteres möglich. An dieser Stelle stellt sich auch die Frage, ob dies denn unbedingt nötig ist. Auch die gefilterten Alltagsfahrten bieten dem Planenden Raum für Erkenntnisse (nähere Informationen zur Definition von Alltagsfahrten finden sich in Detailwissen IV). So ist im Fallbeispiel in Abbildung 12 ein Teil des Elberadweges zu sehen. Dieser wird erkennbar weniger befahren als die gegenüberliegende Seite und die sich anschließenden Elemente. Dabei ist die

Verbindungen in der Stadt

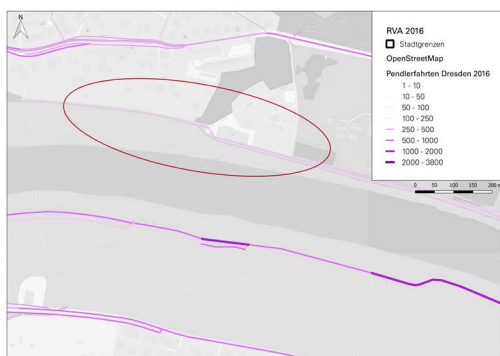


Abbildung 12: Kartenausschnitt des Radverkehrsaufkommens (RVA) aller Strava-Nutzerinnen und -Nutzer im Zeitraum 06/2015 – 06/2016, der die Nutzung des Körnerweges zeigt, sowie ein Foto der Oberflächenbeschaffenheit.

Verbindungsfunktion durchaus die gleiche. Es handelt sich jedoch um einen Abschnitt mit einer deutlich geringeren Oberflächenqualität aufgrund von Randbedingungen des Denkmalschutzes. Dieses Beispiel ist den Planern vor Ort selbstverständlich wohlbekannt, illustriert jedoch sehr gut die Möglichkeiten, die ein Datensatz auf Basis des geäußerten Verhaltens der Radfahrenden bietet. Aufgrund der Plastizität der Darstellung kann diese auch als Argumentationshilfe im Diskurs dienen.

Quelle-Ziel-Matrizen

Die Quelle-Ziel-Beziehungen können in zwei unterschiedlichen Detailgraden betrachtet werden. So ist sowohl eine Betrachtung der gesamten Daten, als auch eine Betrachtung direkter Quelle-Ziel-Beziehungen einzelner Bereiche möglich. Da die

Quellverkehre

Quelle-Ziel-Matrizen nicht direkt aus dem ursprünglichen Datensatz ableitbar sind (siehe Kapitel 2.5), ist im Detailwissen V die Datenaufbereitung der Quelle-Ziel-Matrizen in Dresden exemplarisch dargestellt.

In Abbildung 13 dargestellt sind die Quellverkehre aller von Strava als ‚pendeln‘ gekennzeichneten Fahrten innerhalb der Stadtgrenzen von Dresden. Dafür wurden innerhalb der Quelle-Ziel-Matrix Verkehre betrachtet, die im Stadtgebiet von Dresden starten oder enden, um auch Einpendler räumlich erfassen zu können. Gut zu erkennen ist die starke Konzentration auf das mittig in der Karte gelegene Stadtzentrum, mit einer hohen Dichte an Wohnbebauung. Etwas nördlich des Zentrums tritt das Stadtviertel „Äußere Neustadt“ hervor. Dieses kennzeichnet sich als Szene- und Ausgehviertel in Dresden. Die überwiegend jüngere Bevölkerung legt die These nahe, dass hier auch ein erhöhter Anteil an Radfahrenden vorliegt. Diese These kann anhand der Quellverkehre bestätigt werden. Weitere starke Quellen im Dresdner Norden, beziehungsweise im nördlichen Zentrumsbereich (in Violett dargestellt), sind deckungsgleich mit dem Standort großer Arbeitgeber (Infineon mit 2.000 und Globalfoundries mit 3.700 Beschäftigten). Im Stadtzentrum und südlich davon sind mit dem Campus der Technischen Universität Dresden und den nahe gelegenen Studentenwohnheimen weitere starke Radverkehrsquellen ersichtlich. Ebenfalls auffällig ist abermals die Sonderstellung des Stadtwaldes. Die Dresdner Heide stellt im Nordosten der Stadt eine starke Senke für den Quellverkehr dar. Generell sind starke Quellverkehre also eher im Stadtzentrum als im städtischen Umland zu verorten. Dies ist auch aufgrund der höheren Einwohnerzahlen innerhalb der zentrumsnahen Stadtviertel sowie der Konzentration von Einzelhandel und Arbeitsplätzen folgerichtig.

Kritische Dateninterpretation

Quellverkehrszellen für Pendler sind vor allem in den Städten Radeberg, Pirna und Radebeul im Dresdner Umland zu erkennen. Generell sind jedoch die erfassten absoluten Zahlen für die Pendler mit der gebotenen Vorsicht zu interpretieren, so können Werte diesseits von 200 Fahrten durchaus von einer sehr aktiven Einzelperson erbracht worden sein.

Kritische Betrachtung der Nutzungskorridore

Direkte Quelle-Ziel-Beziehungen ermöglichen zum einen die Identifikation bestimmter, ggf. zu bevorzugender, Nutzungskorridore. Zum Zweiten bieten sie eine Möglichkeit, anhand der absoluten Anzahl der Relationen besonders aktive Radfahrende als alleinige Verursacher zu identifizieren. Dies ist nicht zweifelsfrei möglich, da Informationen zu den Radfahrenden nicht zur Verfügung stehen und es theoretisch auch möglich ist, dass mehrere weniger aktive Fahrende häufig die gleiche Relation nutzen.

Abbildung 14 zeigt für ein ausgewähltes Stadtviertel in Dresden alle aufgezeichneten Quelle-Ziel-Relationen. Zum Einen spricht die Vielfalt der unterschiedlichen übrigen Ziele und Quellen für eine sehr heterogene Nutzermischung und eine Vielzahl aktiver Nutzerinnen und Nutzer, zum Anderen wird auch der vermutete Einfluss einer sehr aktiven Nutzerin bzw. eines sehr aktiven Nutzers zwischen den Stadtteilen Laubegast im Südosten Dresdens und Äußere Neustadt deutlich. Auf dieser Relation allein wurden im beobachteten Zeitraum ca. 450 Fahrten erbracht. Statistisch ist dies zumindest aufgrund der Einwohnerzahl gegenüber anderen Stadtteilen nicht zu rechtfertigen. Eine mögliche Erklärung wäre daher ein sehr regelmäßiger Pendler, oder eine regelmäßig stattfindende Trainingsrunde beziehungsweise eine gemeinsam aktive Gruppe von Radfahrenden. Dieses Problemfeld steht beispielhaft für alle auf Öffentlichkeitsbeteiligung basierenden Verfahren. Sehr aktive Nutzerinnen und Nutzer können Verzerrungen herbeiführen. Hier ist unbedingt anbieterseitig eine Normalisierung der Nutzerzahlen anzuregen, oder der Zeitraum ggf. einzuschränken.

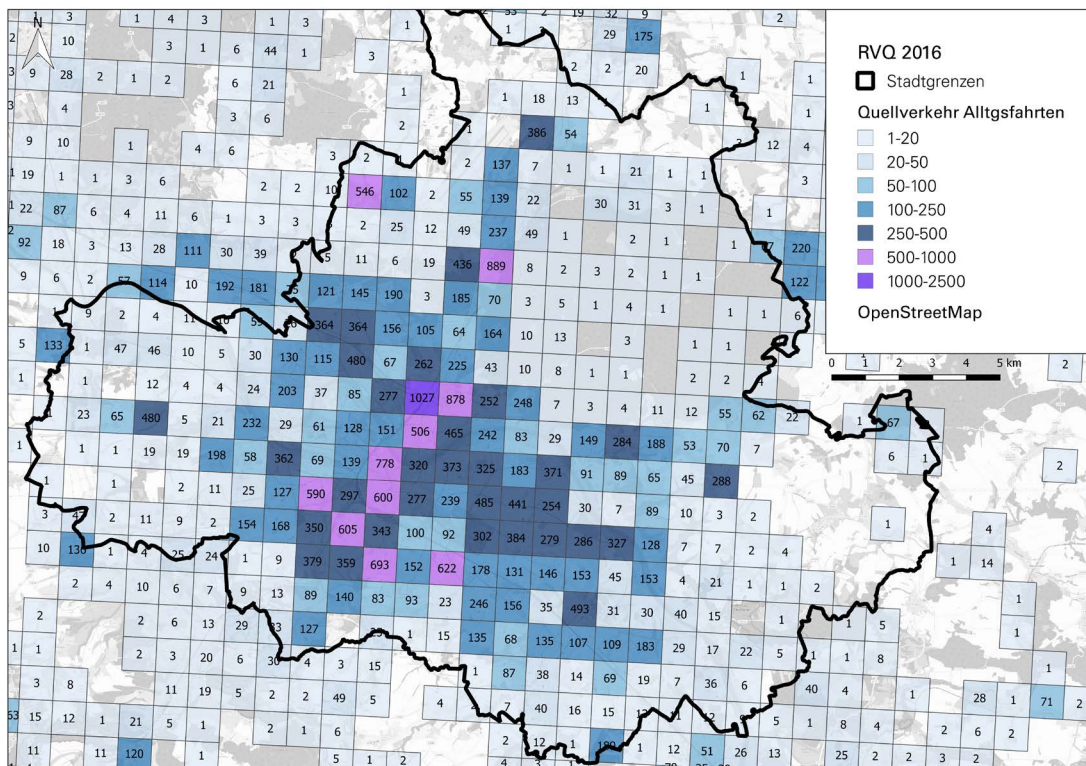


Abbildung 13: Quellverkehrsaufkommen (RVQ) der Strava-Nutzerinnen und -Nutzer für die Stadt Dresden im Zeitraum 06/2015-06/2016

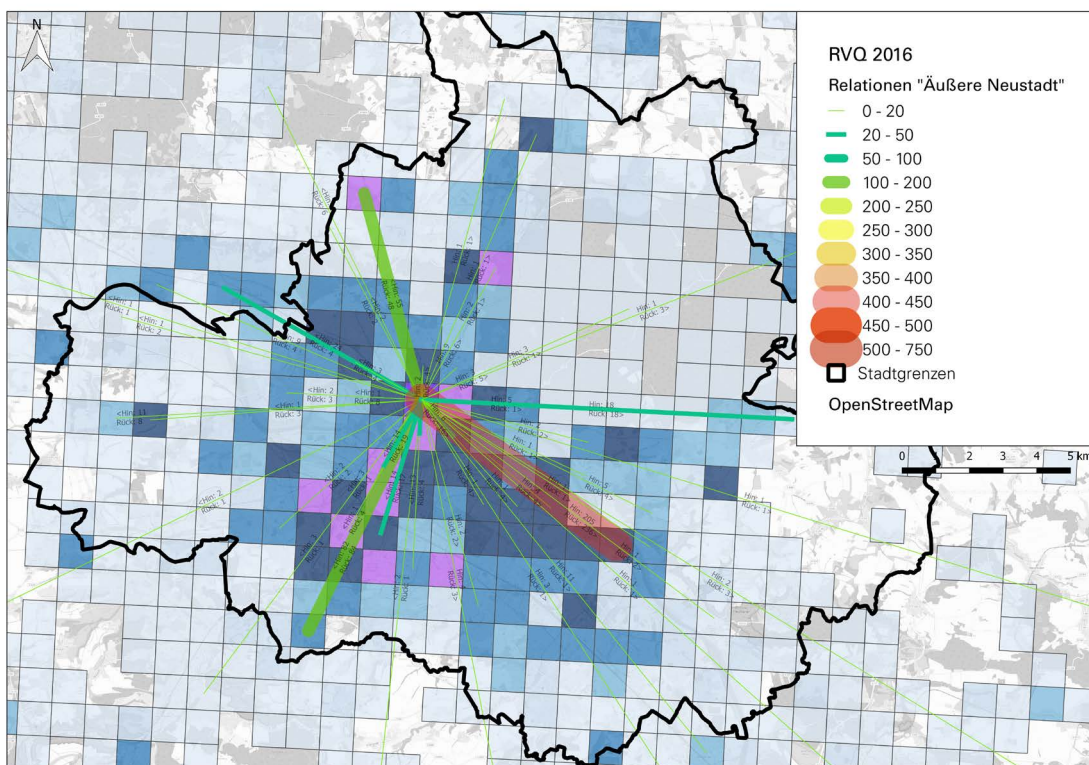


Abbildung 14: Genutzte Wegerelationen eines ausgewählten Stadtviertels in Dresden Äußere Neustadt im Zeitraum 06/2015 - 06/2016

Geschwindigkeiten

Die in Abbildung 15 dargestellten Geschwindigkeiten stellen den Durchschnitt aller von 05/2015 bis 12/2015 aufgezeichneten Fahrten auf den jeweiligen Netzelementen dar. Hier ist zu beachten, dass für jede Kante im System, natürlich richtungstrennt, zwei unterschiedliche Geschwindigkeiten vorliegen. Während der zugrundeliegenden Datenverarbeitung werden diese lediglich ‚mit‘ und ‚entgegen‘ der Digitalisierungsrichtung der entsprechenden Kanten in der GIS-Kartengrundlage

Durchschnittsgeschwindigkeiten

berechnet. Der dargestellte Sachverhalt entspricht dem über die Nutzungshäufigkeit gewichteten Mittelwert beider Richtungen. Für den Innenstadtbereich und auch den Elberadweg lässt sich im Allgemeinen eine erhöhte Radverkehrsgeschwindigkeit, beispielsweise im Vergleich zu Reisezeitmessungen, feststellen. Dies gibt Anlass zu einer erweiterten Prüfung der in diesen Bereichen reell gefahrenen Geschwindigkeiten. Grundsätzlich sind die Geschwindigkeiten jedoch relativ zu interpretieren. Bereiche mit geringeren Geschwindigkeiten im Verlauf eines Straßenzuges deuten auf Behinderungen bzw. Verbesserungsmöglichkeiten in der Führungsform der genutzten Infrastruktur hin.

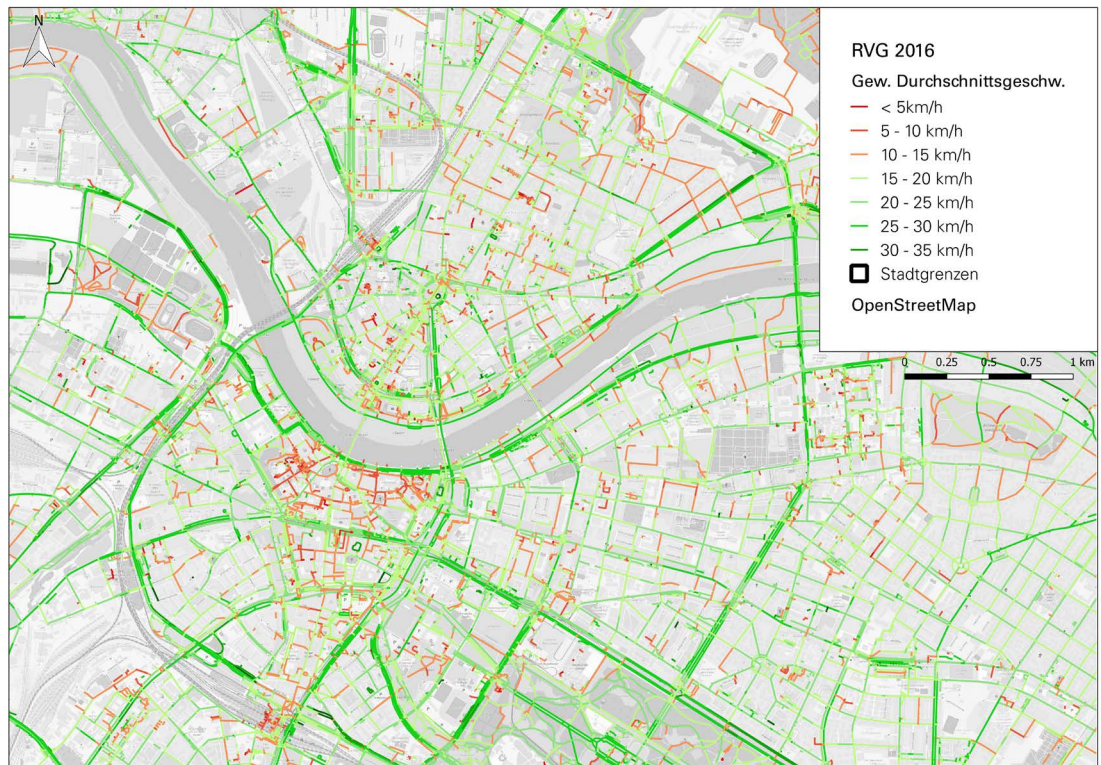


Abbildung 15: Gewichtete Durchschnittsgeschwindigkeiten (RVG) aller Strava-Nutzerinnen und -Nutzer über beide Digitalisierungsrichtungen der Kante im Zeitraum 06/2015 – 06/2016

Schritt 5: Sind die Daten repräsentativ?

Um die Repräsentativität der Daten einschätzen zu können, wurde eine umfangreiche Prüfung zu den einzelnen Parametern durchgeführt. Hintergrund ist die Stichprobe des Dresdner Strava-Datensatzes mit rund 3.200 Personen mit 70.500 Fahrten, die verglichen mit der Grundgesamtheit der Radfahrenden in Dresden vergleichsweise gering ist und eine soziodemographische Verzerrung aufgrund der Nutzerstruktur der App zumindest vermuten lässt. Das genaue Vorgehen sowie Möglichkeiten zum Umgang mit systematischen Fehlern werden im Detailwissen VI und VII beschrieben. Die Einschätzungen basieren auf einer umfangreichen wissenschaftlichen Validierung zwischen GPS-Daten der Firma Strava und Daten aus entsprechenden etablierten Feldmessungen in der Pilotstadt Dresden (siehe auch Projektbericht). Hier ist auch die Repräsentativität von Quelle-Ziel-Matrizen, Wartezeiten und Pendlerfahrten untersucht worden. Dies ist gerade für kleinere Kommunen nicht in dieser Tiefe praktikabel, wohl aber sollten eigene empirische Daten immer in die Datenauswertung einbezogen werden.

Prüfung einzelner Parameter auf Repräsentativität

Einschränkungen

Die Voraussetzungen für die Nutzung von GPS-Daten sind abhängig vom genutzten Anbieter und der gewünschten Auswertungstiefe. Sowohl Strava, als auch andere Firmen, wie z.B. BikeCitizens, bieten Zugang zu Online-Portalen mit bereits fest definierten Auswertungen und Darstellungen. Hierfür benötigt der Nutzende lediglich einen Internetzugang an seinem Arbeitsplatz. Diese Variante der technischen Umsetzung ist insbesondere für kleinere Kommunen und solche mit geringeren personellen Ressourcen attraktiv, grundsätzlich aber für alle Kommunen anzustreben. Notwendig ist dafür die konkrete Absprache mit den Datenanbietern im Vorfeld, vor allem hinsichtlich der Visualisierungswünsche.

Für tiefer gehende, eigene Auswertungen werden hingegen mindestens eine GIS-Software, das zugehörige Fachpersonal sowie die Möglichkeit der Arbeit mit Geodatenbanken benötigt. In der Regel sind die benötigten technischen Voraussetzungen und personellen Kapazitäten zumindest in größeren Kommunen vorhanden, wenn auch meist nicht in einer Person bzw. einer Verwaltungseinheit vereint.

Auswertungen auf Online- Portalen



Zusammenfassung der Ergebnisse aus der Repräsentativitätsprüfung

- » *Verkehrsmenge:* Die durchschnittliche Tagesganglinie lässt sich genauso wie der durchschnittliche tägliche Verkehr ungefähr abbilden. Eine Hochrechnung der Verkehrsmenge mittels eines konstanten Faktors kalibriert an Dauerzählstellendaten ist möglich.
- » *Geschwindigkeit:* Die Geschwindigkeiten liegen im Durchschnitt um etwa 5,5 km/h höher als bei Vergleichsmessungen. Eine Reduktion um 5,5 km/h ist daher sinnvoll. Relative Änderungen sind jedoch auch ohne diesen Schritt sichtbar.
- » *Quelle-Ziel-Matrizen:* Die Verteilung der Abfahrten korreliert mit den Einwohnerzahlen, jedoch sind Exklusionseffekte in einzelnen Bezirken mit geringem Einkommen und Bildungsniveau beobachtbar.
- » *Wartezeiten:* Die Wartezeiten sind ohne massive korrigierende Eingriffe nicht nutzbar.
- » *Pendlerfahrten/
Alltagsfahrten* Der Algorithmus erkennt zuverlässig Alltagsfahrten, solange diese nicht als Rundweg in Form einer Wegekette ausgeführt werden und die einzelnen Abschnitte nicht zwischengespeichert werden. Diese sind jedoch, verglichen mit der üblichen Wegeweite, sehr lang. Der Anteil der Alltagsfahrten am Strava-Datensatz wird also tendenziell überschätzt.

Schritt 6: Ist die Fragestellung beantwortet?

Verschiedene Fragestellungen bezüglich der Zielnetzplanung und der Analyse des bisherigen Angebotes sowie der Maßnahmenpriorisierung können anhand von Verkehrsmengen, Geschwindigkeiten und Wartezeiten beantwortet werden. Dabei gilt es für eine hohe Radverkehrsqualität Geschwindigkeitseinbrüche, beispielsweise durch Barrieren wie Treppen, zu vermeiden, die Verkehrsmengen auf den gewünschten Routen zu erhöhen und Wartezeiten an Knotenpunkten zu senken. Folglich sind diese Parameter auch jene, welche bei einer Datenanalyse im Fokus stehen sollten:

Wo fahren Radfahrende (nicht)?

Wo fahren Radfahrende langsamer als üblich?

Wo warten Radfahrende sehr lange?

Die durchgeführte Datenanalyse sollte Aufschluss über die obigen Fragestellungen geben können. Die resultierenden Antworten können anschließend gezielt in Maßnahmenkonzepte überführt werden.

Voraussetzung zur Datenauswertung



Zusammenfassung in 6 Schritten

Alle Schritte für die praktische Anwendung

- Schritt 1** Fragestellung definieren: Zu Beginn des Prozesses steht eine Fragestellung. Beispiele für mögliche Themen finden sich in Kapitel 3.
- Schritt 2** Anbieter-Auswahl: Eine Übersicht über die Anbieter findet sich in Kapitel 2.
- Bezug eines Beispieldatensatzes: Zur Einschätzung der Qualität der Daten ist ein Beispieldatensatz hilfreich. Eine Analyse auf Verzerrungen ist so bereits in einer frühen Projektphase möglich.
 - Analyse der Nutzerschicht: Einen kritischen Punkt stellt die Nutzerschicht der Stichprobe dar. Abhängig vom gewählten Anbieter kann diese durchaus variieren und sollte deshalb unbedingt berücksichtigt werden.
- Schritt 3** Bereitstellung der Daten: Entscheidung für ein Wegenetz
- Schritt 4** Ergebnisse der Datenanalyse: Datenaufbereitung, explorative Datenanalyse
- Schritt 5** Repräsentativitätsprüfung und Anpassung der Daten: Eine Repräsentativitätsprüfung der in der späteren Analyse berücksichtigten Parameter ist notwendig, um systematische Fehler auszuschließen. Dies ist insbesondere bei verzerrten Stichproben (z.B. sportliche, männliche App-Nutzer) zu beachten.
- Schritt 6** Beantwortung der Fragestellung

Eine ausreichende Netzabdeckung stellt das wichtigste Kriterium für oder gegen den Datenbezug dar (siehe Kapitel 5.1). Bei der Interpretation der Ergebnisse sind dann vor allem die soziodemographischen Merkmale der Nutzergruppe des Anbieters zu beachten. So kann eine Anpassung der Daten im Nachhinein, wie im Anwendungsbeispiel (siehe Kapitel 4), erforderlich sein. Daneben gilt es vor allem die Kartengrundlage zu beachten (siehe Kapitel 5.2).

5.1 Wann ist der Datensatz nutzbar?

Grundsätzlich ist der Datensatz dann nutzbar, wenn eine sinnvolle, qualitative Netzabdeckung möglich wird. Dies kann mit wenigen sehr aktiven Nutzerinnen und Nutzern ebenso gelingen, wie mit deutlich mehr sporadischen Nutzungen. Im Zweifelsfall ist dieser Sachverhalt vorab mit den Datenanbietern zu klären. Auch sollte für eine sinnvolle Datenaggregation ein längerer Zeitraum, wie z.B. ein Jahr gewählt werden. In der Stadt Dresden beispielsweise war für den im Leitfaden vorgestellten Anwendungsfall (siehe Kapitel 4) mit 3.000 Nutzern und Nutzerinnen pro Jahr eine sehr gute Netzabdeckung zu verzeichnen und die Korrelation mit den Dauerzählstellen auf einem sehr hohen Niveau. Diese Nutzerzahl entspricht dabei lediglich 0,68% der Einwohnerinnen und Einwohner der Stadt, durchschnittlich wurden jedoch 22 Fahrten pro Jahr und Nutzenden aufgezeichnet. Das resultiert in 70.500 Fahrten im Stadtgebiet bei einem Anteil der Alltagsfahrten von ca. 50%.

Grundsätzlich gilt: Je weniger sportiv der Hintergrund der Nutzerinnen und Nutzer, desto besser wird eine breite Grundgesamtheit von Radfahrenden abgebildet. Allerdings steht hier die Nutzerzahl dem Bestreben nach der ausschließlichen Auswertung von Alltagsfahrern entgegen. Vor allem sportliche, aktive Radfahrende erbringen hohe Radverkehrsleistungen, sowohl sportlich als auch auf Alltagsfahrten. Hier kann also durchaus ein größerer Datensatz mit höherem Anteil an sportlicher Motivation mehr Informationen liefern als ein Datensatz, der ausschließlich aus Alltagsfahrern besteht, aber deutlich weniger Fahrten enthält.

Neben der Nutzerzahl sollte beim Datenbezug auch auf eine möglichst heterogene Alters- und Geschlechtsverteilung der Nutzerinnen und Nutzern geachtet werden. Falls sich eine homogene Verteilung oder Konzentrationen nicht vermeiden lassen, muss dies in der Interpretation der Daten stets berücksichtigt werden. Zum jetzigen Zeitpunkt sind keine idealen Datensätze auf dem Markt verfügbar. So müssen im einen oder anderen Bereich stets Einschränkungen in Kauf genommen werden. Wichtig ist hier, dass diese sorgfältig evaluiert werden und in die Interpretation mit einfließen.

5.2 Was kann schiefgehen?

Mögliche Fehlerquellen können überwiegend in zwei Ursachen gefunden werden. Einerseits ist es denkbar, dass der Datensatz für die eigene Kommune zu klein ist. Andererseits kann ein Datensatz, wie der der Firma Strava, in der eigenen Kommune auch deutlich weniger repräsentativ sein als in anderen. Kann beispielsweise ein Datensatz mit eher sportiven, jüngeren, männlichen Nutzern die Bevölkerung in einer Universitätsstadt wie Dresden noch eingeschränkt repräsentieren, ist dies bei einer Stadt mit einem wesentlich höherem Anteil an älterer Bevölkerung weniger aussagekräftig. Auch die Verbindung mit den genutzten städtischen GIS-Daten kann Probleme bergen.

Nutzerzahl

Radverkehrsleistung der Nutzenden

Einschränkungen bei der Interpretation

Stichprobenumfang

Das **Mapmatching** ist ein wesentlicher Punkt des Datenverarbeitungsprozesses, wobei zu detaillierte Kartengrundlagen für Duplikate sorgen können. Zu wenige digitalisierte Kanten hingegen (beispielsweise im Verlauf einer Brücke), können auch dazu führen, dass eine große Anzahl von Fahrten nicht auf die Kartengrundlage projiziert werden kann (siehe Kapitel 4.2, Schritt 3). Dies ist grundsätzlich abhängig vom Datenanbieter und sollte mit diesem unbedingt vorab geklärt werden.

Für die Auswertung der Daten kann ganz grundlegend auf mehrere Angebotstiefen zurückgegriffen werden – fast alle Anbieter bieten zu unterschiedlichen Konditionen eine Auswertung via Analyseplattform an. Diese sind unterschiedlich detailliert in ihrer Ausgestaltung. Möglich ist, wie auch bei Strava, der Bezug der Minutendaten für die einzelnen Netzkanten. Damit lassen sich beliebig viele einzelne Auswertungen erstellen. Dies bindet jedoch auch immer Ressourcen und bietet sich eigentlich nur für Kommunen mit ausreichender personeller Ausstattung an. Es wäre schade, wenn der „Datenschatz“ nicht gehoben werden kann, weil die Hürden bei der Auswertung zu hoch sind. Eine Datengrundlage ist schnell gekauft, aber vor allem in der Arbeit mit ihr liegt der große Mehrwert für die kommunale Radverkehrsplanung. Diese Arbeit erfordert Zeit, die sich aber schnell in den für die Planung nutzbaren Daten und Aussagen auszahlt.

Einen Überblick über häufig auftretende Fragen finden Sie außerdem im Anhang dieses Leitfadens.

– Zukünftige Entwicklungen des Marktes

Der Datensatz der Firma Strava ist, wie eingangs erwähnt, in Erwartung einer zukünftigen Entwicklung des Feldes mobiler Massendaten ein wahrscheinlich vergleichsweise anspruchsvoller Anwendungsfall. Die eher homogene Stichprobe läuft den etablierten Erhebungsmethoden in der Verkehrsplanung zunächst einmal entgegen. Dem gegenüber steht jedoch ein hoher Stichprobenumfang zu einem vergleichsweise sehr geringen Kostensatz. Die Datenbasis von Strava hat neben der angesprochenen Nutzergruppe zwei weitere Schwachstellen: Zum einen sind die Ergebnisse der Wartezeitenanalyse an Knotenpunkten aktuell als nicht zufriedenstellend zu bezeichnen. Daran arbeitet der Anbieter jedoch, ebenso wie an differenzierten Mapmatching-Verfahren, die auch bei komplexeren Netzkanten zuverlässig funktionieren und das bisherige Problem doppelt zugeordneter Radfahrender lösen. Fälschlich zu hoch angegebene Radverkehrsmengen an Stellen mit hoher Infrastrukturdichte sollten dann weniger häufig auftreten. Großes Potenzial bieten hier auch weitere Funktionen wie die Möglichkeit, Fahrten aus einem bestimmten Quellbereich zusammen zu visualisieren.

Das diesem Leitfaden zugrundeliegende Projekt hatte die Analyse dieses Datensatzes zum Ziel. Es wird unbedingt empfohlen, den Blick auch für weitere Entwicklungen und Anbieter zu öffnen und die Entwicklung des Marktes regelmäßig zu überprüfen (siehe Kapitel 3). Die Radwende-App von Scholz&Volkmer (siehe Kapitel 3) bietet beispielsweise insbesondere die Chance, eine breite Zielgruppe und die notwendige Netzabdeckung durch eine begleitende Kampagne zu erreichen. Zum Stand der Veröffentlichung bildet der vorliegende Leitfaden zudem nur eine Momentaufnahme des vorliegenden Angebotes ab. Zukünftig kann dieses um Daten anderer Anbieter aus den Bereichen Sport, Navigation, Fahrradverleih oder auch Bonus- und Community-basierten Apps sowie von weiteren Akteuren erweitert werden. Zu beachten ist, dass sich der Bezug von ‚fertigen‘ GPS-Daten nicht lohnt, wenn die Nutzerdichte im betreffenden Gebiet nicht hoch genug ist und so eine Netzabdeckung nicht gewährleistet werden kann. Für diesen Fall kann jedoch in Absprache mit dem Anbieter eine begleitende Kampagne lanciert werden, um die entsprechenden Nutzerzahlen zu erreichen. Derartige Leistungen werden beispielsweise aktuell von der Firma BikeCitizens angeboten, die auch auf dem Gebiet der Analyse und Auswertung schon weiter fortgeschritten sind. Mit der Veröffentlichung des Bike-Analyse-Tools wird eine Analyse von Wartezeiten, Verkehrsmengen und Geschwindigkeiten knoten- und kantenfein ebenso möglich wie eine Analyse der Alternativrouten und Umwege zwischen einem Start- und einem Zielpunkt. Zukünftig ist auch eine Ergänzung der Kampagne STADTRADELN um eine Datenerhebung und -auswertung für Kommunen angedacht. Ein entsprechendes Forschungsprojekt im Rahmen der Forschungsinitiative mFUND des BMVI läuft seit 07/2017 in Zusammenarbeit mit der TU Dresden. Ziel ist die automatisierte Datenerhebung und -auswertung für alle teilnehmenden Kommunen. Mit ersten Ergebnissen ist im Jahr 2019 zu rechnen.

Fehlerbearbeitung

Alternativen zu Strava

Anhang

Im folgenden Anhang werden weiterführende technische Details zusammengefasst, verwendete Fachbegriffe erklärt und häufig gestellte Frage beantwortet.



Detailwissen I: Kantenüberschneidungen vermeiden

Um Kantenüberschneidungen zu vermeiden, wurde versucht die **Kanten** so zu filtern, dass jede Verbindung nur durch eine Kante dargestellt wird. Gleichzeitig sollte vermieden werden, dass Wege gar nicht mehr durch eine Kante repräsentiert werden. Diese Gefahr bestand beispielsweise bei eigenständigen Rad- und Fußwegen, die für den Langsamverkehr eine große Bedeutung als Lückenschluss im Netz besitzen. Auch wurden Kanten entfernt, die sicher nicht für Radfahrende befahrbar sind, um Fehlzuordnungen zu vermeiden. Der verwendete Filter wurde über eine Definitionsabfrage in ArcGIS erarbeitet. Ausgeschlossen wurden so beispielsweise Attribute, die auf ÖPNV oder Autobahnen hinwiesen. Nach weiteren Tests mit dem so generierten Wegenetz konnten Vor- und Nachteile offengelegt werden (siehe Abbildung 16 und 17). So hat das Netz zwar einen hohen Informationsgehalt,

aber bei der erstmaligen Zuweisung von Verkehren auf die jeweiligen Kanten konnte es noch immer zu Konflikten kommen. Als Konsequenz fiel die Entscheidung für den Wegenetzexport des Anbieters geofabrik.de. Dieses Netz ist schon von Beginn an so vereinfacht, dass es meistens jeweils nur eine Kante pro Weg enthält, was aber mit einem Informationsverlust einhergeht. Mittels der identischen Variable „OSM-ID“ lassen sich fehlenden Informationen zu einem späteren Zeitpunkt allerdings wieder anfügen. Dies erwies sich im Projektverlauf jedoch als nicht notwendig. Auf städtischer Ebene ist es hingegen unbedingt sinnvoll in die Erstellung des Netzes einige Zeit zu investieren und ggf. auch manuelle Korrekturen vorzunehmen. Tabelle 2 zeigt einen Vergleich der analysierten Wegenetze, hat aber keinen Anspruch auf Vollständigkeit, sondern soll nur die Vielfalt der unterschiedlichen Datengrundlagen verdeutlichen.



Abbildung 16: Kanten in GIS nach Mapzen.com



Abbildung 17: Kanten in GIS nach Geofabrik.de. Deutlich zu erkennen sind die reduzierten Kanten im Vergleich zu Mapzen.com (in Abbildung 16)

Vorerst ungelöst bleiben bei der OSM-Netzgrundlage Kanten, bei denen beispielsweise Fahrradwege sehr nah und parallel an Kanten des Straßennetzes entlang verlaufen. Dort kann aufgrund der Ungenauigkeit von GPS-Daten in der Praxis nicht unterschieden werden, ob

die Kfz-Fahrbahn oder die Radverkehrsanlage genutzt wurde. Mögliche fehlerhafte Zuordnungen sind eher durch ungenaue GPS-Positionsdaten verursacht. Hier sind weitere Untersuchungen notwendig, um adäquate Lösungen zu finden.



Detailwissen II: Netzkantendaten

Der Beispieldatensatz in Tabelle 3 zeigt den Zeitstempel, die Anzahl der Nutzer, deren Aktivität und die benötigte

Zeit. Hierbei ist die Edge-ID eindeutig referenziert auf die Indizes der Kanten im zugrunde liegenden GIS-Netz.

	edge_id integer	year integer	day integer	hour integer	minute integer	athlete_count integer	rev_athlete_count integer	activity_count integer	rev_activity_count integer	total_activity_count integer	activity_time double precision	rev_activity_time double precision	commute_count integer
1	6291272	2014	305	13	53	1	0	1	0	1	3.34325994805218		0
2	6291300	2015	169	18	34	1	0	1	0	1	1.16634166515055		0
3	9749260	2015	316	15	30	1	0	1	0	1	1.23008924058472		0
4	8582488	2014	305	12	51	0	1	0	1	1		1.41511103287362	0
5	9749263	2015	103	16	31	0	1	0	1	1		2.13238128748964	1
6	9749279	2015	237	8	2	1	0	1	0	1	0.354159852233221		0
7	9749282	2015	135	10	8	0	1	0	1	1		12.7332161530339	1
8	6291466	2015	123	11	53	1	0	1	0	1	1.13122306296737		0
9	6291499	2014	315	8	44	0	1	0	1	1		6.20219514092036	0
10	6291548	2015	104	6	28	0	1	0	1	1		1.27923814209955	1

Tabelle 3: Beispieldatensatz SQL für Kanten des GIS-Netzes, Strava 2016

Die einzelnen Spalten beinhalten dabei folgende Werte:

<i>athlete_count</i>	Anzahl der Radfahrenden in Digitalisierungsrichtung der Kante
<i>rev_athlete_count</i>	Anzahl der Radfahrenden entgegen der Digitalisierungsrichtung der Kante
<i>activity_count</i>	Anzahl der Aktivitäten in Digitalisierungsrichtung der Kante, eine Aktivität kann mehrere Nutzerinnen und Nutzer beinhalten
<i>rev_activity_count</i>	Anzahl der Aktivitäten entgegen der Digitalisierungsrichtung der Kante, eine Aktivität kann mehrere Nutzerinnen und Nutzer beinhalten
<i>total_activity_count</i>	Gesamtanzahl der Aktivitäten über beide Fahrtrichtungen
<i>activity_time</i>	Benötigte Zeit, um die Netzkante in Digitalisierungsrichtung zu passieren, in Verbindung mit der Länge des Netzelementes lässt sich die Geschwindigkeit errechnen
<i>commute_count</i>	Anzahl der als Pendelfahrten identifizierten Radfahrten, entweder 0 = Sport- oder Freizeitfahrt oder >0 <= total_activity_count, dann Pendlerfahrt



Detailwissen III: Netzknotendaten

Ähnlich wie bei den auf die Netzkanten referenzierten Streckendaten werden auch die Knotenpunktdaten behandelt. Auch für diesen Datensatz sind anhand der Zeitstempel und der übrigen Parameter unterschiedliche

Aggregationen möglich. Hierfür wird der Index der Knotenpunkte des GIS-Basisnetzes mit dem indizierten Parameter *node_id* verknüpft.

	node_id integer	year integer	day integer	hour integer	minute integer	athletes integer	activities integer	median_wait integer	max_wait integer	min_wait integer	commute_count integer
1	2752	2015	157	12	9	1	1	1	1	1	1
2	2752	2015	165	21	8	1	1	0	0	0	0
3	2752	2015	181	12	9	1	1	0	0	0	0
4	8254	2015	164	8	35	1	1	4	4	4	0
5	2752	2015	183	19	15	1	1	1	1	1	0
6	11005	2015	158	16	36	1	1	0	0	0	0
7	8254	2015	174	21	21	1	1	4	4	4	0
8	2752	2015	185	11	28	1	1	2	2	2	0
9	11005	2015	159	10	10	1	1	8	8	8	1
10	8254	2015	183	15	48	1	1	3	3	3	1

Tabelle 4: Beispieldatensatz SQL für Kanten des GIS-Netzes, Strava 2016

Wie in Tabelle 4 ersichtlich, werden für die Knotenpunkte einzelne Durchfahrten und deren Wartezeiten angegeben. Die Berechnungsmethodik und Probleme bei der Betrachtung der Wartezeiten werden im Projektbericht genauer betrachtet. Für Knotenpunkte kann generell die Annahme des Knotenpunktes als restriktives Element gelten. Eine Unterscheidung in Pendlerfahrten und Sportfahrten ist hier zwar möglich, aber nicht zwingend nötig. Die einzelnen Parameter gleichen denen des

Datensatzes für die Netzkanten. Einzige Ausnahme bildet dabei die angegebene Wartezeit statt der Fahrtzeit über ein Kantelement. Unterschiedliche Werte für median_wait, max_wait und min_wait sind nur zu erwarten, falls eine Gruppe von Radfahrenden, in der alle die App nutzen, den Knotenpunkt zur gleichen Zeit passiert. Dies ist aufgrund der auf die Gesamtpopulation der Radfahrenden bezogen geringen Verdichtungen der App Strava jedoch ein eher seltenes Ereignis.



Detailwissen IV: Wie identifiziert Strava Pendlerfahrten?

Die Firma Strava identifiziert „Commutes“ vor allem durch ein „Point-to-point“-Matching-Verfahren. Hierbei werden besonders häufig gefahrene Relationen zwischen Quelle und Ziel als regelmäßig und demzufolge auch als Pendler- bzw. Alltagsfahrt eingestuft“ (Strava LLC, 2016). Mit diesem Verfahren können laut Herstellerangaben circa 98% aller durch die Nutzerinnen und Nutzer als Pendelfahrten angegebenen Fahrten erkannt werden sowie eine größere Anzahl an Fahrten, die durch sie nicht als solche gekennzeichnet wurden, aber dennoch die entsprechenden Kriterien erfüllen. Dieses Vorgehen entspricht formell nicht der in Europa gängigen Definition

einer Pendelfahrt in Sinne eines regelmäßigen Weges zwischen Wohnort und Arbeits- oder Ausbildungsplatz, schließt diese allerdings vollumfänglich ein. Ebenso eingeschlossen werden weitere, häufig wiederkehrende Fahrten, wie beispielsweise beliebte Trainingsstrecken von Rennradfahrern oder die Etappen von Radrennen. Hier scheint es die Unternehmensstrategie zu sein, möglichst viele Alltagsfahrten als solche zu markieren und etwaiges ‚Rauschen‘ zu Gunsten einer möglichst hohen Anzahl von Pendelfahrten in Kauf zu nehmen. Dies sollte unbedingt bei der Interpretation der Daten als Pendlerfahrten beachtet werden.



Detailwissen V: Von Strava gelieferte Datengrundlage der Quelle-Ziel-Matrizen

Im vorliegenden Projekt wurde für die Bildung eines Quelle-Ziel-Netztes das einheitliche europäische Gittersystem für amtliche Statistik verwendet. Dazu wurden regelmäßige Quadrate (Polygone) mit einer Kantenlänge von 1.000m gebildet, auf die dann von Strava die diesem Bereich

zugehörigen Daten projiziert wurden. Diese entsprechen dem Standard des neuen statistischen Polygonnetzes. Die durch Strava auf die gelieferten Polygone projizierten Daten haben die in Tabelle 5 dargestellte Form.

	polygon_id integer	year integer	day integer	hour integer	minute integer	commute integer	dest_polygon_id integer	intersected_polygons integer[]
1	339081	2015	191	7	30	1	344108	{339528, 339970, 340403, 340404, 340824, 341243, 341658, 342069, 342477, 342478, 342888, 343297, 343703, 339081, 339082, 339527}
2	340399	2015	223	16	58	1	341238	{340399, 340819, 341238}
3	340403	2015	224	16	5	0	340403	{340404, 340405, 340825}
4	341648	2015	328	17	42	0	341648	{341648, 342059, 342467, 342058, 341647}

Tabelle 5: Schematische Darstellung Datensatz polygons_ride, SQL-Datenbankauszug

Die einzelnen Spalten beinhalten dabei folgende Daten:

<i>Polygon_id</i>	Startpolygon der gefahrenen Strecke; eindeutiger Wert für die Polygonbezeichnung, im Projekt fortlaufende Nummerierung
<i>year, day, hour, minute</i>	Spalten, die den Zeitstempel markieren
<i>commute</i>	1 = Pendlerfahrt; 0 = Freizeit-/Sportfahrt
<i>dest_polygon_id</i>	Zielpolygon der gefahrenen Strecke, eindeutiger Wert für die Polygonbezeichnung, im Projekt fortlaufende Nummerierung
<i>Intersected_polygons</i>	Bezeichnet die zwischen Start und Ziel durchfahrenen Polygone

Dabei können Start- und Zielpolygon den gleichen Wert aufweisen, wenn es sich um ‚Rundfahrten‘ handelt. Neben einer Darstellung der am häufigsten genutzten Quellen und Ziele für Radfahrten können auch die Wegelängen und die genutzten Korridore abgeschätzt werden. Dabei

wird näherungsweise die Kantenlänge eines Polygons mit der Anzahl der durchfahrenen Polygone multipliziert. Die Fahrtweite wird dabei unter Umständen etwas überschätzt, aber es lässt sich so auf diese schließen.



Detailwissen VI:

Ansätze zur Repräsentativitätsprüfung – 1. Verkehrsmengen

Zur Reliabilitätsprüfung wurden Zähldaten aus Dresden mit den Strava-Daten verglichen. Es wurden sowohl Daten aus automatischen Zählstellen als auch Daten aus Kurzzeitzählungen verwendet.

Während die Daten der Zählstellen stundenfein vorliegen, sind die realisierten Fahrten innerhalb des Strava-Datensatzes für die jeweiligen Netzelemente natürlich entsprechend geringer verteilt. Dabei kann der Fall eintreten, dass an mehreren Stunden eines Tages keine Radfahrenden erfasst werden. Das macht es damit zunächst nötig, die Verteilung der Radfahrenden über einen längeren Zeitraum zu betrachten. Aufgrund des über einen längeren Zeitraum vorliegenden Datensatzes wurde hierfür das Straßennetz der vereinfachten OpenStreetMap-Karte verwendet (siehe Kapitel 4.2, Schritt 3). Grundlage der Untersuchung sind die Verkehrsstärken aller Radfahrenden der Kanten, die zunächst minutenfein vorliegen.

Im Zuge der Datenaufbereitung fanden die folgenden Arbeitsschritte statt:

1. Codierung einer Zeit- und Datumsvariable
2. Kombination der Radfahrerinnen und Radfahrer von mehreren parallelen Kanten und Ausschluss von doppelten Fällen anhand des Kriteriums „mehrere Fälle in der gleichen Minute auf zwei parallelen Kanten“

3. Aggregation der minutenfeinen Daten auf Stunden

4. Ergänzung von Werten 0 bei h ohne Radfahrende

Der Vergleich zwischen den GPS-basierten Daten und den Dauerzählstellen erfolgt für exakt miteinander korrespondierende Zeiträume. Dafür wird der gesamte Zeitraum, für den Daten der Dauerzählstellen vorliegen, einbezogen.

Zuerst sollen die Ergebnisse des Vergleichs der Dauerzählstellen mit den Strava-Daten betrachtet werden. Daten von sechs Dauerzählstellen der Landeshauptstadt Dresden im Zeitraum vom 01.09.2015 bis zum 31.05.2016 wurden dafür mit den GPS-Daten verglichen. Die Tagesganglinie auf Basis der Dauerzählstellen ist in Abbildung 18 dargestellt. Die Verkehrsstärken liegen zwischen 400 und 2.000 Radfahrenden, wobei sich die Unterschiede gut anhand der Funktion der Straßen erklären lassen. Die Stichprobe innerhalb des Strava-Datensatzes ist deutlich geringer. In dem neunmonatigen Zeitraum liegt die Summe aller Radfahrenden mit Strava-Nutzung zwischen 500 und 5.500. Die sich daraus ergebenden Tagesganglinien sind in Abbildung 19 dargestellt.

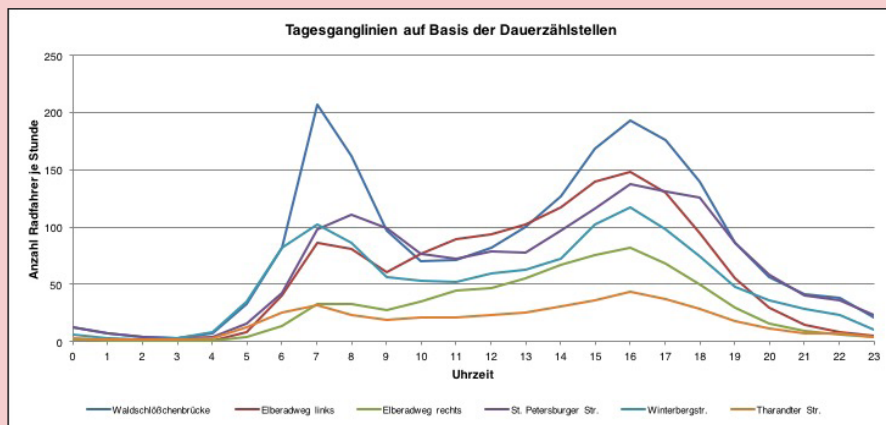


Abbildung 18: Tagesganglinien der Dauerzählstellen in Dresden im Zeitraum 09/2015 – 05/2016

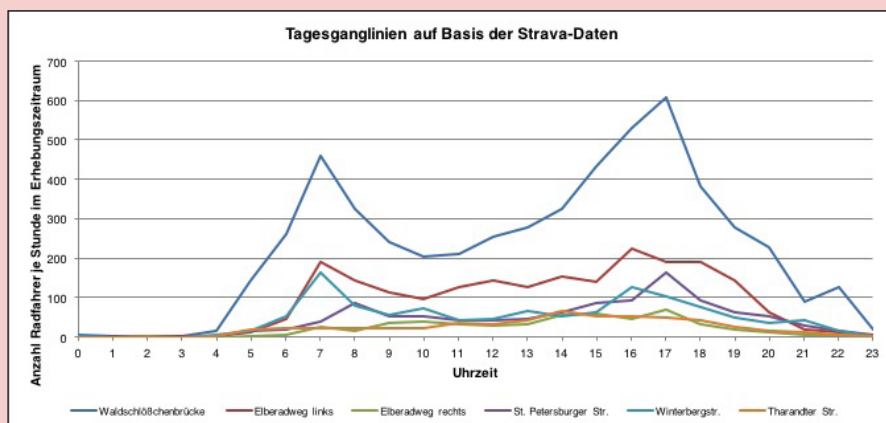


Abbildung 19: Tagesganglinien auf Basis der Strava-Daten in Dresden im Zeitraum 09/2015 – 05/2016



Fortsetzung Detailwissen VI: Ansätze zur Repräsentativitätsprüfung – 1. Verkehrsmengen

Die Hochrechnung der Strava-Daten erfolgte anhand von zwei unterschiedlichen Verfahren. Die Bezugsgröße ist hierbei, abweichend vom in der Analyse betrachteten Zeitraum, jeweils die Hochrechnung von Strava-Daten eines durchschnittlichen Tages auf die Werte der städtischen Dauerzählstellen eines einzelnen, durchschnittlichen Tages. Zuerst wurden mittels mehrerer linearer Regressionen die Stichproben der Verkehrsmengen an den jeweiligen

Dauerzählstellen der Landeshauptstadt Dresden hochgerechnet. Anschließend wurde aus den einzelnen Regressionsgleichungen eine für das ganze Stadtgebiet gültige Regression errechnet (siehe Abbildung 20). Die Regression verwendet die Strava-Daten als Ausgangsdaten (x) und leitet daraus die Werte eines Tages f(x) ab. Die entsprechende Regressionsformel auf Stundenebene lautet daher $f(x) = 0,5087 * x + 16,9719$.

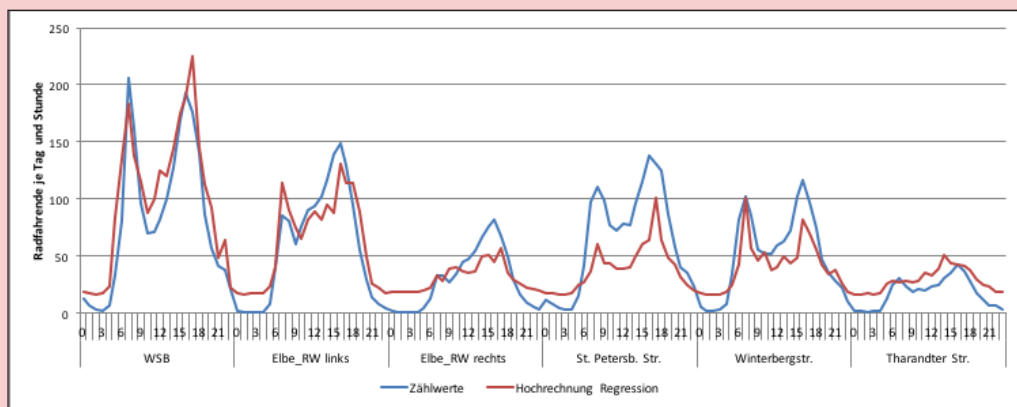


Abbildung 20:
Hochrechnung der Strava-Daten auf das Zählstellenniveau der Dauerzählstellen in Dresden für den durchschnittlichen Tagesgang, auf Basis einer linearen Regression

Zum Zweiten wurde als alternativer Ansatz ein konstanter Faktor ermittelt. Dieser ergibt sich aus den Mittelwerten des Quotienten aus den Werten der Dauerzählstelle und jenen von Strava auf Stundenbasis (siehe Abbildung 21). Das arithmetische Mittel liegt bei 0,00371 und dieser Wert wird für die Prognose in Abbildung 24 verwendet. Der Median liegt bei 0,00306, Minimum und Maximum bei 0 und 0,1449 und die Standardabweichung beträgt 0,00262. Für die Hochrechnung von Jahreswerten auf Tageswerte ergibt sich daraus ein Quotient von 1,3. Das Radverkehrsaufkommen eines Jahres aus dem Datensatz von Strava lässt sich also mit dem Faktor 0,76 auf den DTV herunterrechnen. Selbiges Verfahren liefert für die Berliner Dauerzählstellen den Faktor 0,77. Auffällig bei der Prognose anhand des Mittelwerts sind die hohen Abweichungen bei der Waldschlösschenbrücke. Die prognostizierte Radverkehrsstärke liegt in der Summe

etwa doppelt so hoch wie die an der Dauerzählstelle gemessenen Werte. Eine mögliche Ursache auf Seiten der Strava-Daten könnte beim Mapmatching liegen. Die Brücke besteht aus vier digitalisierten Achsen. Mit der Duplikatesuche wurde zwar bereits etwa die Hälfte der Fälle ausgeschlossen, aber zusätzlich erscheint die Gesamtmethodik bei parallelen Achsen mit Unsicherheiten behaftet zu sein. Aufgrund der Nähe zu einem Knotenpunkt können auch abbiegende Verkehre nicht gänzlich in der Strava-Stichprobe ausgeschlossen werden. Eine andere Ursache könnten die Werte der Dauerzählstelle sein. Diese sind zwar im Hinblick auf Tages- und Jahresganglinie plausibel, aber eine systematische Untererfassung kann nicht ausgeschlossen werden, da der Zugang zur Brücke auch ohne Befahrung der Dauerzählstelle über eine Treppe vom Elberadweg aus möglich ist.

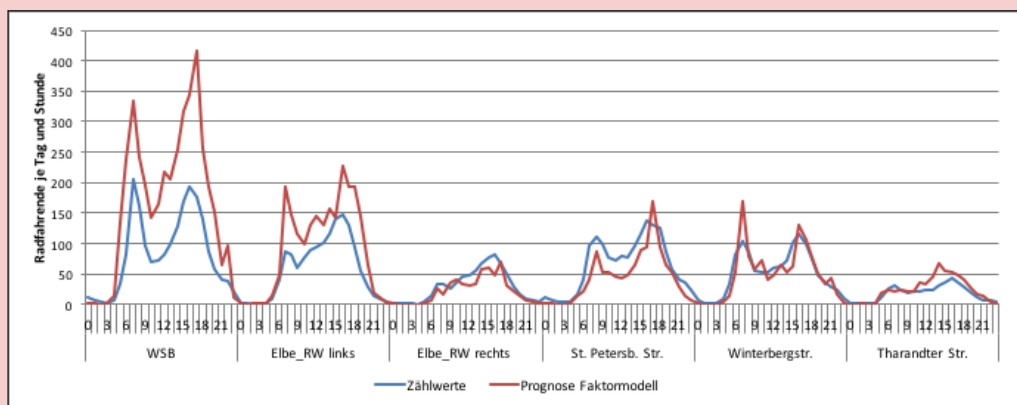


Abbildung 21:
Prognose für Dauerzählstellen auf Basis eines konstanten Faktors berechnet aus den Mittelwerten des Quotienten aus den Werten der Dauerzählstelle und jenen von Strava auf Stundenbasis

Bei der Beurteilung, welches Verfahren besser zur Hochrechnung geeignet ist, sollten die folgenden Aspekte berücksichtigt werden. Bei der linearen Regression gibt es innerhalb der Gleichung, welche die Form

$$f(x) = ax+b$$

aufweist, stets einen konstanten Wert (hier: $b = 16,9719$). Dies bedeutet, dass auf Basis dieser Hochrechnung, auch bei Straßen ohne jegliche Strava-Nutzer, von einer Mindestverkehrsstärke von etwa 17 Radfahrern je Stunde auszugehen wäre. Während der Nachtstunden ist dieser

Effekt in Abbildung 23 deutlich zu erkennen an der Differenz zwischen der roten und der blauen Linie in den nächtlichen Intervallen. Inhaltlich ist dies schwer zu interpretieren. Statistische Kenngrößen zur Sicherheit von Prognosen bieten darüber hinaus eine quantitative Einschätzung der Prognosesicherheit des Radverkehrshauptnetzes in Dresden (siehe Tabelle 6). In der Gesamtbewertung zeigt die intuitiv naheliegende Hochrechnung mittels Faktor eine deutlich höhere Prognosesicherheit, da der mittlere absolute Fehler sowie der relative Fehler geringere Werte aufweisen.

Tabelle 6: Statistische Kennwerte zum Vergleich der Prognosen mittels linearer Regression und konstantem Faktor für Dresden

Statistischer Kennwert	Lineare Regression	Konstanter Faktor
Relativer Fehler	0,363	0,023
Korrelation	0,869	0,866
Varianzaufklärung (r^2)	0,434	0,433

Kurzzeitzählungen wurden von der Professur Verkehrsökologie der TU Dresden im Zeitraum vom 22.05.2015 bis zum 12.06.2015 dienstags bis donnerstags von 14:00-18:00 Uhr durchgeführt. In die Auswahl der Orte für Kurzzeitzählungen flossen die Kriterien verkehrliche Bedeutung, Parallelrouten zum Hauptverkehrsstraßennetz und Verdacht auf überproportionalen Anteil von Sportradfahrern ein.

Das Verhältnis der gezählten Radfahrenden und der im exakt gleichen Zeitraum aufgezeichneten Strava-Nutzerinnen und Nutzer beträgt an den zehn untersuchten Querschnitten 7.309 : 7. Danach nutzten nur 0,096 % der erfassten Radfahrenden in den jeweils vierstündigen Zählperioden die Strava-App. Schon statistisch stimmen also Stichprobe und Grundgesamtheit hier nicht überein. Die Ursachen für die Abweichungen zum Mittelwert der Dauerzählstellen sind vielschichtig. Während bei den Strava-Daten ein gesamtes Jahr in die Analyse einfließt, fanden die Kurzzeitzählungen nur im Mai und Juni bei überwiegend ‚gutem Radfahrwetter‘ statt. Die vergleichsweise hohen Zählraten können somit saisonal bedingt sein.

Daher wird für eine netzweite Hochrechnung der Daten auf die Grundgesamtheit aktuell die Variante eines Faktors, kalibriert an den mittleren Stundenwerten von Dauerzählstellen, empfohlen, da diese wohl die tatsächlichen Belegungen in einem Radverkehrsnetz am stabilsten wiedergeben. Es ist jedoch vor allem im Nebennetz mit Abweichungen zu rechnen.

Werden dagegen nur Kurzzeitzählungen als Vergleichsgröße verwendet, wächst die Spannbreite der Faktoren für die einzelnen Zählstandorte – der mittlere Hochrechnungsfaktor wird ungenauer und die Prognosesicherheit sinkt. Gleiches gilt für die Verwendung von Daten aus einem Verkehrsmodell. Dies wurde für die Daten der Stadt Leipzig getestet. Statistisch gesehen handelt es sich auch hier um zwei unabhängige Stichproben, wobei die Daten aus dem Verkehrsmodell ihrerseits abhängige Sekundärdaten sind. Auf dieser Basis konnten keine sinnvollen Hochrechnungen durchgeführt werden.



Fortsetzung Detailwissen VI:

Ansätze zur Repräsentativitätsprüfung – 2. Geschwindigkeiten

Aufgrund der durch die Inhalte der App mutmaßlich eher angesprochenen sportaffinen, männlichen Nutzergruppe war bei der Fahrgeschwindigkeit ein Unterschied zwischen der Strava-Stichprobe und der Grundgesamtheit ‚Radfahrende in Dresden‘ zu erwarten. Dieser wurde durch umfangreiche Vergleichsmessungen bestätigt (n = 1.000, Messung an fünf Querschnitten in Dresden im gesamtem Stadtgebiet). Für die Vergleichswerte innerhalb des Strava-Datensatzes wurde, um einer sinnvollen Stichprobengröße gerecht zu werden, der Datensatz nur anhand der Zeitscheiben, nicht jedoch anhand des konkreten Datums an die Messspezifika angeglichen. Dies liegt zum einen an der erwähnten Stichprobengröße: ein statistischer Vergleich erfordert gleiche Fallzahlen. Zum anderen nimmt die Wahl einer möglichst niedrigen Disaggregationsform Einfluss auf die Verwendung in der kommunalen Praxis. In diesem Fall wäre dies ein einfacher Mittelwert über einen Erhebungszeitraum. Zu erwähnen ist weiterhin, dass es sich bei den Werten der Strava-Stichprobe nicht um Messwerte als solche handelt, sondern ebenfalls um berechnete Werte aus der benötigten Zeit für das Passieren der betreffenden Kante im GIS-Netz.

Die GPS-Daten weisen im Durchschnitt eine höhere Geschwindigkeit auf als der in den Vergleichsmessungen ermittelte Geschwindigkeitswert. Dies wird in der Betrachtung der Mittelwerte der Geschwindigkeiten bei Strava und der Vergleichsmessungen in Tabelle 7 deutlich. Die Höhe des Unterschiedes differiert jedoch in geringem Maße mit Ausreißern in den Querschnitten Zellescher Weg R2 und Grundstraße R2. Dabei stellt der Querschnitt am Zelleschen Weg den mit der geringsten gemessenen Durchschnittsgeschwindigkeit dar. Dies ist sehr wahrscheinlich durch die hohe Frequenz an Studierenden bedingt, die durch die Nähe zum Campus und die Lage der Universitätsbibliothek unmittelbar am Querschnitt viele sehr kurze Quell- und Zielverkehre realisieren. Der Querschnitt Grundstraße ist, anders als die übrigen, eine Strecke mit starker Steigung (Länge insgesamt 2,9km, 124Hm, Ø-Steigung 4,2%, Steigung max. 7%). Entscheidend für die Interpretation der Daten ist jedoch nicht nur zwingend der Mittelwert der gefahrenen Geschwindigkeiten, sondern auch deren Verteilung. Diese erweisen sich vor allem auf ebenen Streckenelementen als ähnlich normalverteilt (siehe Abbildung 23).

Tabelle 7: Mittelwerte und Standardabweichungen der Geschwindigkeiten an den untersuchten Querschnitten, ermittelt aus GPS-Daten (Strava) und Vergleichsmessungen

Querschnitt	Mittelwert (in km/h)		Standardabweichung (in km/h)	
	Strava	Messung	Strava	Messung
Elberadweg WSB R1	26,07	21,17	4,18	4,72
Elberadweg WSB R2	26,25	22,08	5,12	4,70
Elberadweg ALB R1	26,26	21,70	5,06	3,72
Elberadweg ALB R2	26,29	20,20	4,60	3,41
Zellescher Weg R1	26,33	21,56	5,73	5,00
Zellescher Weg R2	27,00	19,34	5,74	4,20
Chemnitzer Straße R1	29,25	23,56	6,07	5,06
Chemnitzer Straße R2	21,14	21,30	5,59	4,46
Grundstraße R1	41,89	33,68	8,41	5,55
Grundstraße R2	17,17	16,13	4,38	4,44

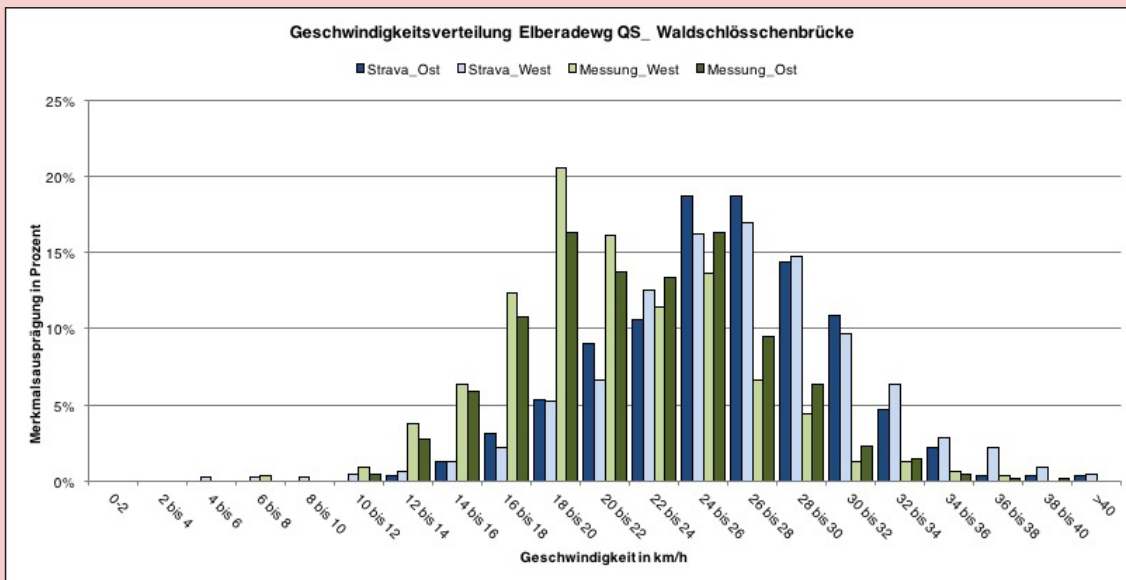


Abbildung 22: Geschwindigkeitstagesgangzeiten und Tagesganglinien auf Basis der Strava-Daten in Dresden im Zeitraum 09/2015 – 05/2016

Basierend auf den aggregierten Verteilungen lässt sich für alle Messstellen eine mittlere Abweichung von rund 5,5 km/h feststellen (siehe Abbildung 22). Die betrachteten Datensätze weisen auch innerhalb statistischer Analysen eine hohe Korrelation auf. Überschlüssig ist es somit für die planerische Praxis praktikabel, die Geschwindigkeiten der

App-Daten um 5,5 km/h zu vermindern, um das tatsächliche gemessene Niveau, d.h. das reale Geschwindigkeitsniveau des Durchschnittsradfahrenden in Dresden abzubilden und eine höhere Aussagekraft der Werte zu erreichen (siehe Abbildung 23).

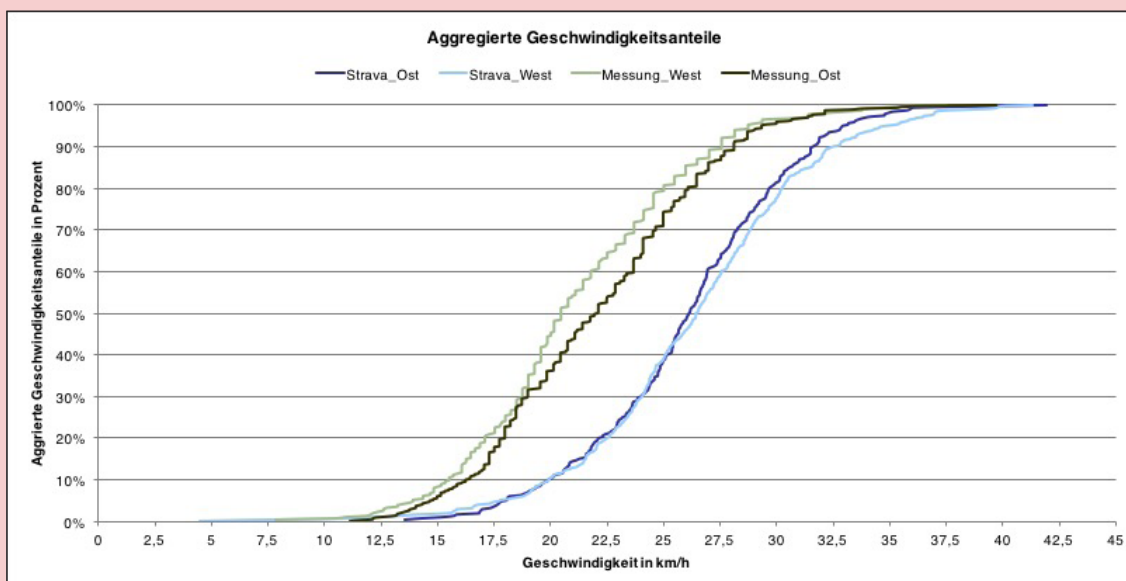


Abbildung 23: Geschwindigkeitsverteilung der Radfahrenden in Dresden, aggregiert, Hauptnetz, Elberadweg (n = 316/473)

Das Vorliegen einer normalverteilten Stichprobe ermöglicht eine Regressionsanalyse der unabhängigen Größe. Diese Voraussetzungen sind sowohl im Haupt- als auch im Nebennetz für ebene Messquerschnitte gegeben. Eine Regressionsanalyse ist also möglich. Gewählt wurde diese Herangehensweise, da eine Korrektur-Konstante (-5,5 km/h) zwar die Richtung der Korrektur deuten kann, aber nicht zwangsläufig für alle Messstellen eine hohe Ergebnisqualität liefern würde. In diesem Fall stellt die Stichprobe der App-Nutzer die unabhängige Größe dar, und die Messwerte sollen zukünftig mittels Regressionsgleichung geschätzt werden. Es wurde eine einfache lineare Regression gerechnet. Alle weiteren Parameter wurden vereinfachend als gleich, und damit vernachlässigbar angenommen. Weitere Forschung

in diesem Bereich kann jedoch die Ergebnisqualität nochmals erhöhen. Dies bedarf jedoch umfangreicher Feldversuche, die im Rahmen des vorliegenden Projektes nicht durchgeführt werden konnten.

Die Regressionsrechnung ergab eine Korrelation der Werte nach Pearson von 0,922, sowie ein Bestimmtheitsmaß von $R^2 = 0,85$. Das heißt, 85% der Variationen der realen Geschwindigkeiten können mit diesem linearen Modell erklärt werden. Die Modellgüte ist demzufolge als hoch einzuschätzen. Es ist in diesem Fall also möglich, auf Basis von per App generierten GPS-Daten die tatsächlichen Fahrgeschwindigkeiten von Radfahrenden in ebenen Bereichen der Stadt zu prognostizieren.



Fortsetzung Detailwissen VI:

Ansätze zur Repräsentativitätsprüfung – 3. Quelle-Ziel-Beziehungen

In den vorangegangenen Abschnitten wurden häufig die Quelle-Ziel-Matrizen als eine mögliche Darstellungsform zitiert. Auch diese Daten bergen Möglichkeiten durch eine heterogene Stichprobe beeinflusst zu werden. So wird zunächst die räumliche Verteilung des Quellverkehrs anhand der Einwohnerzahlen der Quellverkehrszellen validiert (siehe Abbildung 24). Diese basieren auf den Daten des Zensus 2011. Einschränkend ist hier zunächst zu

erwähnen, dass Quellverkehre natürlich auch beispielsweise vom Arbeitsplatz ausgehen können. Hierfür wären weitere vertiefende Untersuchungen anzuraten. Zum Zeitpunkt der vorliegenden Studie standen die Arbeitsplatzdaten jedoch nicht zur Verfügung. Mit dieser Einschränkung ist die statistische Auswertung der Quellverkehrszahlen nur eingeschränkt gültig, liefert jedoch einen ersten Überblick über die Datengüte.

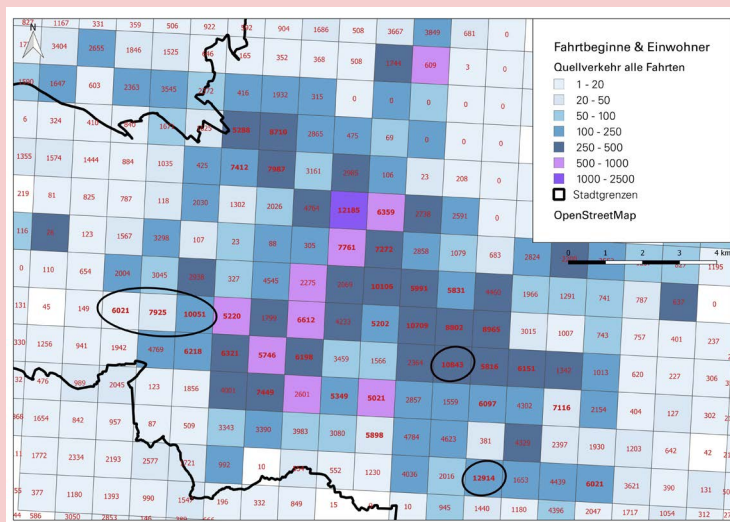


Abbildung 24: Vergleich der Quellverkehre (Alltagsverkehr) mit den Einwohnerzahlen der Stadt Dresden (Quelle: Zensus 2011)

Hierbei zeigt sich, dass für einige Quellverkehrszellen ein hoher Zusammenhang zwischen der Anzahl der Einwohner und der Anzahl der ausgehenden Fahrten zu bestehen scheint. Dies lässt sich jedoch nicht für alle Quellverkehrsbezirke generalisieren. Einige Quellverkehrszellen mit hoher Einwohnerzahl, jedoch eher schwächerem sozialem Milieu weisen vergleichsweise geringe Fahrtenzahlen auf. Hier ist ein hoher Einfluss der homogenen App-Nutzergruppe zu vermuten. Dieser resultiert in diesem Falle eventuell in einer Art sozialer Exklusion für Menschen mit geringerem Bildungsstand oder Einkommen. An dieser Stelle sind geeignete sozialwissenschaftliche Studien zu empfehlen.

Die statistische Auswertung liefert auf Basisniveau eine hohe positive Korrelation der Werte Einwohnerzahl und `start_activity_count`, welcher die Anzahl der Fahrtbeginne beschreibt. Dafür wurde der gesamte Datensatz für das Untersuchungsgebiet im Zeitraum 01/2015 bis 06/2016 ausgewertet. Es kann also davon ausgegangen werden, dass ein Zusammenhang der Werte besteht.

Die Werteausgaben für die Städte Dresden, Chemnitz, Berlin und Leipzig in Tabelle 8 zeigen, dass diese Korrelation (r) lokal unterschiedlich stark ausgeprägt ist und stabil bei ca. 0,7 liegt.

Tabelle 8: Wertekorrelation zwischen Fahrtbeginn und Einwohnerzahl für die untersuchten Städte

Ort	Steigung	Korrelation r	Bestimmtheitsmaß (r^2)	Korrigiertes r^2	Kovarianz	Anzahl Datensätze
Dresden	0,103	0,689	0,474	0,472	583.818	394
Chemnitz	0,050	0,672	0,452	0,448	153.053	276
Leipzig	0,0617	0,702	0,494	0,491	407.306	361
Berlin	0,060	0,678	0,460	0,458	1.110.467	1.017
Gesamt	0,052	0,660	0,435	0,435	159.337	377.048

Dabei ist die Korrelation allerdings stark durch jene Gebiete mit sehr geringen Einwohnerzahlen und Abfahrtszahlen determiniert. Filtert man dementsprechend nach Abfahrtszahlen größer 100 und Einwohnerzahlen größer 0, dann

sinkt die Korrelation auf ca. 0,6. In diesem Fall ist jedoch immer noch von einem starken positiven Zusammenhang zu sprechen.



Detailwissen VII: Weiterführende Auswertungsmöglichkeiten der Datensätze

Die selbstständige Arbeit mit den GPS-Daten ermöglicht es den Nutzerinnen und Nutzern beispielsweise mit einem PostGIS-Aufsatz für PostgreSQL-Datenbanken eigene Visualisierungen, wie Differenznetze, zu erzeugen und fortlaufend in den Online-Themenstadtplan einzuspielen (siehe Abbildung 25). Auch ist so eine Plausibilitätskontrolle der erhaltenen Daten sinnvoll möglich.

Eine Alternative dazu bietet die Vergabe der Datenaufbereitung als Dienstleistung an ein externes Ingenieurbüro. Hierfür sind vor allem die Formate der Ergebnisübergabe im Vorfeld genau zu definieren.

Die Datenübergabe erfolgt ebenso nach zu vereinbarenden Kriterien – so bietet beispielsweise Strava sowohl monatliche Datentranchen als auch rückblickend die Daten für ein Jahr an. Aus Sicht der Autoren ist eine monatliche Datentranche aufgrund der im Verhältnis zur Gesamtbevölkerung noch eher geringen Nutzerzahlen aktuell noch nicht aussagekräftig genug, um ihrem eigentlichem Zweck, z.B. der schnellen Überprüfung von Maßnahmenwirkungen, gerecht zu werden.

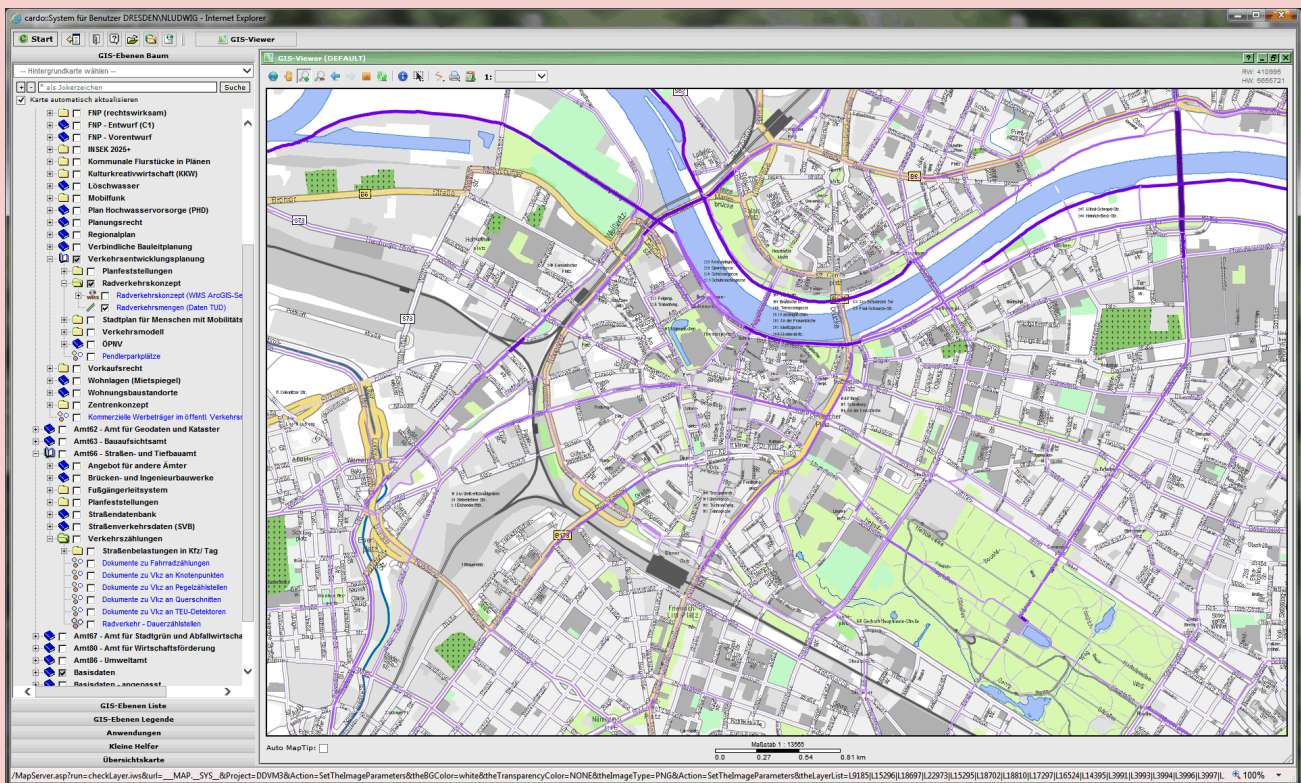


Abbildung 25: Online-Themenstadtplan für die Stadt Dresden (Quelle: Stadt Dresden)

- App:** Apps, oder Applikationen, beschreiben Anwendungsprogramme für Smartphones. Im vorliegenden Leitfaden liegt dabei der Fokus auf Anwendungen, die die Aufzeichnung und Speicherung von GPS-Daten ermöglichen.
- Datenbanken (SQL, PostgreSQL):** Relationale Datenbanken dienen zur Datenverwaltung in Computersystemen und basieren auf einem tabellenbasierten Datenbankmodell. Structured Query Language (SQL) ist dabei eine Datenbanksprache zur Definition von Datenstrukturen und zur Bearbeitung und Abfrage von Datenbeständen. Bei PostgreSQL handelt es sich um ein freies Datenbankmanagementsystem, das weitgehend konform mit dem SQL-Standard ist. Es verfügt über eine Erweiterung zur Verwaltung von Geo-Daten (PostGIS).
- Differenznetze** Differenznetze beschreiben eine Darstellung in Geoinformationssystemen, welche den gleichen Parameter innerhalb eines Netzes, beispielsweise zu zwei unterschiedlichen Zeiten, darstellt. Diese Darstellungsform wird in der Regel genutzt um Entwicklungen zu visualisieren.
- GIS:** Geographische Informationssysteme sind Informationssysteme zur Erfassung, Bearbeitung und Organisation sowie zur Analyse von Daten mit Raumbezug. So werden unter anderem auch die Daten des Straßennetzes in derartigen Informationssystemen abgelegt und gepflegt.
- GIS-Element:** In einem Straßenknotennetz ist der Netzgraph durch die Elemente Knoten und Kante aufgebaut. Knoten werden dabei durch Kanten verbunden. Eine Darstellung von Verkehrsströmen benötigt einen konsistenten, lückenfreien Netzaufbau. Den einzelnen Elementen können dabei unterschiedliche Eigenschaften wie beispielsweise zugelassene Höchstgeschwindigkeiten oder Vorfahrtsregelungen zugewiesen werden.
- GPS-Daten:** GPS-Daten stehen synonym für durch das Global Positioning System (GPS) und weitere Dienste, wie GLONASS oder Galileo, erhobene Positionsdaten. Dabei wird ein globales Navigationssatellitensystem zur Positionsbestimmung genutzt. Die Satelliten strahlen dabei kontinuierlich ihre Position sowie die genaue Uhrzeit aus. Das ermöglicht es, für spezielle Empfänger die Position und Geschwindigkeit zu berechnen. Die Daten beinhalten stets Angaben über Längengrad, Breitengrad und Uhrzeit. Aktuell wird eine horizontale Genauigkeit von 7,8m garantiert.
- Kantendaten:** Siehe auch GIS-Element. Unter Kantendaten sind synonym alle Eigenschaften einer Kante in einem Netzgraph zusammengefasst, wie zum Beispiel die zulässige Höchstgeschwindigkeit, die Straßenklasse, das Vorhandensein von Radverkehrsanlagen oder Radverkehrsmengen.
- Knotenpunkt-daten:** Siehe auch GIS-Element. Unter Knotendaten sind synonym alle Eigenschaften eines Netzknötens zusammengefasst. Das können beispielsweise die Formen der kreuzenden Verkehrswege sein, oder der daraus resultierende Knotenpunkttyp. Ebenso können Wartezeiten für Verkehrsströme hinterlegt werden.
- Mapmatching:** Mit Mapmatching, oder Kartenabgleich, wird ein Verfahren bezeichnet, welches die durch eine Ortung, beispielsweise via GPS, gemessene Position mit den Ortsinformationen einer Karte abgleicht. Im vorliegenden Projekt ist damit der Abgleich und die Zuordnung der GPS-Daten mit bzw. zu einem GIS-Netz gemeint.

<i>Metro Extracts:</i>	Produkt der Firma Strava. Es handelt sich um eine Online-Plattform für Kommunen, über die Radverkehrsmengen, Wartezeiten und weitere Angebote einzusehen sind.
<i>MiD:</i>	Mobilität in Deutschland ist eine bundesweite Haushaltsbefragung zum Verkehrsverhalten der Bundesbürger. Sie findet etwa im 5-Jahresturnus statt. Die Ergebnisse der aktuellen Erhebung stehen voraussichtlich ab 2018 zur Verfügung.
<i>MIV:</i>	Motorisierter Individualverkehr beschreibt Pkw, leichte Nutzfahrzeuge, Motorräder, Mofas oder ähnliche motorgetriebene Verkehrsmittel, die sich im Besitz einer Privatperson befinden und überwiegend der Realisierung der eigenen Mobilität dienen.
<i>Nebennetz:</i>	Das Nebennetz im Radverkehr beschreibt nicht zwingend das Nebenstraßennetz oder vergleichbare Klassifizierungen nach Netzkategorie, wie es für den MIV logisch erscheint, sondern Netzelemente von untergeordneter Netzbedeutung für den Radverkehr. Das können, aber müssen nicht notwendigerweise, Nebenstraßen sein.
<i>OSM:</i>	OpenStreetMap ist ein freies Projekt, dessen Kern eine öffentlich zugängliche Datenbank aller freien Geoinformationen ist. Aus diesen Daten können beispielsweise die im vorliegenden Projekt genutzten Landkarten erstellt werden. Die Nutzung ist dabei kostenfrei.
<i>Point-to-Curve Matching:</i>	Point-to-Curve-Matching beschreibt ein Verfahren der Zuordnung von gemessenen Positionsdaten zu einer Karte, in dem der kürzeste Abstand des Positionspunktes zu einer Kurve (Kante) verwendet wird.
<i>Point-to-Point:</i>	Siehe auch Map-Matching. Point-to-Point Matching beschreibt eine Methode des Mapmatching, bei der GPS-Datenpunkte Punkten des vorhandenen Verkehrsnetzes zugeordnet werden. Dies können für einen Netzgraphen entweder Knotenpunkte, Anfangs- oder Endpunkte von Kanten oder auch erzeugte Punkte entlang von Kanten sein.
<i>Quelle-Ziel-Matrizen:</i>	Quelle-Ziel-Matrizen beschreiben Parameter, die entlang einer Ortsveränderung von einer Quelle zu einem Ziel realisiert werden. Das können Verkehrsmengen unterschiedlicher Verkehrsarten oder auch zeitliche Aufwände sein.
<i>Quelle-Ziel-Relationen:</i>	Siehe Quelle-Ziel-Matrizen. Quelle-Ziel-Relationen beschreiben den räumlichen Zusammenhang zwischen einer Quelle (Beginn) und einem Ziel (Ende) einer Ortsveränderung. Das muss jedoch nicht notwendigerweise ein Weg sein.
<i>Reliabilität:</i>	Auch Zuverlässigkeit, ist ein Maß für formale Genauigkeit bzw. Verlässlichkeit von Messungen.
<i>Repräsentativität:</i>	Repräsentativität (einer Zufallsstichprobe) ist die Charakteristik bestimmter Datenerhebungen, die es ermöglicht, mit einer kleinen Stichprobe Aussagen über größere Mengen (Grundgesamtheit) treffen zu können. Hierbei ist es bei personenbezogenen Daten wichtig, dass die Schätzung von Alter, Bildung und Familienstand in etwa der amtlichen Statistik entsprechen.

- Routendaten:** Routendaten beschreiben die Summe aller GPS-Punkte und deren entsprechenden Geschwindigkeiten in einem meist linienhaften Verlauf von einer Quelle über eine Ortsveränderung zu deren Ziel. Darüber hinaus können Routendaten auch Angaben zu Wegezwecken, der Person oder ähnlichem enthalten.
- SrV:** Das System repräsentativer Verkehrsbefragungen der TU Dresden ist eine aller fünf Jahre stattfindende Verkehrserhebung im Stadtverkehr. Die letzte Feldphase fand im Jahr 2013 statt.
- Städtische Geoinformationssysteme (AKTIS):** Das Amtliche Topographisch-Kartographische Informationssystem wird zum Zweck der digitalen Führung der Ergebnisse der topographischen Landesaufnahmen und der amtlichen topographischen Karten von den Vermessungsämtern und dem Bundesamt für Kartographie und Geodäsie aufgebaut.
- Tagesganglinie:** Die Tagesganglinie beschreibt die Verteilung der stündlichen Verkehrsmengen eines Modus an einem Querschnitt. Dabei sind an Arbeitstagen Spitzenstunden jeweils am Morgen (ca. 7:00 – 9:00Uhr) und am Nachmittag (15:00-17:30Uhr) zu verzeichnen.
- Validierung:** Die Validierung beschreibt die Prüfung der Nutzereingaben oder statistischer Daten auf die Eignung für eine bestimmte Anwendung.
- Verdrängungsverkehre:** Verkehre, die durch eine Maßnahme oder ein anderes Ereignis im Straßennetz, welches den Widerstand bzw. die Reisezeit erhöht, statt auf ihrer üblichen Route über eine eigentlich schlechter geeignete oder unerwünschte Alternativroute abfließen.
- Verkehrszelle:** Eine Verkehrszelle ist eine theoretische Raumeinheit, die einen Teil einer Stadt oder Siedlung verkörpert. Verkehrszellen werden vor allem als Bezugseinheit für die Modellierung verwendet. Dafür werden ihnen Daten, wie Fahrzeugbesitz, Einwohnerzahl, Arbeitsplatzzahl und weitere, zugeordnet.
- Zeitscheiben:** Einteilung eines Tages in verschiedene Zeiteinheiten. Im Projekt wurden überwiegend stundenfeine Zeitscheiben genutzt.

Häufige Fragen

Welche GPS-Daten gibt es?

GPS-Daten sind grundsätzlich Positionsdaten und bilden somit die Bewegung von Radfahrenden im Raum ab. Mögliche Datenquellen sind Tracking-Apps für Smartphones (beispielsweise GPS-Tracker), Sportapps (beispielsweise Strava) oder auch Navigationsapps (beispielsweise BikeCitizens). Entscheidend ist dabei die Frequenz der Datenaufzeichnung. Auch einige Fahrradverleihsysteme verfügen über ein Aufzeichnungssystem für GPS-Daten. Hier wird jedoch meist in sehr großen Intervallen aufgezeichnet (meist sogar nur Start und Zielort). Das erschwert ein Nachvollziehen der genutzten Routen, oder macht es unmöglich (Startpunkt=Zielpunkt).

Was beinhalten die Daten?

GPS Daten beinhalten nach dem Matching auf das genutzte GIS-Netz zumeist einen Datensatz für die Knoten von ebenjenem und einen Datensatz für die Kanten. Der Datensatz für die Kanten enthält Verkehrsmengen und Geschwindigkeiten, der Knotendatensatz zusätzlich auch Wartezeiten. Darüber hinaus sind Quelle-Ziel-Matrizen möglich. Dafür muss jedoch zusätzlich zum GPS-Datensatz ein Raster geliefert werden. Die Daten können für die Städte je nach Anbieter individuell aufbereitet werden. Mindestens lieferbar sind jedoch Abgrenzungen nach Wochentag oder Zeitscheiben.

Was können welche Daten leisten?

Es gibt verschiedene Datentypen, denen unterschiedliche Nutzergruppen oder Erhebungsmethoden zugrunde liegen (siehe Kapitel 2.2). Daten aus Verleihsystemen bieten zwar eine breite Nutzergruppe, aber auch eingeschränkte Wegezwecke bzw. bilden die Nutzung innerhalb von Wegekettten ab, was die Interpretation erschwert. Daten aus Navigationsapps sind als etwas repräsentativer anzusehen. Das liegt in der etwas heterogeneren Stichprobe begründet (z.B. ein etwas höherer Frauenanteil) sowie der weniger sportorientierten Nutzerklientel. Dies wiederum erhöht die Vergleichbarkeit mit dem durchschnittlichen Radfahrenden deutlich. Auch die dabei angesprochene Zielgruppe ist potenziell sehr groß, allerdings muss deren Motivation über einen längeren Zeitraum durch durchdachte Kampagnen erhalten werden. Anders verhält es sich bei der Klientel der Sportappnutzer. Diese sind überwiegend männlich und legen ein deutlich sportiveres Fahrverhalten an den Tag. Hier ist eine Angleichung nötig. Da die Nutzenden eher intrinsisch motiviert sind, kann auch ohne begleitende Kampagne eine hohe Teilnehmerzahl und somit eine gute Netzabdeckung erreicht werden.

Worauf muss bei der Arbeit mit GPS-Daten geachtet werden?

Es gibt eine Vielzahl zu beachtender Punkte bei der Arbeit mit GPS-Daten. Die wichtigsten Punkte betreffen dabei die Auswahl und Verifizierung des GIS-Netzes. Hierbei ist vor allem bei Planung eines kontinuierlicher Datenbezugs auf eine gleiche Netzgrundlage zu achten, da sonst die Vergleichbarkeit innerhalb der Zeitreihen stark erschwert wird. Zudem sollte man sich im Vorhinein klar darüber sein, welche Ziele man mit Datenerwerb und -analyse verfolgt und welche Grundlagen bereits vorliegen, um das passfähigste Angebot auszuwählen. So kann bei einer starken Radfahrercommunity ein sportlich orientierter Datensatz einem navigationsbasiertem Ansatz vorzuziehen sein. Siehe dazu auch die Frage: „Für welche Zwecke können die Daten verwendet werden?“.

Sind die Daten repräsentativ?

Legt man die klassische Form einer statistischen Repräsentativität innerhalb der Grundgesamtheit der Radfahrenden zu Grunde, kann diese durch GPS-Daten nicht erreicht werden. Dies ist im Fehlen statistisch relevanter Zielgruppen, wie die der älteren Menschen oder von Kindern, begründet. Gleichwohl kann unter gewissen Gesichtspunkten und mit einigen Korrekturen selbst ein durch die Nutzer von Sportapps erzeugter Datensatz ein sehr gutes Abbild des Radverkehrsverhaltens liefern. Letztendlich ist das Problem der fehlenden statistischen Repräsentativität bezogen auf die Grundgesamtheit der Radfahrenden auch bei Dauerzählstellen ein relevanter Kritikpunkt.

Für welche Zwecke können die Daten verwendet werden?

Die Daten sind prinzipiell geeignet um folgende Punkte in der Radverkehrsplanung zu bedienen:

- Zielnetzplanung (mit Einschränkungen),
- Maßnahmenplanung,
- Evaluation von durchgeführten Radverkehrsmaßnahmen,
- Marketing,
- Visualisierung der Radfahrenden.

Siehe dazu auch Kapitel 3 im Leitfaden.

Welche Probleme hinsichtlich des Datenschutzes können auftreten?

Der Datenschutz obliegt in den meisten Fällen dem Datenlieferanten. Die bekannten Angebote liefern zudem keine einzelnen Routen sondern aggregierte Radverkehrsstärken auf den Knoten und Kanten des gelieferten GIS-Netzes. Zusätzlich werden von den Daten Providern vor der Aggregation im Zuge eines Mapmatching-Verfahrens noch die ersten und letzten 100m der einzelnen Tracks „abgeschnitten“ um einen Bezug zum Einzelnutzer schon in dieser Stufe der Datenverarbeitung zu erschweren. Das aggregierte Endprodukt macht eine Nachvollziehbarkeit von einzelnen Personen aus dem Gesamtdatensatz damit unmöglich. Somit liegt man formell nicht mehr im Bereich der Datenschutzgesetze der Länder.

Wie sollte man mit der freien Eingabe der Daten, wie z.B. Alter und Geschlecht, durch die Nutzerinnen und Nutzer umgehen, beziehungsweise wie wird sichergestellt, dass die Daten tatsächlich von Radfahrenden stammen?

Hierbei gilt, ähnlich wie bei Haushaltsbefragungen auch, das Prinzip des Vertrauens in eine ausreichend große Stichprobe. Es gibt keine Garantie für die Richtigkeit der Daten. Es entstehen für die Nutzenden jedoch auch keine Vorteile durch eine Falscheingabe. Es wird trotzdem einige wenige geben, die wissentlich falsche Angaben machen. So tauchte im Strava-Datensatz der Stadt Dresden beispielsweise ein 90jähriger Sportradfahrer auf. Die Richtigkeit dieser Angabe ist zumindest anzuzweifeln, jedoch ist dies im Bereich von mehreren tausend Nutzenden kein besonders gravierender Aspekt. Gleiches gilt für eine abweichende Moduswahl. Die Daten können auch durch die Nutzung anderer Modi erhoben worden sein. Dies ist zum einen durch das Mapmatching erkennbar (z.B. auf Bahnlinien) zum anderen kann im Vorhinein zusammen mit dem Datenprovider beispielsweise eine maximale Durchschnittsgeschwindigkeit definiert werden.

Wer sind die App-Nutzenden?

Grundsätzlich gibt es, außer dem Smartphonebesitz, keine Barrieren für eine Nutzung. Das Nutzerfeld der getesteten Apps (BikeCitizens und Strava) war jedoch klar männlich dominiert, weniger bei BikeCitizens und stärker bei Strava. Darüber hinaus sind die Alterskategorien jedoch ähnlich verteilt, mit einer etwas stärker ausgeprägten Gruppe der 20- bis 40-jährigen. Betrachtet man das Radverkehrsverhalten, ist aufgrund der unterschiedlichen Motivation der Nutzerinnen und Nutzer jedoch ein deutlicher Unterschied zu erwarten. Die sportlich motivierten Nutzenden fahren deutlich länger und schneller Rad, als durchschnittliche Radfahrende.

Kann man die Daten erwerben?

Ja, bisher sind drei privatwirtschaftliche Anbieter bekannt: Der Sportappprovider Strava sowie BikeCitizens und Scholz&Volkmer. Die Modalitäten für den Dateneinkauf sind jedoch unterschiedlich. Während Strava aktuell seine Preisgestaltung je Nutzerin/Nutzer aufbaut, ist bei BikeCitizens und Scholz&Volkmer der Preis abhängig vom gewünschten Angebot. Dieses kann beispielsweise eine Begleitkampagne, wie ein BikeBenefit-Programm, Navigation oder spezielle Funktionalitäten in der Datenauswertung, umfassen.

Bekomme ich einzelne Routen und wie viele sind das?

Aus Datenschutzgründen wird auf die Weitergabe von Routendaten verzichtet. Geliefert werden Aggregate aus allen verfügbaren Routen. Zusätzlich besteht die Möglichkeit eine erweiterte Quelle-Ziel-Matrix zu beziehen. Anhand dieser lassen sich zumindest Routenkorridore festlegen. Mehr

Informationen dazu finden sich unter „Datenschutz“. Die Anzahl der zugrundeliegenden Routendaten hängt stark von der Appverbreitung in der jeweiligen Kommune ab. Kennzahlen für die Stadt Dresden und den Provider Strava sind dabei ca. 3.200 Nutzerinnen und Nutzer mit ca. 75.200 Fahrten pro Jahr. Die Nutzer- und Fahrtenzahl lässt sich in diesem Fall jedoch nicht wirklich steuern. Anders ist dies beim Angebot von BikeCitizens. Hier ist eine Zusammenarbeit mit den Kommunen vorgesehen, um einen Nutzerstamm zu erreichen.