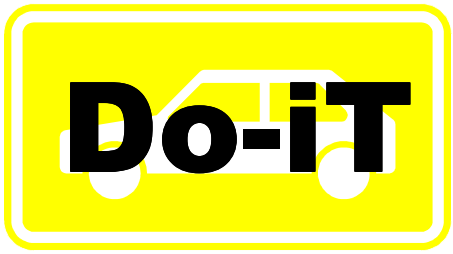
Schlussbericht

Datenoptimierung für integrierte Telematik

Verbundabschlussbericht



Förderkennzeichen

19 B 5007A
19 B 5007B
19 B 5007C
19 B 5007D
19 B 5007E
19 B 5007F

Dezember 2009

DDG Gesellschaft für Verkehrsdaten mbH

Landeshauptstadt Stuttgart

Stadt Karlsruhe

Innenministerium Baden-Württemberg

Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen

Universität Stuttgart, Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik

Verbundpartner

ddg Gesellschaft für Verkehrsdaten mbH	DDG Gesellschaft für Verkehrsdaten mbH Dr. Markus Becker Jürgen Elberich Dr. Ulrich Fastenrath (Projektleitung) Thorsten Rauch Ulf Wagemann Dr. Andreas Wörner
STUTTGART	Landeshauptstadt Stuttgart Dirk Herrmann Ulrich Steimer
Karlsruhe	Stadt Karlsruhe Wilfried Schreiber Stefan Schwartz Stefan Knapp
Baden-Württemberg INNENMINISTERIUM	Innenministerium Baden-Württemberg Reinhard Schult Roland Bettermann
igb	Institut für Anwendungen der Geodäsie im Bauwesen PD Dr.-Ing. Volker Schwieger Dipl.-Ing. Ralf Laufer Dr.- Ing. Renate Czommer Dipl.-Ing. Alexander Beetz Dipl.-Ing. Ralf Schollmeyer
Universität Stuttgart	Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik Prof. Dr.-Ing. Markus Friedrich Katrin Immisch Prokop Jehlicka Thomas Otterstätter Johannes Schlaich

Inhalt Error! Bookmark not defined.

1 Einführung 6

1.1 Aufgabenstellung 6

1.2 Projektstruktur 8

1.2.1 Projektvoraussetzungen 8

1.2.2 Projektplanung und -ablauf 10

1.2.3 Zusammenarbeit mit anderen Stellen 12

2 Wissenschaftlicher und technischer Stand 12

2.1 Stand zu Projektbeginn 12

2.2 Ergebnisse anderer Projekte während der Projektlaufzeit 13

2.2.1 Traffic Online 13

2.2.2 TomTom 14

2.2.3 Generierung von Quelle-Ziel-Matrizen mit Mobilfunkdaten 14

3 Erzielte Ergebnisse 17

3.1 Technische Anbindung an das Mobilfunknetz 17

3.2 Mobilfunkortung aus Abis-Daten (IAGB) 19

3.3 Trajektoriengenerierung aus A und Abis-Daten (IAGB) 25

3.4 Identifikation aktiver Verkehrsteilnehmer (IAGB) 32

3.5 Generierung von Trajektorien aus A-Daten (VuV) 36

3.6 Schätzung der Parameter des Routenwahlverhaltens (VuV) 40

3.7 Generierung von Quelle-Ziel-Matrizen (VuV) 43

3.8 Durchführung von Tagesprognosen (VuV) 47

3.9 Monitoring der Angebotsqualität (VuV) 49

3.10 Aufbau des Qualitätsmanagements (IAGB) 54

3.11 Evaluierung der FPD-Trajektorien (IAGB) 60

3.12 Evaluierung der verkehrstechnischen Anwendungen (DDG, Ka) 67

4 Voraussichtlicher Nutzen, Verwertungspläne der Partner 77

4.1 Verwertungsplan DDG 77

4.2 Verwertungsplan IM 77

4.3 Verwertungsplan Stadt Stuttgart 77
4.4 Verwertungsplan Stadt Karlsruhe 78
4.5 Verwertungsplan IAGB 78
4.6 Verwertungsplan VuV 78

5 Veröffentlichungen von Projektergebnissen 80
5.1 Partnerübergreifende Veröffentlichung 80
5.2 Stadt Karlsruhe 80
5.3 Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen 80
5.4 Universität Stuttgart, Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik 81

6 Anlagen 84

Dokument-Informationen

Version	Datum	Bearbeiter	Bemerkung
00	**24.03.09**	**Schlaich**	• Erstellung der Vorlage
01	**31.03.09**	**Schlaich**	• Füllen der Kapitel 1.1, 1.2.1-2, 2.1, 5.1 und 5.7 • Vorschlag für Zuständigkeiten bei anderen Kapiteln
02	**30.04.09**	**Schwieger**	• Ergänzung der Gliederung und der Literatur um IAGB-Anteil
03_VuV	**09.06.09**	**Schlaich**	• Kleine Anpassungen in den Kap. 1 und 2 • Füllen des Kapitels 2.2.1 (Q-Z-Matrizen Israel) • Füllen der Kap. 3.6-3.10
04	**13.09.09**	**Schwartz**	• Füllen Verbundpartner KA und Kap. 5.4
05	**08.10.09**	**Laufer / Schwieger**	• Füllen der Kapitel des IAGB
06	**14.10.09**	**Schlaich**	• Änderungen im Kap 2.2.4, 3.6-3.10, 5.7 • Streichen von VuV aus dem Kap. 3.14
07	**27.11.09**	**Fastenrath**	• Vervollständigung
08	**28.11.09**	**Schlaich**	• Aktualisierungen in Kap. 5.4 • Vereinheitlichung Format im gesamten Dokument
09	**03.12.09**	**Schwartz**	• Ergänzung im Kapitel 3.12: Demonstrator KA
10	**10.12.09**	**Schwieger**	• Letzte Ergänzungen und Korrekturen IAGB
11	**15.12.09**	**Schlaich**	• Formatanpassungen
12	**16.12.09**	**Fastenrath**	• Schlußredaktion

1 Einführung

Do-iT war ein spannendes Projekt voller Herausforderungen – organisatorischer, technischer, zuweilen sogar politischer Natur. Diese konnten alle soweit bewältigt werden, dass am Ende eine im direkten Vergleich mit der Verkehrsrealität erfolgreiche Demonstration gelang. Daher ist es keine Übertreibung zu sagen, dass Do-iT den Boden bereitet hat für die kommerzielle Anwendung einer neuen Technologie zur Echtzeit-Erfassung von Verkehrsdaten. Wir freuen uns, diesen Abschlussbericht vorlegen zu können, welcher die erzielten Ergebnisse im Überblick und ausführliche technische Darstellungen, Offenlegungsschriften und Evaluierungen in den Anlagen enthält.

1.1 Aufgabenstellung

Zur Umsetzung moderner Verkehrsmanagementmethoden ist es erforderlich, die aktuelle Verkehrslage zu erfassen, um darauf aufbauend zuverlässige Verkehrsprognosen erstellen zu können. Nur wenn die Qualität der Verkehrslageerfassung und der Verkehrsprognose gesichert ist, können verkehrsabhängige Steuerungsstrategien bewertet bzw. umgesetzt und Verkehrsteilnehmer zuverlässig informiert werden. Für die Verkehrslageerfassung sind neben lokalen Messdaten an Streckenquerschnitten insbesondere Floating Car Data (FCD) von Bedeutung. Lokale Messdaten liefern ein Bild des Verkehrszustandes für die Menge aller Verkehrsteilnehmer an ausgewählten festen Orten im Netz. FCD liefern im Gegensatz dazu Informationen über die Fortbewegungsgeschwindigkeit im Netz für eine Teilmenge der Verkehrsteilnehmer. Außerdem können aus FCD Informationen über die Quellen und Ziele der erfassten Verkehrsteilnehmer abgeleitet werden. Die Kombination beider Datenquellen bietet eine gute Basis für Methoden zur Ermittlung und Prognose der Verkehrslage. Allerdings ist derzeit nur ein kleiner Teil der Fahrzeugflotte so ausgerüstet, das FCD erhoben werden können. Für die Ermittlung von Quelle-Ziel-Daten ist dieser Teil der Fahrzeugflotte nicht ausreichend, da nur ein kleiner Teil der Oberklassefahrzeuge damit ausgestattet ist.

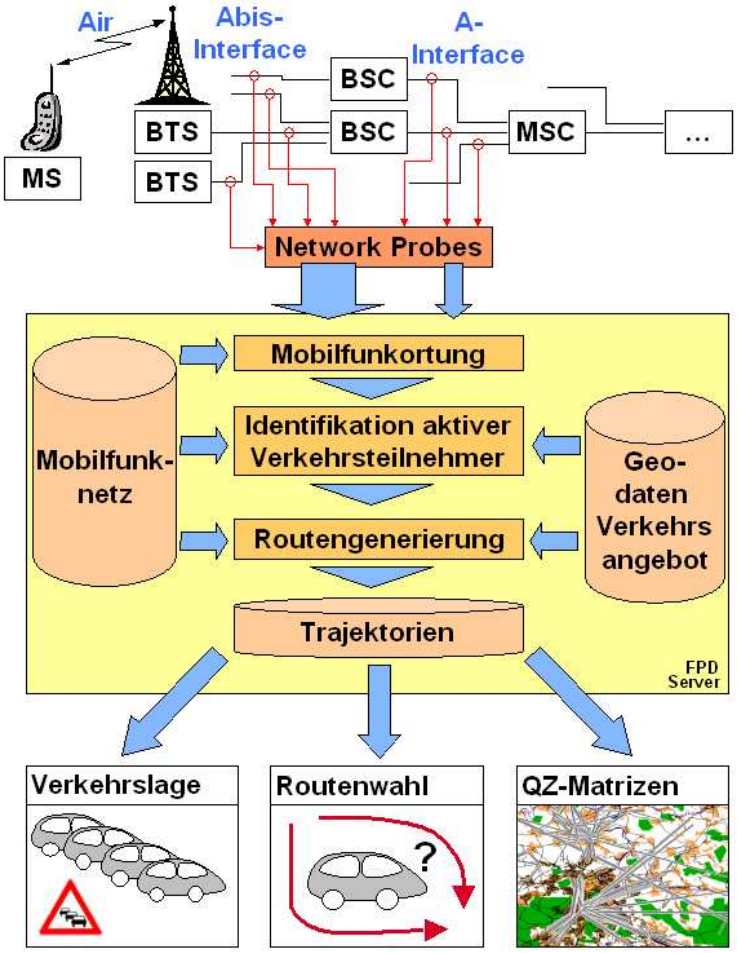


Abbildung 1-1: Do-iT Gesamtsystem

Zentrales Ziel des Projektes Do-iT (Datenoptimierung für integrierte Telematik) ist es, Mobilfunkgeräte als mobile Sensoren zu verwenden und die Brauchbarkeit der gewonnenen Daten für verkehrliche Anwendungen zu prüfen. Dabei sollen drei Nutzungsmöglichkeiten von Mobilfunkdaten untersucht werden (vgl. Abbildung 1-1):

- *Verkehrslageerfassung und –prognose:*
  Es soll gezeigt werden, dass Mobilfunkdaten als Ergänzung zu Streckenmesswerten, Verkehrsmeldungen und Floating Car Data maßgeblich zu einer preiswerten und flächendeckenden Verkehrslageerfassung und –prognose beitragen.

- *Erfassung von Routenwahl und Befolgungsgraden:*
  Es soll analysiert werden, ab welcher Fahrtweite in städtischen und außerörtlichen Netzen Routen oder Teilrouten von Verkehrsteilnehmern im Straßenverkehr aus den Mobilfunkdaten abgeleitet werden können. Auf der Basis der beobachteten Fahrtrouten werden Aussagen über das Routenwahlverhalten der Verkehrsteilnehmer in Abhängigkeit der Reisezeit, der gemeldeten Verkehrslage und der Schaltzustände von Wechselwegweisungen abgeleitet.

- *Generierung von Quelle-Ziel-Matrizen:*
  Es soll geprüft werden, ob sich aus den Mobilfunkdaten die Quellen und Ziele der Verkehrsteilnehmer so bestimmen lassen, dass daraus Quelle-Ziel-Matrizen für die Netzsteuerung und für Planungszwecke abgeleitet werden können. Dazu sollen die aus den Mobilfunkdaten abgeleiteten Fahrtrouten mit den Zählwerten stationärer Zählstellen hochgerechnet werden.

1.2 Projektstruktur

1.2.1 Projektvoraussetzungen

Mobilfunktelefone haben sich in den letzten Jahren in Deutschland und weltweit zu einem ständigen Begleiter der Menschen entwickelt. Mobilfunk-Module werden zudem verstärkt in technischen Systemen wie z.B. Navigationssystemen oder Onboard-Units zur Mauterfassung verbaut. Im Projekt Do-iT sollen die in den Mobilfunknetzen anfallenden Daten für die Planung und Steuerung des Verkehrsablaufes im Verkehrwesen erschlossen werden.

In Deutschland wird heute und in näherer Zukunft der Großteil des Mobilfunkverkehrs über GSM-Netze abgewickelt. GSM steht dabei für Global System for Mobile Communications. Der für die mobile Kommunikation zuständige Teil eines GSM-Netzes besteht vereinfacht dargestellt aus den in Abbildung 1-2 abgebildeten Komponenten:

- Mobile Stations (MS): Dies sind die Mobilfunktelefone bzw. -endgeräte.
- Base Tranceiver Station (BTS): Sie bilden die kleinste Zelleneinheit des GSM-Netzes. Eine BTS hat in den meisten Fällen 3 Sektor-Antennen mit einem Strahlwinkel von 120°. Jede Antenne entspricht einer Funk zelle. Das Mobiltelefon nimmt über die BTS mit dem Netz Verbindung auf.
- Base Station Controller (BSC): Eine BSC verwaltet mehrere BTS. Daten zwischen der BTS und dem BSC werden über eine spezielle Schnittstelle, das sogenannte Abis-Interface übertragen
- Mobile Switching Center (MSC): Ein MSC ist eine mobile Vermittlungsstelle, die für mehrere BSC zuständig ist. Der Datenaustausch erfolgt über das sogenannte A-Interface.

Jeweils mehrere Funkzellen eines GSM-Netzes werden zu einer Location Area zusammengefasst. Der Aufenthaltsort eines eingeschalteten Mobilfunkgerätes ist netzseitig nur auf der Ebene von Location Areas bekannt. Erst bei einem Verbindungsaufbau wird die aktive Funkzelle (Serving Cell) ermittelt, über die dann die weitere Kommunikation abläuft. Dieses zweistufige Vorgehen minimiert den Aufwand an Verwaltungsdaten im Netz sowie den Energieverbrauch im Endgerät.

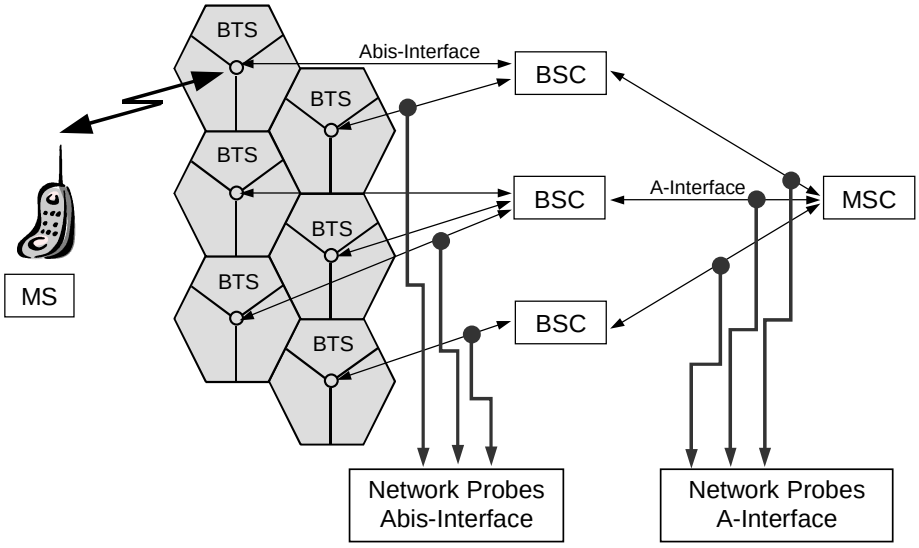


Abbildung 1-2: Vereinfachte Systemarchitektur eines Mobilfunknetzes mit den Schnittstellen für die Network Probes

Die Verwaltungsdaten, die das Mobilfunkendgerät mit den verschiedenen Komponenten im Netz (BSC, MSC) austauscht, können an den Schnittstellen zu diesen Komponenten mitprotokolliert werden. Dies erfordert momentan die Ausrüstung der Mobilfunknetzinfrastruktur mit sogenannten Network Probes. Dabei handelt es sich um spezielle Rechner, die den Datenverkehr zwischen BTS und BSC (Abis-Interface) oder BSC und MSC (A-Interface) protokollieren. Mittelfristig könnten diese Probes durch spezielle Softwareerweiterungen in den BSC bzw. MSC ersetzt werden.

Da in der Regel ein BSC für mehrere BTS zuständig ist, benötigt man für eine Datenerfassung am Abis-Interface deutlich mehr Network Probes als bei einer Erfassung auf der A-Ebene. Dafür lassen sich aus den Daten der Abis-Ebene genauere Informationen über den aktuellen Standort eines Mobilfunkgerätes ableiten.

Art und Umfang der Verwaltungsdaten variieren, je nachdem ob ein Mobiltelefon lediglich im Netz eingebucht ist (Standby-Mode) oder ob gerade eine aktive Gesprächs- oder Datenverbindung besteht (Dedicated-Mode).

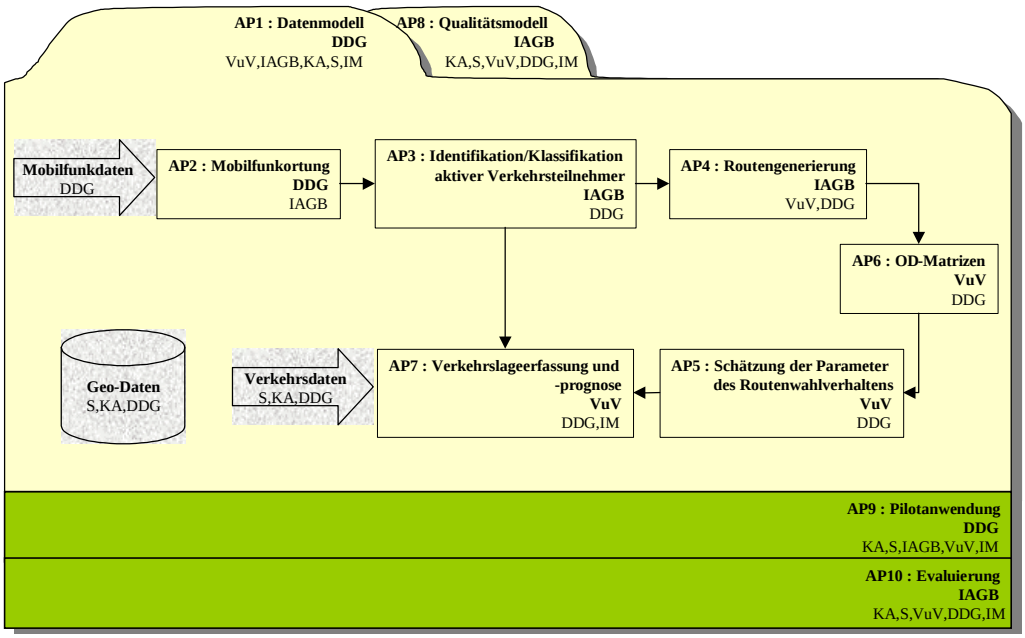
Beim Standby-Mode wird im Wesentlichen nur das Wechseln einer Location Area über einen sogenannte Location Area Update erfasst. Dies kann sowohl am A- als auch am Abis-Interface protokolliert werden.

Im Dedicated-Mode sendet das Endgerät dagegen alle 480ms einen sogenannten Measurement-Report an den BSC. Dieser enthält detaillierte Daten (u.a. die Empfangsfeldstärke) über die aktive Funkzelle und bis zu sechs Nachbarzellen und kann am Abis-Interface protokolliert werden. Am A-Interface kann im Dedicated-Mode lediglich der Wechsel von einer in eine andere Funkzelle (sogenannter Handover) erfasst werden.

Im Projekt Do-iT werden sowohl die Daten der A-Ebene als auch Abis-Ebene auf ihre Nutzbarkeit für die genannten Projektziele hin untersucht.

1.2.2 Projektplanung und -ablauf

Das Projekt Do-iT gliedert sich in 10 Arbeitspakete (AP). Die folgende Abbildung 1-3 stellt den logischen Zusammenhang zwischen den Arbeitspaketen dar. Das Arbeitspaket an einer Pfeilspitze baut auf den Ergebnissen seines jeweiligen Vorgängers auf.



DDG : DDG Gesellschaft für Verkehrsdaten mbH
S : Landeshauptstadt Stuttgart
KA: Stadt Karlsruhe
IAGB: Universität Stuttgart, Institut für Anwendung der Geodäsieim Bauwesen
VuV: Universität Stuttgart, Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik
IM: Innenministerium Baden-Württemberg -

Abbildung 1-3: Überblick über die Arbeitspakete

- *AP1: Spezifikation von Datenmodellen und Austauschformaten:*
  Große Mengen an anwendungsunspezifischen Daten aus dem Mobilfunknetz werden referenziert, mit anderen Verkehrsdaten und Infrastrukturdaten verknüpft, weiterverarbeitet und anderen Projektpartnern und Nutzern zur Verfügung gestellt.
- *AP2: Verkehrsdatenerfassung durch Mobilfunkortung:*
  Durch Nutzung von Daten, die auf den Ort von Endgeräten schließen lassen wie z.B. Laufzeit- und Feldstärkemessungen, sowie die Netzgeometrie sollen Geschwindigkeiten auf Netzkanten ermittelt werden. Die erforderlichen Rohdaten werden durch geeignete Schnittstellen aus dem operativen Mobilfunknetz extrahiert.

- *AP3: Identifikation aktiver Verkehrsteilnehmer:*
  Aus der Menge aller im Testgebiet befindlichen Mobiltelefone werden diejenigen D1-Netz-Mobiltelefone, die sich momentan aktiv im Verkehr bewegen, herausgefiltert und anhand der einzelnen Bewegungsprofile dem Individualverkehr und dem Öffentlichen Verkehr zugeordnet.
- *AP4: Routengenerierung und –attributierung:*
  Ausgehend von den Rohdaten des Mobilfunkbetreibers wird ein Map-Matching gestütztes Ortungsverfahren entwickelt, das Straßennetzdaten und Mobilfunkdaten fusioniert und so die Fahrtrouten der Verkehrsteilnehmer bereitstellt.
- *AP5: Informationen über das Routenwahlverhalten:*
  Auf der Basis der beobachteten Fahrtrouten der Verkehrsteilnehmer werden Aussagen über das Routenwahlverhalten der Verkehrsteilnehmer in Abhängigkeit von der Reisezeit, der Verkehrszustände und aktueller Verkehrssteuerungsmaßnahmen abgeleitet.
- *AP6: Generierung von Quelle-Ziel-Matrizen:*
  Aus den Mobilfunkdaten werden die Quellen und Ziele der Verkehrsteilnehmer bestimmt und mit Zählwerten stationärer Zählstellen zu Quelle-Ziel-Matrizen hochgerechnet. Diese Matrizen stehen für Planungszwecke, für die Steuerung von Netzbeeinflussungsanlagen und für die Bewertung verkehrlicher Wirkungen zur Verfügung.
- *AP7: Verkehrslageerfassung und –prognose:*
  Es soll gezeigt werden, dass Mobilfunkdaten als Ergänzung zu Streckenmesswerten, Verkehrsmeldungen und Floating Car Daten maßgeblich zu einer preiswerten und flächendeckenden Verkehrslageerfassung und –prognose beitragen.
- *AP8: Aufbau eines Qualitätsmodells*:
  Zur Überprüfung und Überwachung der Qualität von Verkehrsinformationen z. B. im Hinblick auf Verfügbarkeit und Aktualität werden quantifizierbare Qualitätsmerkmale definiert, erforscht und erprobt. Das dabei entstandene Qualitätsmodell dient dem Nachweis der Wirksamkeit der implementierten Maßnahmen zur Qualitätssicherung sowie der Evaluierung der Mobilfunktrajektorien in den Demonstratoren.
- *AP9:* Pilotanwendung*:*
  Eine vollständige Verarbeitungskette vom Monitoringsystem des Mobilfunknetzes über die Datenextraktion, -filterung, verkehrstechnische Verarbeitung bis hin zur Generierung von Verkehrsmeldungen, Reisezeitschätzung und „Level of Service"-Darstellung wird aufgebaut, in Echtzeit betrieben und im Rahmen einer Internetapplikation demonstriert.
- AP10: *Evaluierung*:
  Zur Evaluierung der Qualität der aus Mobilfunkdaten generierbaren Trajektorien werden GPS-Testfahrten durchgeführt und die Messwerte der stationären Verkehrsdatenerfassungssysteme der Projektpartner als Referenzdaten herangezogen.

Die Verwendbarkeit der Datenquelle Mobilfunk in typischen verkehrstechnischen Anwendungen wird in den Demonstratoren evaluiert.

Im Laufe der Projektbearbeitung zeigte sich, dass der Zugang zu den Daten des Mobilfunknetzes deutlich länger benötigte als erwartet. Dies lag sowohl an administrativen als auch technischen Schwierigkeiten beim Netzzugang.

- T-Mobile war nicht als Verbundpartner, sondern als „assoziierter Partner“ in das Projekt eingebunden (vgl. Abschnitt 1.2.3).
- Die Erstellung der Schnittstellen zum Mobilfunknetz ist eine hochspezialisierte Aufgabe, die nicht durch die Verbundpartner geleistet werden konnte. Die Umsetzung erforderte die Unterstützung durch externe Firmen.

1.2.3 Zusammenarbeit mit anderen Stellen

T-Mobile spielte als Betreiber des Mobilfunknetzes eine tragende Rolle im Projekt, war aber kein Projektpartner, sondern wurde im Rahmen eines Kooperationsvertrages mit DDG eingebunden.

Des weiteren hat die Firma Optel als Unterauftragnehmer der DDG eine Schnittstellen-Software für die Datenextraktion aus dem Monitoring-System der T-Mobile entwickelt.

2 Wissenschaftlicher und technischer Stand

2.1 Stand zu Projektbeginn

Bereits vor Beginn des Projektes ließen erste Erfahrungen mit FPD im Ausland vermuten, dass Mobilfunkgeräte als mobile Sensoren eine preiswerte Möglichkeit darstellen können, Informationen über den aktuellen Verkehrszustand zu erhalten.

Die britische Firma Applied Generics hatte das System RoDIN24 entwickelt, dass die Positionen aktiver Mobilfunkteilnehmer, d.h. die gerade eine Telefonat führen, bestimmt und so Informationen über Geschwindigkeiten im Straßennetz ableitet. Laut Applied Generics (http://www.appliedgenerics.com/) hatten erste Feldversuche ergeben, dass bis zu 95 % des Hauptstraßennetzes mit dem System detektiert werden können. Im Projekt „Intermezzo“ (Auftraggeber: Rijkswaterstaat) hat Applied Generics gemeinsam mit LogicaCMG ihr System angewendet. Dabei wurden nach den im Do-iT Konsortium vorliegenden Erkenntnissen allerdings nur Geschwindigkeitsklassen ermittelt, ohne die Qualität des Ergebnisses zu validieren, ohne Datenfusion mit anderen Quellen durchzuführen und ohne Verkehrszustände zu berechnen. In einem parallelen Projektstrang innerhalb von Intermezzo hat die DDG als Unterauftragnehmer von Siemens die Möglichkeiten konventioneller Verkehrsdatenerfassung in nachgeordneten

Straßennetzen demonstriert. LogicaCMG und Applied Generics vermarkten das Projekt als Erfolg, systematische Untersuchungen der erreichten Informationsqualität oder ein offizielles Projektergebnis von Rijkswaterstaat waren den Do-iT Partnern jedoch nicht bekannt.

Gemeinsam mit der finnischen Straßenverwaltung versuchte der finnische Netzbetreiber Radiolinja Finnland ebenfalls Verkehrsinformationen aus Mobilfunkdaten zu gewinnen. Der Ansatz von Radiolinja Finnland basierte auf speziellen Messstellen, die in einem Abstand von rund 5 Kilometern einzelne Mobilfunkgeräte verfolgen und daraus Geschwindigkeiten ableiten.

Im Jahre 2001 wurde das EU-Projekt STRIP in Frankreich abgeschlossen. Dabei wurden vielversprechende Ergebnisse zum Durchdringungsgrad und zu Korrelationen zwischen FPD und Schleifendaten erzielt, aber weiterer Forschungsbedarf in Bezug auf engmaschigere Straßennetze und Geschäftsmodelle erkannt. Der in diesem Projekt eingeschlagene Weg erscheint dem Do-iT Konsortium zielführend. Er sollte daher in Do-iT weiter verfolgt und um die Map-Matching-Algorithmen erweitert werden. Es deutete vieles darauf hin, dass auch die bei Applied Generics entwickelten Algorithmen auf den Ansätzen von STRIP beruhen. Als wesentliche Erweiterung in Do-iT sollten Informationen nicht nur für einzelne Strecken oder Netzabschnitte sondern darüber hinaus Routen und Quelle-Ziel-Beziehungen abgeleitet werden können. Außerdem werden die FPD mit lokalen Messwerten fusioniert, um so Informationen über Verkehrsmengen einzubeziehen.

2.2 Ergebnisse anderer Projekte während der Projektlaufzeit

2.2.1 Traffic Online

Parallel zu Do-iT wurde im Rahmen von VM2010 das Projekt Traffic Online durchgeführt, welches ebenfalls zum Ziel hatte, Verkehrsinformationen aus Mobilfunkdaten zu gewinnen, allerdings mit einem anderen technischen Ansatz. Im Ergebnis wurden die Möglichkeiten, dieses Ziel insbesondere im nachgeordneten Netz zu erreichen, eher zurückhaltend bewertet. Dies ist allerdings nach Ansicht der Do-iT Projektpartner eine Bewertung, die nicht ohne weiteres verallgemeinert werden kann, sondern auf die speziellen im Rahmen von Traffic Online entwickelten Verfahren gesehen werden muss. Entscheidend ist, dass Mobilfunkdaten nicht zum Zwecke der Verkehrslageerfassung erzeugt werden, und somit keine kanonische Art der Datenverarbeitung existiert. Vielmehr müssen verschiedene mögliche Herangehensweisen unabhängig voneinander beleuchtet werden.

2.2.2 TomTom

TomTom hat während der Projektlaufzeit unter der Bezeichnung „HD Traffic“ einen kommerziellen Dienst auf den Markt gebracht, der Verkehrsinformationen aus dem Mobilfunknetz von Vodafone generiert. Dazu hat sich TomTom vermutlich der Technik der im Januar 2006 übernommenen Firma „Applied Generics“ bedient, welche außer den Daten der A-Ebene auch den „Timing Advance Report“, nicht jedoch Feldstärkeinformationen auswertet. Außer den Patentschriften aus der Zeit vor der Übernahme von Applied Generics durch TomTom sind keine weiteren technischen Einzelheiten öffentlich bekannt. Die Produktqualität von HD Traffic wurde durch den TÜV Süd zertifiziert, Einzelheiten zu den eingesetzten Messmethoden und quantitative Ergebnisse der Untersuchung sind aber auch in diesem Fall nicht öffentlich.

2.2.3 Generierung von Quelle-Ziel-Matrizen mit Mobilfunkdaten

In Israel wurden Mobilfunkdaten für die Erstellung des Nachfragemodells für das nationale Verkehrsmodell verwendet. Dafür wurden 16 Wochen lang jeweils 10.000 Mobilfunkteilnehmer des Mobilfunknetzbetreibers „Orange“ zufällig ausgewählt und eine ID für jeden Mobilfunkteilnehmer, die aktuelle Funkzelle sowie Zeitstempel aufgezeichnet. Der genaue Aufbau des Orange-Netzes ist nicht bekannt, somit kann keine Aussage über die Größe von Location Areas und damit die Häufigkeit von Location Area Updates (falls sie in derselben Weise existieren) gemacht werden. Wie im Projekt Do-iT wurden aber auch in Israel nur die Daten des Mobilfunknetzes ausgewertet, die ohnehin während des normalen Betriebs anfallen. In Israel haben die Verkehrsplaner allerdings den Vorteil, dass das periodische Update von Orange auf zwei Stunden eingestellt ist, so dass alle zwei Stunden eine Ortungsinformation vorliegt.

Eine automatisierte Telefonumfrage hat ergeben, dass Orange als einer von drei Mobilfunknetzbetreibern vor allem junge Menschen, allerdings relativ wenig Araber, als Kunden hat. Desweiteren wurde ermittelt, wie die Mobilnetzgerät-Ausstattungsgrade in verschiedenen Gruppen sind. So haben 83 % aller Personen über 8 Jahren mindestens ein Mobilfunkgerät, während es bei Studenten 93 % sind. Dies führt zu Verzerrungen des Samples, die durch Gewichtungsfaktoren berücksichtigt werden müssen.

Darüber hinaus kann eine ungünstige Stichprobenwahl dazu führen, dass bestimmte Regionen unverhältnismäßig oft berücksichtigt werden. Dafür wurden aus dem Sample die Anzahl der Übernachtungen in den 10 untersuchten Distrikten der tatsächlichen Bevölkerungsverteilung gegenübergestellt. Auch wenn sie erwartungsgemäß eine klare Korrelation zwischen den Werten zeigen, ist auch hier eine Korrektur nötig.

Für das Verkehrsmodell wurden die 2.200 Funkzellen zu 600 Verkehrszellen zugeordnet. Anschließend werden über verschiedenen Regeln aus den Daten der Mobilfunkteilnehmer Fahrten ermittelt sowie unbrauchbare Daten gelöscht. Ähnlich wie in Do-iT

zeigten sich hier Schwierigkeiten, dass z.B. die Funkzelle, bei der das Mobilfunkgerät angemeldet ist, nicht zwingend die nächste Funkzelle sein muss. Daraus ergaben sich vor allem Schwierigkeiten bei der genauen Ermittlung der Start- und Zielzellen.

Durch die Verfolgung eines Mobilfunkteilnehmers über eine Woche und in einem Untersuchungsgebiet, das Gesamt-Israel umfasst, können Wegeketten einzelner Mobilfunkteilnehmer erstellt werden. Dies ist in Do-iT nur sehr eingeschränkt möglich, da die Wahrscheinlichkeit, dass ein Mobilfunkteilnehmer den Untersuchungsraum verlässt höher ist und außerdem Mobilfunkteilnehmer nach einigen Stunden ohne Meldung aus technischen Gründen bei der nächsten Meldung nicht wiedererkannt werden und eine neue User-ID vergeben wird.

Die Erstellung der Quelle-Ziel-Matrizen erfolgt in den folgenden drei Schritten:

1. *Verkehrserzeugung:*
   Mit Hilfe der um die Korrekturfaktoren angepassten Wege aus den Mobilfunkdaten wurden die Parameter der Funktionen für die Produktion und die Attraktion von Fahrten der Verkehrsbezirke geschätzt und anschließend die Verkehrszeugung und -anziehung je Zelle und Fahrtzweck berechnet.
2. *Verkehrsverteilung:*
   In einem zweistufigen Verfahren wird die Verkehrsverteilung erst für 34 Distrikte und im Anschluss für die 600 Verkehrszelle ermittelt. Basis sind wiederum die beobachteten Wege der Mobilfunkteilnehmer.
3. *Verkehrsaufteilung:*
   Die Verkehrsaufteilung erfolgt mit einem klassischen Aufteilungsmodell in die drei Modi Kfz-, Zug- und Busverkehr.

Die wesentlichen Unterschiede zu Do-iT sind damit die folgenden:

- Der Untersuchungsbereich in Israel besteht aus der gesamten Nation mit 600 Verkehrszellen (durchschnittlich ca. 12.000 Einwohner je Bezirk). In Do-iT wird dagegen ein Autobahnviereck untersucht, Verkehrszellen sind dabei im Wesentlichen die Autobahnausfahrten (TMC-Knoten).
- In Israel werden keine Trajektorien erzeugt, es werden nur Start und Ziel einzelner Fahrten ermittelt.
- In Israel erfolgt die Erzeugung der Quelle-Ziel-Matrizen modellbasiert. Eingangsgrößen sind neben der Umfrage und den Mobilfunkdaten auch Strukturgrößen wie z.B. Einwohneranzahl und Einkaufsflächen. In Do-iT werden dagegen die Trajektorien an den Werten der Zählstellen hochgerechnet.

Quellenangabe:

Gur, Y., Bekhor, S., Solomon, C., Kheifits, L. (2009), Use of Massive Cell Phone Data to Obtain Inter-City Person Trip Tables for Nationwide Transportation Planning in Israel, *TRB 88th Annual Meeting Compendium of Papers DVD*, Washington DC, USA.

3 Erzielte Ergebnisse

3.1 Technische Anbindung an das Mobilfunknetz

Je nach Kommunikationsebene des Mobilfunknetzes stehen Daten unterschiedlichen Detaillierungsgrades zur Verfügung. Für das Projekt wurden 2 online-Datenerfassungssysteme realisiert.

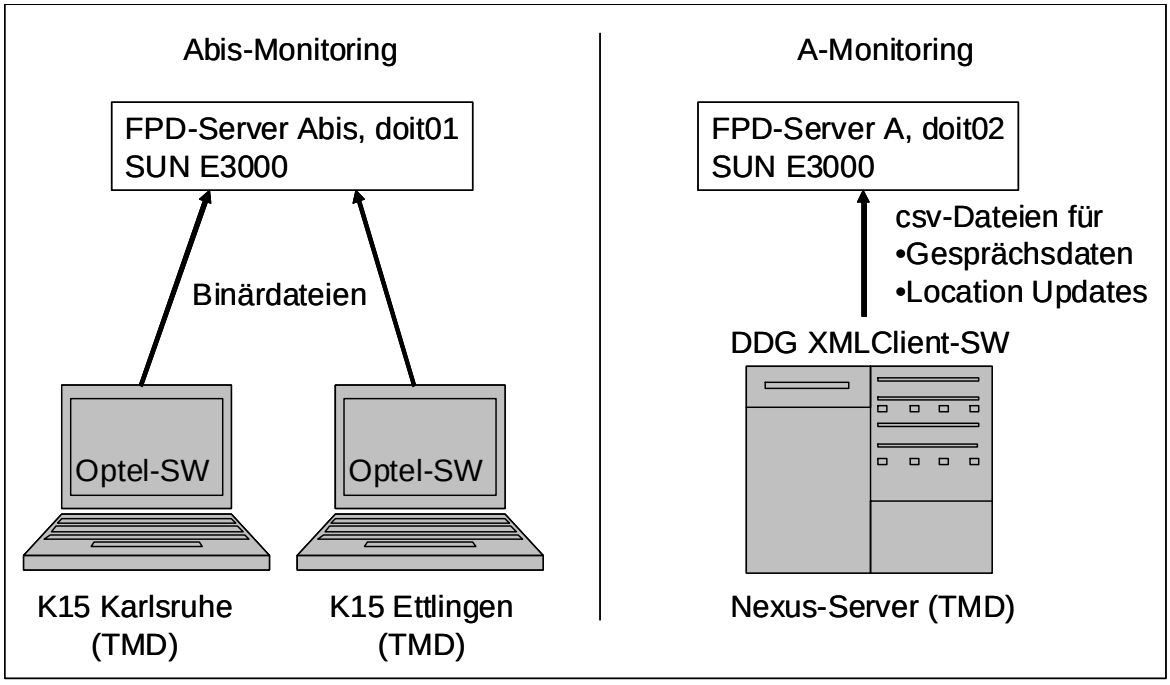


Abbildung 3-1: Erfassungssysteme

- Projektnetzweite Erfassung auf A-Ebene
  Auf der A-Ebene stehen Informationen über Wechsel der Location Area aller Mobilfunkteilnehmer, sowie Beginn, Ende eines Gesprächs und Zellwechsel während eines Gesprächs telefonierender Mobilfunkteilnehmer zur Verfügung. Die A-Datenerfassung erfolgt für 48 BSCs, der Erfassungsbereich deckt dabei das Projektnetz vollständig ab. Die T-Mobile Deutschland (TMD) hat im Rahmen des Kooperationsvertrages hierfür ein eigenständiges Nexus-System aufgebaut, dass die Daten aus dem gesamten Bereich sammelt und über eine XML-Schnittstelle zugänglich macht. Für den Datenzugriff wurde von der Fa. Optel eine von der DDG in Auftrag gegebene XML-Client-Software realisiert, die von der TMD auf ihrem Rechner betrieben wird und die die projektrelevanten Daten minütlich in Dateiform (je eine Datei für Location Updates und Gesprächsinformationen) auf den FPD-Server A (Standort TMD Bonn) überträgt.

- Detailliertere Erfassung auf Abis-Ebene für die Stadt Karlsruhe
  Auf Abis-Ebene stehen während eines Gesprächs zusätzlich Informationen über die Entfernung zur bedienenden Zelle (TA-Wert) und Feldstärkemessungen von und zur bedienenden Zelle und von bis zu 6 Nachbarzellen in hoher zeitlicher Rate (~ 2 Hz)

zur Verfügung. Die Abis-Datenerfassung erfolgt für die beiden BSCs Karlsruhe und Ettlingen, die den Stadtbereich Karlsruhe abdecken. Die Datenerfassung in diesem System geschieht über je einen Tektronix Protocol Analyzer K15 pro überwachter BSC. Die Mobilfunkdaten werden zuerst in einem Rohformat (rf5) zwischengespeichert. Zur Weiterverarbeitung wurde von der Fa. Optel eine von der DDG in Auftrag gegeben Software realisiert, die aus diesen Dateien die projektrelevanten Informationen extrahiert und minütlich in einer Binärdatei speichert. Die Software wird von der TMD auf den K15 betrieben. Die Dateien beider K15 werden dann auf den FPD-Server Abis (Standort TMD BSC Karlsruhe) übertragen.

Beide Systeme sind realisiert, Daten aus dem Karlsruher System (Abis) stehen seit Herbst letzten Jahres online zur Verfügung, Daten aus dem projektnetzweiten A-Daten-Monitoring seit Ende Januar 2008.

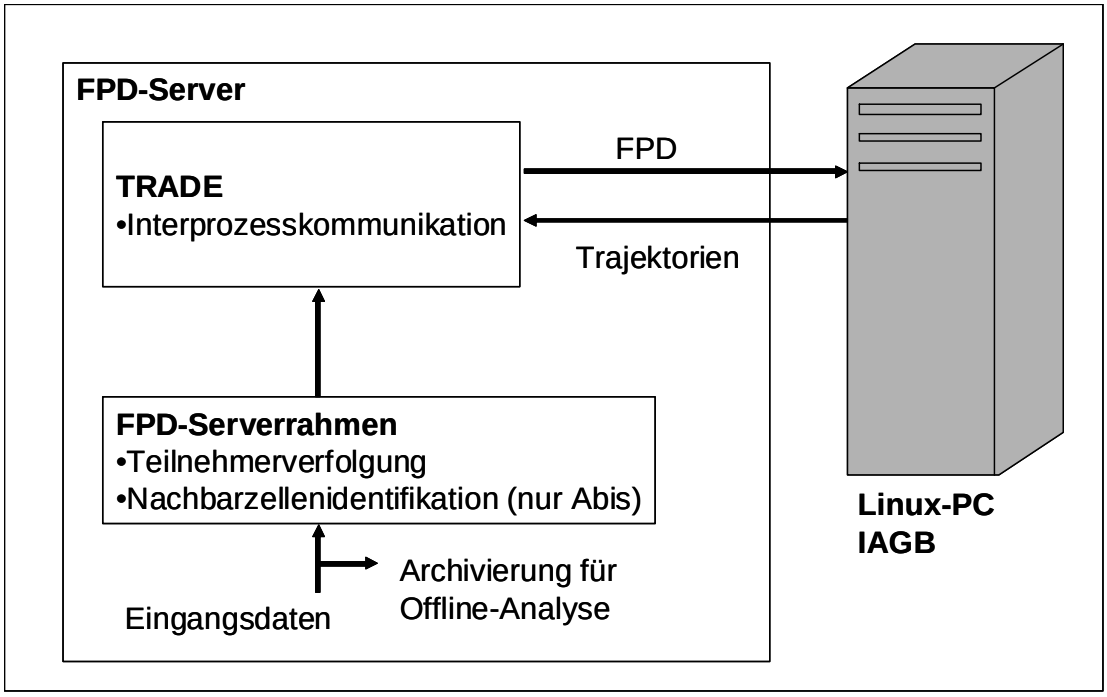


Abbildung 3-2: Systemarchitektur FPD-Server

Auf jedem der beiden FPD-Server läuft ein Prozess FPD-Serverrahmen, der folgende Schritte durchführt:

Die Daten der Dateien beider K15 im Falle des Abis-Monitorings, bzw. der Gesprächs- und der Location Update – Datei im Falle des A-Monitorings, werden in chronologische Reihenfolge gebracht. Dies ist notwendig, um in einem zweiten Schritt die Teilnehmerverfolgung zu realisieren. Hierbei wird ein Mobilfunkteilnehmer über die verschiedenen, wechselnden IDs in den Mobilfunkdaten verfolgt und alle Daten eines Teilnehmers werden mit einer eindeutigen Teilnehmer-ID versehen.

Im Falle des Abis-Servers findet weiter eine Bestimmung der Zell-IDs der in den Measurement Reports enthaltenen Nachbarzellen statt.

Die so prozessierten Daten stehen als Floating Phone Data (FPD) dann zur weiteren Verarbeitung zur Verfügung.

Auf beiden Servern ist eine Middleware (TRADE = „TRAffic Data Engine“, Eigenentwicklung DDG) zur prozessübergreifenden Datenübertragung installiert, über die sowohl die FPD an den Linux-Rechner des IAGB (s. Kapitel 4) als auch zurück die ermittelten Trajektorien übertragen werden.

Auf jedem der beiden FPD-Server werden die Eingangsdaten für eine spätere offline-Analyse archiviert und tagesweise zur DDG übertragen.

Für beide Systeme wurde ein geeigneter Serverrahmen entwickelt und auf dem jeweiligen FPD-Server betrieben. FP-Daten stehen also aus beiden Erfassungssystemen online zur Verfügung.

Die für die Interpretation der FPD notwendigen Netzinformationen werden von der TMD in ausreichendem Umfang bereitgestellt. Täglich wird auf dem FPD-Server eine Stammdatendatei hinterlegt, aus der sich die für den online Betrieb notwendigen Größen (insbesondere Nachbarschaftsbeziehungen zur Nachbarzellidentifikation) entnehmen lassen. Größere Änderungen im Funknetz finden etwa alle zwei Monate zu sogenannten Netzintegrations-Terminen statt. Die Termine werden im Vorfeld von der TMD dem Projekt bekannt gegeben. Zu jedem der Termine stellt die TMD einen Best-Server-Plot über das gesamte Projektgebiet und Feldstärkekarten der für den Bereich mit Abis-Monitoring in Karlsruhe relevanten Zellen zur Verfügung.

Für die offline-Analyse stehen dem IAGB neben den archivierten Daten die zugehörigen FPD-Serverrahmen und TRADE in einer auf Linux lauffähigen Fassung zur Verfügung.

3.2 Mobilfunkortung aus Abis-Daten (IAGB)

Der erste Schritt, um Mobilfunkteilnehmer auf das digitale Straßennetz zu matchen, und damit FPD-Trajektorien zu erzeugen, ist die Bestimmung zeitreferenzierter Positionsfolgen, die sich einzelnen Mobiltelefonen und damit auch Verkehrsteilnehmern zuordnen lassen. Abbildung 3-1 zeigt den Gesamtzusammenhang der FPD-Generierung auf der Abis-Ebene.

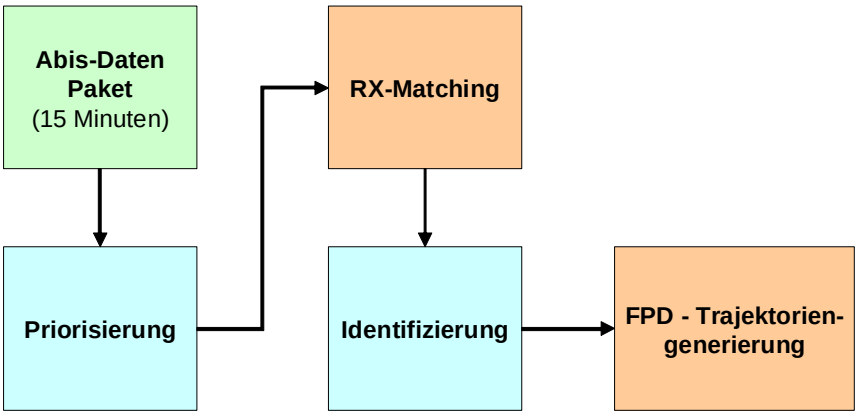


Abbildung 3-1: Grobablauf FPD-Generierung aus Abis-Daten

Hieraus kann man ablesen, dass das Signalstärken-Matching (kurz: RX-Matching) genutzt wird, um Positionen zu bestimmen. Grundsätzlich können beim RX-Matching nur Daten verwendet werden, die **während des Telefonats** im zwei Hz Takt an der Abis-Schnittstelle aufgezeichnet werden. Dabei werden unter anderem folgende, für das RX-Matching, wichtige Daten ausgegeben:

- Teilnehmer ID (anonymisiert)
- ID Location Area (Location Area in der sich die bedienende Zelle befindet)
- ID bedienende Zelle (Zelle in der Mobilfunkteilnehmer angemeldet ist)
- TA-Wert (Timing Advance – Abstand zur Antenne in 550 m - Schritten)
- Signalstärke bedienende Zelle (dBm)
- IDs aller empfangenen Nachbarzellen (bis zu sechs)
- Signalstärken aller empfangen Nachbarzellen (dBm)

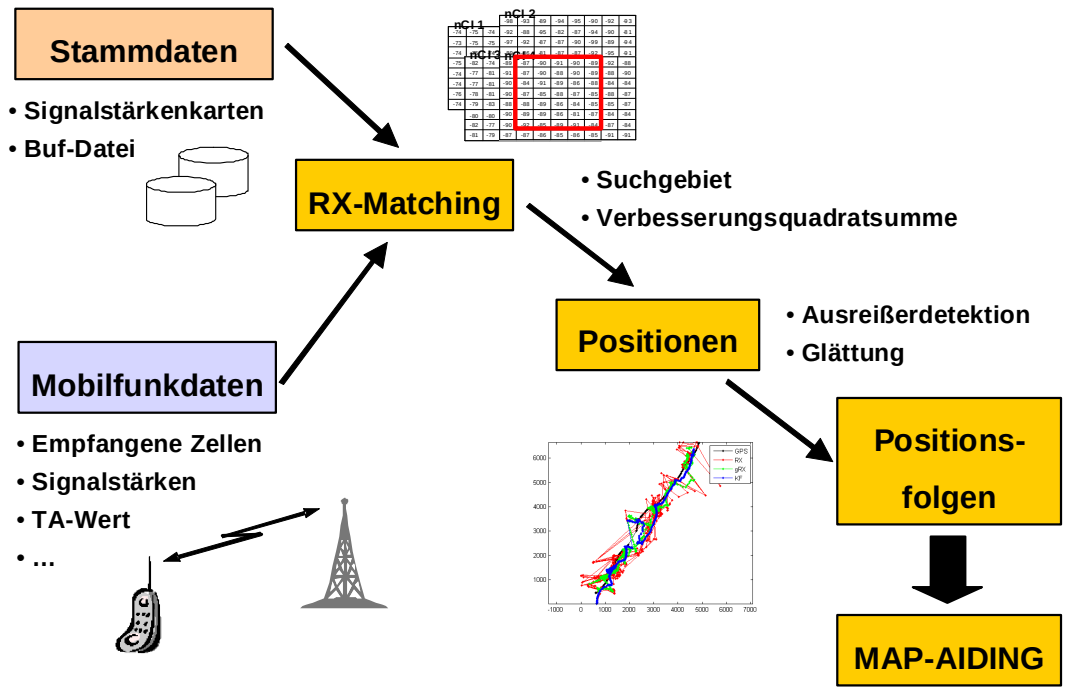
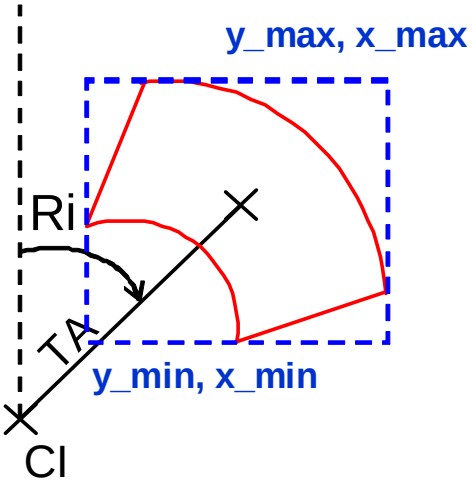


Abbildung 3-2: Ablauf der Mobilfunkortung aus Abis-Daten

Abbildung 3-2 Zeigt den Gesamtablauf zur Generierung der Positionsfolgen. Man erkennt gut, dass das RX-Matching der zentrale Baustein ist. Aus diesem Grunde wird es nachfolgend auch genauer beschrieben. Für eine genauere Dokumentation und für die Beschreibung der Glättungsalgorithmen wie dem Kalman-Filter sei auf den Bericht zur Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren (vgl. Anlage 3) verwiesen.

Im ersten Schritt werden für das RX-Matching aus den Informationen bedienende Zelle (CI), TA-Wert sowie die Abstrahlrichtung und dem Öffnungswinkel der Antenne ein Suchgebiet in Rechteckform berechnet.

Suchgebiet:



Sektormaske:

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Abbildung 3-3 : Berechnung des Zielgebietes und der Sektormaske

Im zweiten Schritt wird eine Sektormaske generiert, die der Größe des Suchgebietes entspricht. Hiermit werden die entsprechenden Gebiete aus den Signalstärkenkarten herausgeschnitten (vgl. Abbildung 3-4).

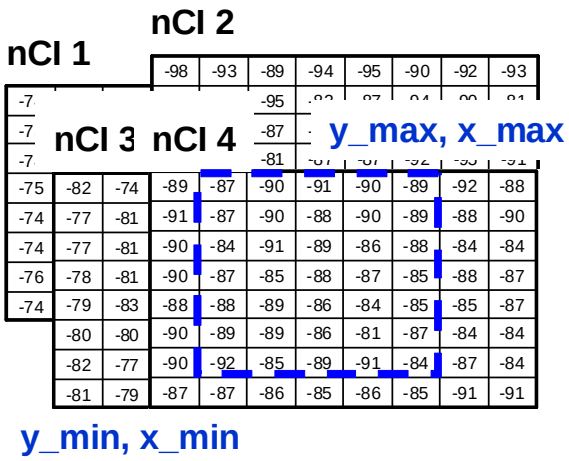


Abbildung 3-4: Extrahieren des Suchgebietes

Es entstehen dadurch, je nach Anzahl der empfangenen Mobilfunkzellen bis zu sieben Kartenauszüge.

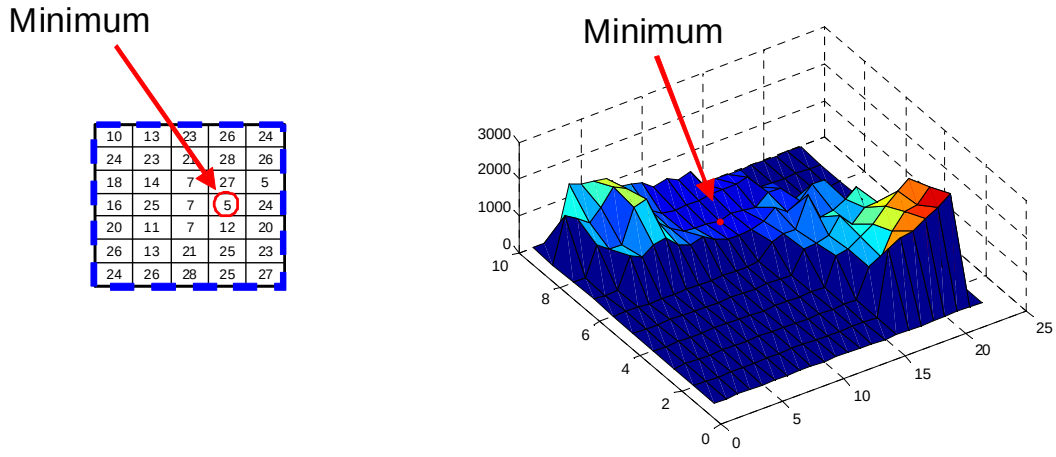
Dem RX-Matching selbst liegt dann folgende Idee zu Grunde. Es wird die Position gesucht, in die sich alle gemessenen Signalstärken im Vergleich zu den Soll-Signalstärken der Signalstärkenkarten optimal einpassen lassen. Das Minimum dieser Quadratsumme ergibt, bezogen auf die koordinierten Rasterwerte der Signalstärkenkarten, die gesuchte Position.

Dieses Verfahren wurde in der Projektlaufzeit erweitert. Da gemessene Daten mit Planungsdaten verglichen werden, kann es zu systematischen Fehlern kommen. Die Planungsdaten entsprechen nicht ganz der tatsächlichen Ausbreitung der Antennensignale. Dieser Unterschied ist jedoch bei allen Planungsdaten gleich, da hinter allen Berechnungen für die Signalstärkenkarten das gleiche Dämpfungsmodell steckt. Dieser systematische Fehler drückt sich näherungsweise durch einen Offset aus, der sich durch Bildung von Differenzen beseitigen lässt und zu einer deutlichen Positionsverbesserung beiträgt. Die Differenzen werden dabei zwischen einer Signalstärkenkarte der bedienenden Zelle und den Signalstärkenkarten aller anderen empfangenen Zellen gebildet. Dabei werden die Signalstärken gleicher Koordinaten voneinander abgezogen. Des Weiteren werden die empfangenen Signalstärken der einzelnen Zellen des Mobilfunkgerätes in gleicher Weise voneinander abgezogen. Anschließend werden aus den gebildeten Differenzen die Abweichungen gebildet und die Verbesserungsquadratsumme berechnet. Das Minimum ergibt dann wieder die gesuchte Position. Der Zusammenhang lässt sich mit folgendere Formel darstellen:

$$\min_{y,x\in A} \sum_{i=1}^{n} ((P_{Zell,BED}(y,x) - P_{Zell,NACH}(y,x)_i) - (P_{Mess,BED} - P_{Mess,NACH,i}))^2$$

A: Suchgebiet
n: Anzahl der gemessenen Zellen
P_{Zell}: Signalstärke aus Signalstärkenkarten auf Position (y,x) bezogen
P_{Mess}: Signalstärke aus Mobilfunkgespräch
BED: Bedienende Zelle
NACH: Nachbarzelle

In Abbildung 3-5 wird zum einen eine Ergebnismatrix der Quadratsummenbildung und zum anderen ein 3D-Bild eines Ergebnissektors gezeigt. Zur Verdeutlichung wurden auch hier sehr hoch gewichtete Werte auf Null gesetzt.

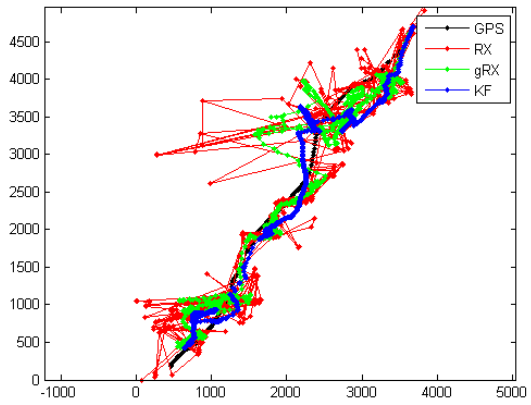


Ergebnis der Quadratsummenbildung 3D-Darstellung eines Ergebnissektors

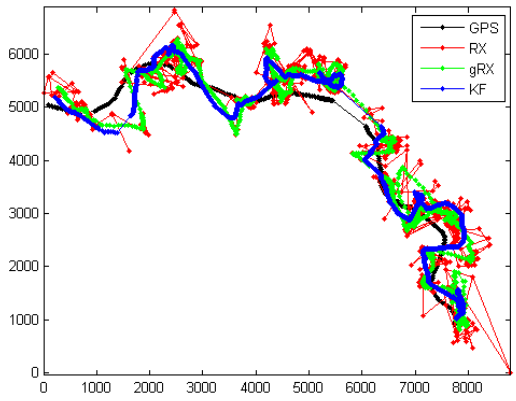
Abbildung 3-5: Ergebnis der Quadratsummenbildung

Mit dieser Vorgehensweise erhält man für jeden Zeitpunkt eine Position. In der Regel liegen zu jedem Verkehrsteilnehmer mehrere Positionen in dichter zeitlicher Reihenfolge vor. Diese werden durch gleitende Mittelbildung oder Kalman-Filterung (vgl. Anlage 3) zu Positionsfolgen zusammengefasst.

In den nachfolgenden Abbildungen sind einige Messfahrten, die die Ergebnisse der unterschiedlichen Auswertemethoden aufzeigen, als Beispiele dargestellt. Schwarz ist dabei die GPS-Trajektorie als Referenz (GPS). Rot sind die ungefilterten Positionen (RX), grün sind die mit einem gleitenden Mittel gefilterten (gRX) und letztendlich blau, die mit einem Kalman Filter gefilterten Positionen (KF). Des Weiteren sind für jede Messfahrt die Standardabweichungen der einzelnen Auswertestrategien bezogen auf die GPS-Solltrajektorie angegeben.

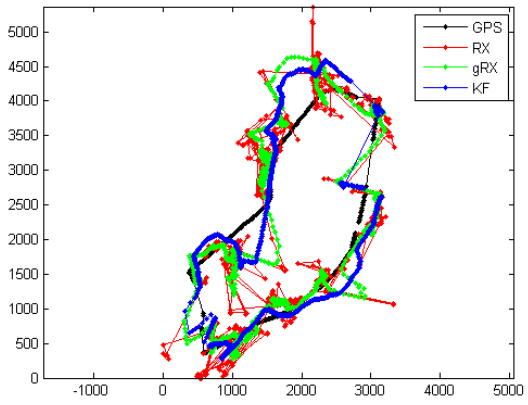


s_{RX}= 496 m; s_{gRX}= 316 m; s_{Kal}= 317 m

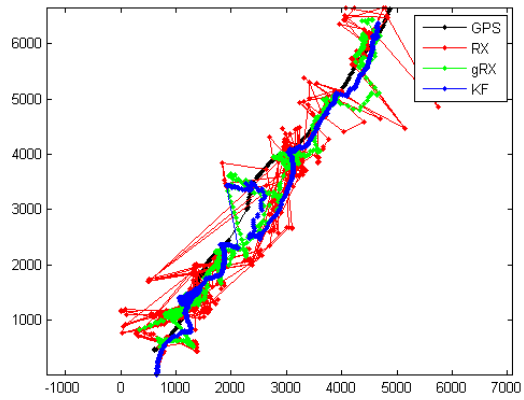


s_{RX}= 479 m; sg_{RX}= 409 m; s_{Kal}= 373 m

Abbildung 3-6: Versuchsfahrten Bundestrasse_B3_4_H2, Nr.72 (links) und Handy_an_Handy_1_H1, Nr.241 (rechts) (Bezeichnung der Messfahrten nach Anlage 5)



s_{RX}= 395 m; sg_{RX}= 333 m; s_{Kal}= 333 m



s_{RX}= 527 m; sg_{RX}= 384 m; s_{Kal}= 374 m

Abbildung 3-7: Versuchsfahrten Kurve_gr_4_H1, Nr.232 (links) und Bundesstrasse_B3_1_H2, Nr.42 (Bezeichnung der Messfahrten nach Anlage 5)

An Hand der gezeigten Beispiele lässt sich sehr gut erkennen, dass die beste Formanpassung an die gefahrene Strecke bei Nutzung des Kalman Filters zu finden ist. Weiter wird deutlich, dass die Genauigkeit der Positionen durch die Filteralgorithmen gesteigert werden kann, was für die weitere Datenprozessierung von großer Bedeutung ist.

3.3 Trajektoriengenerierung aus A und Abis-Daten (IAGB)

Im Projekt Do-iT werden alle für Ortung, Identifikation und Map-Matching erforderlichen Kartendaten aus dem Lokalisierungsnetz (Do-iT Datenmodell V05) abgeleitet, das auf GDF-Daten der Version NavTeq 38 Q3/2005 basiert. Im Gegensatz zum Projektnetz, das nur die für die Weiterverarbeitung interessanten Straßen enthält, beinhaltet das Lokalisierungsnetz alle Straßen im Projektgebiet, so dass auch Routen abseits des Projektnetzes gefunden werden können. Im Lokalisierungsnetz ist das Verkehrsnetz als kantenorientierter Graph dargestellt, wobei für jede Fahrtrichtung eine eigene Kante angegeben wird und der geometrische Verlauf einer Kante durch fortlaufend gespeicherte Formpunkte gegeben ist. In der weiteren Darstellung wird zwischen den beiden Schnittstellen A und Abis unterschieden. Auf eine Beschreibung von Algorithmen zur Mobilfunkortung auf A-Ebene kann verzichtet werden, da die Informationen hierfür nicht vorliegen und nur eine Prioirisierung der Teilnehmer erfolgt. Auf der Abis-Ebene folgt die Trajektoriengenerierung der Mobilfunkortung (siehe vorheriges Kapitel) nach.

Abis-Ebene

Die nachfolgende Abbildung zeigt prinzipielle Vorgehensweise bei der Generierung von FPD auf Abis-Ebene. Für Details sei auf Anlage 3 verwiesen. Hier werden ausgewählte Schritte erläutert und exemplarische Ergebnisse aufgezeigt.

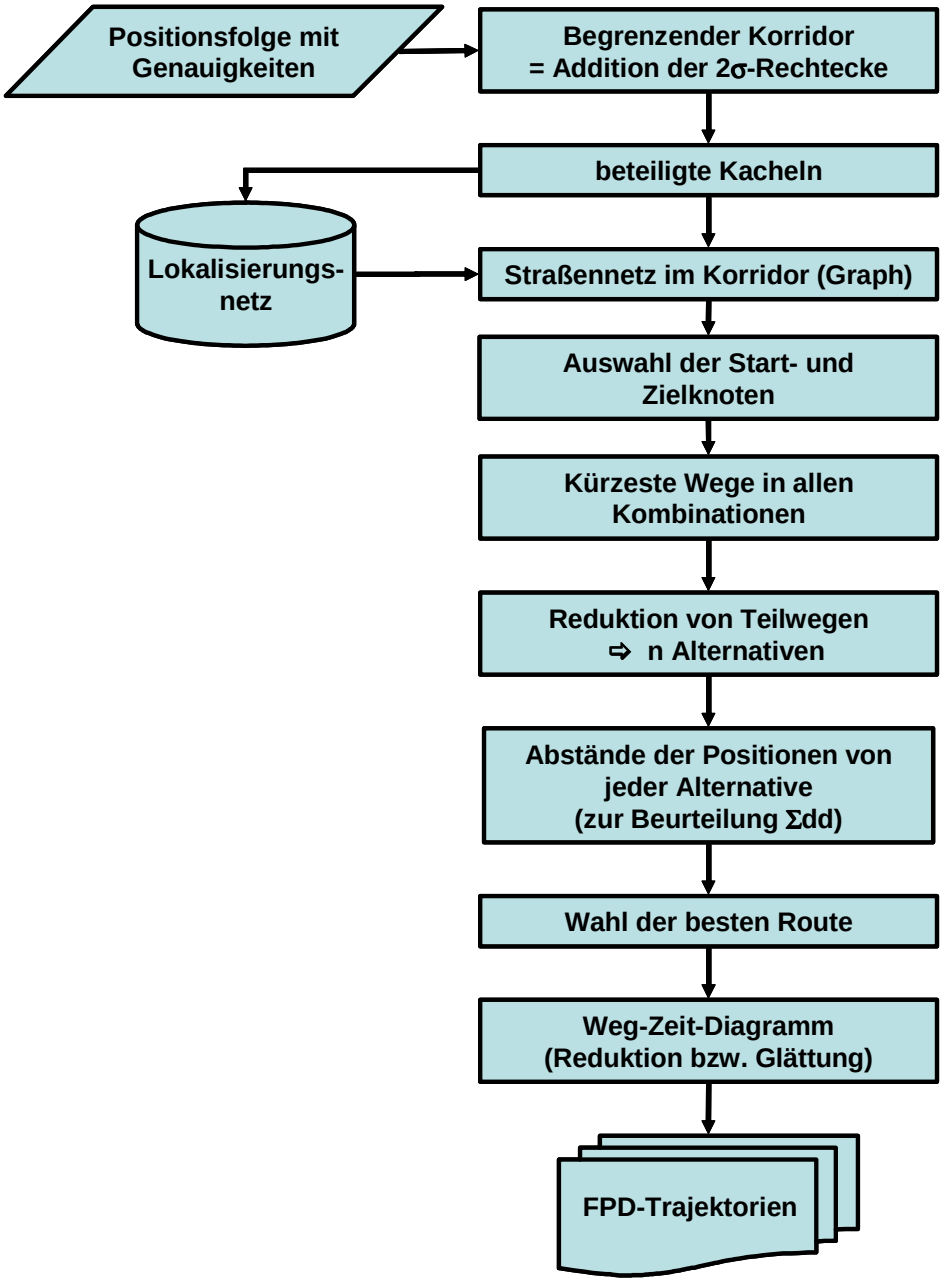


Abbildung 3-8: Schematischer Ablauf der Trajektoriengenerierung auf Abis-Ebene

Nachdem unter Nutzung der aus der Mobilfunkortung erhaltenen Ortungsgenauigkeiten ein Korridor aus dem Lokalisierungsnetz ausgeschnitten ist, werden die möglichen Start-und Zielknoten festgelegt (siehe Abbildung 3-9). In der Folge wird der Djekstra-Algorithmus genutzt, um mögliche Verbindungsrouten zu berechnen. Diese werden in der Folge aufgrund von Doppelbefahrungen von Straßenabschnitten reduziert, so dass eine reduzierte Anzahl von möglichen Verbindungen zur endgültigen Auswahl erhalten bleibt.

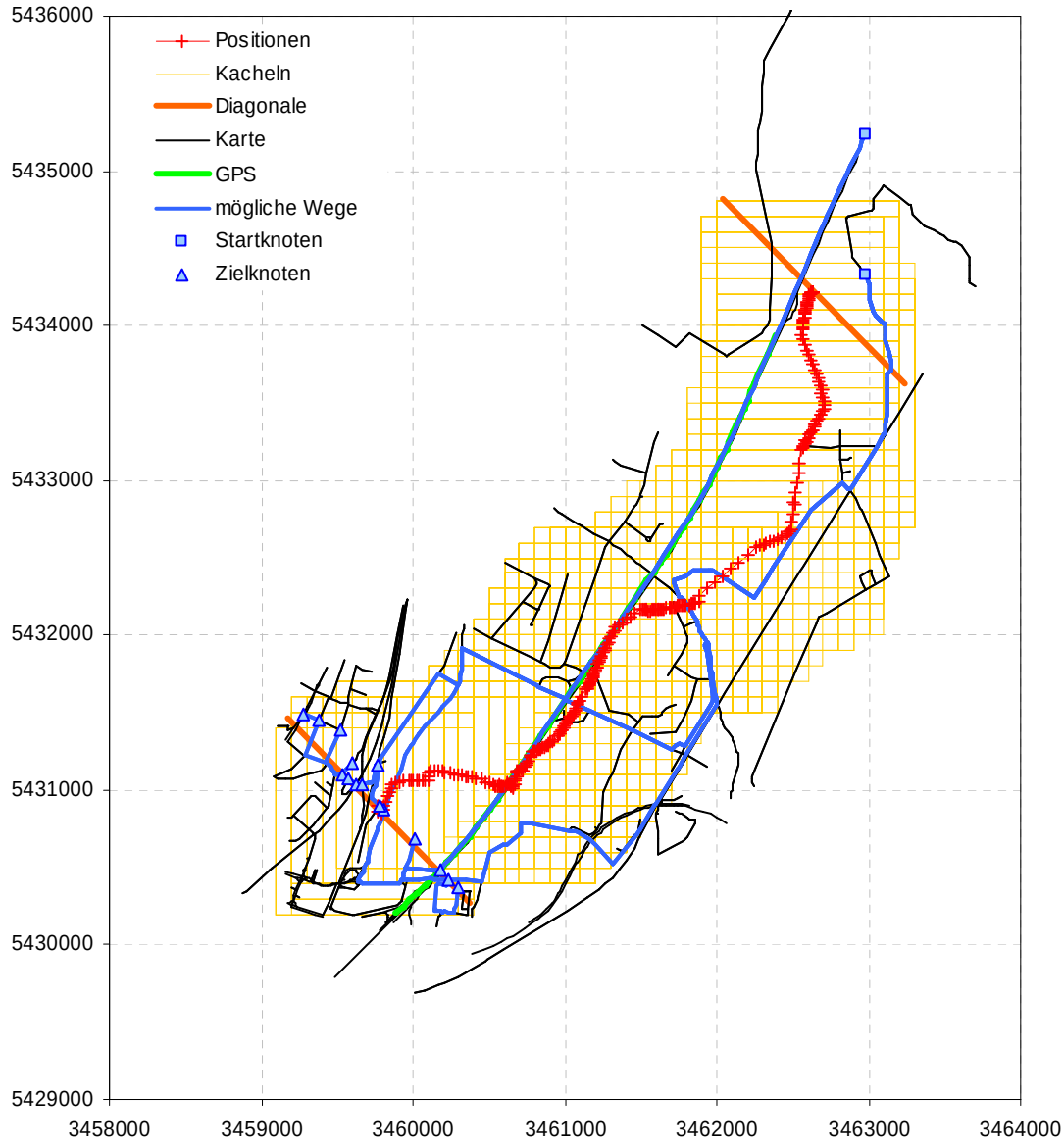


Abbildung 3-9: Wahl der Start- und Zielknoten

Für alle verbliebenen Routen werden die Querabweichungen zu der Mobilfunkroute berechnet. Hierfür werden die Lote auf die jeweilige Route gefällt und der jeweilige Abstand zwischen Lotfußpunkt und zugehörigem Punkt auf der Mobilfunkpositionsfolge als Lotabstand bzw. einzelne Querabweichung bestimmt. Abbildung 3-10 zeigt die Vorgehensweise. Dort wird aber von einer GPS-Referenztrajektorie aus gelotet. Dieses ist nur bei vorliegenden Referenzdaten möglich. Im Normalfall wird abschließend der RMS über alle Lotabstände gerechnet und die Route mit dem geringsten RMS wird als gefahrene Route mit Zeitbezug ausgegeben. Eine Route mit Zeitbezug wird als Trajektorie oder auch FPD bezeichnet.

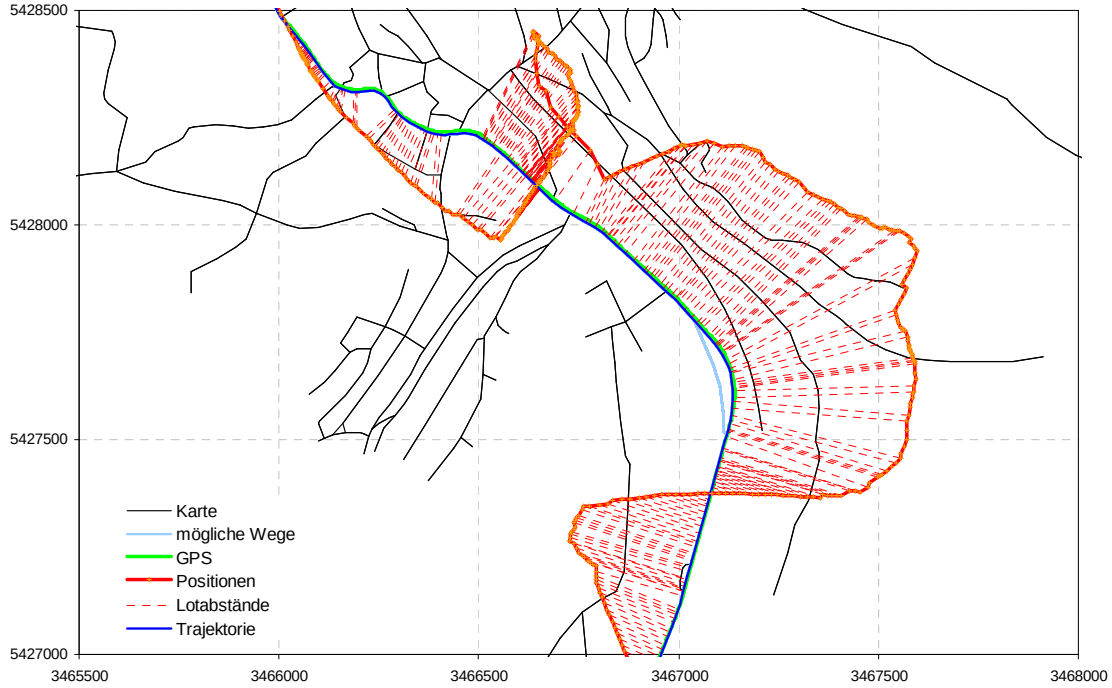


Abbildung 3-10: Lotfußpunkte

Die kartenreferenzierten Trajektorien können auch in der Form von Weg-Zeit- oder Geschwindigkeits-Zeit-Diagrammen dargestellt werden. Abbildung 3-11 zeigt exemplarisch ein Mobilfunk Weg-Zeit-Diagramm nach einigen Glättungsverfahren im Vergleich zu einen GPS-Bestimmung, die als Referenzdiagramm betrachtet werden kann.

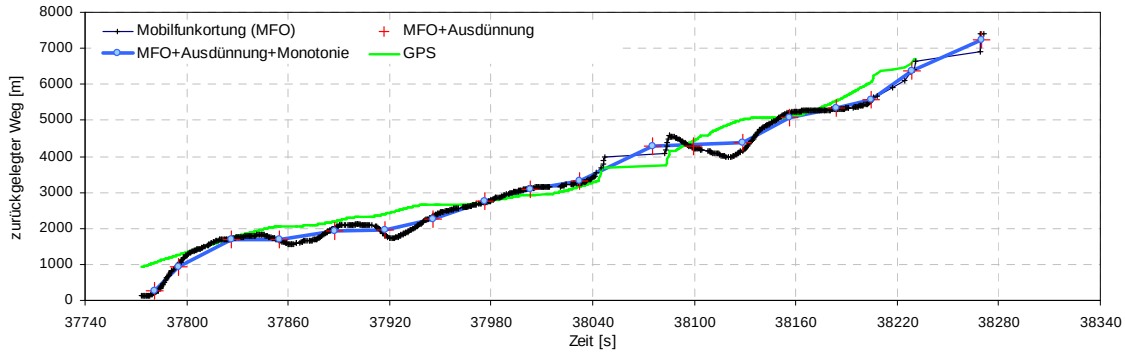


Abbildung 3-11: Weg-Zeit-Diagramm

A-Ebene

Die nachfolgende Abbildung zeigt die prinzipielle Vorgehensweise bei der Generierung von FPD auf A-Ebene. Für Details sei auf Anlage 35 verwiesen. Hier werden ausgewählte Schritte erläutert und exemplarische Ergebnisse aufgezeigt.

Die vorliegenden Informationen auf A-Ebene sind deutlich weniger. Es werden nur Zellwechsel oder sogenannte Location Area Updates (LUP) ausgegeben. Die Zellwechsel sind nur während eines Gesprächs verfügbar und können Positionsbereiche im Abstand von 100 m bis zu 10 km liefern. Die LUPs sind für alle nicht-ausgeschalteten Mobilfunkgeräte verfügbar, die Auflösung und Genauigkeit ist aber noch weiter heruntergesetzt. Folglich ergibt sich automatisch räumlich und zeitlich im Vergleich zur Abis-Ebene deutlich geringere Auflösung.

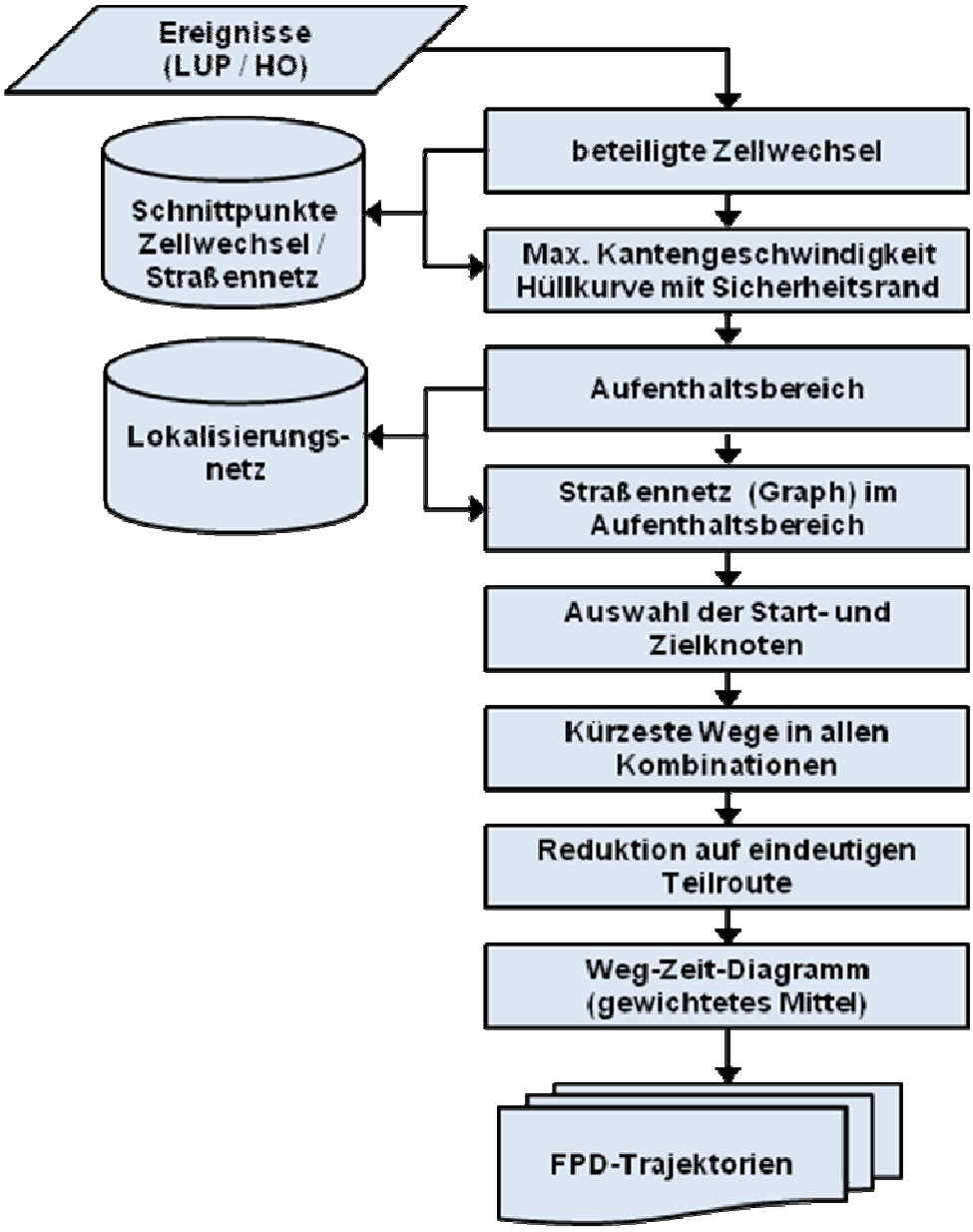


Abbildung 3-12: Schematischer Ablauf der Trajektoriengenerierung auf A-Ebene

Der Algorithmus sucht zunächst anhand der erfolgten Zellwechsel und LUPs und unter Berücksichtigung der maximalen Fahrgeschwindigkeit einen möglichen Aufenthaltsbereich aus dem Lokalisierungsnetz heraus. Dabei wird in Abhängigkeit von der Zellen-

bzw. Location Area Größe ein Sicherheitsrand b definiert, um zu vermeiden, dass auf dem Straßennetz kein Routing möglich ist. Abbildung 3-13 zeigt die prinzipielle Vorgehensweise zur Definition einer Aufenthaltsbereichs zwischen zwei Zellwechseln.

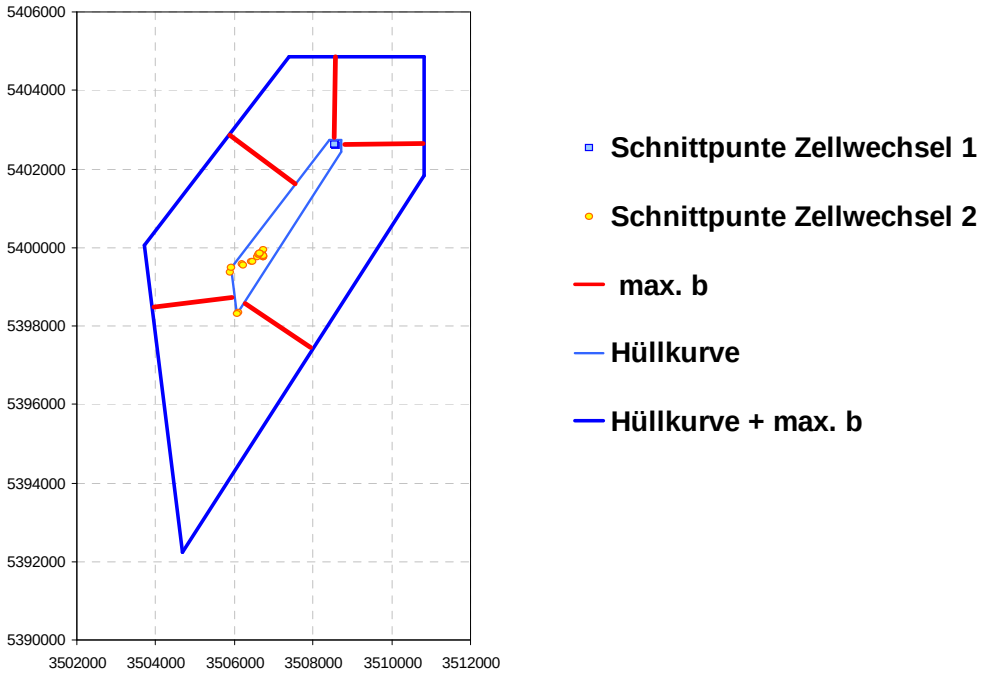


Abbildung 3-13: Hüllkurve und möglicher Aufenthaltsbereich

Verschiedene Aufenthaltsbereiche zwischen zwei aufeinander folgenden Zellwechseln können auf verschiedene Arten zusammengesetzt werden. Hier sei sich auf die Variante konzentriert, bei der zunächst alle Aufenthaltsbereiche zu einem Korridor zusammengefasst werden. Erst danach werden die verschiedenen Routen von dem ersten bis zum letzten Zellwechsel berechnet. Die Reihenfolge der Zellwechsel zwischen dem ersten und letzten Zellwechsel ist zumindest für die Positionsbestimmung ohne Bedeutung. Auf diese Weise ist dieser Algorithmus fehlertoleranter.

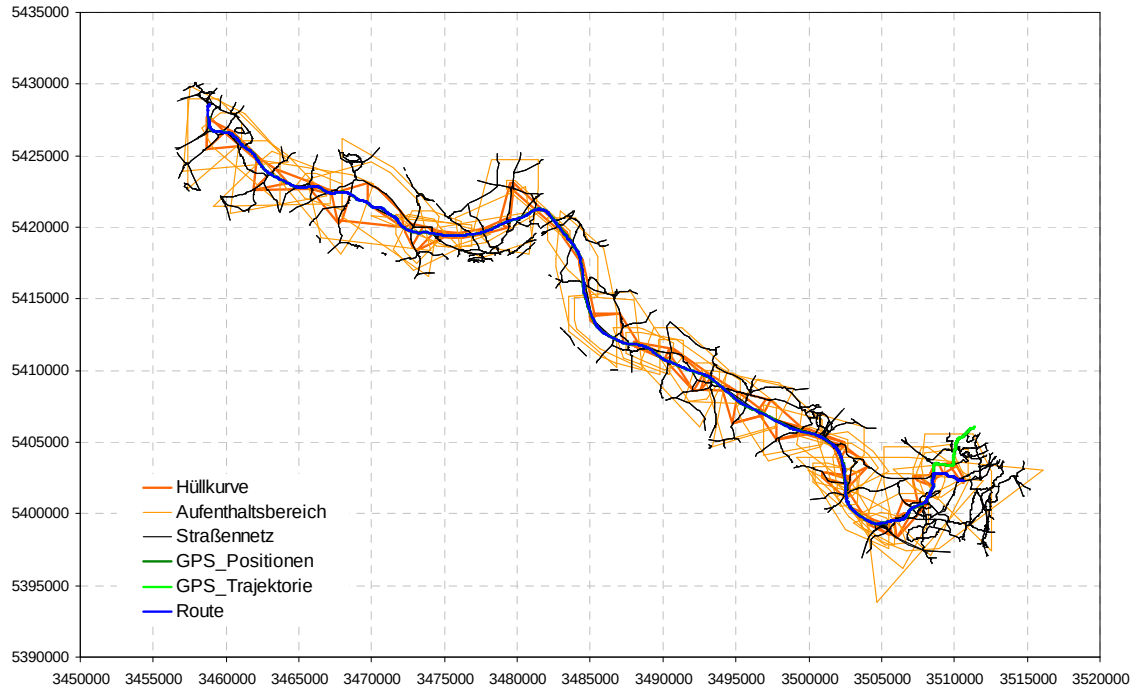


Abbildung 3-14: Testfahrt A8 Stuttgart ⇨ Karlsruhe

Abbildung 3-14 zeigt exemplarisch das aus dem Lokalisierungsnetz extrahierte Straßennetz, die Aufenthaltsbereiche und die endgültige Trajektorie sowie die in diesem Fall vorhandene GPS-Referenztrajektorie. Es ist gut zu erkennen, dass die Referenztrajektorie fast durchgängig getroffen wird. Die prinzipielle Eignung ist hiermit bereits nachgewiesen. Dieses gilt auch für die Weg-Zeit- und Geschwindigkeits-Zeit-Diagramme, die in nachfolgender Abbildung zu sehen sind.

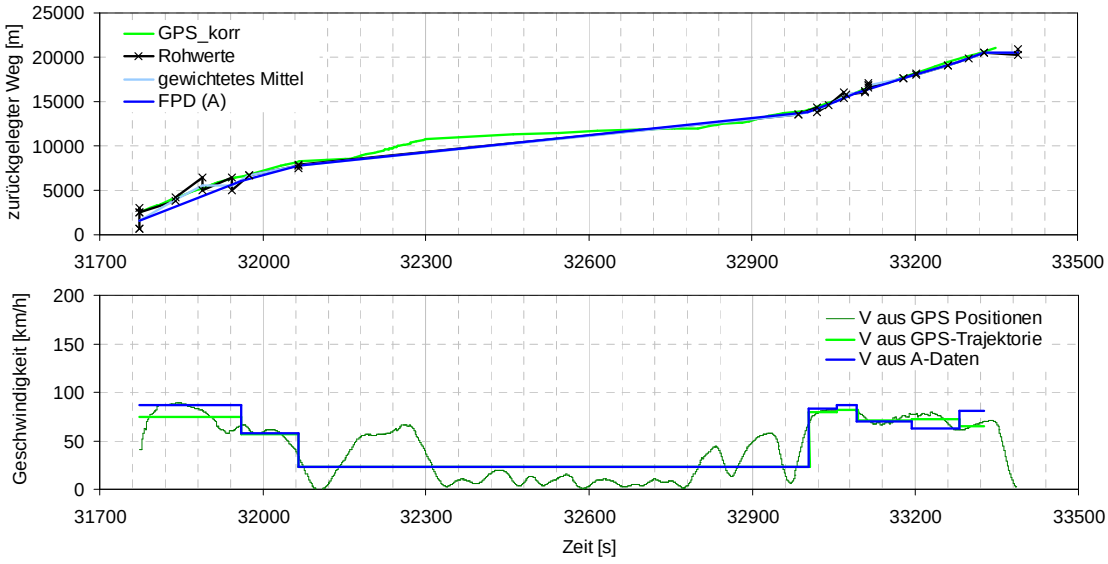


Abb.3-15: Weg-Zeit- und Geschwindigkeits-Zeit-Diagramm

3.4 Identifikation aktiver Verkehrsteilnehmer (IAGB)

Ein wesentlicher Bestandteil der Generierung von FPD ist die Identifikation im Verkehrssinne aktiver Mobilfunkteilnehmer, so dass Fehlinterpretationen z.B. aufgrund stehender –aber nicht im Stau stehender- Telefonierender ausgeschlossen werden können. Im Rahmen des Projekts beschränkt sich die Implementierung der Identifikationsalgorithmen auf die Abis-Schnittstelle, daher wird auch nur diese Vorgehensweise hier erläutert. Für die A-Schnittstelle ist eine einfache Priorisierung implementiert (vergleiche Anlage 3), die hier nicht weiter beschrieben wird. In diesem Bericht kann aufgrund des großen Umfangs der implementierten Methoden nur ein Überblick geboten werden. Für weitere Details sei auf die Anlage 4 und Anlage 5 verwiesen.

Zu Beginn der Identifizierung startet jeder Mobilfunkteilnehmer als passiv bis der Status aktiv mittels der Verfahren eindeutig identifiziert werden konnte.

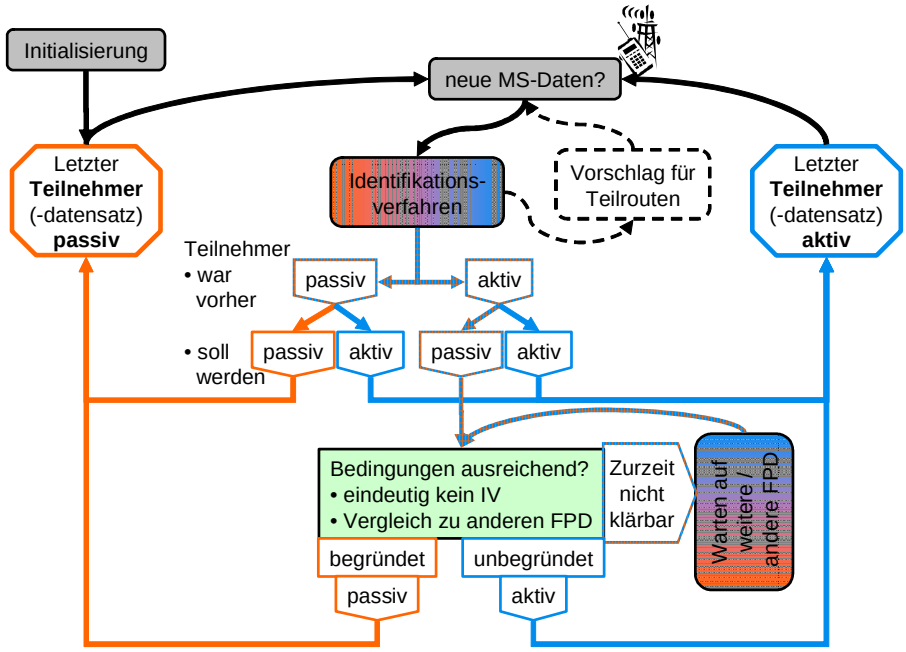


Abbildung 3-16: Ablaufdiagramm für die Online-Identifizierung

Für die weiteren Identifizierungen werden die Koordinatenfolgen genutzt. Die weiteren Analysen werden in kartenunabhängige und kartengestützte Analyse gegliedert. Beginnend mit der kartenunabhängigen Geschwindigkeitsanalyse wird eine Trennung von stehenden und sich bewegenden Mobilfunkteilnehmern durchgeführt, um eine verkehrliche Unterscheidung zu gewährleisten und um nicht erforderliche Analyseschritte einzusparen. Wird ein Teilnehmer als sich bewegend eingestuft, wird anschließend mit der kartenunabhängigen Geschwindigkeitsanalyse die Bewegungsgeschwindigkeit bzw. der Verlauf der Bewegungsgeschwindigkeit sowie die daraus ableitbare Verkehrsteilnehmerklasse ermittelt.
In den kartengestützten Analysen werden weitere Informationsquellen, wie digitale Straßennetze, ÖV-Fahrpläne und Points-of-Interest (POI), zur Verbesserung der Ver-

kehrsklassenidentifikation verwendet. In der „kartengestützten Geschwindigkeitsanalyse“ werden die vom Mobilfunkteilnehmer vermeintlich gewählten Routen mittels map-matching-Technologien zu einer Gesamtroute zusammengeführt, wodurch die Bestimmung des zurückgelegten Weges und damit die Berechnung der Geschwindigkeit verbessert werden kann. Unter der Annahme, dass die Fortbewegungen auf Verkehrswegen von Verkehrsnetzen stattfinden, werden die FPD-Positionen mit dem Vorhandensein und der Nutzungsmöglichkeit dieser Verkehrsnetze in der „verkehrsnetzgestützten Verkehrsteilnehmerklassifizierung“ untersucht. Mittels der Prognose von Fahrzeugpositionen des ÖV wird anschließend geprüft, ob Mobilfunkteilnehmer einen ÖV-Verkehrsmodus gewählt haben. In der „Aufenthaltsanalyse“ wird angenommen, dass an bestimmten POI nur ausgewählte Verkehrsteilnehmer anzutreffen sind. Wird ein FPD in der Nähe eines solchen POI geortet, so kann auf dessen Verkehrsteilnehmerklasse zurück geschlossen werden.
Nutzen mehrere FPD dieselben Fahrzeuge, ist verkehrstechnisch gesehen eine Zusammenfassung dieser FPD, das so genannte Clustern von FPD, anzustreben. In der abschließenden „Teilnehmer übergreifenden Klassifizierung“ werden hierfür Clusterverfahren vorgestellt, die nach ähnlichen Identifikationsmerkmalen der FPD suchen.

In Abbildung 3-17 ist der Ablauf der Identifizierungsverfahren dargestellt. Die einzelnen Verfahren werden in Anlage 4 erläutert.

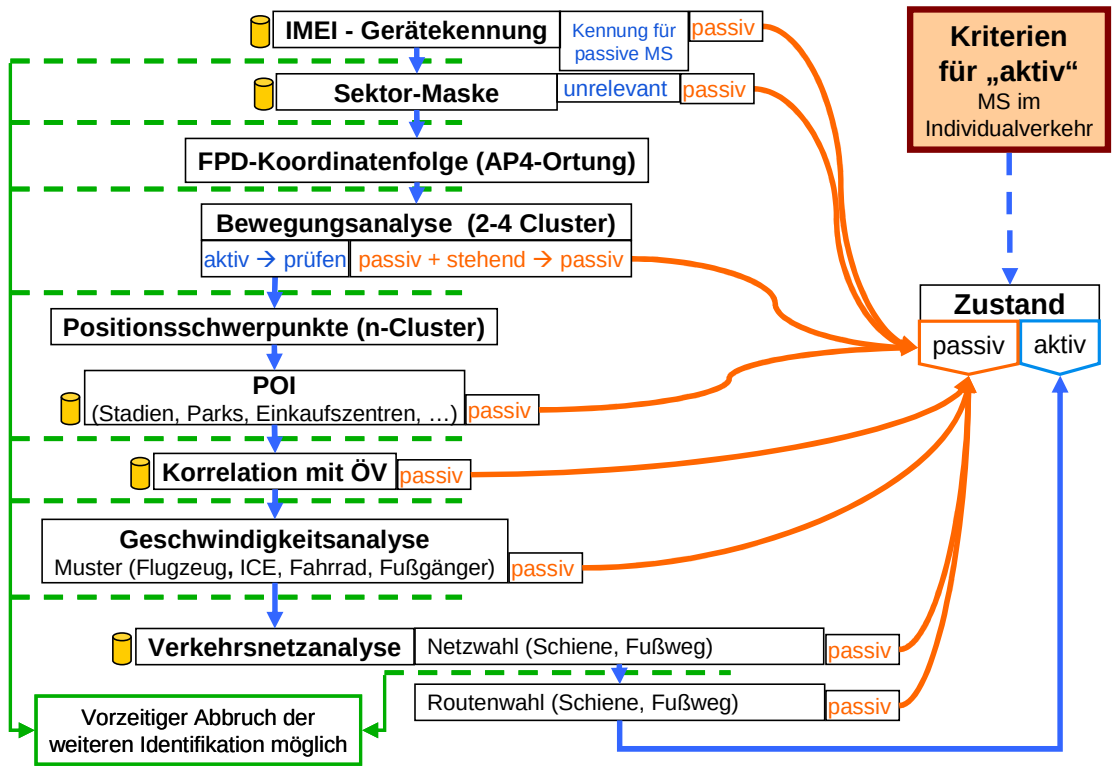


Abbildung 3-17: Ablaufreihenfolge der Identifikationsverfahren im Online-Betrieb

Tabelle 3-1: Stand der im Rahmen von Do-iT implementierten Verfahren

Identifizierungsverfahren (Abschnitt)	**ge-plant**	**imple-mentiert**	**in Betrieb**
IMEI – Gerätekennung	X		
Sektor Maske	X	X	X
xy – Bewegungsanalyse	X	X	X
xy – Positionsschwerpunkte	X	X	X
xy - POI – Analyse	X		
xy - Korrelation mit ÖV	X	X	X
xy - Geschwindigkeitsanalyse	X		
xy - Verkehrsnetzanalyse	X	X	X
xy - Unscharfes Routing	X	X	X
A-Daten - Unscharfes Routing	X	X	

In vorstehender Tabelle 3-1 wird für die verschiedenen Identifikationsverfahren der aktuelle Zustand dokumentiert. Die Spalte „in Betrieb“ dokumentiert, dass das entsprechende Verfahren bereits auf Abis-Ebene auf dem FPD-Server im Online-Betrieb läuft.

Da eine umfangreiche Darstellung aller Identifikationsverfahren den Umfang und die Tiefe dieses Berichts sprengen würde, werden in der Folge die Ergebnisse der Implementierung vorgestellt. Für Details sei auf Anlage 5 verwiesen. Es wurden insgesamt vier Versuchstage mit 89 Referenzfahrten ausgewertet. Es liegen während der Referenzfahrten aufgezeichnete GPS-Daten zur vergleichenden Auswertung zu Grunde. Die Verfahren zur Identifizierung von Mobilfunkteilnehmer (MFT) bestehen bei dieser Qualitätsbewertung im Einzelnen aus der „Bewegungsanalyse“ zur Erkennung des Bewegungszustandes, der „ÖV-Korrelation“ zur Erkennung von ÖV-Verkehren, der Verkehrsnetzzugehörigkeit zur Prüfung der infrage kommenden Verkehrsnetze auf Geometrieebene und einer Routensuche mit zeitlich-topologischen Restriktionen. Die Ergebnisse der Verfahren sowie deren Qualitätsbewertung werden im Folgenden zusammengefasst.

In der Bewegungsanalyse wird der Bewegungszustand der MFT anhand einer Signifikanzprüfung des zurückgelegten Weges mit einer Sicherheitswahrscheinlichkeit geprüft. Als Sicherheitswahrscheinlichkeit stellte sich ein Wert von 68 % als bester Kompromiss heraus. Mit diesem Wert konnten im Falle eines unbewegten Soll-Zustands 100 % aller mit GPS und 45 % aller FPD-Daten korrekt erkannt werden. Befand sich der MFT tatsächlich in Bewegung wurden bei den GPS-Daten 96 % und bei den FPD-Daten 88 % korrekt erkannt. Maßgebliche Einflüsse bei den unkorrekt erkannten Daten sind auf die Fußgängerszenarien und bei GPS auf systembedingte Ortungsfehler in dichter Bebauung und Erfassung in Gebäuden zurückzuführen.

Zur Identifizierung aktiven MFT, im Rahmen von Do-iT ist das die Erkennung von MIV (Motorisierter Individual-Verkehr), sind die nicht eindeutig vom ÖV (Öffentlichen Ver-

kehr) trennbaren MFT auszuschließen. Hierfür werden die ÖV-Verkehre: Linienbusse, Straßenbahnen sowie die Bahnen der DB AG im Identifikationsverfahren ÖV-Korrelation berücksichtigt. Fahrplanabweichungen werden durch eine Fuzzyfizierung der Fahrplandaten modelliert. Die Ergebnisdaten dieser Fuzzyfizierung wurden für den Online-Betrieb vorprozessiert und getrennt nach ÖV-Verkehr in Form effizienter Datenstrukturen bereitgestellt.

In den Analysen konnte für die GPS-Daten eine 73%-ige und für die FPD-Daten eine 61%-ige Erkennungsrate der ÖV-Szenarien nachgewiesen werden. Wurde eine ÖV-Korrelation nachgewiesen, dann handelte es sich bei GPS zu 91% und bei FPD zu 68% tatsächlich um Fahrten in einem ÖV-Fahrzeug oder um Fußgänger im Bahnhof. Das heißt eine zu Unrecht vorgenommene ÖV-Einstufung, und der damit verbundene Ausschluss der FPD-Daten, geschah nur zu 9% bei GPS und zu 32% bei FPD. Maßgebliche Fehler bei der Erkennungsrate sowohl bei GPS als auch bei FPD konnten auf fehlerhafte Geometrien der Schienenwege zurückgeführt werden. Es ist zu erwähnen, dass der Stichprobenumfang für verallgemeinernde statistische Aussagen insgesamt als gering einzustufen ist.

Die eingeschränkte Korrektheit der Schienendaten sowie unvollständige Fußwegenetze sind ebenfalls maßgebliche Ursachen der unkorrekten Einstufungen bei der Verkehrsnetzzugehörigkeit. In diesem Identifikationsverfahren wurden 87 % bei GPS und 94 % der Versuchsszenarien mit den FPD-Daten korrekt identifiziert. Die im Do-iT-Lokalisierungsnetz integrierte Straßenkarte enthält kein vollständiges Fußwegenetz und bedingt dadurch die Fehlerrate von 13 % bei GPS, die vollständig auf unkorrekt erkannte Fußgängerszenarien zurückzuführen ist. Das Identifikationsverfahren Verkehrsnetzzugehörigkeit dient vorrangig dazu, zu überprüfen, ob das nachfolgende Verfahren der Routensuche mit zeitlich-topologischen Restriktionen überhaupt einsetzbar ist. Und zwar kann nur dann von einer möglichen Route ausgegangen werden, wenn zu jedem Schwerpunkt, einschließlich Fehlerfigur, mindestens ein Element des zu untersuchenden Verkehrsnetzes gefunden wird. Ist das nicht der Fall, kann bereits aus topologischen Gründen keine Route gefunden und die entsprechende Verkehrsklasse ausgeschlossen werden. Für die GPS-Daten, die als Referenz aktiver MFT aufgezeichnet wurden, konnten alle Routen des MIV gefunden werden. Bei den FPD-Daten gelang dies für 66 % aller Fahrten. Die restlichen 34 % Fahrten wurden fehlerhaft als passiv eingestuft, da für sie keine den zeitlichen oder topologischen Kriterien genügende Route gefunden werden konnte. D. h. 34 % der aktiven MFT werden durch dieses Verfahren ggf. zu Unrecht ausgeschlossen.

Die aktiv-Einstufung eines MFT erfolgte bislang nur, wenn dies mindestens durch Vorhandensein einer Route auf dem Straßennetz belegt werden konnte. Eine Bewegung des MFT wird hiermit implizit vorausgesetzt. Besonders stehende MFT sind jedoch wesentlich zur Stauerkennung. Mit der Annahme, dass ein als aktiv identifizierter MFT keinen Verkehrsklassenwechsel vornahm, können nachfolgende ggf. als stehend erkannte Datensätze ebenfalls aktiv eingestuft werden. Das verkehrstechnische Verhal-

ten anonymer MFT ist als komplex und ohne Zusatzinformation nicht modellierbar hinzunehmen. Es wurde daher mittels einer maximal zulässigen Zeitdifferenz $dt_{\max\ active}$ ein Zeitraum definiert, innerhalb dessen angenommen wird, dass der bislang aktive MFT keinen Verkehrsmittelwechsel vornahm. Die Verkettung der FPD-Datensätze ermöglicht eine Verbesserung der Korrektheit für unbewegte aktive MFT. Zur Reduzierung von unentdeckten Verkehrsklassenwechseln kann der Parameter $dt_{\max\ active}$ optimiert werden. Eine empirisch begründete Wahl von $dt_{\max\ active}$ wurde anhand von Auswertungen eines 24-Stundendatensatzes gezeigt.

Die einzelnen Identifikationsschritte bzw. die gesamte Identifikationskette sind zwar ermutigend und konnten im Rahmen dieses Projekts verwertbare Resultate liefern, jedoch sind weitere Verbesserungen erforderlich, um höherer Korrektheitsraten zu erzeugen. Dieses ist weitergehenden Forschungsarbeiten vorbehalten.

3.5 Generierung von Trajektorien aus A-Daten (VuV)

Für viele verkehrliche Anwendungen ist es wichtig, dass Trajektorien über die gesamte Fahrt eines Verkehrsteilnehmers zur Verfügung stehen. Daher bietet sich ein Verfahren zur Generierung von Trajektorien an, das auf Basis der Location Area Codes (LAC) die Mobilfunkdaten auswertet, da Wechsel von einer Location Area in eine andere Location Area auch von Mobilfunkteilnehmern ohne Gespräch in den A-Daten aufgezeichnet werden (Location Area Update).

Daher wurde am Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik ein Verfahren zur LAC-basierten Trajektoriengenerierung entwickelt. Die folgende Abbildung 3-18 zeigt den Ablauf des Verfahrens, das im Wesentlichen darauf basiert, A-Daten-LAC-Folgen aus Mobilfunkdaten mit aus dem Straßen- und Schienennetz ermittelten Netz-LAC-Folgen zu vergleichen und bei Übereinstimmung den Mobilfunkteilnehmer einer bzw. bei Mehrdeutigkeit mehreren Routen im Verkehrsnetz zuzuordnen.

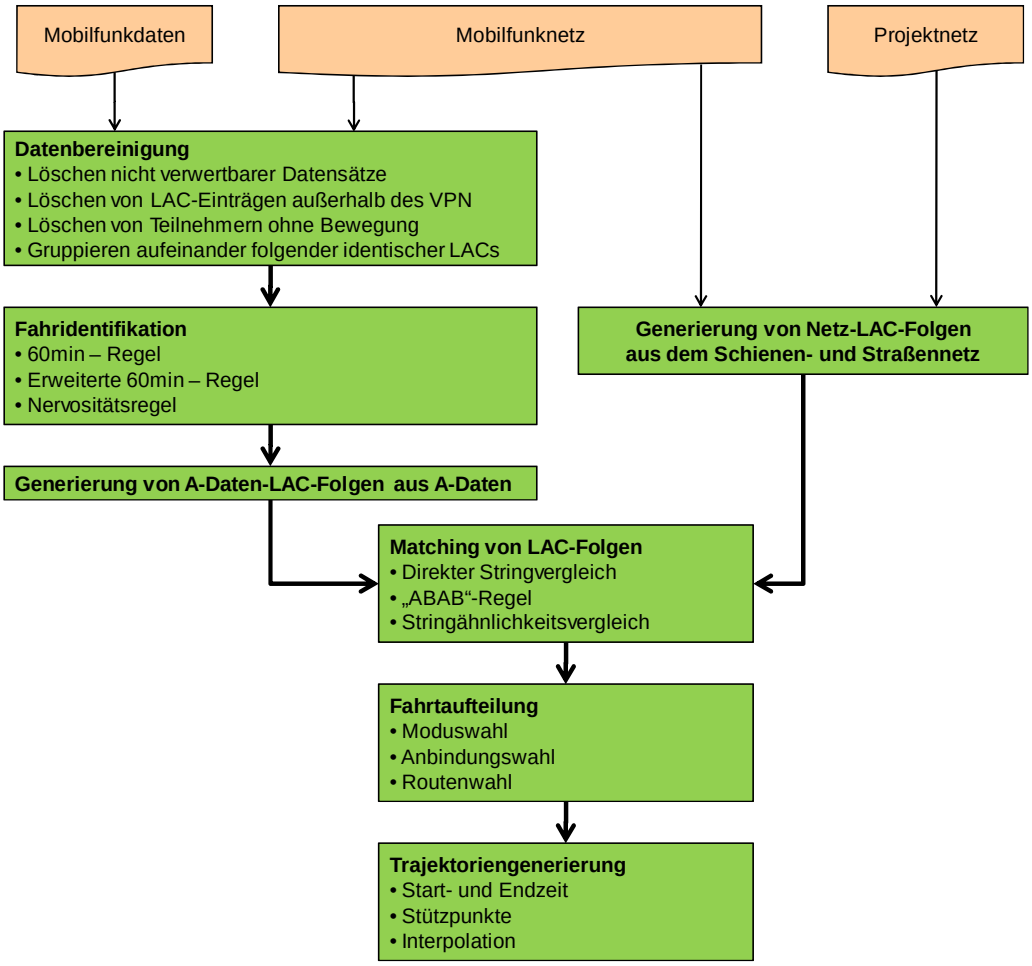


Abbildung 3-18: Ablauf der LAC-basierten Generierung von Trajektorien.

Für die Generierung von A-Daten-LAC-Folgen ist es nach der Datenbereinigung notwendig, einzelne Fahrten aus den Daten der Mobilfunkteilnehmer, die bis zu 24 Stunden am Stück aufgezeichnet werden, zu identifizieren und herauszuschneiden. Ergebnis sind einzelne LAC-Folgen, die bei einer Umwandlung der Location Area Codes in Zeichen als Zeichenfolge (z.B. „ABFREH“) dargestellt werden können.

Für die Netz-LAC-Folgen aus dem Straßen- und Schienennetz werden zunächst über ein Mehrwege-Suchverfahren alle sinnvollen Wege zwischen allen Location Areas gesucht. Anschließend werden den Strecken und Knoten im Verkehrsnetz die jeweils nächste Location Area zugeordnet, so dass auch hier die LAC-Folgen als Zeichenfolge dargestellt werden können.

Über ein dreistufiges Verfahren werden dann für die A-Daten-LAC-Folgen passende Netz-LAC-Folgen gesucht:

1. *Direkter Stringvergleich:*
   Im ersten Schritt wird geprüft, ob es eine direkte Übereinstimmung zwischen einer beobachteten A-Daten-LAC-Folge und einer Netz-LAC-Folge gibt.

2. *ABAB-Regel:*
   Die ABAB-Regel berücksichtigt, dass Mobilfunkteilnehmer aufgrund von Abschattungen o.ä. gelegentlich zwischen Location Areas hin- und herwechseln. Die ABAB-Regel kürzt die A-Daten-LAC-Folgen aus den Mobilfunkdaten um diese mehrfachen Wechsel und vergleicht die daraus entstehenden LAC-Folgen wiederum direkt mit den Netz-LAC-Folgen aus dem Straßen- oder Schienennetz.

3. *Stringähnlichkeitsvergleich:*
   Durch Abschattungen sowie Ungenauigkeiten bei der Berechnung der LAC-Folgen werden etliche LAC-Folgen trotz eindeutiger Bewegung nicht in den ersten beiden Schritten als Fahrt identifiziert. Daher erfolgt im letzten Schritt ein Stringähnlichkeitsvergleich, der sehr ähnliche LAC-Folgen identifiziert und somit eine Routengenerierung ermöglicht.

Tabelle 3-2 zeigt die Anzahl der identifizierten Fahrten in den drei Schritten. Insgesamt werden an einem durchschnittlichen Werktag ungefähr 500.000 Fahrten identifiziert, von denen ca. 420.000 im Straßennetz und 80.000 Schienennetz verlaufen.

Tabelle 3-2: Statistik für das Matching der LAC-Folgen (23. Juni 2008).

Beschreibung	**Identifizierte Fahrten**			**Anteil identifzierter Fahrten**	
	Straße	**Schiene**	**Gesamt**	**je Schritt**	**kumuliert**
Direkter Stringvergleich	210.000	37.000	247.000	39,9%	39,9%
„ABAB“-Regel	73.000	7.000	80.000	13,0%	52,9%
Stringähnlichkeitsvergleich	137.000	38.000	175.000	28,3%	81,2%
Gesamt	**420.000**	**82.000**	**502.000**		

Einige, vor allem kurze, LAC-Folgen repräsentieren sowohl Fahrten im Straßen- als auch Schienennetz. Diese werden genauso wie mehrdeutige Routen innerhalb einer Location Area durch verschiedene Aufteilungsverfahren auf die möglichen Routen verteilt. Die Abbildung 3-19 zeigt die endgültigen Belastungen im Projektnetz für den 23. Juni 2008.

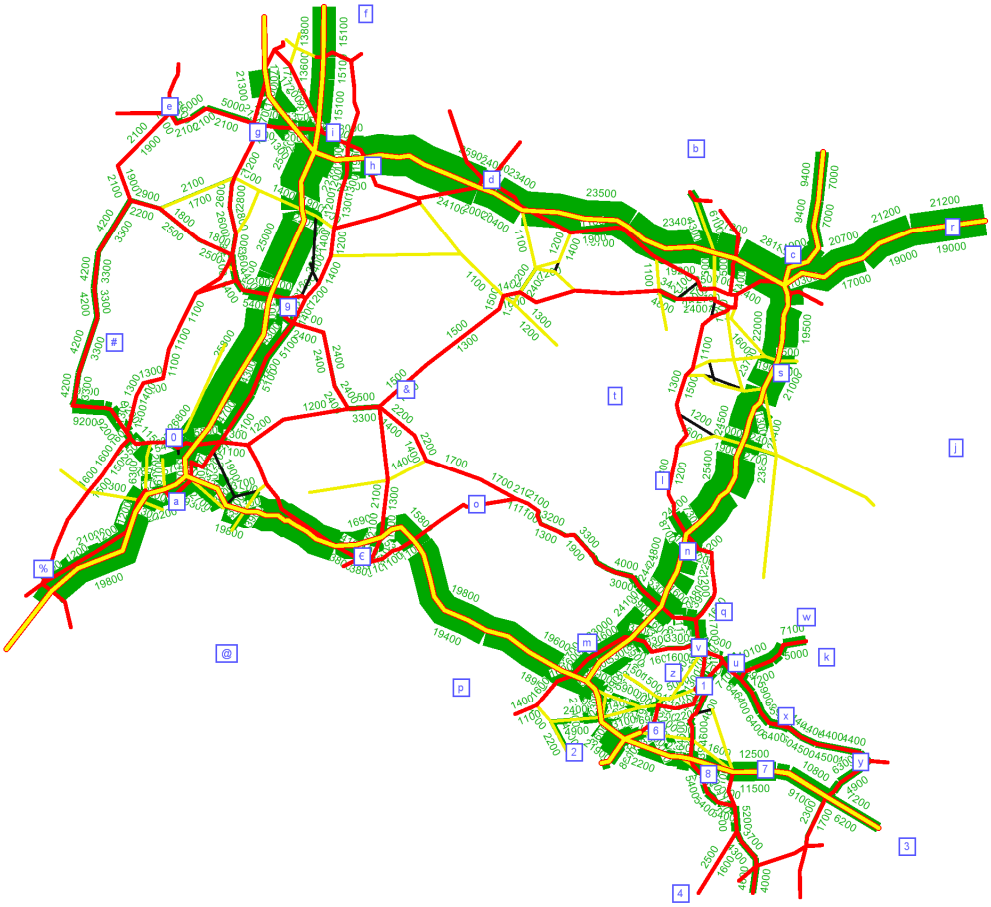


Abbildung 3-19: Belastungen aus FPD-Trajektorien im Projektnetz.

Abschließend zeigt Abbildung 3-20 die Fahrzeiten zwischen Stuttgart und Walldorf (über Karlsruhe), die eine Aufteilung der Fahrzeuge in die Fahrzeugklassen Pkw und Lkw ermöglichen. Diese zeigt wie auch weitere detaillierte Analysen in Kapitel 3 der Anlage 2, dass das Verfahren für Anwendungen auf den Autobahnen im Untersuchungsgebiet sehr gute Trajektorien generiert.

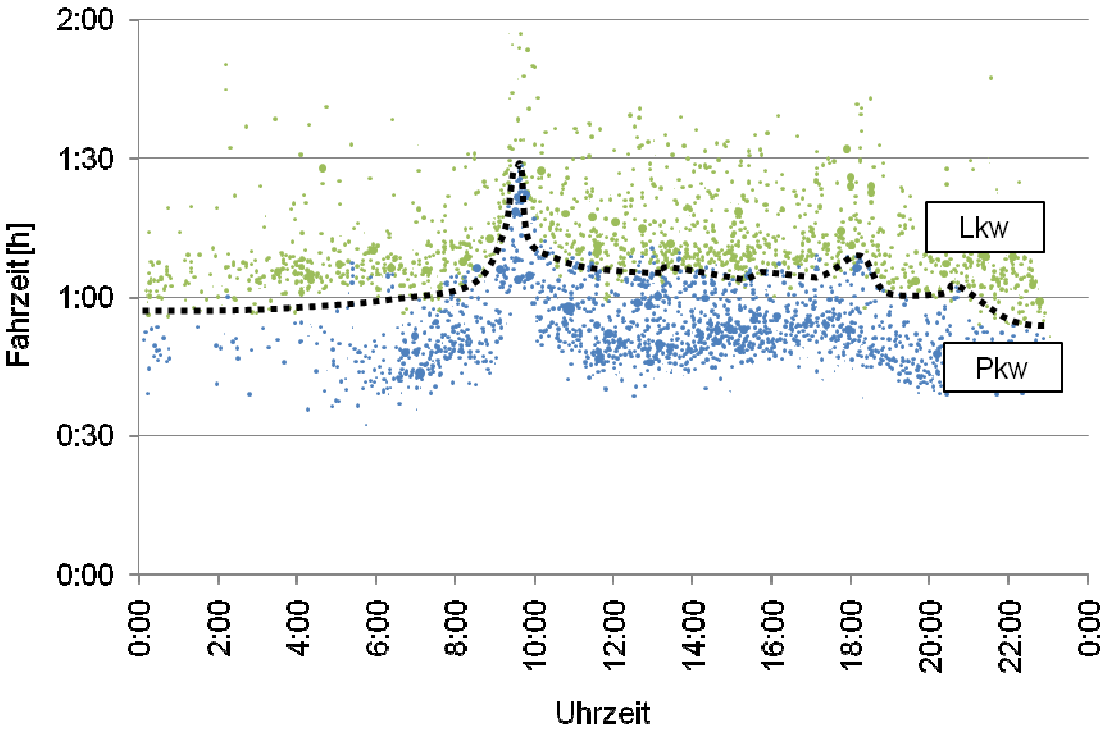


Abbildung 3-20: Fahrzeiten zwischen Stuttgart und Walldorf (über Karlsruhe) und Aufteilung in die Fahrzeugklassen Pkw und Lkw.

3.6 Schätzung der Parameter des Routenwahlverhaltens (VuV)

Aus den Trajektorien, die mit dem Verfahren aus Kapitel 3.5 generiert werden, können Untersuchungen zum Routenwahlverhalten in der BAB-Netzmasche (Autobahnen A5/A6/A8/A81) und im Bereich der Netzbeeinflussungsanlage Stuttgart (Autobahn A81 und Bundesstraßen B10/27/295) durchgeführt werden. Dabei wird mit Maximum-Likelihood-Schätzungen untersucht, welche Kenngrößen (Verkehrsmeldungen, Wechselwegweisung etc.) die Routenwahl überhaupt beeinflussen und, wenn ja, welche Gewichtung die einzelnen Kenngrößen haben. Abbildung 3-21 zeigt die Untersuchungsgebiete sowie die Lage der dynamischen Wegweiser mit integrierten Stauinformationen (dWiSta).

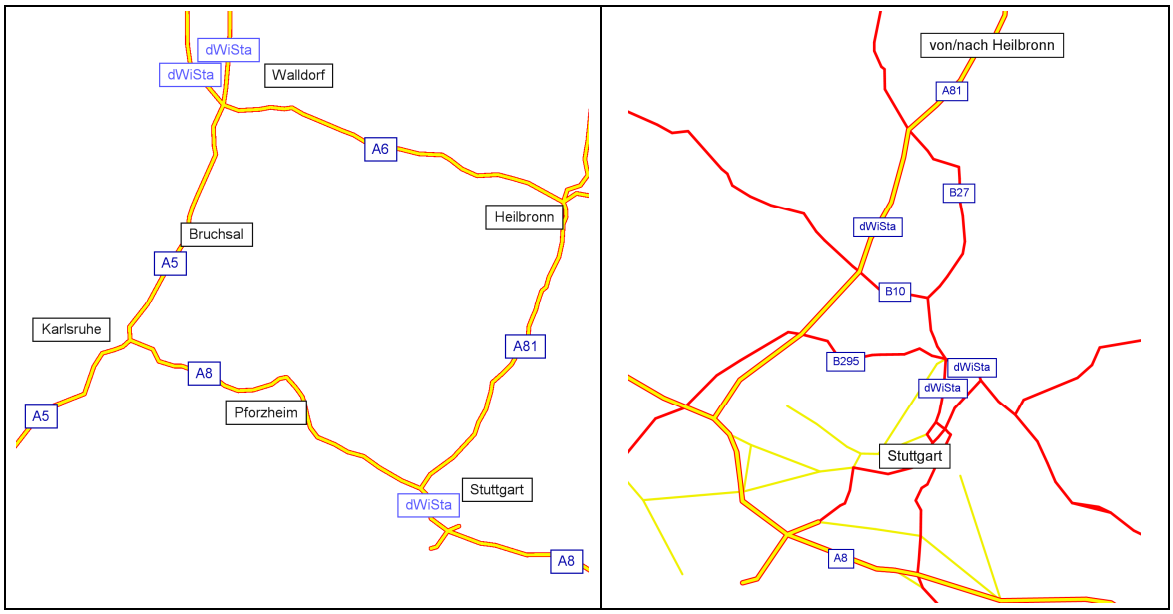


Abbildung 3-21: Untersuchungsgebiete für die Routenwahl
(links: BAB-Netzmasche; rechts: NBA Stuttgart)

In der BAB-Netzmasche können über eine Millionen Fahrten auf den beiden Diagonalen Stuttgart ↔ Walldorf und Heilbronn ↔ Karlsruhe analysiert werden. Für den Bereich der NBA Stuttgart können ca. 350.000 Fahrzeuge des Durchgangsverkehrs analysiert werden.

Für die *BAB-Netzmasche* ergeben sich statistisch abgesicherte Ergebnisse bei den Maximum-Likelihood-Schätzungen. Die Ergebnisse der Schätzungen können mit Hilfe des Logit-Modells grafisch dargestellt werden, so dass für jeden Zustand die zu erwartende Aufteilung abgelesen werden kann. So kann aus Abbildung 3-22 zum Beispiel abgelesen werden, dass bei 5 km mehr Stau über Heilbronn als über Karlsruhe bei inaktiver Wechselwegweisung (WWW) ca. 78 % des Durchgangsverkehrs die Route über Karlsruhe wählen. Eine aktive Wechselwegweisung bewirkt einen zusätzlichen Anteil von 6 %, was einen Befolgungsgrad von ca. 29 % bedeutet. Bei einer auf das Autobahndreieck Leonberg (bei Stuttgart) zufließenden Verkehrsstärke von z.B. 6.000 Fahrzeugen in einer Stunde ergibt sich daraus bei einem mittleren Anteil des Durchgangsverkehrs von ca. 12 % einen Anzahl von ca. 50 Fahrzeugen pro Stunde, die durch die Wechselwegweisung umgelenkt werden.

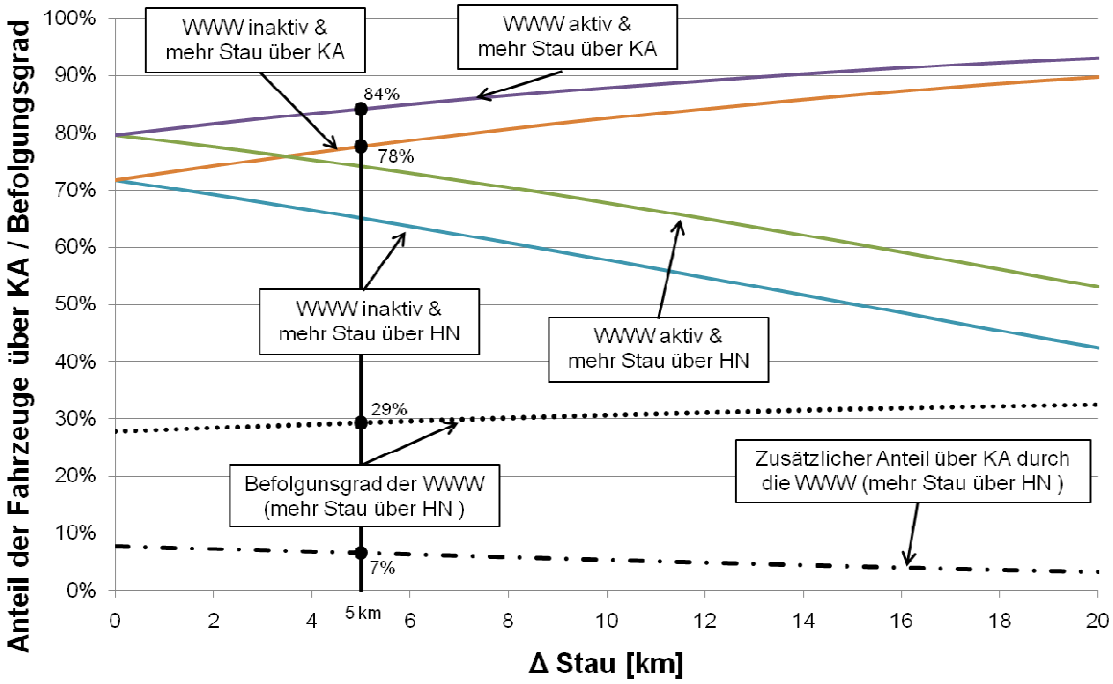


Abbildung 3-22: Grafische Darstellung des Basismodells der Routenwahl für die Richtung Stuttgart → Walldorf.

Weitere ausgewählte Ergebnisse der Routenwahlschätzungen sind:

- Pkw und Lkw zeigen ein vergleichbares Routenwahlverhalten. Dabei tendieren Lkw etwas stärker zur kürzeren Route, während Pkw etwas stärker auf Verkehrsmeldungen und die Wechselwegweisung reagieren.
- Die Verkehrsteilnehmer bewerten Verkehrsmeldungen, die sie früher auf ihrer Route erreichen, höher als Meldungen auf einem späteren Abschnitt ihrer Route. Bei einer Unterteilung in zwei Abschnitte der ca. 100 km langen Alternativen werden die Verkehrsmeldungen des ersten Abschnittes fast doppelt so hoch bewertet wie die des zweiten Abschnitts.
- Die Verkehrsteilnehmer, die der Routenempfehlung der Wechselwegweisung folgen, tun dies unabhängig von der Anzeige der Staulänge bzw. des Stauortes. Dies bedeutet allerdings nicht, dass die zusätzlichen Anzeigen keine Wirkung haben. Dies konnte nicht untersucht werden, da die zusätzlichen Anzeigen zu allen Routenempfehlungen gegeben worden sind.

Für die *NBA Stuttgart* ergeben sich bei den Maximum-Likelihood-Schätzungen weniger statistisch gesicherte Ergebnisse. Dies ist unter anderem auf die seltene Schaltung der NBA zurückzuführen, die dazu führt, dass z.B. in Stadteinwärtsrichtung nur 545 Fahrzeuge während einer Schaltung der NBA erfasst wurden. Abbildung 3-23 zeigt die Routenwahl für die beiden Richtungen in Abhängigkeit der Schaltungen der NBA mit Angabe der jeweiligen Fallzahl unter den Säulen:

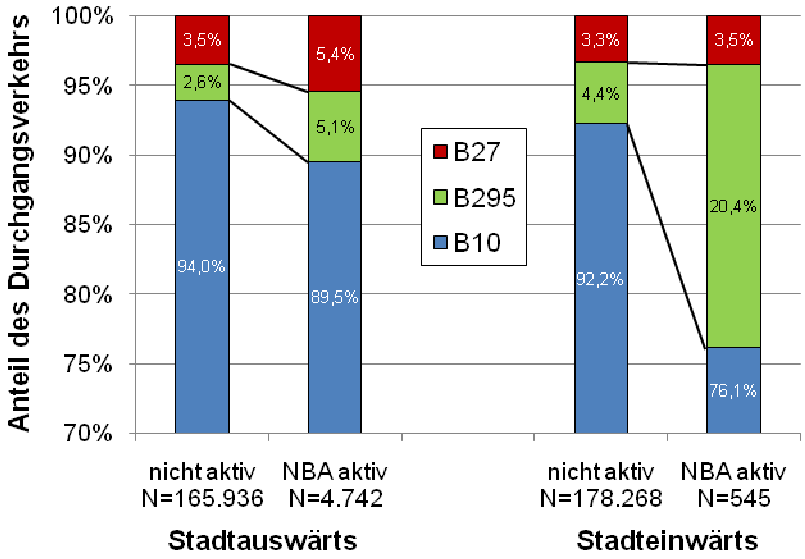


Abbildung 3-23: Routenwahlverhalten im Bereich der NBA Stuttgart.

In Stadtauswärtsrichtung wirkt eine Empfehlung der NBA in kleinerem Umfang auf beide Alternativrouten B27 und B295. In Stadteinwärtsrichtung dagegen erhöht die Routenempfehlung der NBA den Anteil der Route über die B295 deutlich. Aus den Werten ergeben sich bei Vernachlässigung der B27 Befolgungsgrade von 2,7 % (stadtauswärts) bzw. 17,3 % (stadteinwärts). Ein Einfluss der Staumeldungen kann nicht festgestellt werden.

Die detaillierte Verfahrensbeschreibung und die Ergebnisse sind im Kapitel 4 der Anlage 2 zu finden.

3.7 Generierung von Quelle-Ziel-Matrizen (VuV)

In Kapitel 5 der Anlage 2 werden Methodik und Ergebnisse zur Erzeugung von typischen Quelle-Ziel-Matrizen aus Mobilfunktrajektorien beschrieben. Die Methodik gliedert sich dabei zunächst in die folgenden beiden parallel laufenden Prozesse:

- Erzeugen von Clustern typischer Tage (z.B. Mo-Do, So/Feiertag, usw.) aus den Ähnlichkeiten der Ganglinien an Richtungsmessquerschnitten (RMQ).
- Tagesweises Hochrechnen der vorab erzeugten Mobilfunktrajektorien an den Zähldaten der jeweiligen RMQs.
- Zusammenfassen der hochgerechneten Mobilfunktrajektorien eines Clusters zu einem typischen Tag.

Abbildung 3-24 zeigt das Ergebnis der Clusterung, wenn eine Klassifizierung in die drei Gruppen Montag bis Freitag während bzw. außerhalb der Schulferien und Samstag, Sonntag und Feiertage durchgeführt wird. Die Höhe der Balken gibt die Anzahl der

Tage in dem jeweiligen Cluster wieder und die Farbe repräsentiert den Wochentag. Die Abbildung zeigt deutlich, dass die Freitage (grün) in den Werktagsgruppen eine Sonderstellung einnehmen und Samstag und Sonntag eine Trennung in der dritten Gruppe erfahren. Die weiteren Cluster mit sehr wenigen Tagen repräsentieren entweder besondere Tage wie Heiligabend oder Ostermontag oder sind Ausreißer aufgrund besonderer Ereignisse (z.B. Unwetter, Streik im Öffentlichen Verkehr).

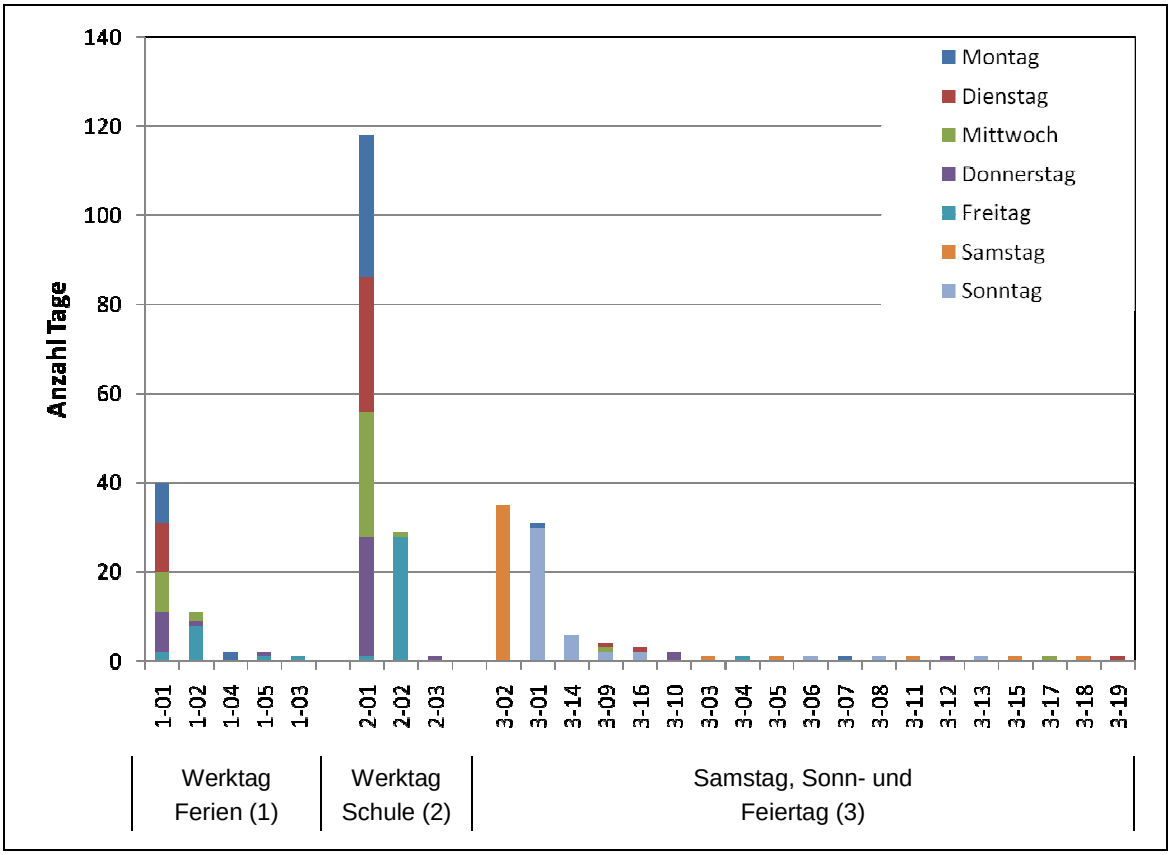


Abbildung 3-24: Clusterverteilung.

Die Belastungen der aus den Mobilfunkdaten erzeugten Trajektorien, die eine Stichprobe aller Kfz-Fahrten darstellen, werden mit einem dynamischen Hochrechnungsverfahren, das die räumlich/zeitliche Fortbewegung im Straßennetz berücksichtigt, hochgerechnet. Die Sollwerte des Verfahrens sind dabei die von den RMQs gezählten Verkehrsstärken, die in einer Qualität reproduziert werden können, die von anderen Verfahren bisher nicht erreicht wird. Abbildung 3-25 zeigt den Verlauf der Verkehrsstärken an einer Strecke im Projektnetz. Dabei ist die von einem RMQ gezählte Sollbelastung in rot, die Belastung der Mobilfunktrajektorien vor der Hochrechnung in schwarz und die Belastung der hochgerechneten Mobilfunktrajektorien in blau dargestellt.

Zu beachten ist, dass aufgrund der Mindestfahrtweite von Mobilfunktrajektorien nur RMQs auf Bundesautobahnen (BAB) verwendet werden und der Anspruch auf eine hohe Qualität des Ergebnisses auch nur auf BAB besteht.

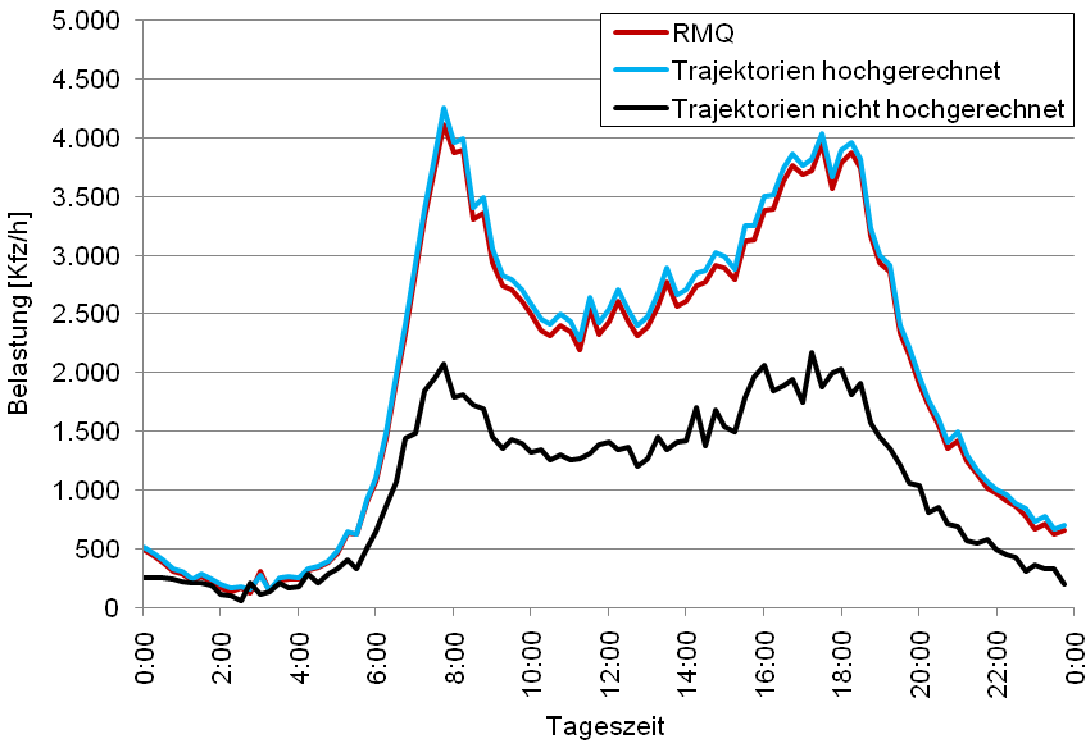


Abbildung 3-25: Verkehrsstärke von RMQ, Mobilfunktrajektorien und hochgerechneten Mobilfunktrajektorien am 19. September 2008 an der A81 von Stuttgart nach Heilbronn.

Abbildung 3-26 zeigt den Verlauf aggregierter Eingangsgrößen und die Ergebnisse der 68 hochgerechneten Tage, wobei die Sommerferien in Baden-Württemberg gelb hervorgehoben sind. Die schwarzen Kreuze repräsentieren die Summe der Belastungen aller eingegangen Mobilfunktrajektorien und die blauen Kreuze entsprechend dazu die Summe der Belastungen der hochgerechneten Trajektorien.

Für die Mobilfunktrajektorien vor und nach der Hochrechnung ist deutlich ein 7-Tageszyklus mit steigenden Belastungen von Montag bis Freitag und schwächerem Wochenendverkehr zu erkennen. Ferner geht die Summe aller Fahrten während der Sommerferien leicht zurück. Tiefergehende Analysen wie die Fahrtweitenverteilung oder aus anderen Untersuchungen bekanntes Routenwahlverhalten weisen ebenfalls plausible Ergebnisse auf (siehe auch Kapitel 5 der Anlage 2).

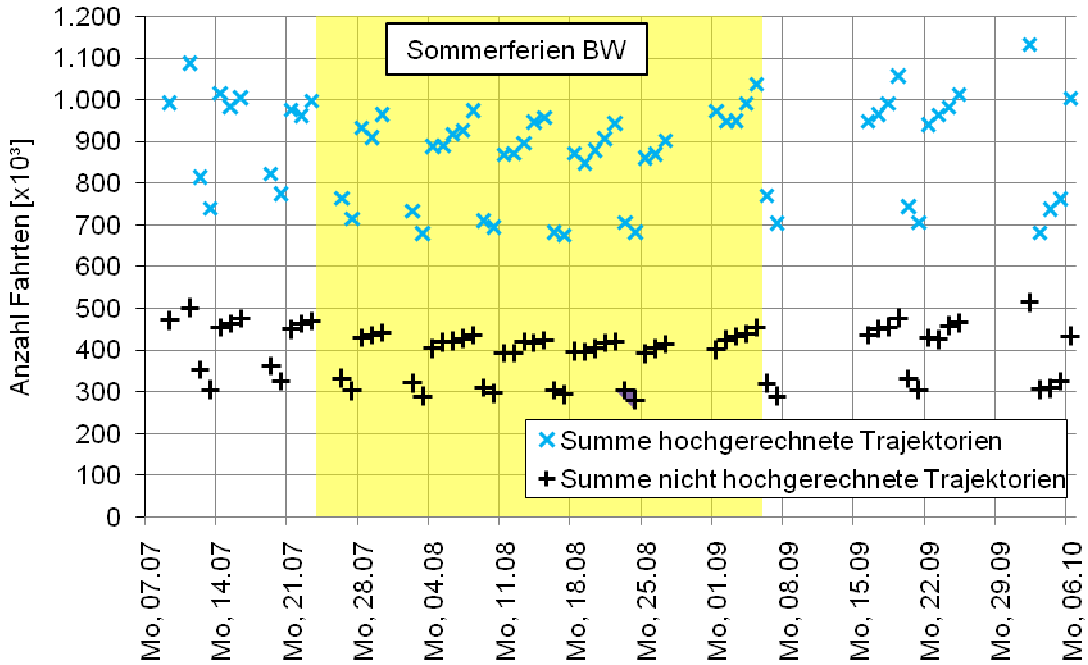


Abbildung 3-26: Anzahl Detektoren, Summe der FPD und hochgerechneten Fahrten.

Die hochgerechneten Trajektorien werden anschließend zu den aus der Clusterung erzeugten Tagesgruppen über eine einfache Mittelwertbildung zusammengefasst. Tabelle 3-3 listet die Cluster auf, die von den 68 hochgerechneten Tagen versorgt werden. Das letzte Cluster ist nur der Vollständigkeit halber aufgeführt, da in den 68 zur Verfügung stehenden Tagen mit dem 3. Oktober nur ein einziges Mal ein Feiertag auf einen Donnerstag oder Freitag fällt.

Tabelle 3-3: Erzeugte typische Verkehrstage.

Beschreibung	Anzahl enthaltener Tage	Anzahl Fahrten
Montag bis Donnerstag (Schule)	16	984.000
Freitag (Schule)	3	1.092.000
Montag bis Donnerstag (Ferien)	22	909.000
Freitag (Ferien)	4	978.000
Samstag	11	745.000
Sonntag oder Feiertag an einem Montag	11	714.000
Sonstige Feiertage	1	681.000

Aus diesen auf diese Weise erzeugten typischen Trajektorien können durch zeitscheibenbasiertes Aufsummieren beliebig feine Quelle-Ziel-Matrizen erzeugt werden. Für die üblichen Planungszwecke sind stundenfeine Matrizen, die auch zur Durchführung der im Folgenden beschriebenen Tagesprognose verwendet werden, allerdings vollkommen ausreichend.

3.8 Durchführung von Tagesprognosen (VuV)

Kapitel 6 der Anlage 2 beschreibt das Verfahren für die Tagesprognose. Ziel einer Tagesprognose ist es, den Verkehrszustand im Untersuchungsnetz für einen in der Zukunft liegenden Kalendertag stundenfein zu prognostizieren. Dabei werden die typische zeitabhängige Verkehrsnachfrage für den Kalendertag und bekannte Kapazitätsreduzierungen durch Baustellen beim Verkehrsangebot berücksichtigt:

- Verkehrsangebot: Als Grundlage für die Tagesprognose wird das auf TMC-Knoten basierende Projektnetz verwendet. In diesem Netz ist jede Strecke mit der zulässigen Geschwindigkeit und einer Stundenkapazität attributiert. Am Kalendertag vorhandene Baustellen werden über eine Reduktion der Kapazität und der Geschwindigkeit berücksichtigt.
- Verkehrsnachfrage: Als Nachfragedaten werden die Quelle-Ziel-Matrizen verwendet, die nach dem in Kapitel 3.7 vorgestellten Verfahren geclustert worden sind. Diese Matrizen bestehen aus 24 stundenfeinen Teilmatrizen. Quellen und Ziele sind TMC-Locations. Damit entspricht die Verkehrsnachfrage einem Erwartungswert der Verkehrsbelastung, mit dem an ähnlichen Tagen wie dem Prognosetag zu rechnen ist.

Die Berechnung der Tagesprognose erfolgt in drei Schritten. Zunächst wird für den zu prognostizierende Tag das Cluster bestimmt, zu dem er gehört. Daraus ergibt sich die für dieses Cluster typische Verkehrsnachfrage.

Im zweiten Schritt wird diese Verkehrsnachfrage auf das verwendete Netz umgelegt. Dabei wird das Routenwahlverhalten der Verkehrsteilnehmer nachgebildet. Routenwahlverfahren verteilen die Verkehrsnachfrage auf die im Netz verfügbaren Routen und ermitteln so Verkehrsstärken für jede Strecke. Um die Routenwahl nachzubilden, wird häufig von einem sogenannten Nutzergleichgewicht (Deterministic User Equilibrium, DUE) ausgegangen. Es ist dann gegeben, wenn sich die Verkehrsteilnehmer so verteilen, dass die objektiv messbare Fahrtdauer auf allen alternativen Routen gleich ist und jeder Wechsel auf eine andere Route die persönliche Fahrzeit erhöhen würde. Ein deterministisches Nutzergleichgewicht kann im Untersuchungsgebiet zumindest bezogen auf die Fahrzeit nicht beobachtet werden. Deshalb wird für die Tagesprognose ein stochastisches Nutzergleichgewicht (Stochastic User Equilibrium, SUE) verwendet, bei dem die Verkehrsteilnehmer ihre subjektive empfundene Fahrtdauer optimieren.

Führt man mit den vorliegenden Quelle-Ziel-Matrizen eine Verkehrsumlegung durch, erhält man für jede Strecke Verkehrsstärken. Da die Routenwahl neu berechnet wird, ist es möglich, Auswirkungen von Baustellen, wie beispielsweise Geschwindigkeitsbegrenzungen und Verringerung der Kapazität, in die Tagesprognose einzubeziehen. Die Routenwahl erfolgt nach den Widerständen, in denen sich die aus den Baustellen resultierenden Behinderungen niederschlagen. Damit wird davon ausgegangen, dass der

Verkehrsteilnehmer über die Behinderung durch die Baustelle auf der entsprechenden Strecke informiert ist und in seine Routenwahl mit einbezieht.

Aus der Gesamtverkehrsstärke der einzelnen Straßenabschnitte wird im dritten Schritt in Abhängigkeit der Kapazität für jede Stunde des Prognosetages die Auslastung berechnet. Daraus wird die zu erwartende Qualitätsstufe des Verkehrsablaufs (Level of Service, LOS) abgeleitet. Der LOS kann als Ergebnis der Tagesprognose wie in Abbildung 3-27 dem Verkehrsteilnehmer für jeden Streckenabschnitt und jede Stunde eines kommenden Tages angezeigt werden. Tabelle 3-4 zeigt die dafür verwendete Klassifizierung des LOS, die auf der Auslastung bzw. dem Verhältnis der Fahrzeit t_{akt} im belasteten Netz zur Fahrzeit t_0 im unbelasteten Netz basiert.

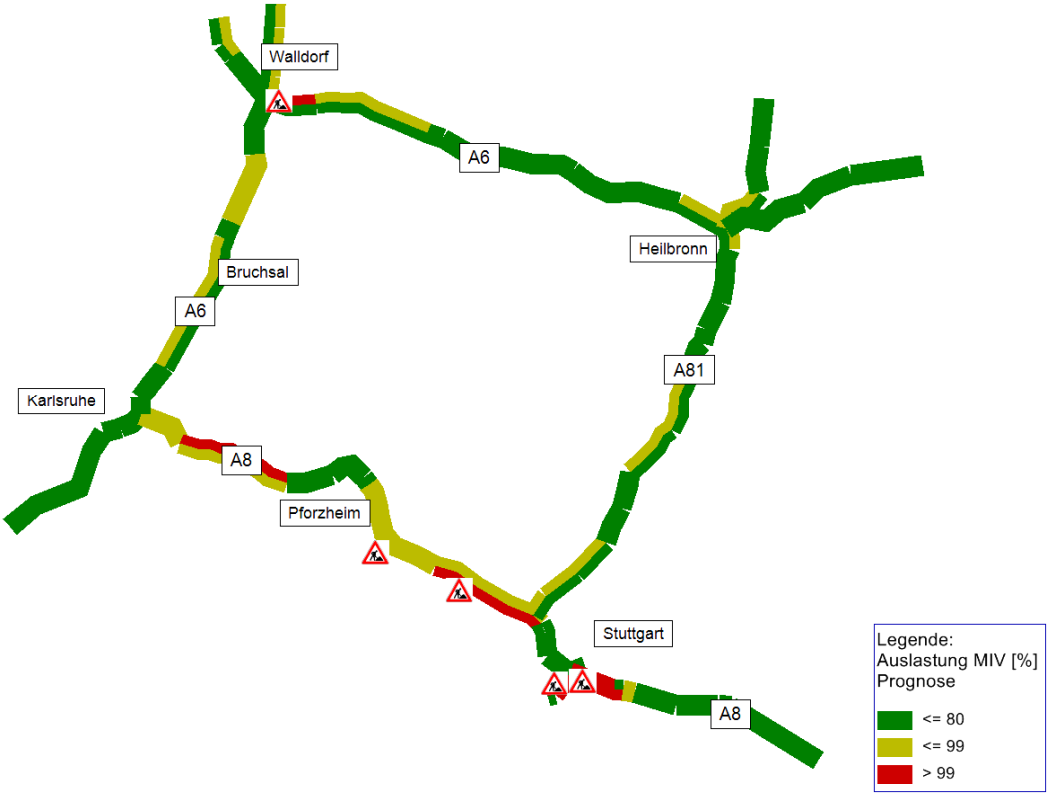


Abbildung 3-27: Level of Service-Darstellung als Ergebnis einer Tagesprognose am 23.06.08 von 7:00 Uhr bis 8:00 Uhr

Tabelle 3-4: Ableitung des LOS

Qualitätsstufe	**Auslastung**	**Verhältnis t_{akt}/t_0**
frei	< 80%	< 1,5
stockend	80% bis 99%	1,5 bis 2,0
gestaut	> 99 %	> 2,0

Um die Qualität der Tagesprognose abschätzen zu können, wurden Tagesprognosen für 68 Tage berechnet. Die daraus auf den Streckenabschnitten ermittelten Verkehrs-

stärken und Auslastungen werden dann mit den tatsächlich an diesem Tag gemessenen Verkehrsbelastungen verglichen. Dies erfolgt für jede Stunde der 68 Testtage. Abbildung 3-28 zeigt die Häufigkeit der Abweichung zwischen der in der Tagesprognose vorhergesagten und der tatsächlich eingetretenen Qualitätsstufe.

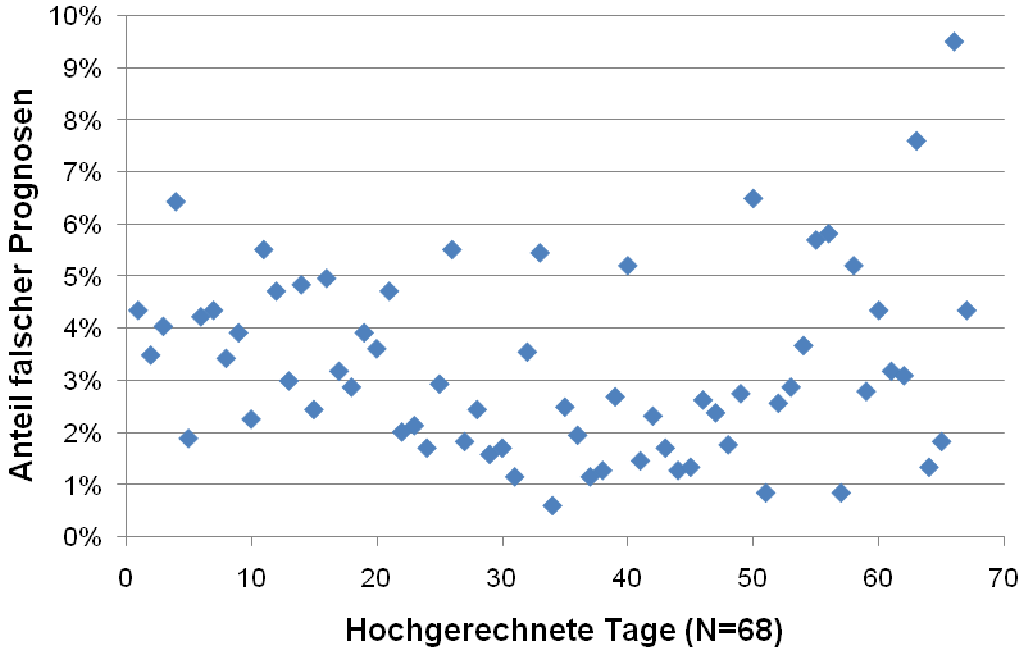


Abbildung 3-28: Anteil der nicht korrekt prognostizierten Qualitätsstufen.

Der durchschnittliche Wert der Abweichung über alle untersuchten Tage liegt bei ca. 3 %. Das bedeutet, dass durchschnittlich je Strecke seltener als einmal am Tag für eine Stunde eine Qualitätsstufe falsch vorhergesagt wird.

Abweichungen zwischen Prognose und Realität sind unter anderem mit unvorhergesehenen Ereignissen wie Unfällen zu begründen. Aufgrund der Unberechenbarkeit dieser Situationen lässt sich deren Einfluss auf den Verkehrsablauf nicht im Voraus in die Tagesprognose einbeziehen.

3.9 Monitoring der Angebotsqualität (VuV)

Die Angebotsqualität in einem Verkehrsnetz beschreibt die Qualität des Verkehrsablaufs aus der Sicht des Nutzers. Das aus Nutzersicht entscheidende Kriterium ist der Zeitaufwand, der durch unterschiedliche Kenngrößen beschrieben werden kann. Im Folgenden werden die Kenngrößen Geschwindigkeit, Verlustzeit und Fahrzeitindex verwendet, zudem werden Verkehrsleistung und Verkehrszeitaufwand angegeben:

- Geschwindigkeit: mittlere Geschwindigkeit in einem Netzabschnitt [km/h]
- Verlustzeit: Summe der Verlustzeiten [Kfzh] in einem Netzabschnitt

- Fahrzeitindex: Quotient aus der mittleren Fahrzeit und der Soll-Fahrzeit
- Verkehrsleistung: Summe der gefahrenen Strecke aller Fahrzeuge in einem Netzabschnitt [Kfzkm]
- Verkehrszeitaufwand: Summe der Fahrzeiten aller Fahrzeuge in einem Netzabschnitt [Kfzh]

Unter Monitoring ist die kontinuierliche Erfassung der Angebotsqualität zu verstehen. Das Monitoring ermöglicht eine Analyse der Angebotsqualität über einen längeren Zeitraum. Es ermöglicht Trends zu erkennen und die Wirksamkeit verkehrlicher Maßnahmen zu quantifizieren. Für das Monitoring stehen sowohl Daten aus stationären Detektoren als auch Mobilfunktrajektorien zur Verfügung.

Dabei werden die Kenngrößen mittlere Geschwindigkeit, Fahrzeitindex, mittlerer Verkehrszeitaufwand sowie Verlustzeitaufwand auf einzelnen Streckenzügen sowie im gesamten Untersuchungsgebiet verwendet. Der Fahrzeitindex ist der Quotient aus der mittleren Fahrzeit und der Soll-Fahrzeit. Für einzelne Tage, an denen Messdaten der stationären Detektoren bzw. den FPD-Daten vorliegen, lassen sich die Kenngrößen Fahrzeit und mittlere Geschwindigkeit aus beiden Datenquellen ableiten.

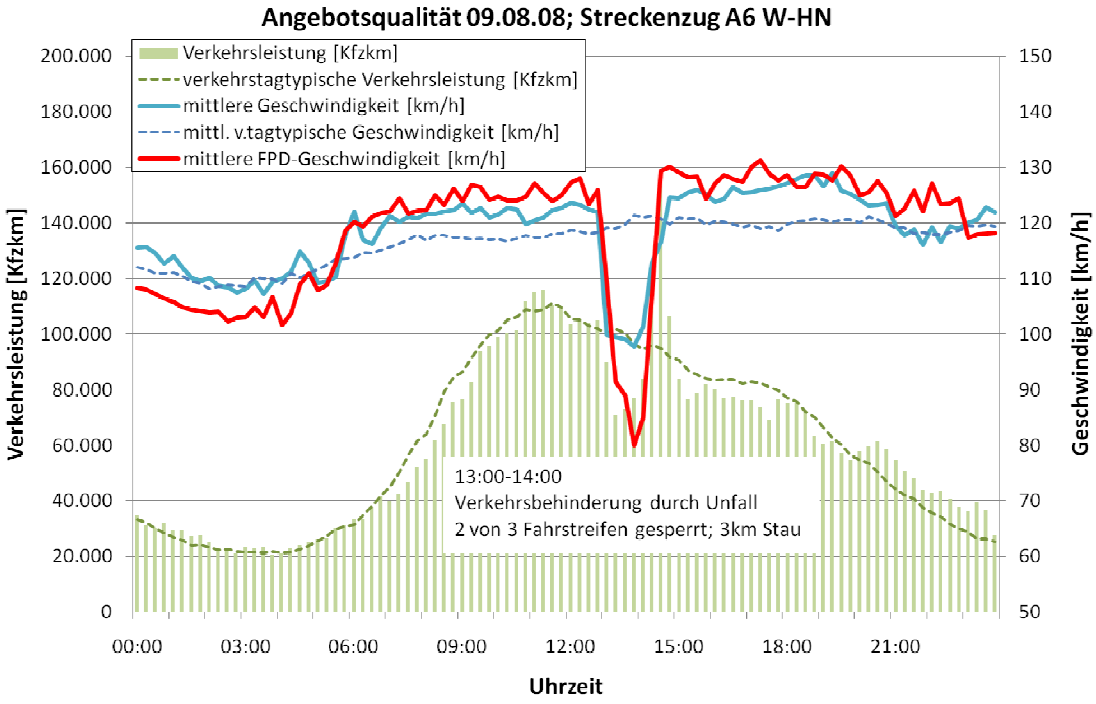


Abbildung 3-29: Tagesübersicht vom 09.08.2008 mit Verkehrsstörung durch Unfall.

In Abbildung 3-29 sind die Kenngrößen Verkehrsleistung sowie die Geschwindigkeitsverläufe für beide Datenquellen dargestellt. In diesem Tagesbericht der Angebotsqualität für die A6 von Walldorf nach Heilbronn ist deutlich zu sehen, wie die interpolierten Geschwindigkeiten der Zähldaten und die mittleren FPD-Geschwindigkeiten einen

ähnlichen Verlauf über den Tag besitzen. Der Geschwindigkeitseinbruch zwischen 13 und 14 Uhr von über 40 km/h ist sowohl bei den FPD-Geschwindigkeiten als auch bei den Zähldaten stationärer Detektoren zu erkennen, jedoch in unterschiedlich stark ausgeprägter Form. Ein Abgleich mit den Verkehrsmeldungen zeigt, dass die Störung auf einen Unfall zurückzuführen ist, der zu einer Sperrung von zwei Fahrstreifen auf der dreistreifigen Autobahn führte. Die mittleren verkehrstagtypische Geschwindigkeit repräsentiert die für diesen Tag typischerweise zu erwartende Geschwindigkeit. Der Vergleich zwischen dieser und der Ist-Geschwindigkeit zeigt, dass es sich hierbei um ein einmaliges Ereignis handelt.

Andere Streckenzüge weisen dagegen stark unterschiedliche Verläufe zwischen den Geschwindigkeiten von FPD- und Zähldaten auf. Abbildung 3-30 zeigt eine Tagesübersicht für den Streckenzug A8. An diesem Tag fanden Instandhaltungsarbeiten statt, die zu einem Stau führten. Die über den Streckenzug laufenden Mobilfunktrajektorien zeigen den aus dem Stau resultierenden Geschwindigkeitseinbruch. Die Geschwindigkeiten aus den stationären Messstellen bleiben jedoch von diesem Ereignis unbeeinflusst. Diese Tatsache lässt sich durch eine schlechte Messstellenversorgung der A8 auf dem Streckenzug der A8 erklären, auf dem es zwischen Karlsbad und Stuttgart einen 35 km langen Streckenabschnitt gibt, für den keine Messdaten vorliegen. Störungen in diesem Abschnitt können dort nicht erkannt werden, wenn sich Verkehrsstärke und Geschwindigkeit ausschließlich in diesem Abschnitt verändern und nicht auf benachbarte Strecken mit Detektoren auswirken. Bei der Erzeugung des Verkehrslageberichts wird dieser Streckenzug mit Messdaten der Detektoren bei Karlsbad und Stuttgart versorgt.

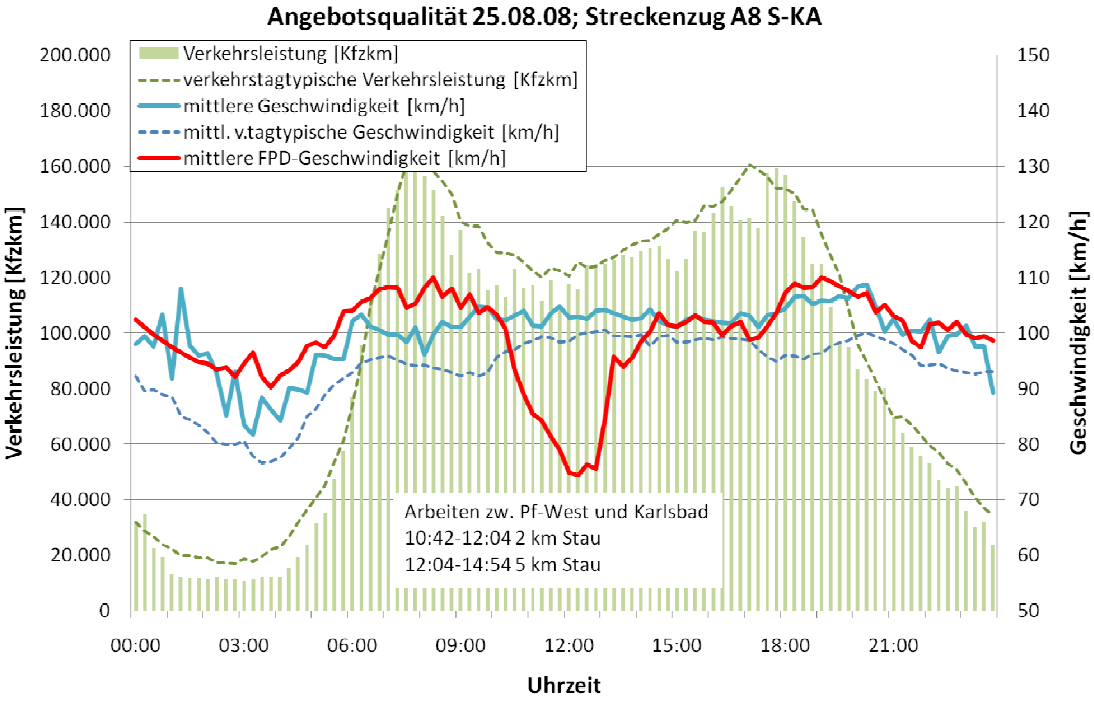


Abbildung 3-30: Tagesübersicht vom Streckenzug A8 S-KA am 25.08.2008 mit den Kenngrößen Verkehrsleistung und den Geschwindigkeiten aus FPD-Trajektorien und Zähldaten

Eine Monatsübersicht des Verkehrslagereports wird durch eine räumliche Aggregierung der Kenngrößen auf das gesamte Untersuchungsgebiet sowie eine zeitliche Aggregierung auf ganze Tage erzeugt. Am Beispiel des März 2008 in Abbildung 3-31 ist deutlich der niedrigere Fahrzeitindex an Wochenenden zu erkennen, der sich durch die geringere Verkehrsnachfrage ergibt. Auch der letzte Arbeitstag vor den Osterferien, Freitag, der 14. März ist deutlich als Tag zu erkennen, der im Vergleich zu den sonstigen Freitagen eine um etwa 8% längere mittlere Fahrzeit besitzt.

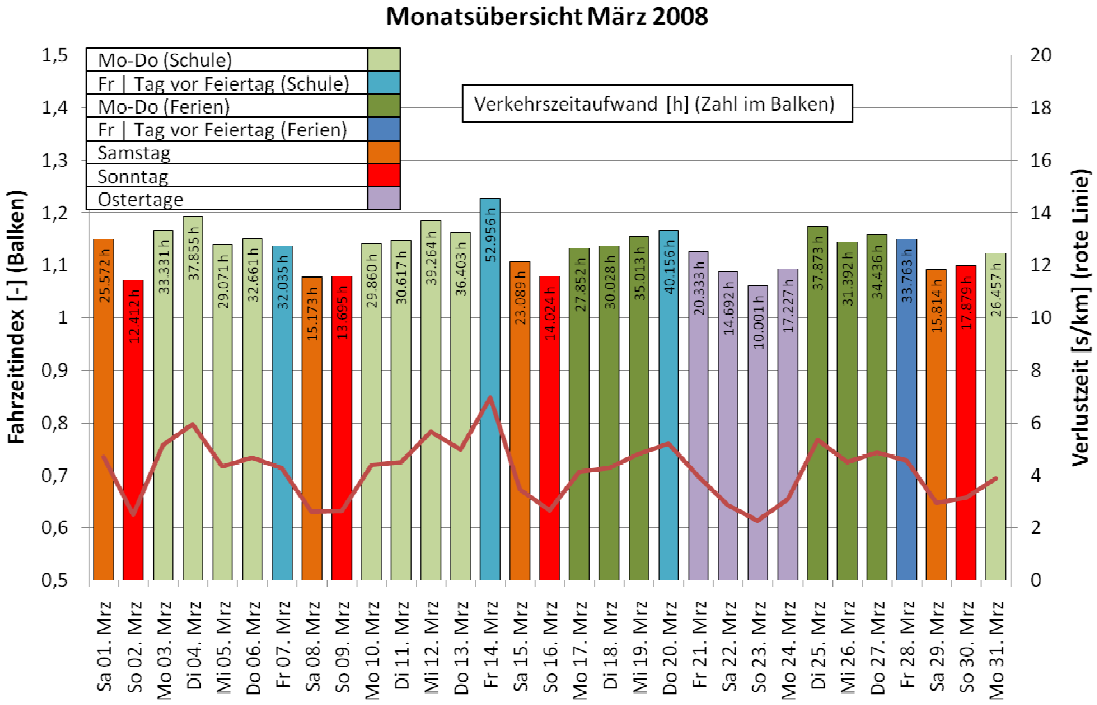


Abbildung 3-31: Übersicht der Kenngrößen des Monats März 2008 aus Messdaten stationärer Detektoren

Liegen Messdaten über einen längeren Zeitraum vor, sind Zeitreihenanalysen möglich. Durch einen Vergleich gemittelter Kenngrößen über ganze Monate hinweg, kann man z.B. eine langzeitbedingte Veränderung in der Angebotsqualität erkennen. Eine entsprechende Übersicht über den Fahrzeitindex und die und ein Jahresvergleich des Erfassungszeitraumes 2007 und 2008 ist in Abbildung 3-32 dargestellt. Tendenzen der Angebotsqualität lassen sich dadurch sowohl das gesamte Netz betreffend, als auch streckenzugfein darstellen.

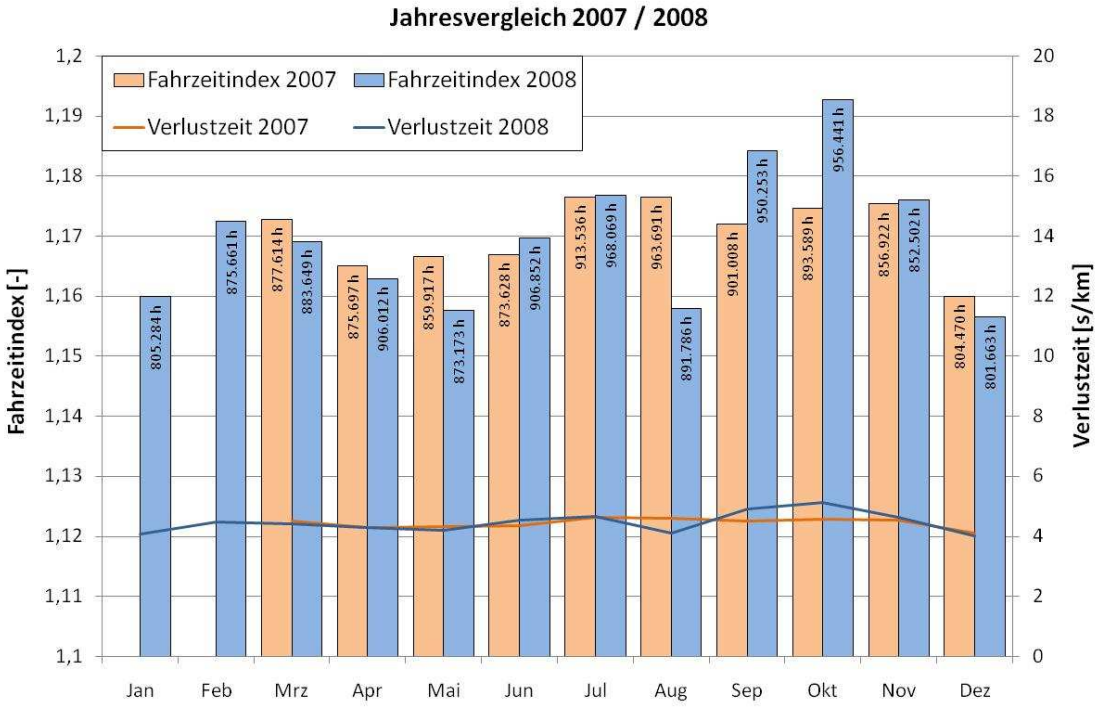


Abbildung 3-32: Vergleich der Fahrzeitindizes, Verkehrszeitaufwände und Verlustzeit der Jahre 2007 und 2008 im Untersuchungsgebiet.

3.10 Aufbau des Qualitätsmanagements (IAGB)

Der Aufbau eines Qualitätsmanagements erfolgte mit Zuarbeit der Projektpartner DDG, der Städte Karlsruhe und Stuttgart sowie des Innenministeriums und des VuV. Es wurden Qualitätsparameter definiert und Qualitätssicherungsmaßnahmen erarbeitet. Zusammen mit den Partnern Karlsruhe, Stuttgart und dem Innenministerium wurden die Prozesse der stationären Verkehrsdatenerfassung analysiert. Auf Basis dieser Analyseergebnisse wurden in Karlsruhe konkrete Qualitätssicherungsmaßnahmen entwickelt und umgesetzt sowie deren Wirksamkeit evaluiert.

Als Grundlage zur Beschreibung der Qualität der unterschiedlichen Daten im Projekt war zunächst ein Qualitätsmodell erforderlich, welches einen einheitlichen Rahmen zur individuellen Beschreibung der Qualität verschiedenartiger Daten bietet. Nach eingehender Recherche wurde der am IAGB entwickelte Qualitätsmerkmalskatalog bestehend aus 6 inhärenten Merkmalen als geeignet befunden. Auf dessen Grundlage wurden in Zusammenarbeit mit allen Projektpartnern Qualitätsparameter zur Beschreibung von

- originären Verkehrsinformationen,
- der Verkehrslage,
- Positionsfolgen aus Mobilfunkdaten sowie
- der Qualität von FPD-Trajektorien

definiert. In der folgenden Tabelle 3-5 ist das entwickelte Qualitätsmodell für FPD-Trajektorien dargestellt, wie es auch zur Evaluierung der Trajektorienqualität verwendet wurde. In der zweiten Spalte sind die 6 inhärenten Qualitätsmerkmale aus dem Qualitätsmerkmalskatalog dargestellt. Ab der dritten Spalte finden sich die definierten Qualitätsparameter.

Tabelle 3-5: Qualitätsmodell für FPD-Trajektorien (vgl. Anlage 7)

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
P1.1	Verfügbarkeit VE	Globale Ausfallrate (offline+relativ)	AUR_{global} [%]	Beschreibt die Verfügbarkeit des FPD-Servers in einem best. Zeitraum
P1.2		Trajektoriendichte (online+relativ)	ρ_{Tra} [1/s]	Summe(Routenlängen)/L/T mit L = Länge des betrachteten Streckenabschnitts und T = Dauer
P2.1	Aktualität AK	Berechnungszeit (online+relativ)	Δt_{ber} [s]	Berechnungszeit für eine FPD-Trajektorie aus der Differenz *FPD_CalculationTime – FPD_Starttime*
P2.2		Mittlere Berechnungszeit (offline+relativ)	$\overline{\Delta t}_{ber}$ [s]	Mittlere Berechnungszeit (Stunde, Tag) aus P2.1
P2.3		Alter der Trajektorie (online+relativ)	Δt_{FPD} [min]	Alter einer Trajektorie, berechnet aus der Differenz zwischen der aktuellen Zeit und dem Zeitstempel der ersten Position *aktuelle Zeit –FPD_Starttime*
P2.4		Mittleres Trajektorienalter (offline+relativ)	$\overline{\Delta t}_{FPD}$ [min]	Mittleres Alter der Daten (aus P2.3)
P3.1 Abis	Vollständigkeit VO	Trajektorienvollständigkeit (online+relativ)	VO_{FPD} [%]	**Abis:** Anteil der verwendeten Positionen einer Positionsfolge eines Teilnehmers, der für die Trajektorie verwendet wird *FPD_Quality (2)*
P3.1 A				**A:** Anteil der verwendeten Zellwechsel eines Teilnehmers der für die Trajektorie verwendet wird *FPD_Quality (2)*
P3.2		Abdeckungsrate (offline+absolut)	ABR_{FPD} [%]	Grad der Abdeckung des betrachteten Straßennetzes mit FPD (Begriff: „vollständige Abdeckung“ ist zu definieren)
P3.3		Durchdringung mit FPD (offline+absolut)	d [%]	Punktueller Anteil der mit FPD detektierten Verkehrsstärke im Vergleich zur tatsächlichen (jeweils an SES-Standorten)
P4.1	Konsistenz KO	Wird durch Einhaltung des Datenmodells gewährleistet		
P5.1	Korrektheit KR	Klassifizierungskorrektheit (online+absolut)	KR_{Kl} [%] oder [0...3]	Wahrscheinlichkeit, mit der die Teilnehmerklassen identifiziert werden (Aggregation auf einen QP "aktiver IVT" geplant) *FPD_Identification*
P5.2		Mittlere Klassifizierungskorrektheit	$\overline{KR}_{Kl}$ [%] oder [0...3]	Mittlere Klassifizierungskorrektheit der Klassen z. B. über einen Tag berechnet (aus P5.1)

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
		(offline+absolut)		
P5.3		Trajektorienlänge (online+relativ)	L_{FPD} [km]	Länge der FPD-Trajektorie (*FPD_RouteLength*)
P5.4		Mittlere Trajektorienlänge (offline+relativ)	$\overline{L}_{FPD}$ [km]	Mittlere Länge der FPD-Trajektorie (aus P5.3)
P5.5		Zuordnungskorrektheit Typ A (offline+absolut)	KR_{ZuA} [%]	Korrekter Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, bezogen auf die Länge der GPS-Route (GPS als Referenz erforderlich)
P5.6		Zuordnungskorrektheit Typ B (offline+absolut)	KR_{ZuB} [%]	Korrekter Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, bezogen auf die Länge der ermittelten FPD-Route (GPS erforderlich)
P5.7		Rangkorrelation nach Spearman (offline+absolut)	r	Punktuelle Korrelation des Tagesgangs der detektierten Verkehrsstärke mit der tatsächlichen aus SES ermittelten (jeweils an SES-Standorten)
P6.1 ABis	Genauigkeit GE	Mittlere Querabweichung (online+relativ)	QA_{FPD} [m]	Mittel der orthogonalen Abweichungen der verwendeten Positionen von der wahrscheinlichsten Route *FPD_Quality (1*) **nur Abis-Daten**
P6.1 A		Anzahl der Zellwechsel eines Teilnehmers (online+absolut)	Anz_{ZW} [-]	Anzahl der Zellwechsel, die einem Teilnehmer zugeordnet werden konnten *FPD_Quality (1)* **nur A-Daten**
P6.2		Standardabweichung der Geschwindigkeit (offline+absolut)	S_V [km/h]	Gibt die Genauigkeit der aus den Trajektorien ermittelbaren Geschwindigkeit an. Referenz erforderlich (Messfahrt oder TEU)
P6.3		Differenz der Durchschnittsgeschwindigkeiten (offline+absolut)	δ [km/h]	Vorzeichenfreie Differenz zwischen Durchschnittsgeschwindigkeit aus GPS und aus MFD

Die entwickelten Qualitätsmodelle dienten als Evaluierungsgrundlage für die FPD-Trajektorien und wurden zum Nachweis der Wirksamkeit der Qualitätssicherungsmaßnahmen herangezogen.

Als wichtiger Beitrag zur Sicherung der Datenqualität im online-Betrieb und damit zum Qualitätsmanagement wurde eine Protokollierung der Tagesergebnisse umgesetzt. Es wurden unter Anderem die Anzahl berechenbarer Trajektorien, deren Lage und Länge sowie eine Reihe weiterer wichtiger Kenngrößen stundenweise protokolliert. Diese Informationen dienen der Überwachung der Prozesse und ermöglichen das schnelle Aufdecken von Änderungen der Datenqualität. So kann das Einhalten einer geforderten Datenqualität über einen längeren Zeitraum gewährleistet werden. Weitere Opti-

mierungen in der Auswertung können so auch quantitativ gemessen und nachgewiesen werden.

Mit Hilfe der durchgeführten Analysen (Ursache-Wirkungs-Diagramme sowie FMEA mit Bestimmung von Risikoprioritätszahlen) der stationären Erfassungssysteme der Projektpartner Stuttgart, Karlsruhe und der Landesstelle für Straßentechnik konnten Schwachstellen in der Datenerzeugung- und Aufbereitung identifiziert werden. Auf Grundlage dieser identifizierten Schwachstellen wurden kurz-, mittel- und langfristige Qualitätssicherungsmaßnahmen entwickelt und exemplarisch in Karlsruhe umgesetzt. Mit Hilfe einer erneuten FMEA und dem Vergleich von Risikoprioritätszahlen sowie geeigneter Qualitätsparameter aus dem Qualitätsmodell konnte der Nachweis der Wirksamkeit einzelner Maßnahmen erbracht werden.

Die folgende Tabelle 3-6 zeigt als Beispiel die geplanten Maßnahmen in Karlsruhe zur Steigerung der Sensordatenqualität und damit zur Senkung der Risikoprioritätszahlen. In der vorletzten Spalte wird die zu erwartende Wirkung der jeweiligen Maßnahme auf die Auftretenswahrscheinlichkeit (A), die Bedeutung des möglichen Fehlers (B) sowie dessen Erkennbarkeit bzw. Behebbarkeit (E) angegeben. In der letzten Spalte ist der Zeithorizont für die Maßnahme angegeben.

Tabelle 3-6: QS-Maßnahmen mit zu erwartender Wirkung (Karlsruhe)

Nr.	System	Bauteil	QS-Maßnahme	Ziel	Wirkung (A/B/E)*	Zeit-horizont
1.1	Schleife (RMQ)	Induktions-schleife	Gute Vorplanung von Baumaßnahmen, Baustellen- und Ereignismanagement "BEM" ist derzeit in der ersten Betriebsphase	Reduzierung der Auftretens-häufigkeit durch rechtzeitiges Hinweisgeben auf Schleifen-kabel bzw. Reduzierung der Entdeckungs- und Behebungszeit	Schleife: A 6 → 5 E 6→ 5 Zuleitung: A 4 → 3 E 6→ 5	kurzfristig
1.2	Schleife (RMQ)	Muffe	Verwendung von Muffen hoher Qualität; wird seit ca. drei Jahren in Karlsruhe praktiziert in Form von gelverfüllten witterungsbeständigen Muffen ("Gelsnap").	Mittel- bis langfristige Reduzierung der Auftretenshäufigkeit zu erwarten	A 5 → 4,5 A 5 → 4	mittel-/langfristig
1.3	Schleife (RMQ)	Zuleitung	Verwendung moderner (fett-verpresster) Kabel; wird seit ca. zehn Jahren in Karlsruhe praktiziert.	Mittelfristige Reduzierung der Auftretenshäufigkeit zu erwarten	A 3→ 2,5	mittelfristig
1.4	Schleife (RMQ)	Steuergerät	Kontinuierliche Modernisierung des Gerätebestandes, insbesondere der "abgekündigten" Gerätetypen, bei denen eine Ersatzteilvorhaltung nicht mehr gewährleistet ist.	Mittel- und langfristige Reduzierung der Auftretenshäufigkeit zu erwarten	A 4→ 3 A 3→ 2	mittel-/langfristig
1.5	Schleife (RMQ)	Spannungs-versorgung	Standardmäßiger Einbau einer USV zur unabhängigen Spannungsversorgung bei Netzausfall	Mittel- und langfristige Reduzierung der Auftretenshäufigkeit zu erwarten	A 4→ 3,5 A 4→ 3	mittel-/langfristig
1.6	Schleife (RMQ)	Steuerkabel	Kabelkataster	Deutliche Reduzierung der Behebungszeit bei Defekten einzelner Adern durch Umrangieren auf Ersatzaderpaar	Bruch, Kurzschluss und sporadischer Kurzschluss: E -1 E -2*	kurz-/mittelfristig

Nr.	System	Bauteil	QS-Maßnahme	Ziel	Wirkung (A/B/E)*	Zeit-horizont
2.3	TEU	Infrarotdetek-tor	Stringente Vorgaben zur Plausibilisierungskontrolle in Concert-Software	Unplausible Daten werden vom System verworfen (Con-cert: Status grau), so dass eine deutliche Reduzierung der Erkennungszeit zu erwar-ten ist	E 7→ 6	kurzfristig
2.4	TEU	Infrarotdetek-tor	Tournusmäßige Plausibilitäts-kontrolle der Messwerte, insbesondere der Geschwin-digkeit, in Bezug zu histori-schen Daten, 1x wöchentlich. Einrichten einer Messwert-kontrolldatei	Deutliche Reduzierung der Erkennungszeit	E 9→ 7	kurzfristig
2.5	TEU	GPRS-Server	Automatischer Restart des Softwareprozesses sonntags 12:00 Uhr	Reduzierung der Behebungs-zeit bei Ausfall am Wochenen-de	E 4→ 3	kurzfristig
2.6	TEU	TEU-Gesamtsys-tem	Ausarbeitung, Abstimmung, Bekanntmachung und Vorhal-ten von Aktionsplänen	Reduzierung der Behebungs-zeit insbesondere wenn Absperrmaßnahmen zur Fehlerbeseitigung zu organi-sieren sind	E -1 für alle Teile der TEU-Außenanlage	kurzfristig
3.1	Schleife / TEU	Concert-Rechner	Automatischer Restart der Verarbeitungsprozesse sonn-tags 12:00 Uhr	Reduzierung der Behebungs-zeit bei Ausfall am Wochenen-de	E 4→ 3	kurzfristig
3.2	Schleife / TEU	Gesamtsys-tem	Tournusmäßige visuelle Kontrolle des Gesamtsystems auf Funktionstüchtigkeit aller Messwerte 2x täglich in Ab-hängigkeit von den vereinbar-ten Reaktionszeiten der Servicekräfte	Reduzierung der Erkennungs-und Behebungszeiten	E -0,5 für alle ausfallrelevanten Störungen, die nicht anderweitig schneller erkannt werden (z.B. Ausfall LSA-Steuergerät) **	kurzfristig

* Bezeichnung entsprechend der durchgeführten System-FMEA: A=Auftrittswahrscheinlichkeit, B=Bedeutung E=Entdeckung und Behebung eines Fehlers

** Eine Entdeckungs- und Behebungszeit der Stufe 3 (entspr. ½ Tag) oder geringer wird bei Schleifen und Concert- bzw. Verkehrsrechner nicht weiter reduziert, bei TEU-Außenanlage eine Entdeckungs- und Behebungszeit der Stufe 4 (entspr. 1 Tag) oder geringer, da in diesen Fällen bereits die minimal erreichbaren Entdeckungs- und Behebungszeiten erreicht wur-den.

Die Qualität der SES-Daten konnte durch die Einführung einer Reihe von kurzfristigen Maßnahmen signifikant gesteigert werden. Die Wirkung der mittel- und langfristigen Maßnahmen dagegen wird erst in den kommenden Jahren nachweisbar sein (vgl. An-lage 6).

Parallel dazu erfolgte der Nachweis der Wirksamkeit der kurzfristig durchgeführten Maßnahmen auf Basis geeigneter Qualitätsparameter. Im Folgenden sind die in Karls-ruhe erzielten Verbesserungen anhand des Vergleichs der vor und nach der Einfüh-rung der Maßnahmen berechneten Qualitätsparameter dargestellt.

Tabelle 3-7: Qualitätsparameter der SES mit und ohne der kurzfristigen Qualitätssicherungsmaßnahmen

Merkmal	Parameter	Abk. und Einheit	Vorher (KW40/2008)	Nacher (KW06/2009)
Verfügbarkeit VE	Globale Ausfallrate	AUR_{global} [%]	93,1% (6,9%AUR)	95,8% (4,2%AUR)
Aktualität AK	Mittlere Meldeverzögerung	$\overline{\Delta t}_M$ [min]	1´13	1´15
Vollständigkeit VO	Abdeckungsrate (Gesamt: 64,3km)	ABR [%]	41,9%	41,9%
Konsistenz KO	Globale Konsistenzrate	KON_{global} [%]	99,3%	99,1%
Korrektheit KR	Korrektheitsrate der Messwerte	KOR_V [%]	n.a.*	n.a.*
Genauigkeit GE	Lagegenauigkeit	G_L [± m]	± 5m	± 5m
	Genauigkeit der Geschwindigkeit	G_V [± km/h]	± 3 km/h / 5%	± 3 km/h / 5%
	Genauigkeit der Verkehrsstärke	G_q [± KFZ/Δt]	± 10 KFZ/ 5min	± 10 KFZ/ 5min
	Genauigkeit der Belegung	G_B [± %]	n.a.*	n.a.*

n.a.*: nicht auswertbar, keine Aussage möglich

Es zeigt sich, dass nach Einführung der QS-Maßnahmen insbesondere die Verfügbarkeit der Messdaten von 93,1 % auf 95,8 % gesteigert werden konnte. Die Ausfallrate der Daten aus stationären Erfassungssystemen konnte entsprechend deutlich von 6,9 % auf 4,2 % zurückgeführt werden – das entspricht einem Rückgang in Höhe von ca. 39,1 %.

Gleichzeitig ist ein leichter Rückgang der Datenkonsistenzrate von 99,3 % auf 99,1 % zu verzeichnen, was durch die stringenteren Vorgaben zur Plausibilitätskontrolle innerhalb der Verkehrsrechnersoftware zum schnelleren Erkennen von Messwert-Implausibilitäten bedingt sein dürfte.

Die Aktualität der Messdaten ist mit einer mittleren Meldeverzögerung von im Schnitt einer Minute und dreizehn Sekunden in der Vorher-Untersuchung zu einer Minute und 15 Sekunden in der Nachher-Untersuchung annähernd gleich geblieben.

3.11 Evaluierung der FPD-Trajektorien (IAGB)

Das zugrunde liegende Evaluierungskonzept wurde unter Mitarbeit aller Projektpartner erarbeitet. In den Demonstratoren der Stadt Karlsruhe und des Innenministeriums BW wurde die Nutzung der FPD-Trajektorien in verkehrstechnischen Anwendungen erprobt. Am VuV wurde die Verwendbarkeit der Trajektorien für verkehrsplanerische Zwecke evaluiert.

Die wesentlichen Ergebnisse der Evaluierung wurden in einem gemeinsamen Bericht zusammengefasst (vgl. Anlage 7). Da die FPD-Trajektorien die Grundlage der verkehrstechnischen Anwendungen in den Demonstratoren darstellen, wurde zunächst die Qualität der FPD-Trajektorien mit Hilfe von GPS und SES-Daten als Referenz untersucht.

Zusammenfassung der Qualitätsbewertung der Abis-Trajektorien

Zur Beschreibung der Qualität der Abis-Trajektorien dienen die aufgeführten Qualitätsmerkmale Verfügbarkeit, Aktualität, Vollständigkeit, Konsistenz, Korrektheit und Genauigkeit mit ihren dort definierten Parametern.

Die Aktualität der Trajektorien wird durch die Qualitätsparameter „Alter der Trajektorie“ Δt_{FPD} [min] und „Berechnungszeit“ Δt_{ber} [s] beurteilt und aus den Messzeitpunkten und der aktuellen Zeit exakt berechnet. Die Berechnungszeit pro Trajektorie liegt bei Abis-Daten bei etwa 70 Sekunden. Dagegen ist der Parameter Δt_{FPD} [min] mit 27 bis 35 Minuten deutlich länger, da zum einen die Rohdaten nur in 15 Minuten-Blöcken übermittelt und zum anderen umfangreiche Sortieralgorithmen vorgeschaltet werden müssen. Des Weiteren ist zu berücksichtigen, dass der Parameter zwischen dem Beginn der Trajektorie und dem Abschluss der Berechnung bestimmt wird. Das spielt insbesondere bei sehr ausgedehnten Gesprächen (15-Minuten-Block-übergreifend) eine Rolle, da hier zwar nach spätestens 15 Minuten eine Berechnung angestoßen wird, die aber dann nicht mehr auf den ersten Trajektorienzeitpunkt bezogen ist.

Die Parameter Korrektheit der Routen und Genauigkeit der aus den Trajektorien bestimmten Geschwindigkeit können mit den GPS-Testfahrten als Referenz bestimmt werden und dienen zur objektiven Qualitätsbeurteilung der Trajektorien.

Wie die Untersuchungen zeigen, hängt die Korrektheit der Routen (sowohl Typ A als auch Typ B) stark von der Länge des zurückgelegten Wegs ab. Tabelle 3-8 gibt dieses auch für andere Prozentzahlen an. Dabei bezieht sich die Spalte „GPS“ auf die mit GPS bestimmte Länge der Route und stellt somit den Korrektheitsratentyp A dar. Die Spalte „MFO“ bezieht sich dagegen auf die mittels Mobilfunkortung bestimmte Länge der Trajektorie. Es handelt sich folglich um den Korrektheitsratentyp B, der als deutlich kritischer einzustufen ist.

Tabelle 3-8: Korrektheit Typ A und Typ B für verschiedene Routenlängen

Länge	Zuordnung zu mindestens *% richtig					
	90%		70%		50%	
	GPS Typ A	**MFO Typ B**	GPS Typ A	**MFO Typ B**	GPS Typ A	**MFO Typ B**
1-2 km	12%	**15%**	33%	**33%**	48%	**48%**
2-3 km	23%	**27%**	50%	**51%**	63%	**66%**
über 3 km	23%	**81%**	71%	**87%**	83%	**92%**

Bei weniger als einem Viertel der über 3 km langen Positionsfolgen können Routen bestimmt werden, die mit der GPS-Route zu über 90 % übereinstimmen. Dies bedeutet, dass nicht über die komplette Gesprächszeit zuverlässige Informationen bestimmt werden können. Die ermittelten Routen enthalten jedoch in über 80 % der Fälle weniger als 10 % falsche Informationen. Will man 70 % Korrektheit Typ B erreichen, so gelingt das für 1-2 km lange Trajektorien nur für 33 % der Trajektorien, bei 2-3 km für 51 % und für mehr als 3 km lange Trajektorien für immerhin 87 % der Trajektorien. Es sollten folglich nur Trajektorien, die länger als 3 km sind, einer Auswertung zugeführt werden. Als Kompromiss zwischen Performance, Genauigkeit und Abdeckungsrate wurden im Online-Betrieb aber Trajektorien, die länger als 2 km sind und gleichzeitig den Bedingungen „mehr als 15 Positionen“, „mindestens 60 s Gesprächsdauer“ und „berechnete Geschwindigkeit kleiner als 200 km/h“ genügen, einer Auswertung zugeführt.

Erfreulich ist jedoch in jedem Fall, dass die nur unter Nutzung von Referenzdaten ableitbare Korrektheitsrate Typ B mit dem im Online-Betrieb verfügbaren Parametern „Trajektorienvollständigkeit“ und „durchschnittliche Querabweichung“ in einen empirisch ermittelten Zusammenhang gebracht werden kann. Wenn man 70 % richtig im Sinne der Korrektheit Typ B zugeordnet haben will, so sind nur Trajektorien mit einer Vollständigkeit > 97 % und einer durchschnittlichen Querabweichung < 250 m auszuwerten.

Die aus den FPD-Trajektorien abgeleiteten Standardabweichungen weisen für 30 s Diskretisierungsintervalle Werte von rund 20 km/h auf. Die Durchschnittsgeschwindigkeit kann insbesondere bei längeren Fahrten mit ca. 8 km/h deutlich genauer bestimmt werden. Abweichungen von der tatsächlichen Durchschnittsgeschwindigkeit lassen sich mit den im Online-Betrieb verfügbaren Parametern („Trajektorienvollständigkeit > 97 %“ und „durchschnittliche Querabweichung < 250 m“) so gering wie möglich halten.

Die Untersuchungen haben deutlich gezeigt, dass die FPD-Routen am Anfang und Ende deutlich unzuverlässiger sind als in der Mitte. Um die gefundenen Routen auf den mittleren Teil zu beschränken, wird auf den ersten und letzten 30 % der Strecke eine Geschwindigkeitskontrolle durchgeführt. In diesem Bereich werden Teile abgeschnitten, wenn sie nur mit sehr hohen Geschwindigkeiten zu erreichen sind (größer 170 km/h bzw. größer als die zweifache Kantengeschwindigkeit). Hierbei handelt es sich um eine abschließende Plausibilitätskontrolle.

Die Qualität der Verkehrsstärken aus Abis-Trajektorien konnte nur im Raum Karlsruhe untersucht werden. Durch die Verwendung von SES als Referenz, war in erster Linie der Vergleich mit der Verkehrsstärke und deren Änderung möglich. Dabei handelt es sich um eine der wesentlichen Informationen, die als Grundlage für diverse verkehrstechnische Anwendungen erforderlich ist.

Zur Evaluierung standen in Karlsruhe zum einen die stadteigenen SES zur Verfügung. Die Sensoren sind insbesondere auf der verkehrstechnisch wichtigen Südtangente als Ost-West-Transversale, sowie auf wichtigen Ein- und Ausfallsstraßen installiert. Zum anderen fielen 5 Sensoren des IM sowie 6 weitere Sensoren der DDG in das mit Abis-Daten abgedeckte Gebiet. Diese Sensoren lagen hauptsächlich auf der Autobahn A5 zwischen der Anschlussstelle Ka-Durlach im Süden und Ka-Nord im Zentrum des Abis-Testgebiets. Hinzu kamen 3 Sensoren am Rand des Abis-Gebiets, einer auf der A5 kurz vor Bruchsal im Norden von Ka und zwei auf der A8 im Südosten des Testgebiets. Die folgende Abbildung 3-33 zeigt alle verwendeten SES im Gebiet Karlsruhe.

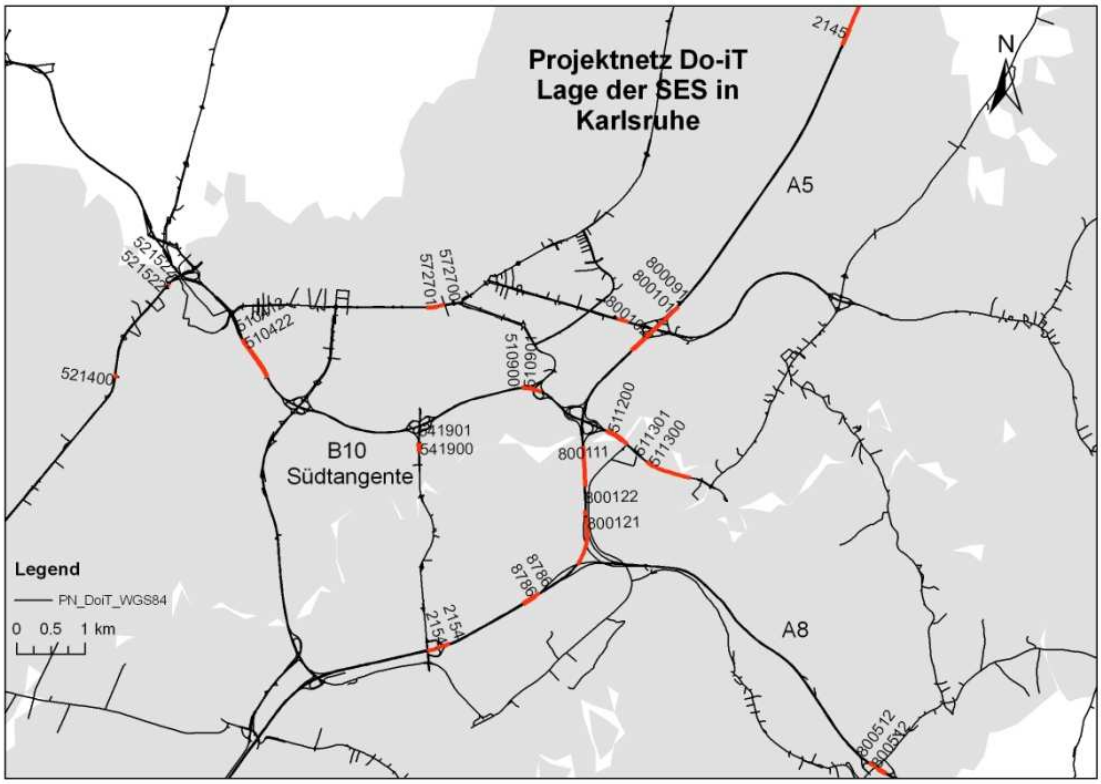


Abbildung 3-33: Alle SES im Abis-Testgebiet Karlsruhe

Die Evaluierung fand Ende März statt und es wurden insgesamt 9 Tage vom Mo. 23.03.09 bis zum Projektende am Dienstag 31.03.09 ausgewertet. In diesem Zeitraum lag die durchschnittliche Durchdringung mit FPD (Quotient aus der mit Mobilfunkdaten ermittelten Verkehrsstärke und der Zählwerte eines SES) innerhalb 24 h bei 0,4 bis 3,3 ‰ bei einer relativ starken Streuung von 0,2 bis 2,6 ‰. Die Standardabweichung innerhalb dieser 9 Tage lag bei maximal 0,8 ‰. Am Rand des Abis-Gebiets in Fahrtrichtung Karlsruhe nimmt die Durchdringung erwartungsgemäß stark ab, da in das Gebiet einfahrende Fahrzeuge erst nach einer gewissen Fahrtstrecke auswertbar sind.

Die Durchdringungen auf der Autobahn sowie auf der weitgehend kreuzungsfreien Südtangente erreichen mit knapp 2 bis 3 ‰ höhere Durchdringungen als auf dem untergeordneten Straßennetz. In der Innenstadt ist die Berechnung von brauchbaren Trajektorien sehr schwierig, es können hier lediglich 0,4 ‰ bei einer Streuung von ca. 0,6 ‰ detektiert werden.

Zur Beurteilung, in welchem Maße die Abis-Daten zur Detektion von Veränderungen der Verkehrsstärke geeignet sind, wurde die Rangkorrelation zwischen den Tagesganglinien aus Abis-Daten und aus der jeweiligen Referenzquelle berechnet. Die Korrelationen blieben neben einzelner tageweiser Ausnahmen im Mittel unter 0,5 bei einer Streuung von über 0,5, bei 6 von 25 SES lag die Streuung sogar über 1,0. Die zugehörigen Standardabweichungen bewegen sich zwischen 0,1 und 0,6.

Die Betrachtung der absoluten Anzahl berechneter Trajektorien in 24 h erklärt diese relativ schlechten Werte. Nur an 3 SES konnten im Mittel mehr als 96 Trajektorien am Tag bzw. 4 Trajektorien pro Stunde berechnet werden. Meist lag die Anzahl weit darunter. Daher führen einzelne berechnete Trajektorien zu sehr großen Streuungen und stark schwankenden Durchdringungen. Eine befriedigende Abbildung der realen Verkehrsstärkeänderungen eines Tages scheint mit dieser geringen Trajektoriendichte nur schwer möglich.

An den SES auf der Autobahn wurden trotz teilweise relativ hoher Trajektorienanzahl nur mäßige Korrelationen ermittelt. Eine mögliche Erklärung für diesen Umstand könnte die fehlende Berücksichtigung des tageszeitabhängigen Telefonierverhaltens der Verkehrteilnehmer sein. Die Auswertung der Teilnehmer hängt sehr stark von der Anzahl und Länge der Telefonate ab, die in den Fahrzeugen geführt werden. Bislang wird von einem konstanten Telefonierverhalten der Teilnehmer ausgegangen. Diese Annahme ist jedoch falsch. Dazu sind jedoch umfassende Untersuchungen zum Telefonierverhalten erforderlich. Nur so ist es möglich, diesen Effekt aus den Tagesganglinien zu eliminieren.

Des Weiteren reicht die Auflösung der Abis-Daten im Innenstadtbereich derzeit nicht aus, um in größerem Umfang korrekte Fahrzeugtrajektorien zu generieren. Es ist nicht zu erwarten, dass eine weitere Optimierung der Auswertesoftware alleine diesen technisch bedingten Mangel in den originären Mobilfunkdaten auf Abis-Basis lösen kann.

Zusammenfassung der Qualitätsbewertung der A-Trajektorien

Zur Beschreibung der Qualität der A-Trajektorien dienen die in Tabelle 3-5 aufgeführten Qualitätsmerkmale. Hier wird sich auf die Korrektheit der Trajektorien und die Genauigkeit der Geschwindigkeitsbestimmung beschränkt, da sie zur unmittelbaren Bewertung der Trajektorien dienen.

Die Aktualität der Trajektorien wird auch hier wieder durch die Qualitätsparameter „Alter der Trajektorie“ Δt_{FPD} [min] und „Berechnungszeit“ Δt_{ber} [s] beurteilt. Δt_{ber} liegt bei A-Daten bei etwa 8 Sekunden; also deutlich geringer als bei Abis-Daten, da weniger Berechnungsschritte notwendig sind. Demgegenüber ist der Parameter Δt_{FPD} [min] mit 26 bis 37 Minuten wie auf Abis-Ebene deutlich länger. Eine Verbesserung in Relation zur Abis-Ebene gelingt hier nicht, da der selbe Bereitstellungsalgorithmus in 15 Minuten-Blöcken umgesetzt wird. Außerdem gibt es mehr Trajektorien, die 15-Minuten-Block-übergreifend berechnet werden.

Die Korrektheit der Trajektorienzuordnung (Typ B) ergibt im Idealfall (durchgängiges Telefonat) einen prozentualen Wert von etwa 73 % in Bezug auf die Trajektorienanzahl. Bezieht man den prozentualen Wert auf die gefahrenen Strecken, so steigert er sich sogar auf etwa 86 %. Bei mehrfach unterbrochenen Telefonaten fällt der Wert auf 67 % beziehungsweise 80,5 %. Bleibt das Mobiltelefon passiv, so fällt die Verfügbarkeit

von 100 % auf 33 % und die Korrektheit auf 14,2 bzw. 41,9 %. Für diesen Fall ist eine zuverlässige FPD-Generierung nicht möglich.

Die Standardabweichung der Genauigkeit und die mittlere Abweichung derselben liefern ein anderes Bild. Bei durchgängigen Telefonaten werden durchschnittlich 20 km/h für die Standardabweichung und 10 km/h für die mittlere Abweichung erreicht. Das ändert sich für unterbrochene Telefonate und passive Telefone nur in geringem Umfang; nur die Verfügbarkeit wird natürlich herabgesetzt.

Mit anderen Worten kann die Korrektheit der Trajektorienzuordnung bei aktiven Mobilfunkteilnehmer zu mehr als 80 % gewährleistet werden; bei passiven Teilnehmern werden deutlich geringere Werte erreicht. Die Geschwindigkeit wird aber bei deutlich herabgesetzter Verfügbarkeit bei aktiven und passiven Telefonen fast gleichgenau geschätzt.

Die Beurteilung der Verkehrsstärken aus A-Trajektorien erfolgte ebenfalls im Zeitraum vom Montag, 23.03. bis zum Dienstag, den 31.03.2009. Aus ungeklärter Ursache standen am Samstag, den 28.03. keine A-Daten zur Verfügung, daher beschränkt sich die Evaluierung auf insgesamt 8 Tage.

Die Durchdringung mit A-Daten erreicht bis zu 15 ‰, sinkt in der Innenstadt jedoch auf bis zu 0,4 ‰ ab. Die zugehörigen Streuungen liegen außerorts bei bis zu 9 ‰ und innerorts etwa auf dem Niveau der Durchdringung. Die Standardabweichungen erreichen außerorts bis zu 3 ‰ und Innerorts 1 bis 2 ‰. Auf der Kriegstraße in unmittelbarer Nähe des Stadtzentrums in Karlsruhe liegt die Durchdringung und die Streuung bei 0,4 bis 0,6 ‰. Hier ist aufgrund der begrenzten Auflösung der A-Daten, die noch großmaschiger ist als bei den Abis-Daten, eine ausreichende Abbildung des Verkehrs derzeit nicht möglich. Auf der Südtangente sowie auf der Autobahn in Stadtnähe hingegen können durchschnittlich 6 – 20 Trajektorien pro Stunde berechnet werden.

Die Korrelationen der A-Trajektorien mit den Verkehrsstärken aus SES liegen auf der Südtangente und auf der Autobahn durchweg über 0,5 und erreichen bis zu 0,8. Die zugehörigen Streuungen liegen meist unter dem halben Wert für die Durchdringung. Die Standardabweichungen von unter 0,2 deuten auf ein homogenes Verhalten hin.

Innerorts scheint eine ausreichende Detektion von Veränderungen der Verkehrsstärke nur schwer möglich zu sein. Das zeigen die Korrelationen von unter 0,35 bei Streuungen von 0,7 und mehr. Bei Betrachtung der absoluten Anzahl berechenbarer Trajektorien erklären sich diese schlechten Werte sofort. Im Mittel konnten nur 2 oder weniger Trajektorien je Stunde berechnet werden. Die Hochrechnung der Verkehrsstärke sowie die Feststellung von signifikanten Änderungen der Verkehrsstärke wird damit nahezu unmöglich.

In Stuttgart konnten aufgrund des Ausfalls der A-Daten-Aufzeichnung am zweiten Tag der Evaluierung bis zum Rest der Projektlaufzeit keine Mobilfunktrajektorien mehr be-

rechnet werden. Die Evaluierung in Stuttgart beschränkt sich daher auf einen bzw. zwei Tage (die Durchdringung konnte für den 24.03. nach dem Ausfall um ca. 11 Uhr noch extrapoliert werden).

Die Durchdringungen, die in Stuttgart ermittelt werden konnten, bestätigen mit weniger als 2 ‰ die Werte in Karlsruhe. Lediglich auf den Ausfallstraßen B295 und B10 nördlich vom Stadtzentrum werden höhere Durchdringungen bis 6,2 ‰ erreicht. An diesen Sensorstandorten ist jedoch die Infrastrukturdichte wesentlich kleiner als bei den anderen Sensoren, die in Stuttgart zur Verfügung standen.

Die Korrelation mit der tatsächlichen Verkehrsstärke erreichen bei den Sensorstandorten außerhalb des Stadtzentrums bis zu 0,83, innerstädtisch liegen sie mit einer Ausnahme bei teilweise weit unter 0,6. Wie bereits in Karlsruhe lässt sich dieser Umstand teilweise durch die geringe Anzahl auswertbarer Teilnehmer auf Grund der geringen Auflösung der A-Daten erklären. Einige SES-Standorte z. B. in der Nähe des Pragsattels zeigen jedoch trotz hoher Straßendichte mit 2 bis 3 Trajektorien pro Stunde und Korrelationen von 0,6 bessere Werte. Vermutlich ist auf diesen Strecken ein sehr hoher Transitanteil vorhanden, der vor der Einfahrt bzw. nach der Ausfahrt in den Innenstadtbereich berechnet werden kann. Dadurch gelingt es gelegentlich, die mangelhafte Auflösung im Mittelteil durch Verbindung der beiden Anfangs- und Endstücke zu überwinden und die Gesamttrajektorie zu berechnen.

Für Teile des Autobahnnetzes standen ab dem zweiten Tag aufgrund des LAC-Ausfalls ebenfalls keine Mobilfunkdaten mehr zur Auswertung zur Verfügung. Der Ausfall betraf 44 von 136 SES und damit ca. 1/3 der Sensorstandorte. Des Weiteren wurden 11 SES des Innenministeriums identifiziert, über die an keinem der 9 Evaluierungstage FPD-Trajektorien verliefen. Da benachbarte SES hohe Durchdringungen mit A-Trajektorien aufwiesen, liegt die Ursache darin, dass die entsprechenden REID der digitalen Straßenkarte nicht im Lokalisierungsnetz vorhanden waren. Es handelt sich damit nicht um einen Verfahrensfehler. Die entsprechenden Sensoren wurden daher aus der Auswertung ausgeschlossen.

An 82 % aller evaluierbaren SES Standorte auf dem Autobahnnetz wurde eine durchschnittliche Durchdringung von mehr als 5 ‰ ermittelt. Bei der Hälfte der Sensoren liegt die Korrelation über 0,75 bei einer relativ kleinen Streuung von <= 0,4 sowie einer Standardabweichung von < 0,16. Die absolute Anzahl berechneter Fahrzeugtrajektorien erreicht im Mittel 8 bis 25 pro Stunde. Wie die im Evaluierungsbericht (vgl. Anlage 7) dargestellten Beispiele von gegenübergestellten Tagesganglinien bestätigen, ist mit diesen Werten eine Abbildung der Verkehrslage möglich. Ungewöhnliche Schwankungen in der Verkehrsstärke, die auf Stauungen hindeuten, sind innerhalb weniger Minuten detektierbar.

Die Kriterien Durchdringung, Rangkorrelation und absolute Trajektorienanzahl sind zur Beurteilung einzelner Straßenabschnitte hinsichtlich der Überwachung mit A-Daten

geeignet. Eine exakte Festlegung der Grenzwerte erscheint jedoch nur in Verbindung mit der Evaluierung weiterer Anwendungen sinnvoll.

3.12 Evaluierung der verkehrstechnischen Anwendungen (DDG, Ka)

Verkehrslageerfassung auf Basis von A-Daten (DDG)

Consider a road segment, connection two consecutive intersections like entrances and exists on motorways or federal highways. If the complete enumeration (population) of the individual travel times or journey speeds experienced by all vehicles, which exit the road segment during a fixed averaging time interval Δt (i.e. 5 minutes) is denoted by $(x_i)_{1\le i\le N}$ the population mean is $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$ and its standard deviation is $\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)^2}$. The individual travel times or journey speeds measured by all mobiles, which exit the road segment during the time interval Δt and which provide floating phone data, is regard as a random sample $(x_{i_j})_{1\le j\le n}$ which is drawn from the underlying population $(x_i)_{1\le i\le N}$. The sample mean is $\overline{X} = \frac{1}{n}\sum_{j=1}^{n} x_{i_j}$ and its standard deviation is $S = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(x_{i_j}-\overline{X}\right)^2}$.

As is known from inferential statistics an (asymptotic) interval estimate (i.e. confidence interval) of the population mean with confidence level $1-\alpha$ is

$$\Pr\left[\overline{X} - \frac{u_{1-\alpha/2}\cdot\sigma}{\sqrt{n}}\cdot\sqrt{\frac{N-n}{N}} \le \mu \le \overline{X} + \frac{u_{1-\alpha/2}\cdot\sigma}{\sqrt{n}}\cdot\sqrt{\frac{N-n}{N}}\right] = 1-\alpha\,.$$

Values for the $(1-\alpha/2)$-th percentiles $u_{1-\alpha/2}$ of the standard normal distribution, which correspond to frequently used values for the confidence level $1-\alpha$, are given in the next table.

α	confidence level	$u_{1-\alpha/2}$
0.1	90 %	1.64
0.05	95 %	1.96
0.01	99 %	2.57

Consequently, an interval estimate for the relative error $\delta = \left| \frac{\overline{X} - \mu}{\mu} \right|$ of the estimate $\overline{X}$ is $\Pr\left[\delta \le \frac{u_{1-\alpha/2}}{\sqrt{n}} \cdot C_v \cdot \sqrt{\frac{N-n}{N}} \right] = 1-\alpha$. The coefficient of variation $C_v = \frac{\sigma}{\mu}$ is the ratio of the standard deviation σ to the mean μ of the population.

The formula gives the relationship $\delta(\alpha, N, n, C_v)$ between

- the statistical accuracy of the segment travel time or journey speed estimate (maximal relative error δ),
- the statistical significance of the accuracy statement (confidence level $1-\alpha$) ,
- the traffic flow conditions, i.e. vehicles per hour, which determines in combination with the averaging time interval Δt the absolute number N of vehicles traversing the road segment during the time period under consideration,
- the amount of travel time or journey speed measurements available from mobiles, which provide floating phone data (sample size n),
- the travel time variability (coefficient of variation C_v).

The amount of data (sample size), needed to guarantee, that the maximal estimation error δ is below a maximal threshold ε, is at least $n \ge \frac{N}{1 + N \cdot \left(\frac{\varepsilon}{u_{1-\alpha/2} \cdot C_v} \right)^2}$ (minimal sample size criterion).

In the following table the minimal sample size corresponding to different values of the coefficient of variation and characteristic traffic flow conditions but fixed values for the confidence level (90 %), the averaging time interval Δt (1min) and the maximal estimation error (10 %) are given.

coefficient of variation C_v [-]	**traffic flow q [veh/h]**	**population size $N = q \cdot \Delta t$ [veh]**	**required sample size n [mobiles]**
0.05	300	5	1
0.05	1200	20	1
0.05	1800	30	1
0.1	300	5	2
0.1	1200	20	3
0.1	1800	30	3
0.2	300	5	4
0.2	1200	20	7
0.2	1800	30	8

The values given in the last table could be compared to the statistics of real floating phone travel time data as shown in the next figures. Floating phone travel time data which relates to two consecutive and heavily travelled road sections on the German motorway (Autobahn) A8 (cf. figure 1) on 4th of December 2008 including travel time measurements based on location area updates and cell-handover at A-data-level is shown in figures 2 and 3.

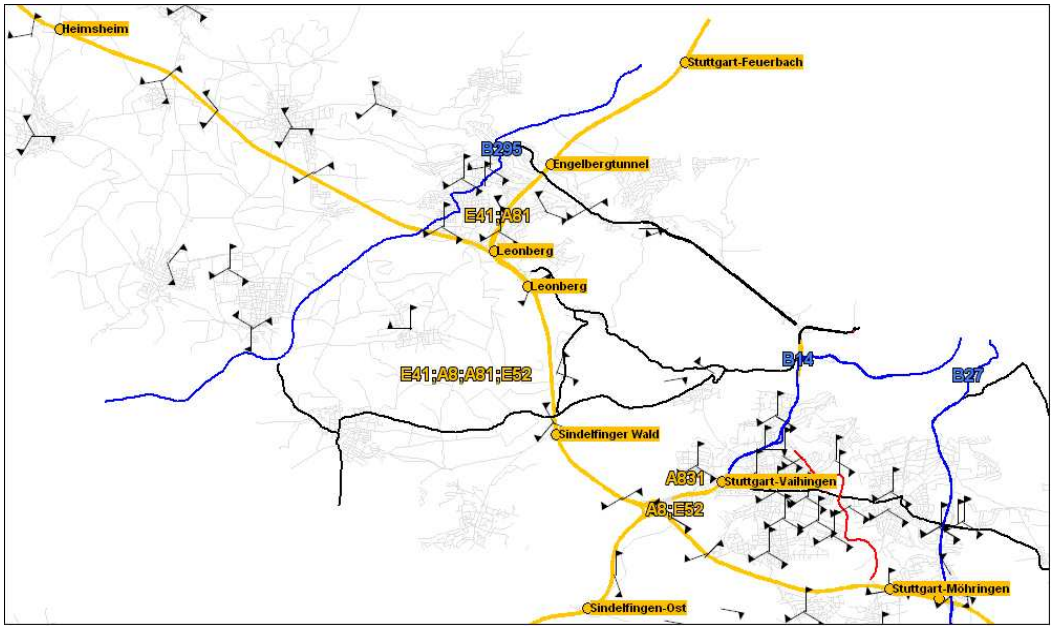


Figure 1: Road Sections on German motorway A8 between TMC locations Stuttgart-Möhringen (52a) and Heimsheim (46).

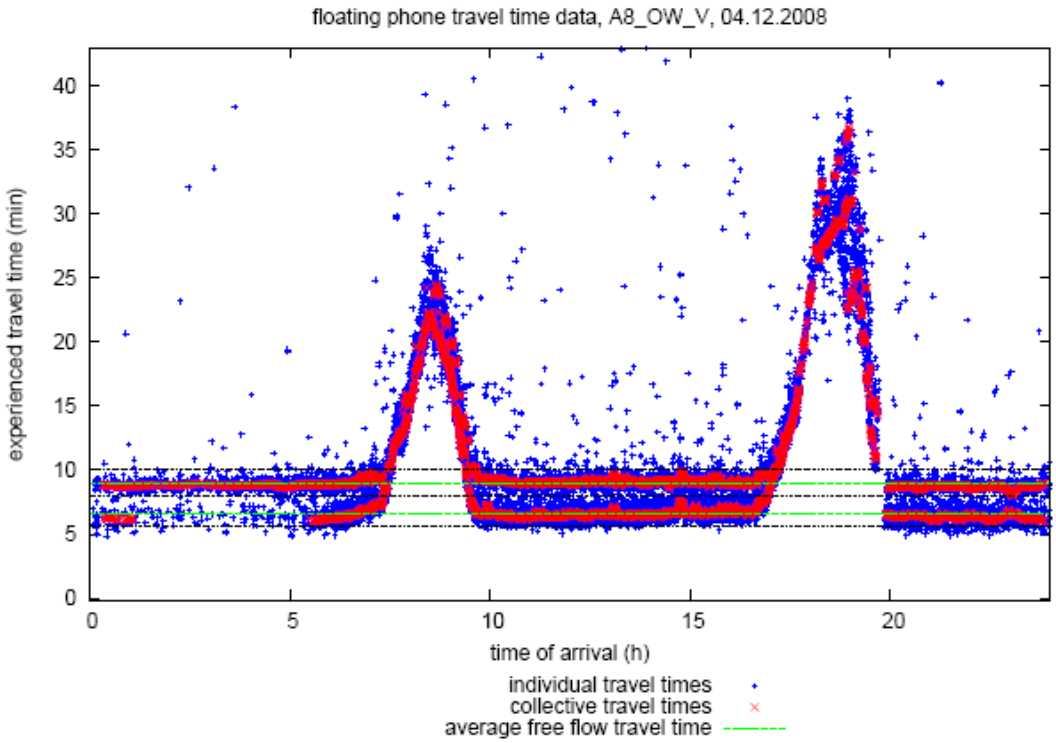


Figure 2: Floating phone travel time data versus time of day for the road section on the German motorway A8 between the TMC locations Stuttgart-Möhringen (52a) and Leonberg (50) on Thursday the 4th of December 2008. The length of this road section is approximately 13 km.

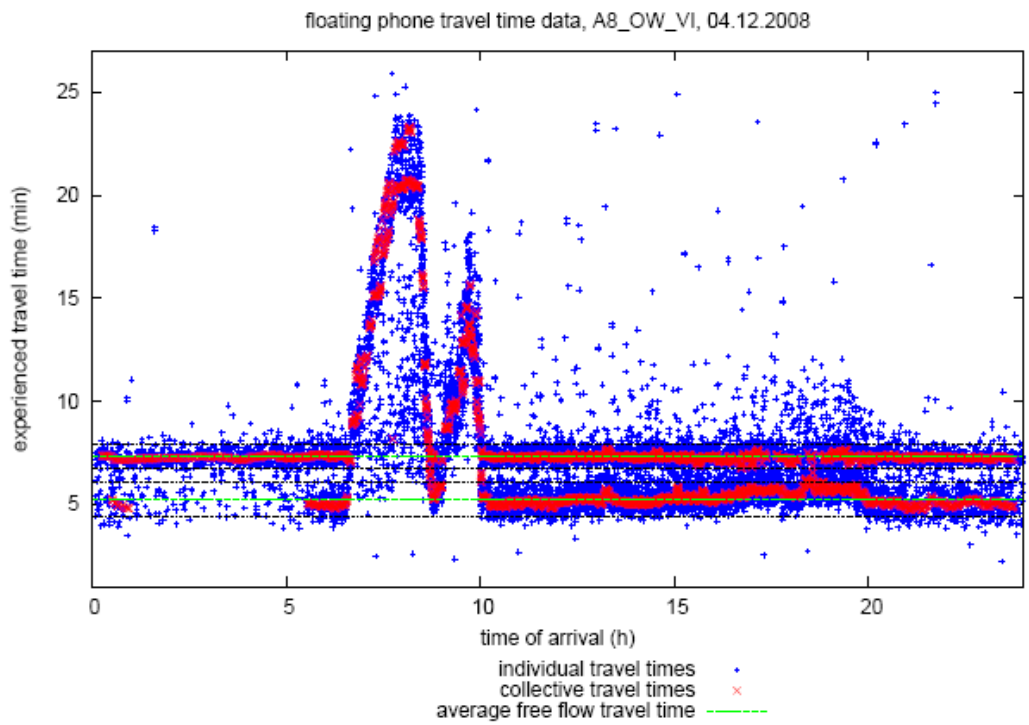


Figure 3: Floating phone travel time data versus time of day for the road section on the German motorway A8 between the TMC locations Leonberg (50) and Heimsheim (46) on Thursday the 4th of December 2008. The length of this road section is approximately 10 km.

On German Autobahns trucks are only allowed to drive 80 km/h, which leads to separate travel times for cars and trucks and yields in general a bimodal structure of the travel time distribution (cf. figure 4).

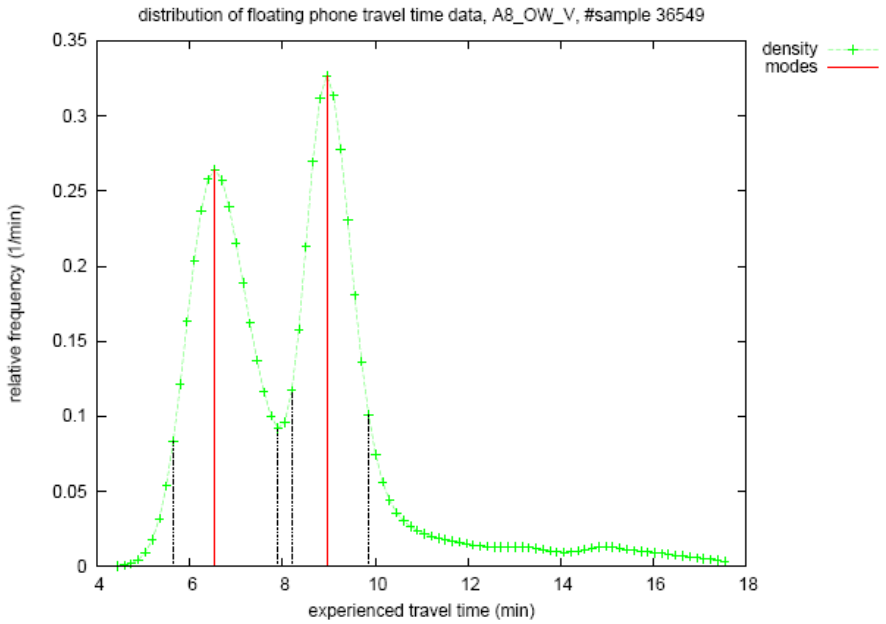


Figure 4: Distribution of floating phone travel time data for the road section on the German motorway A8 between the TMC locations Stuttgart-Möhringen (52a) and Leonberg (50) on Thursday the 4th of December 2008.

As is shown in figures 2 and 3 it is possible to monitor change of travel time on different vehicle types evolving with time of day. During the night hours the percentage of trucks is significantly higher than the percentage of cars and consequently a separate travel time cluster for cars is missing often. A traffic disturbance affecting the road sections under consideration leads to large but uniform time losses during the morning (figure 2 and 3) and afternoon rush hour (figure 3 only). Once the traffic disturbance has occurred the travel time clusters for cars and trucks merge into each other and after the return to undisturbed traffic two separate travel time clusters emerge. There will be different traffic streams, i.e. on different roads or on different modes, which have to be distinguished.

The amount of travel time or journey speed measurements available from mobiles, which provide floating phone data, is shown in figure 5. During the morning and afternoon rush hour the rate of travel time measurements based on floating phone location area updates is of magnitude 200 measurements per 5 minutes whereas the rate of travel time measurements based on floating phone cell-handover events is one order of magnitude below this value.

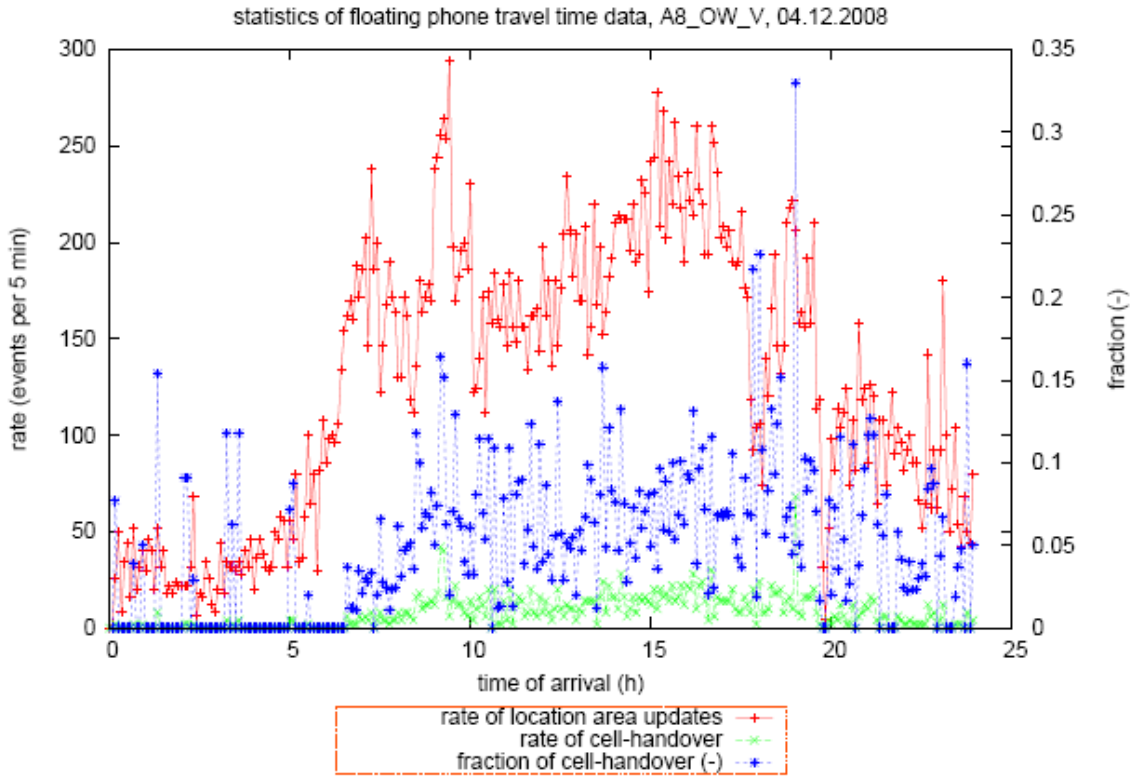


Figure 5: Statistics of floating phone travel time data versus time of day for the road section on the German motorway A8 between the TMC locations Stuttgart-Möhringen (52a) and Leonberg (50) on Thursday the 4th of December 2008.

For the road section under consideration the amount of travel time measurements, available from floating phone *location area updates*, is sufficient to estimate the mean

travel time of *all* vehicles traversing the section during each of the 5 minute intervals with less than 2 % estimation error assuming that the confidence level is 90 % and the coefficient of variation is 10 % (cf. figure 6). This means, that travel time measurements based on location area updates, sample the underlying travel time population more or less completely.

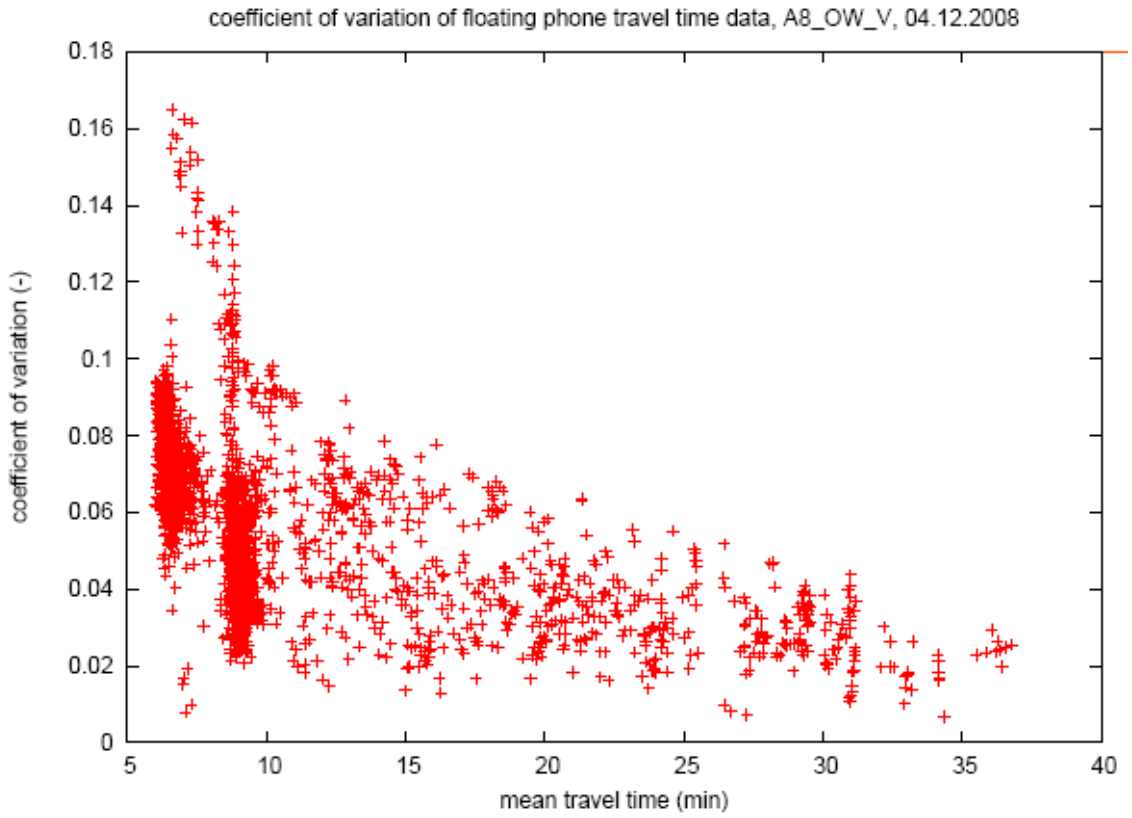


Figure 6: Coefficient of Variation with respect to the principal component of travel time data for the road section on the German motorway A8 between the TMC locations Stuttgart-Möhringen (52a) and Leonberg (50) on Thursday the 4th of December 2008.

In general, it is necessary to incorporate floating phone cell-data (at least at A-level) into the travel time estimation process in order to achieve travel time or journey speed estimate with spatial resolution in accordance with the length of road sections, defined by consecutive intersections on motorways or federal highways. As is shown in figure 7 this will lead to an increase of estimation error. During the morning and afternoon rush hour, when there is need to estimate travel time with high spatial resolution in order to decide, which road sections maybe suffer from traffic disturbances and consequently time losses, the estimation error is of magnitude 5 %. Whereas at night, an estimate based on cell-data alone, if often not possible at all.

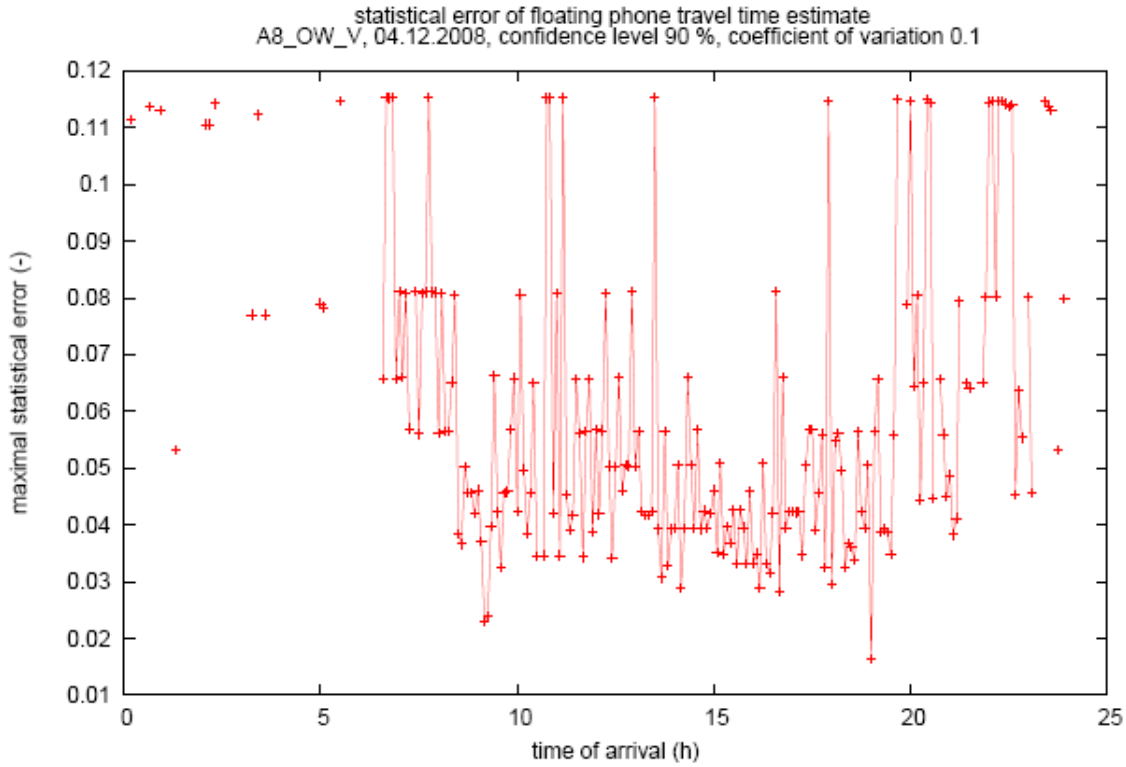


Figure 7: Statistical error of travel time estimate based of floating phone cell-data (at A-level) versus time of day for the road section on the German motorway A8 between the TMC locations Stuttgart-Möhringen (52a) and Leonberg (50) on Thursday the 4th of December 2008.

According to the method described above the estimation error δ is depending on the coefficient of variation C_v. From theoretical considerations it follows, that the coefficient of variation is inverse proportional to the length of the considered road section and proportional to the square root of the number of vehicles traversing the road section during the averaging time interval. By taking account of this, it is possible to make a statement concerning the accuracy of floating phone based travel time data for different road sections and traffic conditions.

Demonstrator LOS auf Abis-Daten in Karlsruhe (Ka)

Die Pilotanwendung von FPD im städtischen Netz Karlsruhe sollte in erster Linie Aufschlüsse über die Qualität und Verwertbarkeit der Verkehrsinformationen in Form von Verkehrslage- bzw. –zustandsberichten liefern. Es sollte evaluiert werden, in wie weit die Informationen aus den anfallenden Mobilfunkdaten geeignet sind, auf innerstädtischer Ebene als Informationsquelle für die Verkehrsteilnehmer dienen zu können.

Es konnte der Nachweis erbracht werden, dass eine vollautomatisierte Übernahme von Floating-Phone-Daten in ein städtisches Verkehrsmanagementsystem möglich ist und

insofern diese Daten grundsätzlich zur Auslösung bestimmter verkehrsbeeinflussender Maßnahmen herangezogen werden können.

Die Ergebnisse der Evaluierung basieren ausschließlich auf Abis-Daten des Mobilfunknetzes, für die bereits in vorangegangenen Arbeitsschritten auf die Verfügbarkeitsproblematik hingewiesen wurde. Ein durchschnittlicher innerstädtischer Durchdringungsgrad von ca. 0,4 bis 1,0 ‰ aller Fahrzeuge führt zu einer zeitlichen Verfügbarkeit aktueller Verkehrslageinformationen von projektnetzweit ca. 12,8%. Als ausschließliche Datenbasis wird diese Verfügbarkeit nicht für eine zuverlässige Verkehrslageschätzung ausreichend sein können.

Die Datenqualität konnte ausschließlich auf der Karlsruher Südtangente evaluiert werden, einer vierstreifigen, teilweise angebauten Hauptverkehrsstraße mit einer Verkehrsbelastung von ca. 80.000 Kfz/Tag und einer Verkehrslageverfügbarkeit von knapp 20%. Andere innerörtliche Hauptverkehrsstraßen wiesen zu wenige Verkehrslageinformationen auf, als dass hier eine sinnvolle Evaluierung der Datenqualität möglich gewesen wäre.

Tabelle 3-9: Verfügbarkeit Abis-Daten im städtischen Netz Karlsruhe

Verfügbarkeit	**0-24h**	**6-22h**
Autobahnen	39,2%	45,8%
Südtangente	18,5%	21,8%
Innerstädtische Hauptverkehrsstraßen	2,6%	3,2%
Raum-Zeitliche Verfügbarkeit (Abis-Netz)	12,8%	15,1%

Es hat sich erwiesen, dass die Verkehrslageinformationen aus FPD in ca. 88% aller Fälle mit den jeweiligen Referenzdaten aus stationären Erfassungssystemen übereingestimmt haben.

Insgesamt wurden in einem neuntägigen Auswertezeitraum, der detailliert untersucht wurde, 28 verschiedene Verkehrsbehinderungen unterschiedlicher Ausdehnung durch die Referenzmessstellen auf der Südtangente detektiert. Davon konnten sieben durch die Floating-Phone Daten auf Abis-Ebene ebenfalls erkannt werden. In 17 Fällen waren keine FP-Daten verfügbar und in vier Fällen wurden die Störungen nicht detektiert und fälschlicherweise ein freier Verkehrsfluss angezeigt. Es ergaben sich die folgenden Qualitätsparameter der aus FPD erzeugten Verkehrslage:

Tabelle 3-10: Qualitätsparameter der Verkehrslage (Abis-Daten)

Nr.	Merkmal	Parameter	Wert
P1.1	Verfügbarkeit VE	Zeitliche Verfügbarkeit	s.o.
P1.2		Raum-Zeitliche Verfügbarkeit	12,8%
P2.1	Aktualität AK	Anmeldeverzögerung	30,9 Min. +/- 26,6 Min.
P2.2		Abmeldeverzögerung	-8,4 Min+/- 26,1 Min.
P3.1	Vollständigkeit VO	Detektionsrate	25%
P5.1	Korrektheit KR	Falschalarmrate	91,3 %
P5.2		Korrekter Positionierungsanteil	42,9 %

Die Detektion von Verkehrsbehinderungen kann als grundsätzlich nicht so unzuverlässig eingestuft werden, wie die ermittelte Detektionsrate von 25% dies suggeriert, da alle größeren Verkehrsstörungen detektiert werden konnten und nur die zeitlich und lokal sehr begrenzten Behinderungen in den meisten Fällen nicht detektiert wurden. Weil in über 80% der nicht detektierten Behinderungen die Verkehrslage nach FPD „unbekannt“ war, beruht letztlich die niedrige Detektionsrate größtenteils auf den niedrigen Verfügbarkeitswerten. Das gleiche gilt auch für die gegenüber den Referenzsystemen ermittelten teilweise recht hohen An- bzw. Abmeldeverzögerungen.

Wenn es daher gelänge, die Verfügbarkeit der Abis-Daten noch weiter zu erhöhen, könnte diese Datenquelle tatsächlich einen wichtigen Beitrag zu einer Verkehrslageschätzung auf hochbelasteten innerstädtischen Hauptverkehrsstraßen leisten. Enormes zusätzliches Potenzial bietet auch die Integration der aus A-Daten gewonnenen Verkehrslage, da auf dieser Ebene von einer wesentlich höheren, wenn auch weniger detaillierten Erfassung ausgegangen werden kann.

4 Voraussichtlicher Nutzen, Verwertungspläne der Partner

4.1 Verwertungsplan DDG

Die heutigen DDG-Produkte sind verschiedene Formen von Verkehrsinformation, die sich fast ausschließlich auf bundesdeutsche Autobahnen beziehen. Aus dem Markt ist jedoch auch ein Wunsch nach Informationen für das nachgeordnete Netz sowie für Nachbarländer spürbar. Diesem Wunsch kann nicht entsprochen werden, da es an einer wirtschaftlich tragfähigen Datenquelle für diese Netze mangelt, welche eine Informationsqualität zu erreichen gestattet, die mit der heute bereits auf Bundesautobahnen mit konventionellen Datenquellen erreichten vergleichbar ist. Die konventionellen Datenquellen (stationäre Datenerfassung, Floating Car Data) selbst eignen sich dafür nicht, da die Refinanzierung ihres Aufbaus bereits für Bundesautobahnen durch Datenverkauf bis heute nicht gelungen ist.

„Floating Phone Data“ eröffnen die Aussicht auf eine wirtschaftlich tragfähige, flächendeckende, international einsetzbare Quelle für hochwertige Verkehrsdaten. DDG beabsichtigt, diese Technologie europaweit zum Einsatz zu bringen und so zur Grundlage für alle Produkte von DDG-Kunden und deren durch „Floating Phone Data“ erst mögliche Erweiterungen zu machen, die aktuelle Verkehrsinformationen enthalten. Derzeit werden diesbezüglich intensive Gespräche mit mehreren Mobilfunkbetreibern geführt.

4.2 Verwertungsplan IM

Die beteiligte Straßenbauverwaltung für Außerortsstraßen erwartet eine Verwendbarkeit der von der DDG künftig angebotenen Quelle-Ziel-Daten für den Betrieb ihrer eigenen Anlagen bzw. für angedachte Verkehrsinformationsdienste. Dazu werden die Daten gegen Entgelt zu erwerben sein.

4.3 Verwertungsplan Stadt Stuttgart

Die Landeshauptstadt Stuttgart beabsichtigt die Verknüpfung der FPD-Daten mit den für die Verkehrslagedarstellung genutzten Daten (Level of Service), um, durch Anreicherung dieser Daten mit den Reisezeiten, eine verbesserte und gesicherte Kenntnis über die Verkehrslage zu erhalten. Weiterhin ermöglichen die FPD-Daten die Verkehrserfassung auch in Netzbereichen die bisher nicht mit Erfassungseinrichtungen ausgestattet sind, ohne diese mit aufwendigen Erfassungseinrichtungen auszustatten. Die Nutzung der Reisezeiten im Rahmen der Verkehrssteuerungsstrategie an Knotenpunkten mittels Signalprogrammauswahl ermöglicht einen verbesserten Verkehrsfluss an diesen Knotenpunkten und einen entsprechenden messbaren Nachweis.

4.4 Verwertungsplan Stadt Karlsruhe

Die Stadt Karlsruhe erwartet eine Verwertbarkeit der von der DDG künftig angebotenen Daten für den Betrieb der Verkehrsinformationszentrale Karlsruhe.

Sie sollen einen wichtigen Beitrag zur angestrebten flächenhaften und aktuellen Abbildung der Verkehrslage (Level of Service) sowie zur Prognose auf den Hauptverkehrsstraßen des Stadtgebietes (Region) leisten, und somit eine weitere Datengrundlage für ein aktives Verkehrsmanagement bilden.

In einer angedachten späteren Ausbaustufe kann das oben beschriebene Szenario auf die grenzüberschreitende PAMINA-Region übertragen werden.

Hierbei ist insbesondere an den Aufbau einer Verkehrsbeeinflussungsanlage (Streckenbeeinflussungsanlage) im Bereich des häufig überlasteten Streckenabschnittes der B 10 zwischen dem Wörther Kreuz (A 65/B 9) und Karlsruhe-Mühlburg gedacht.

Des Weiteren kann der Gedanke der Umstiegshilfen MIV/ÖV auf das Gebiet des gesamten Karlsruher Verkehrsbundes übertragen werden.

Hierzu sollen die Nutzungsrechte an den Datenbeständen durch entsprechende Vereinbarungen mit der DDG analog zum MOBIN-Projekt erworben bzw. gesichert werden.

4.5 Verwertungsplan IAGB

Es wurden neue Algorithmen für die folgenden Bereiche entwickelt und umgesetzt:

- Ortungsverfahren mittels wenig präzise Mobilfunkortungsergebnisse,
- Map Aiding für nicht-kontinuierliche und wenig präzise Ortungsinformation,
- Identifikation verschiedener Verkehrsteilnehmerklassen anhand ihres Bewegungsverhaltens aus den genannten Ortungs- und Map-Aiding-Ergebnissen sowie verschiedener digitaler Karteninformationen.

Im Rahmen des Projekts Do-iT hat das IAGB eine Kopie der FPD-Server mit allen Algorithmen offline am Institut betrieben. Diese Server werden seit Beendigung des Projekts für die nicht-kommerzielle Nutzung in Lehre und Forschung eingesetzt und stellen die Basis für weitere Forschungsanträge dar.

4.6 Verwertungsplan VuV

Als universitäre Forschungseinrichtung hat der Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik keine wirtschaftlichen Verwertungspläne. Aus wissenschaftlicher Sicht

lassen sich die Projektergebnisse sowohl in allgemeiner Form im Bereich der Verkehrsmodelltechnik als auch speziell für das Untersuchungsgebiet im Bereich der konkreten Planung und Steuerung verwerten:

- Bereitstellung der Routenwahlparameter für die Übernahme in Verkehrsplanungsmodelle durch Dritte.
- Bereitstellung einer Methodenbeschreibung für eine kontinuierliche Erfassung des Verkehrsverhaltens und der Verkehrsqualität in Straßennetzen zur Übernahme von Dritten.
- Bereitstellung der Kenngrößen der Angebotsqualität und von Quelle-Ziel-Matrizen des Untersuchungsgebietes für Planungszwecke.

5 Veröffentlichungen von Projektergebnissen

5.1 Partnerübergreifende Veröffentlichung

Do-iT (2008): Nutzung von Mobilfunkdaten für das Verkehrsmanagement - das Projekt Do-iT, *Beilage Nr. 5 zur Straßenverkehrstechnik*, Heft 4/2008, herausgegeben vom Projektträger Mobilität und Verkehr, Bauen und Wohnen (TÜV Rheinland Group), April 2008.

5.2 Stadt Karlsruhe

Schreiber, W., Knapp, S., Schwartz, S. (2009): Verkehrsmanagement in einer mittleren Großstadt, *Straßenverkehrstechnik*, Heft 1/2009, S. 13-18, Kirschbaum Verlag, Köln.

5.3 Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen

Beetz, A.; Czommer, R.; Schwieger, V. (2009): Signalstärken-Matching und Map-Aiding Methoden zur Positionsbestimmung von Mobilfunkteilnehmern in Echtzeit. *Beiträge* zum POSNAV 2009, Dresden, 27.-28.10.2009.

Laufer, R.; Beetz, A.; Czommer, R.; Schwieger, V. (2009): Evaluierung der Qualität von Verkehrsinformationen aus Mobilfunkdaten. *Beiträge* zum POSNAV 2009, Dresden, 27.-28.10.2009.

Laufer, R.; Beetz, A.; Czommer, R.; Schwieger, V. (2009): Managing data quality – exemplified für traffic data generated from mobile phone data. *Proceedings on 16th Word Congress and Exhibition on Intelligent Transport Systems and Services,* Stockholm 21.-25.09.2009.

Laufer, R. (2008): Usage of a quality management concept for data exemplified with stationary traffic data acquisition systems. *Proceedings on 4th International Symposium "Networks for Mobility",* Stuttgart, 25.-26.09.2008.

Schwieger, V. (2008): Generierung von Verkehrsdaten auf Basis verfügbarer Mobilfunkdaten innerhalb der GSM-Nerzinfrastruktur. *Beiträge zur Fachkonferenz „Verkehrsmanagement und Verkehrstechnologien“ des BMWi*, Halle, 20.- 21. Mai 2008.

Ramm, K., Schwieger, V. (2008): Mobile Positioning for Traffic State Acquisition. *Journal of Location Based Services,* Taylor & Francis, London, UK, 2008.

Wiltschko, T., Schwieger, V. (2008): Floating Phone Data – Ein Ansatz zur Generierung von Verkehrsdaten aus Mobilfunkdaten. *Beiträge zur HEUREKA 08*, Stuttgart, 5.-6. März 2008.

Wiltschko T., Schwieger, V., Möhlenbrink, W. (2008): Zum Einsatz von Mobilfunkortungsverfahren zur Erfassung von Verkehrsdaten aktiver Verkehrsteilnehmer. *Beiträge zum DGON-Symposium POSNAV 2007*, Magdeburg, 6.-7. November 2007.

Schwieger, V. Ramm, K., Czommer, R., Möhlenbrink, W. (2007): Mobile Phone Positioning for Traffic State Acquisition. *Journal of Applied Geodesy*, Walter de Gruyter, Berlin – New York, Volume 1, Heft 3, 2007.

Wiltschko. T., Schwieger, V., Möhlenbrink, W. (2007): Acquisition of traffic state information by mobile phone positioning. *Proceedings on 6th European Congress on Intelligent Transport Systems*, Aalborg, Dänemark, 18.-20.06.2007.

Schollmeyer, R., Wiltschko, T. (2007): Classification of public transport vehicles using cellular mobile radio data. *Proceedings on 6th European Congress and Exhibition on Intelligent Transport Systems and Services,* Aalborg, Denmark, 18.-20.06.2007.

Ramm, K., Schwieger, V. (2007): Mobile Positioning for Traffic State Acquisition. *Proceedings on 5th International Symposium on Mobile Mapping Technology*, Padua, Italien, 28.-31.05.2007.

Ramm, K., Czommer, R., Schwieger, V. (2006): Map-based Positioning using Mobile Phones. *Proceedings on XXIII International FIG Congress*, München, 8.-13.10.2006.

Wiltschko, T., Schwieger, V, Möhlenbrink, W. (2006): Generating Floating Phone Data for Traffic Flow Optimization. *Proceedings on 3rd International Symposium "Networks for Mobility",* Stuttgart, 5.-6.10.2006.

5.4 Universität Stuttgart, Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik

Schlaich, J. (2010): Analyses of the Dynamic Route Choice Behavior with Mobile Phone Trajectories, *TRB 89th Annual Meeting Compendium of Papers*, Transportation Research Board of the National Academies, Washington, D.C. (im Druck).

Schlaich, J., Otterstätter, T., Friedrich, M. (2010): Generating Trajectories from Mobile Phone Data, *TRB 89th Annual Meeting Compendium of Papers*, Transportation Research Board of the National Academies, Washington, D.C. (im Druck).

Friedrich, M., Immisch, K., Jehlicka, P., Otterstätter, T., Schlaich, J. (2010): Generating OD matrices with Mobile Phone Trajectories, *TRB 89th Annual Meeting Compendium of Papers*, Transportation Research Board of the National Academies, Washington, D.C. (im Druck).

Schlaich, J., Friedrich, M. (2009), Mobilfunkdaten für den rollenden Verkehr, *GIS-Business*, Heft 7, S. 44-47, abc Verlag, Heidelberg.

Schlaich, J. (2009), Analyse der Routenwahl mit Mobilfunkdaten, Kurzfassung des Vortrags beim Fortbildungsseminar des VSVI Baden-Württemberg am 14.05.2009, *Mitgliederzeitschrift des VSVI Baden-Württemberg*, S. 16.

Schlaich, J., Friedrich, M. (2008): Staumeldungen und Routenwahl in Autobahnnetzen – Teil 2: Routenwahl in Autobahnnetzen, *Straßenverkehrstechnik*, Heft 11, 706-712, Kirschbaum Verlag, Köln.

Friedrich, M., Jehlicka, P., Otterstätter, T., Schlaich, J. (2008): Mobile Phone Data for Telematic Applications, *Proceedings of International Multi-Conference on Engineering and Technological Innovation: IMETI 2008: International Institute of Informatics and Systemics (IIIS)*, Orlando, Florida, USA

Friedrich, M., Jehlicka, P., Otterstätter, T., Schlaich, J. (2008): Monitoring Travel Behaviour and Service Quality in Networks with Floating Phone Data, *Proceedings of the 4th International Symposium "Networks for Mobility" (FOVUS)*, Stuttgart.

Friedrich, M., Jehlicka, P., Otterstätter, T., Schlaich, J. (2008): Monitoring Travel Behaviour and Service Quality in Networks with Floating Phone Data, *Proceedings of the 18. ptv vision user group meeting*, Karlsruhe.

Friedrich, M., Jehlicka, P., Otterstätter, T., Schlaich, J. (2008): Monitoring Travel Behaviour and Service Quality in Networks with Floating Phone Data, *International Conference on Transportation decision making: issues, tools, models and case studies*, Venedig, Italien.

Schlaich, J., Friedrich, M. (2008): Staumeldungen und Routenwahl in Autobahnnetzen – Teil 1: Analyse von Staumeldungen, *Straßenverkehrstechnik*, Heft 10, S. 621-627, Kirschbaum Verlag, Köln.

Schlaich, J. (2007): Routenwahl in Autobahnnetzen, Beitrag zum Workshop „Innovationen in der Verkehrstelematik“, *Tagungsband zur Hochschultagung 2007*, Rust am See, Österreich.

Schlaich, J. (2005): Nutzungsmöglichkeiten von Floating Phone Data, in: Institut für Straßen- und Verkehrswesen (ed.), *Kurzfassungen der Vorträge zum 9. Fachkolloquium Straße und Verkehr 2005*, Stuttgart.

6 Anlagen

Anlage 1: Offenlegungsschrift DE 10 2007 013 220 A1: „Verfahren und Vorrichtungen zur Generierung von Verkehrsinformation und ihrer Lokalisierung innerhalb eines räumlichen Bereichs (DDG)

Anlage 2: Bericht zur LAC-basierten Trajektoriengenerierung aus A-Daten und zu den Arbeitspaketen 5-7 (VuV)

Anlage 3: Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren auf Abis- und A-Ebene sowie Evaluierung der Positionsbestimmung durch höherwertige Ortungsverfahren. Abschlussbericht zu AP4-6 und AP10-3 (IAGB)

Anlage 4: Methoden zur Identifikation aktiver Verkehrsteilnehmer. Abschlussbericht zum AP3-5 (IAGB)

Anlage 5: Qualitätsbewertung zur Identifikation aktiver Verkehrsteilnehmer. Abschlussbericht zum AP3-6 (IAGB)

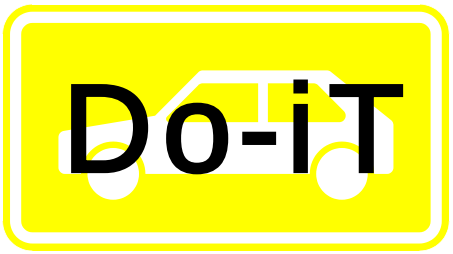
Anlage 6: Nachweis der Wirksamkeit der Qualitätssicherungsmaßnahmen. Abschlussbericht zum AP8-7 (IAGB)

Anlage 7: Evaluierungsbericht. Abschlussbericht zum AP 10-2 (IAGB)

Anlage 2 (VuV)

Datenoptimierung für integrierte Telematik

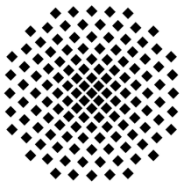
Bericht zur Trajektoriengenerierung und den Arbeitspaketen 5-7



Förderkennzeichen: 19 B 5007B

Prof. Dr.-Ing. Markus Friedrich
Dipl.-Ing. Katrin Immisch
Dipl.-Inf. Prokop Jehlicka
Dipl.-Ing. Thomas Otterstätter
Dipl.-Ing. Johannes Schlaich MBA (USQ)

Oktober 2009



Universität Stuttgart
Institut für Straßen- und Verkehrswesen
Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik

Inhalt

Kurzfassung **6**

1 Einleitung **24**

2 Eingangsdaten und Grundlagen **26**

2.1 Vereinfachtes Projektnetz (VPN) 26

2.2 Zähldatenaufbereitung 29

2.3 Binnenverkehrsproblematik 32

2.4 Routenbaumgenerator 34

2.4.1 Übersicht über bestehende Ansätze 35

2.4.2 „Vollständige“ Enumeration 36

3 Trajektoriengenerierung **40**

3.1 Beschreibung der A-Daten 41

3.2 A-Daten-Datenbereinigung 41

3.2.1 Löschen von nicht verwertbaren Datensätzen 42

3.2.2 Löschen von Mobilfunkteilnehmern ohne ausreichende Bewegung 43

3.2.3 Gruppieren aufeinander folgender identischer LACs 43

3.2.4 Verlassen des relevanten Netzes 44

3.3 Fahrtidentifikation 46

3.3.1 60min-Regel 47

3.3.2 Erweiterte 60min-Regel 51

3.3.3 Nervositätsregel 52

3.4 Generierung von LAC-Folgen aus A-Daten 53

3.5 Netz-LAC-Folgen aus Routenbaumgenerator 55

3.6 Matching mit LAC-Folgen aus Routenbaum 59

3.6.1 Direkter Stringvergleich 59

3.6.2 Berücksichtigung häufiger LAC-Wechsel („ABAB-Regel“) 60

3.6.3 Stringähnlichkeitsvergleich 60

3.7 Aufteilung der identifizierten LAC-Folgen aus VPN-Routen 71

3.7.1 Moduswahl 71

3.7.2 Anbindungswahl 73
3.7.3 Routenwahl 75
3.8 Fahrzeiten einer Route 78
3.9 Ergebnisse für das Testfeld 81
3.9.1 Ergebnisse für einen Tag 81
3.9.2 Auswertung im Längsschnitt 94
3.10 Diskussion der Einbeziehung der Cell-IDs 96
3.11 Verfahrensanpassungen für die Netzbeeinflussungsanlage Stuttgart 98
3.12 Bewertung 99

4 Schätzung der Parameter des Routenwahlverhaltens (AP5) 104
4.1 Grundlagen 105
4.1.1 Nutzenfunktion 105
4.1.2 Entscheidungsmodell (Logit-Modell) 106
4.1.3 Maximum-Likelihood-Schätzung 107
4.1.4 Verkehrsmeldungen 110
4.1.5 Wechselwegweisungen im Untersuchungsbiet 113
4.2 Routenwahlanalysen in der BAB-Netzmasche mit ANPR-Messungen 117
4.2.1 Datenaufbereitung für Routenwahlanalysen 117
4.2.2 Stichprobenbeschreibung 119
4.2.3 Untersuchung einzelner Einflussgrößen mittels linearer Regression 121
4.2.4 Entwicklung des Basismodells der Routenwahl 124
4.2.5 Einfluss der tatsächlichen Fahrzeit 130
4.2.6 Einfluss der Lage von Störungsmeldungen 133
4.3 Routenwahlanalysen in der BAB-Netzmasche mit Mobilfunktrajektorien 134
4.3.1 Datenaufbereitung für Routenwahlanalysen 135
4.3.2 Stichprobenbeschreibung 138
4.3.3 Basismodell der Routenwahl in der BAB-Netzmasche 140
4.3.4 Einfluss der Fahrzeugklasse (Lkw-Pkw) 142
4.3.5 Einfluss der untersuchten Relation 144
4.3.6 Einfluss der typischen, tageszeitabhängigen Fahrzeit 145
4.3.7 Prüfung der Aufnahmefähigkeit des Verkehrsteilnehmers 149

4.4 Routenwahlanalysen in der NBA Stuttgart mit Mobilfunktrajektorien 152
4.4.1 Stichprobenbeschreibung 152
4.4.2 Basismodell der Routenwahl in der NBA Stuttgart 154
4.4.3 Weitere Analysen der Routenwahl in der NBA Stuttgart 156
4.5 Zusammenfassung und Bewertung 157
4.5.1 BAB-Netzmasche 157
4.5.2 Netzbeeinflussungsanlage Stuttgart 161
4.5.3 Beurteilung der Wirksamkeit von dynamischen Routenempfehlungen 162
4.5.4 Beurteilung von FPD-Trajektorien zur Analyse der Routenwahl 165

5 Generierung von Quelle-Ziel-Matrizen (AP6) 166
5.1 Durchführung von Clusterverfahren 167
5.1.1 Vorklassifizierung 167
5.1.2 Abstandsfunktion 168
5.1.3 Clusterung von Netzganglinien 176
5.1.4 Ergebnisse 183
5.2 Hochrechnung von Quelle-Ziel-Matrizen 199
5.2.1 QZ-Matrizen (Kfz) aus Korrektur von QZ-Matrizen (FPD) 199
5.2.2 QZ-Matrizen (Kfz) aus Hochrechnung von Trajektorien 201
5.2.3 Langzeitergebnisse 231

6 Tagesprognose (AP7-B) 239
6.1 Ansätze für die Tagesprognose 239
6.1.1 Verkehrszustände aus historischen Ganglinien 239
6.1.2 Verkehrszustände aus einer statischen Umlegung 240
6.1.3 Verkehrszustände aus einer statischen Umlegung mit Rückstau 241
6.1.4 Verkehrszustände aus einer dynamischen Umlegung 241
6.2 Umsetzung der Tagesprognose 242
6.2.1 Verwendete Software 242
6.2.2 Widerstände bei der Routenwahl 242
6.2.3 Ermittlung der belastungsabhängigen Fahrzeit bei freier Strecke 243
6.2.4 Ermittlung der belastungsabhängigen Fahrzeit bei Baustellen 245
6.2.5 Verwaltung von Baustellen 247

6.2.6 Ableitung des LOS 248
6.2.7 Netzkalibrierung 249
6.3 Ergebnisse 249
6.3.1 Bewertung der Ergebnisse 249
6.3.2 Ergebnisse der Zustandsprognose aus historischen Ganglinien 254
6.3.3 Ergebnisse der Zustandsprognose aus statischem Umlegungsverfahren 255
6.3.4 Ergebnisse aus statischen Verfahren mit Rückstauberechnung 262
6.3.5 Ergebnisse aus dynamischen Verfahren 264
6.4 Grenzen der Tagesprognose 264

7 Monitoring der Angebotsqualität 266
7.1 Vorgehen 266
7.2 Kenngrößen der Angebotsqualität 267
7.2.1 Kenngrößen für eine Strecke 268
7.2.2 Aggregierte Kenngrößen 269
7.3 Ermittlung der Kenngrößen der Angebotsqualität aus stationären Detektoren 270
7.3.1 Verfügbare Daten 270
7.3.2 Aggregierung von Messdaten zu VPN-Streckendaten 270
7.3.3 Interpolation von nicht versorgten VPN-Strecken 272
7.4 Ermittlung der Kenngrößen der Angebotsqualität aus stationären Detektoren und Mobilfunkdaten 275
7.4.1 Verfügbare Daten 275
7.4.2 Anpassung der FPD-Geschwindigkeit 275
7.5 Verkehrslagebericht aus RMQ-Daten 278
7.5.1 Tagesbericht 278
7.5.2 Monatsbericht 280
7.5.3 Jahresbericht 282
7.6 Verkehrslagebericht aus RMQ- und Mobilfunkdaten. 284

8 Glossar 288

9 Literatur 295

Kurzfassung

Generierung von Trajektorien aus A-Daten (Kapitel 3)

Für viele verkehrliche Anwendungen ist es wichtig, dass Trajektorien über die gesamte Fahrt eines Verkehrsteilnehmers zur Verfügung stehen. Daher bietet sich ein Verfahren zur Generierung von Trajektorien an, das auf Basis der Location Area Codes (LAC) die Mobilfunkdaten auswertet, da Wechsel von einer Location Area in eine andere Location Area auch von Mobilfunkteilnehmern ohne Gespräch in den A-Daten aufgezeichnet werden (Location Area Update).

Daher wurde wird ein Verfahren zur LAC-basierten Trajektoriengenerierung entwickelt. Die folgende Abbildung 1 zeigt den Ablauf des Verfahrens, das im Wesentlichen darauf basiert, A-Daten-LAC-Folgen aus Mobilfunkdaten mit aus dem Straßen- und Schienennetz ermittelten Netz-LAC-Folgen zu vergleichen und bei Übereinstimmung den Mobilfunkteilnehmer einer bzw. bei Mehrdeutigkeit mehreren Routen im Verkehrsnetz zuzuordnen.

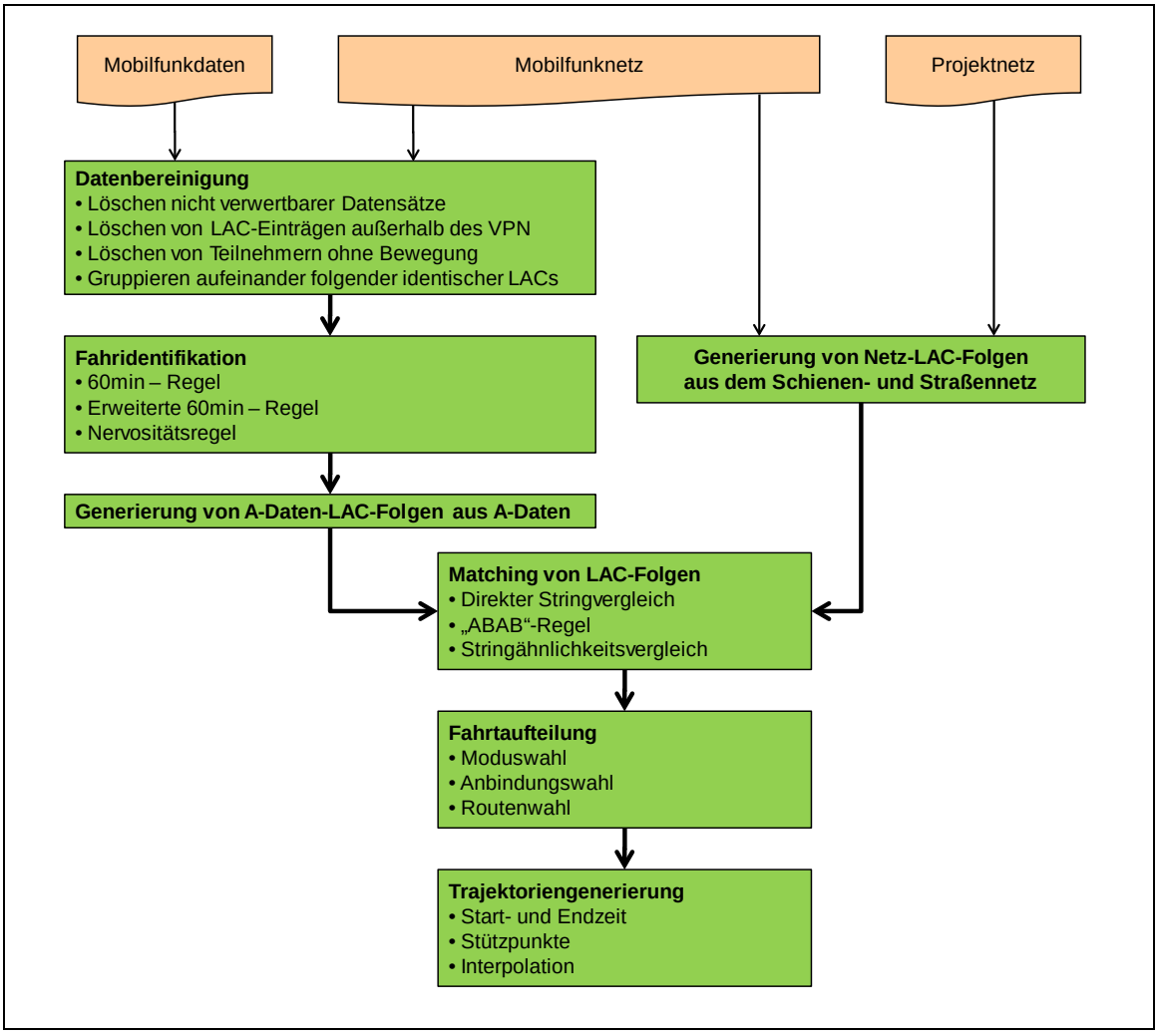


Abbildung 1: Ablauf der LAC-basierten Generierung von Trajektorien.

Für die Generierung von A-Daten-LAC-Folgen ist es nach der Datenbereinigung notwendig, einzelne Fahrten aus den Daten der Mobilfunkteilnehmer, die bis zu 24 Stunden am Stück aufgezeichnet werden, zu identifizieren und herauszuschneiden. Ergebnis sind einzelne LAC-Folgen, die bei einer Umwandlung der Location Area Codes in Zeichen als Zeichenfolge (z.B. „ABFREH“) dargestellt werden können.

Für die Netz-LAC-Folgen aus dem Straßen- und Schienennetz werden zunächst über ein Mehrwege-Suchverfahren alle sinnvollen Wege zwischen allen Location Areas gesucht. Anschließend werden den Strecken und Knoten im Verkehrsnetz die jeweils nächste Location Area zugeordnet, so dass auch hier die LAC-Folgen als Zeichenfolge dargestellt werden können.

Über ein dreistufiges Verfahren werden dann für die A-Daten-LAC-Folgen passende Netz-LAC-Folgen gesucht:

1. *Direkter Stringvergleich:*
   Im ersten Schritt wird geprüft, ob es eine direkte Übereinstimmung zwischen einer beobachteten A-Daten-LAC-Folge und einer Netz-LAC-Folge gibt.

2. *ABAB-Regel:*
   Die ABAB-Regel berücksichtigt, dass Mobilfunkteilnehmer aufgrund von Abschattungen o.ä. gelegentlich zwischen Location Areas hin- und herwechseln. Die ABAB-Regel kürzt die A-Daten-LAC-Folgen aus den Mobilfunkdaten um diese mehrfachen Wechsel und vergleicht die daraus entstehenden LAC-Folgen wiederum direkt mit den Netz-LAC-Folgen aus dem Straßen- oder Schienennetz.

3. *Stringähnlichkeitsvergleich:*
   Durch Abschattungen sowie Ungenauigkeiten bei der Berechnung der LAC-Folgen werden etliche LAC-Folgen trotz eindeutiger Bewegung nicht in den ersten beiden Schritten als Fahrt identifiziert. Daher erfolgt im letzten Schritt ein Stringähnlichkeitsvergleich, der sehr ähnliche LAC-Folgen identifiziert und somit eine Routengenerierung ermöglicht.

Tabelle 1 zeigt die Anzahl der identifizierten Fahrten in den drei Schritten. Insgesamt werden an einem durchschnittlichen Werktag ungefähr 500.000 Fahrten identifiziert, von denen ca. 420.000 im Straßennetz und 80.000 Schienennetz verlaufen.

Beschreibung	**Identifizierte Fahrten**			**Anteil identifzierter Fahrten**	
	Straße	**Schiene**	**Gesamt**	**je Schritt**	**kumuliert**
Direkter Stringvergleich	210.000	37.000	247.000	39,9%	39,9%
„ABAB“-Regel	73.000	7.000	80.000	13,0%	52,9%
Stringähnlichkeitsvergleich	137.000	38.000	175.000	28,3%	81,2%
Gesamt	**420.000**	**82.000**	**502.000**		

Tabelle 1: Statistik für das Matching der LAC-Folgen (23. Juni 2008).

Einige, vor allem kurze, LAC-Folgen repräsentieren sowohl Fahrten im Straßen- als auch Schienennetz. Diese werden genauso wie mehrdeutige Routen innerhalb einer Location Area durch verschiedene Aufteilungsverfahren auf die möglichen Routen verteilt. Die Abbildung 2 zeigt die endgültigen Belastungen im Projektnetz für den 23. Juni 2008.

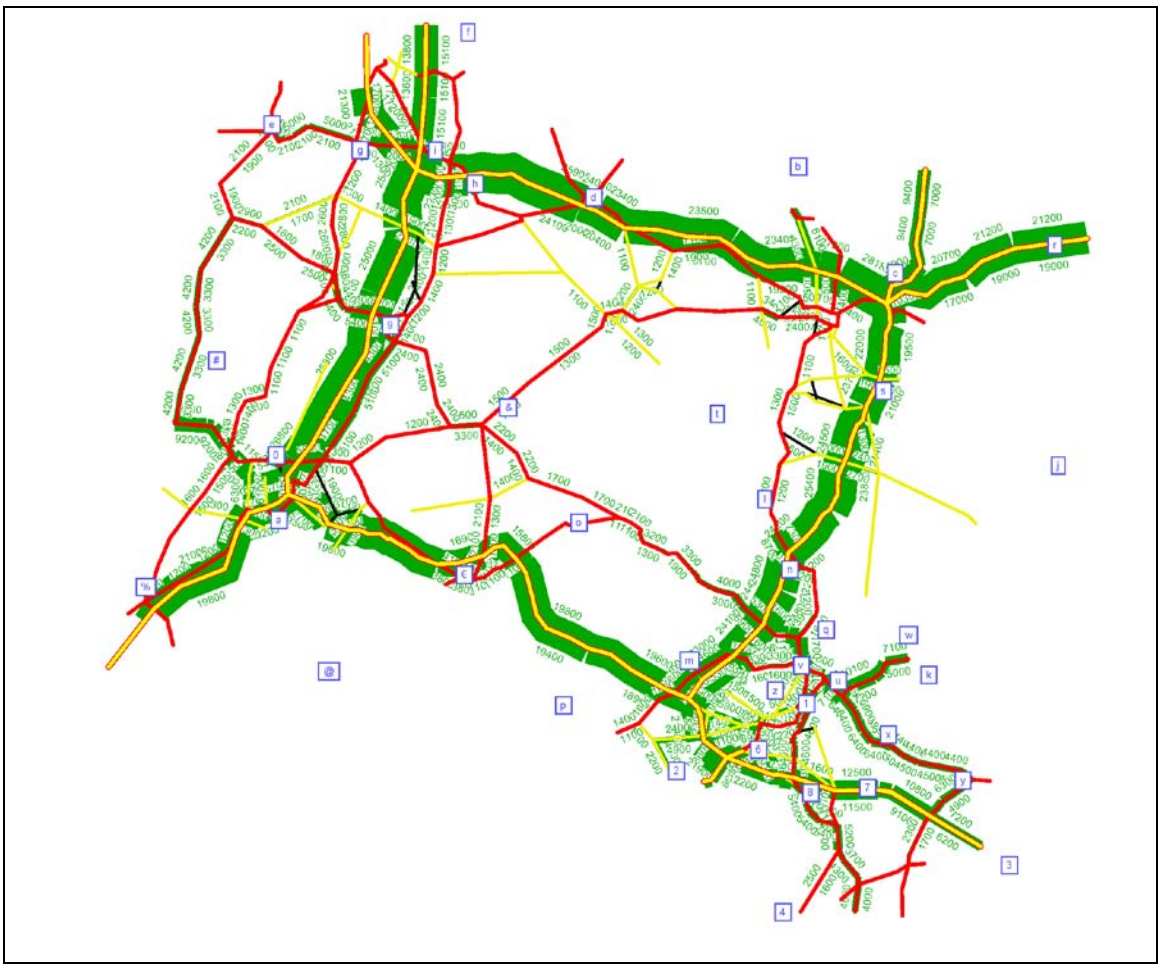


Abbildung 2: Belastungen aus FPD-Trajektorien im Projektnetz.

Abschließend zeigt Abbildung 3 die Fahrzeiten zwischen Stuttgart und Walldorf (über Karlsruhe), die eine Aufteilung der Fahrzeuge in die Fahrzeugklassen Pkw und Lkw ermöglichen. Diese zeigt wie auch weitere detaillierte Analysen im Kapitel 3, dass das Verfahren für Anwendungen auf den Autobahnen im Untersuchungsgebiet sehr gute Trajektorien generiert.

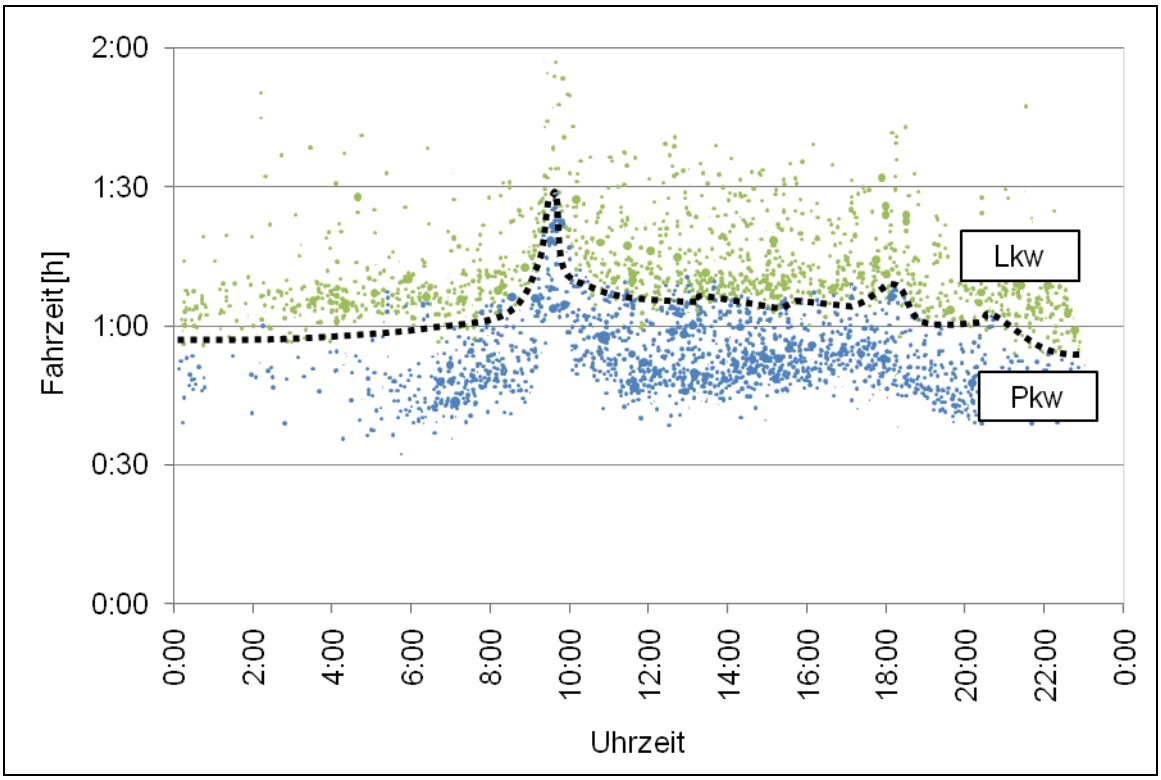


Abbildung 3: Fahrzeiten zwischen Stuttgart und Walldorf (über Karlsruhe) und Aufteilung in die Fahrzeugklassen Pkw und Lkw.

Schätzung der Parameter des Routenwahlverhaltens (Kapitel 4)

Mit den FPD-Trajektorien können Untersuchungen zum Routenwahlverhalten in der BAB-Netzmasche (Autobahnen A5/A6/A8/A81) und im Bereich der Netzbeeinflussungsanlage Stuttgart (Autobahn A81 und Bundesstraßen B10/27/295) durchgeführt werden. Dabei wird mit Maximum-Likelihood-Schätzungen untersucht, welche Kenngrößen (Verkehrsmeldungen, Wechselwegweisung etc.) die Routenwahl überhaupt beeinflussen und, wenn ja, welche Gewichtung die einzelnen Kenngrößen haben. Abbildung 4 zeigt die Untersuchungsgebiete sowie die Lage der dynamischen Wegweiser mit integrierten Stauinformationen (dWiSta).

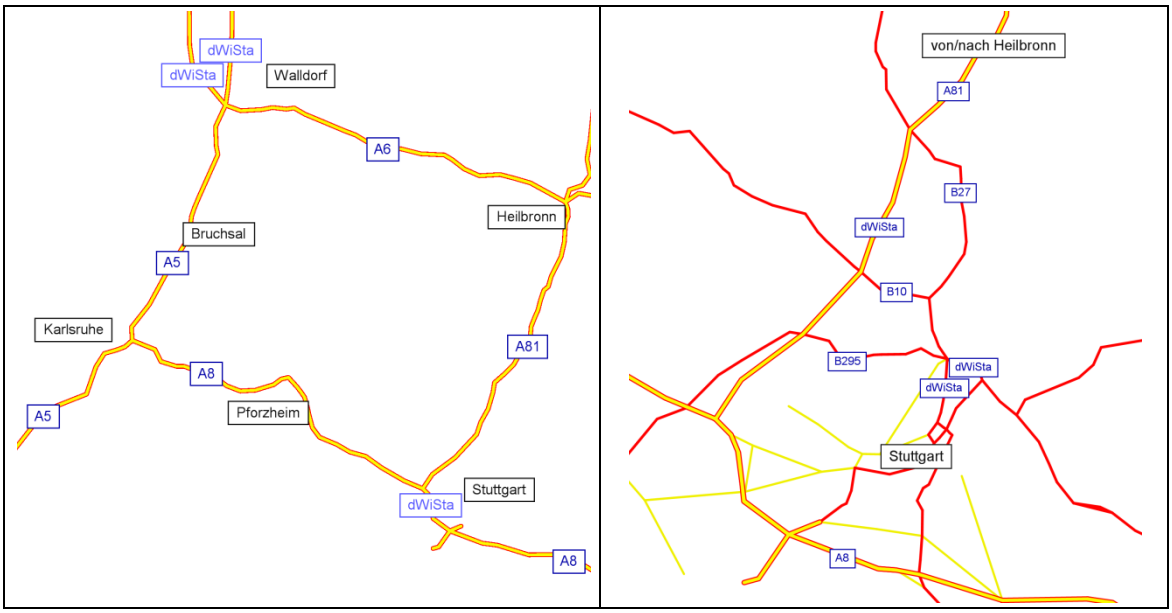


Abbildung 4: Untersuchungsgebiete für die Routenwahl (links: BAB-Netzmasche; rechts: NBA Stuttgart)

In der BAB-Netzmasche können über eine Millionen Fahrten auf den beiden Diagonalen Stuttgart ↔ Walldorf und Heilbronn ↔ Karlsruhe analysiert werden. Für den Bereich der NBA Stuttgart können ca. 350.000 Fahrzeuge des Durchgangsverkehrs analysiert werden.

Für die *BAB-Netzmasche* ergeben sich statistisch abgesicherte Ergebnisse bei den Maximum-Likelihood-Schätzungen. Die Ergebnisse der Schätzungen können mit Hilfe des Logit-Modells grafisch dargestellt werden, so dass für jeden Zustand die zu erwartende Aufteilung abgelesen werden kann. So kann aus Abbildung 5 zum Beispiel abgelesen werden, dass bei 5 km mehr Stau über Heilbronn als über Karlsruhe bei inaktiver Wechselwegweisung (WWW) ca. 78 % des Durchgangsverkehrs die Route über Karlsruhe wählen. Eine aktive Wechselwegweisung bewirkt einen zusätzlichen Anteil von 6 %, was einen Befolgungsgrad von ca. 29 % bedeutet. Bei einer auf das Autobahndreieck Leonberg (bei Stuttgart) zufließenden Verkehrsstärke von z.B. 6.000 Fahrzeugen in einer Stunde ergibt sich daraus bei einem mittleren Anteil des Durchgangsverkehrs von ca. 12 % einen Anzahl von ca. 50 Fahrzeugen pro Stunde, die durch die Wechselwegweisung umgelenkt werden.

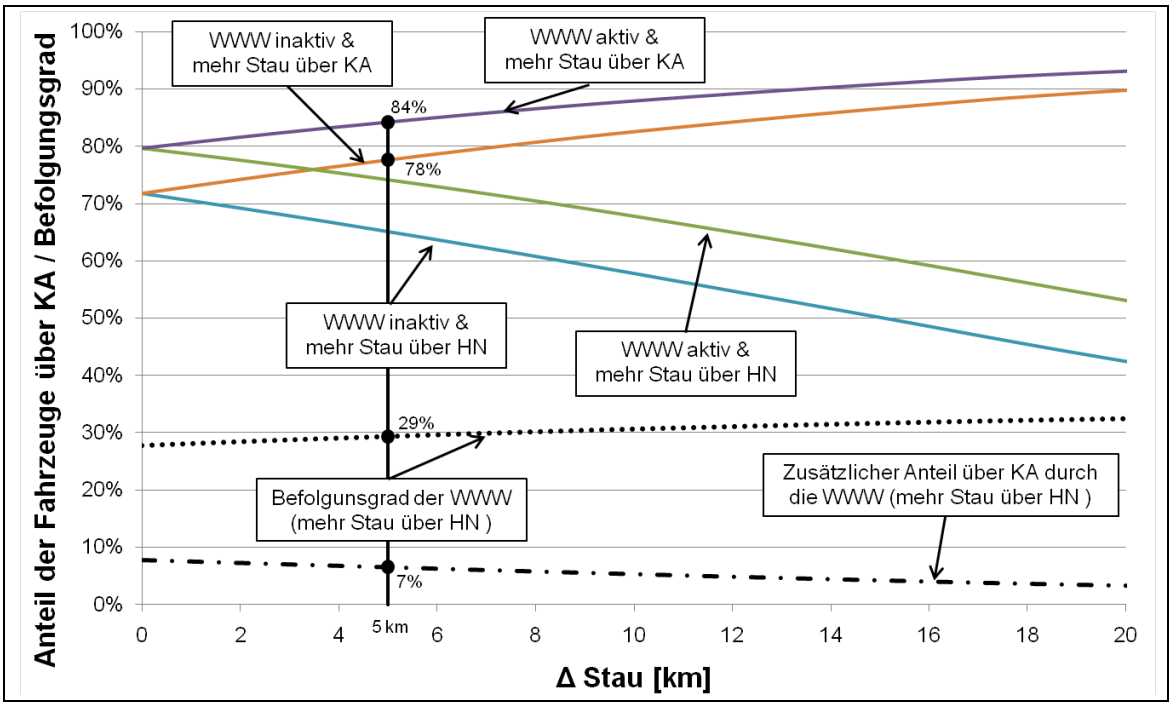


Abbildung 5: Grafische Darstellung des Basismodells der Routenwahl für die Richtung Stuttgart → Walldorf.

Weitere ausgewählte Ergebnisse der Routenwahlschätzungen sind:

- Pkw und Lkw zeigen ein vergleichbares Routenwahlverhalten. Dabei tendieren Lkw etwas stärker zur kürzeren Route, während Pkw etwas stärker auf Verkehrsmeldungen und die Wechselwegweisung reagieren.
- Die Verkehrsteilnehmer bewerten Verkehrsmeldungen, die sie früher auf ihrer Route erreichen, höher als Meldungen auf einem späteren Abschnitt ihrer Route. Bei einer Unterteilung in zwei Abschnitte der ca. 100 km langen Alternativen werden die Verkehrsmeldungen des ersten Abschnittes fast doppelt so hoch bewertet wie die des zweiten Abschnitts.
- Die Verkehrsteilnehmer, die der Routenempfehlung der Wechselwegweisung folgen, tun dies unabhängig von der Anzeige der Staulänge bzw. des Stauortes. Dies bedeutet allerdings nicht, dass die zusätzlichen Anzeigen keine Wirkung haben. Dies konnte nicht untersucht werden, da die zusätzlichen Anzeigen zu allen Routenempfehlungen gegeben worden sind.

Für die *NBA Stuttgart* ergeben sich bei den Maximum-Likelihood-Schätzungen weniger statistisch gesicherte Ergebnisse. Dies ist unter anderem auf die seltene Schaltung der NBA zurückzuführen, die dazu führt, dass z.B. in Stadteinwärtsrichtung nur 545 Fahrzeuge während einer Schaltung der NBA erfasst wurden. Abbildung 6 zeigt die Routenwahl für die beiden Richtungen in Abhängigkeit der Schaltungen der NBA mit Angabe der jeweiligen Fallzahl unter den Säulen:

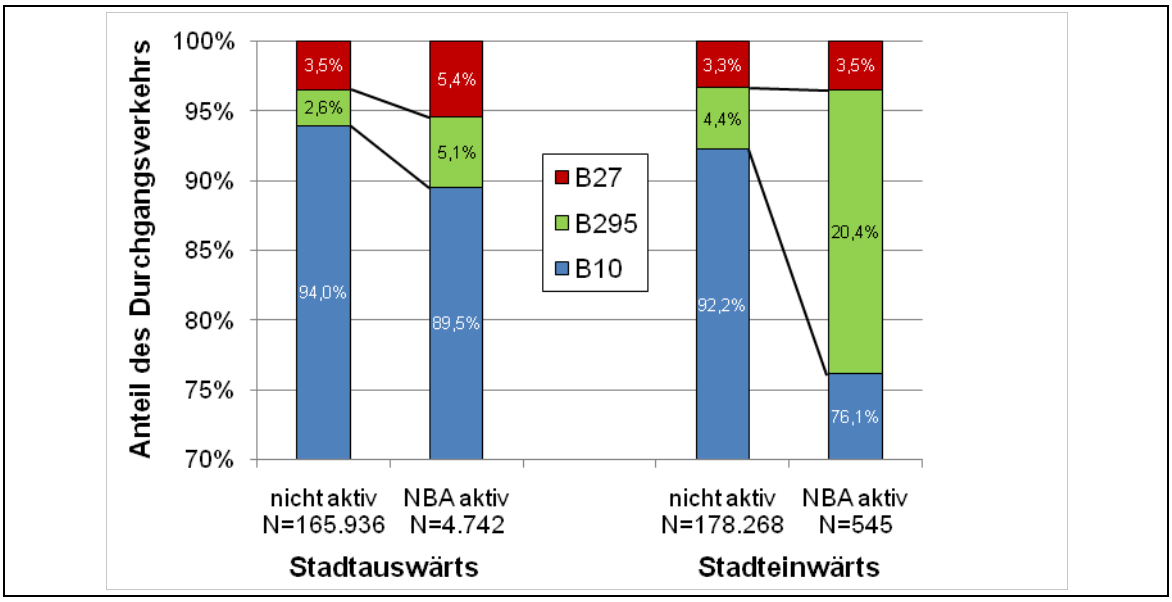


Abbildung 6: Routenwahlverhalten im Bereich der NBA Stuttgart.

In Stadtauswärtsrichtung wirkt eine Empfehlung der NBA in kleinerem Umfang auf beide Alternativrouten B27 und B295. In Stadteinwärtsrichtung dagegen erhöht die Routenempfehlung der NBA den Anteil der Route über die B295 deutlich. Aus den Werten ergeben sich bei Vernachlässigung der B27 Befolgungsgrade von 2,7 % (stadtauswärts) bzw. 17,3 % (stadteinwärts). Ein Einfluss der Staumeldungen kann nicht festgestellt werden. Die detaillierte Verfahrensbeschreibung und die Ergebnisse sind im Kapitel 4.

Generierung von Quelle-Ziel-Matrizen (Kapitel 5)

In Kapitel 5 werden Methodik und Ergebnisse zur Erzeugung von typischen Quelle-Ziel-Matrizen aus Mobilfunktrajektorien beschrieben. Die Methodik gliedert sich dabei zunächst in die folgenden beiden parallel laufenden Prozesse:

- Erzeugen von Clustern typischer Tage (z.B. Mo-Do, So/Feiertag, usw.) aus den Ähnlichkeiten der Ganglinien an Richtungsmessquerschnitten (RMQ).
- Tagesweises Hochrechnen der vorab erzeugten Mobilfunktrajektorien an den Zähldaten der jeweiligen RMQs.
- Zusammenfassen der hochgerechneten Mobilfunktrajektorien eines Clusters zu einem typischen Tag.

Abbildung 7 zeigt das Ergebnis der Clusterung, wenn eine Klassifizierung in die drei Gruppen Montag bis Freitag während bzw. außerhalb der Schulferien und Samstag, Sonntag und Feiertage durchgeführt wird. Die Höhe der Balken gibt die Anzahl der Tage in dem jeweiligen Cluster wieder und die Farbe repräsentiert den Wochentag. Die Abbildung zeigt deutlich, dass die Freitage (grün) in den Werktagsgruppen eine

Sonderstellung einnehmen und Samstag und Sonntag eine Trennung in der dritten Gruppe erfahren. Die weiteren Cluster mit sehr wenigen Tagen repräsentieren entweder besondere Tage wie Heiligabend oder Ostermontag oder sind Ausreißer aufgrund besonderer Ereignisse (z.B. Unwetter, Streik im Öffentlichen Verkehr).

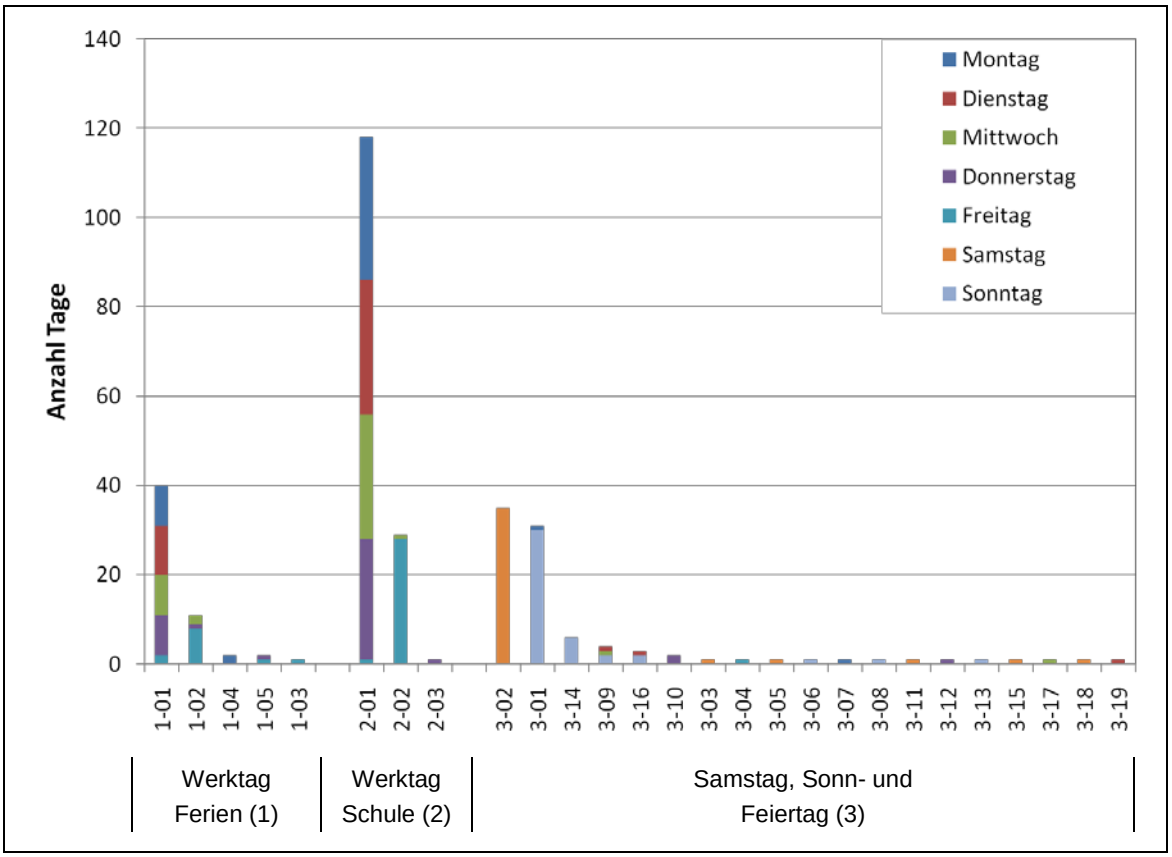


Abbildung 7: Clusterverteilung.

Die Belastungen der aus den Mobilfunkdaten erzeugten Trajektorien, die eine Stichprobe aller Kfz-Fahrten darstellen, werden mit einem dynamischen Hochrechnungsverfahren, das die räumlich/zeitliche Fortbewegung im Straßennetz berücksichtigt, hochgerechnet. Die Sollwerte des Verfahrens sind dabei die von den RMQs gezählten Verkehrsstärken, die in einer Qualität reproduziert werden können, die von anderen Verfahren bisher nicht erreicht wird. Abbildung 8 zeigt den Verlauf der Verkehrsstärken an einer Strecke im Projektnetz. Dabei ist die von einem RMQ gezählte Sollbelastung in rot, die Belastung der Mobilfunktrajektorien vor der Hochrechnung in schwarz und die Belastung der hochgerechneten Mobilfunktrajektorien in blau dargestellt.

Zu beachten ist, dass aufgrund der Mindestfahrtweite von Mobilfunktrajektorien nur RMQs auf Bundesautobahnen (BAB) verwendet werden und der Anspruch auf eine hohe Qualität des Ergebnisses auch nur auf BAB besteht.

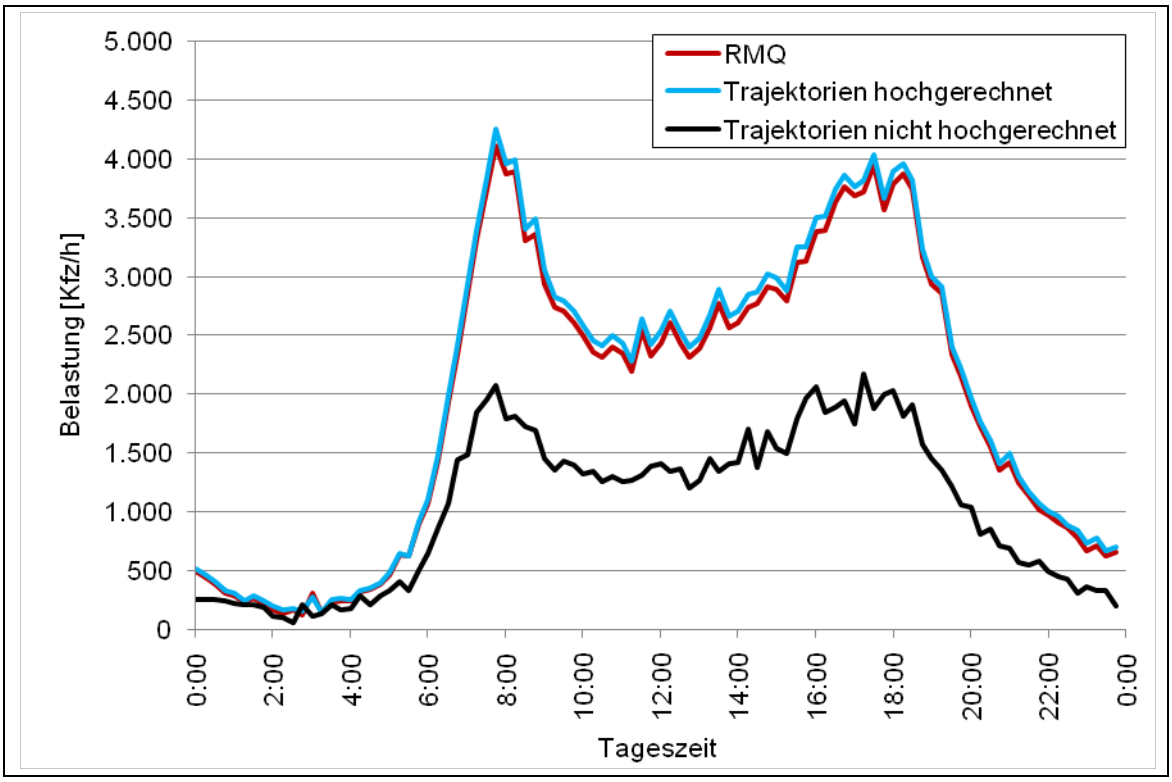


Abbildung 8: Verkehrsstärke von RMQ, Mobilfunktrajektorien und hochgerechneten Mobilfunktrajektorien am 19. September 2008 an der A81 von Stuttgart nach Heilbronn.

Abbildung 9 zeigt den Verlauf aggregierter Eingangsgrößen und die Ergebnisse der 68 hochgerechneten Tage, wobei die Sommerferien in Baden-Württemberg gelb hervorgehoben sind. Die schwarzen Kreuze repräsentieren die Summe der Belastungen aller eingegangen Mobilfunktrajektorien und die blauen Kreuze entsprechend dazu die Summe der Belastungen der hochgerechneten Trajektorien.

Für die Mobilfunktrajektorien vor und nach der Hochrechnung ist deutlich ein 7-Tageszyklus mit steigenden Belastungen von Montag bis Freitag und schwächerem Wochenendverkehr zu erkennen. Ferner geht die Summe aller Fahrten während der Sommerferien leicht zurück. Tiefergehende Analysen wie die Fahrtweitenverteilung oder aus anderen Untersuchungen bekanntes Routenwahlverhalten weisen ebenfalls plausible Ergebnisse auf.

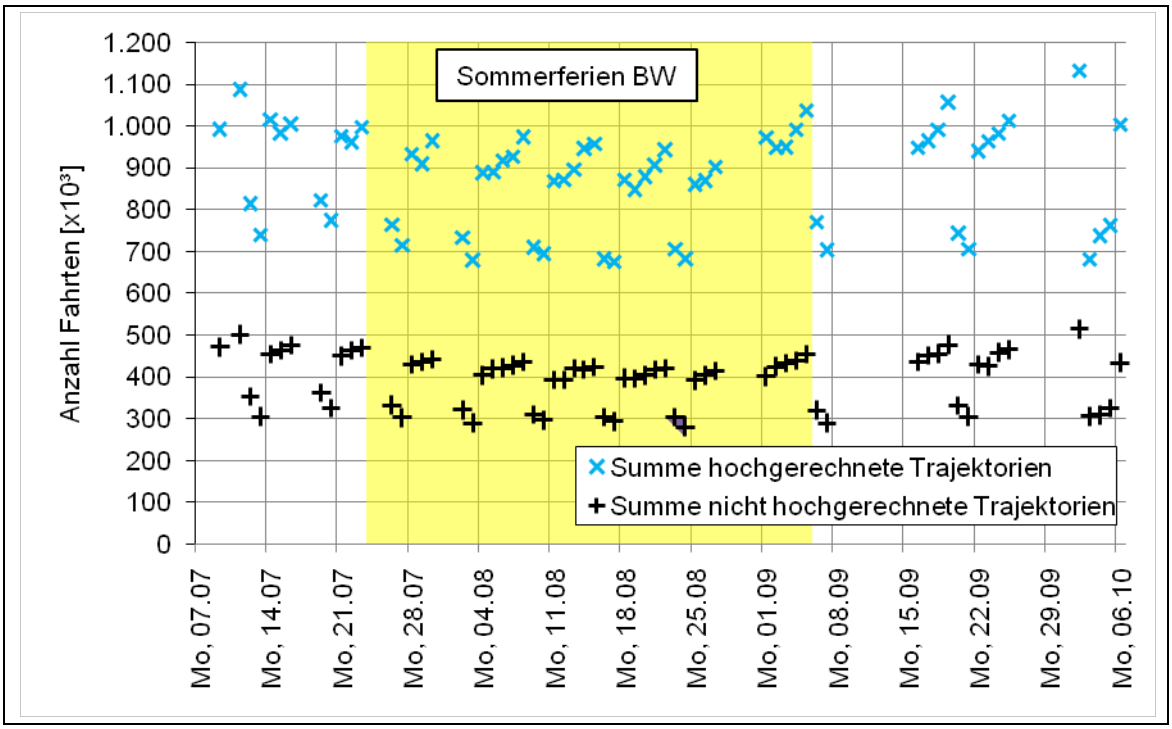


Abbildung 9: Anzahl Detektoren, Summe der FPD und hochgerechneten Fahrten.

Die hochgerechneten Trajektorien werden anschließend zu den aus der Clusterung erzeugten Tagesgruppen über eine einfache Mittelwertbildung zusammengefasst. Tabelle 2 listet die Cluster auf, die von den 68 hochgerechneten Tagen versorgt werden. Das letzte Cluster ist nur der Vollständigkeit halber aufgeführt, da in den 68 zur Verfügung stehenden Tagen mit dem 3. Oktober nur ein einziges Mal ein Feiertag auf einen Donnerstag oder Freitag fällt.

Beschreibung	Anzahl enthaltener Tage	Anzahl Fahrten
Montag bis Donnerstag (Schule)	16	984.000
Freitag (Schule)	3	1.092.000
Montag bis Donnerstag (Ferien)	22	909.000
Freitag (Ferien)	4	978.000
Samstag	11	745.000
Sonntag oder Feiertag an einem Montag	11	714.000
Sonstige Feiertage	1	681.000

Tabelle 2: Erzeugte typische Verkehrstage.

Aus diesen auf diese Weise erzeugten typischen Trajektorien können durch zeitscheibenbasiertes Aufsummieren beliebig feine Quelle-Ziel-Matrizen erzeugt werden. Für die üblichen Planungszwecke sind stundenfeine Matrizen, die auch zur Durchführung der im Folgenden beschriebenen Tagesprognose verwendet werden, allerdings vollkommen ausreichend.

Durchführung von Tagesprognosen (Kapitel 6)

Kapitel 6 beschreibt das Verfahren für die Tagesprognose. Ziel einer Tagesprognose ist es, den Verkehrszustand im Untersuchungsnetz für einen in der Zukunft liegenden Kalendertag stundenfein zu prognostizieren. Dabei werden die typische zeitabhängige Verkehrsnachfrage für den Kalendertag und bekannte Kapazitätsreduzierungen durch Baustellen beim Verkehrsangebot berücksichtigt:

- Verkehrsangebot: Als Grundlage für die Tagesprognose wird das auf TMC-Knoten basierende Projektnetz verwendet. In diesem Netz ist jede Strecke mit der zulässigen Geschwindigkeit und einer Stundenkapazität attributiert. Am Kalendertag vorhandene Baustellen werden über eine Reduktion der Kapazität und der Geschwindigkeit berücksichtigt.
- Verkehrsnachfrage: Als Nachfragedaten werden die Quelle-Ziel-Matrizen verwendet, die nach dem im vorherigen Abschnitt vorgestellten Verfahren geclustert worden sind. Diese Matrizen bestehen aus 24 stundenfeinen Teilmatrizen. Quellen und Ziele sind TMC-Locations. Damit entspricht die Verkehrsnachfrage einem Erwartungswert der Verkehrsbelastung, mit dem an ähnlichen Tagen wie dem Prognosetag zu rechnen ist.

Die Berechnung der Tagesprognose erfolgt in drei Schritten. Zunächst wird für den zu prognostizierende Tag das Cluster bestimmt, zu dem er gehört. Daraus ergibt sich die für dieses Cluster typische Verkehrsnachfrage.

Im zweiten Schritt wird diese Verkehrsnachfrage auf das verwendete Netz umgelegt. Dabei wird das Routenwahlverhalten der Verkehrsteilnehmer nachgebildet. Routenwahlverfahren verteilen die Verkehrsnachfrage auf die im Netz verfügbaren Routen und ermitteln so Verkehrsstärken für jede Strecke. Um die Routenwahl nachzubilden, wird häufig von einem sogenannten Nutzergleichgewicht (Deterministic User Equilibrium, DUE) ausgegangen. Es ist dann gegeben, wenn sich die Verkehrsteilnehmer so verteilen, dass die objektiv messbare Fahrtdauer auf allen alternativen Routen gleich ist und jeder Wechsel auf eine andere Route die persönliche Fahrzeit erhöhen würde. Ein deterministisches Nutzergleichgewicht kann im Untersuchungsgebiet zumindest bezogen auf die Fahrzeit nicht beobachtet werden. Deshalb wird für die Tagesprognose ein stochastisches Nutzergleichgewicht (Stochastic User Equilibrium, SUE) verwendet, bei dem die Verkehrsteilnehmer ihre subjektive empfundene Fahrtdauer optimieren.

Führt man mit den vorliegenden Quelle-Ziel-Matrizen eine Verkehrsumlegung durch, erhält man für jede Strecke Verkehrsstärken. Da die Routenwahl neu berechnet wird, ist es möglich, Auswirkungen von Baustellen, wie beispielsweise Geschwindigkeitsbegrenzungen und Verringerung der Kapazität, in die Tagesprognose einzubeziehen. Die Routenwahl erfolgt nach den Widerständen, in denen sich die aus den Baustellen resultierenden Behinderungen niederschlagen. Damit wird davon ausgegangen, dass

der Verkehrsteilnehmer über die Behinderung durch die Baustelle auf der entsprechenden Strecke informiert ist und in seine Routenwahl mit einbezieht.

Aus der Gesamtverkehrsstärke der einzelnen Straßenabschnitte wird im dritten Schritt in Abhängigkeit der Kapazität für jede Stunde des Prognosetages die Auslastung berechnet. Daraus wird die zu erwartende Qualitätsstufe des Verkehrsablaufs (Level of Service, LOS) abgeleitet. Der LOS kann als Ergebnis der Tagesprognose wie in Abbildung 10 dem Verkehrsteilnehmer für jeden Streckenabschnitt und jede Stunde eines kommenden Tages angezeigt werden. Tabelle 3 zeigt die dafür verwendete Klassifizierung des LOS, die auf der Auslastung bzw. dem Verhältnis der Fahrzeit t_{akt} im belasteten Netz zur Fahrzeit t_0 im unbelasteten Netz basiert.

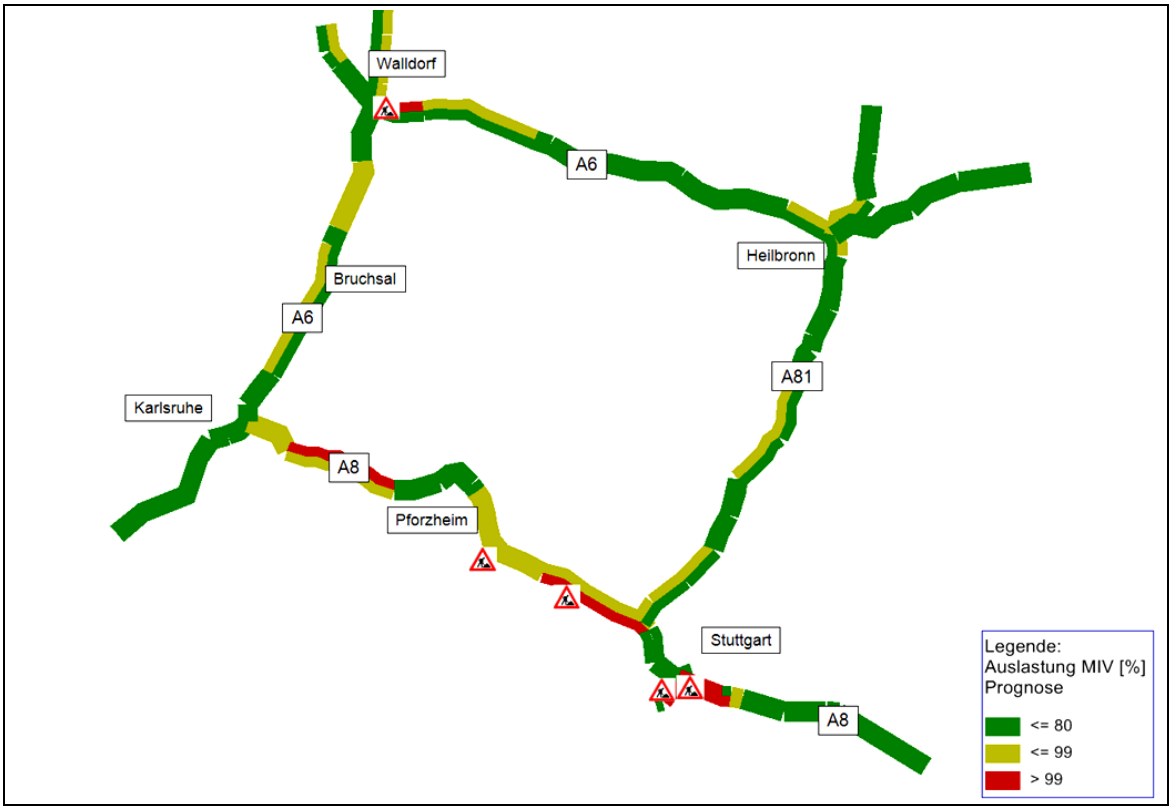


Abbildung 10: Level of Service-Darstellung als Ergebnis einer Tagesprognose am 23.06.08 von 7:00 Uhr bis 8:00 Uhr

Qualitätsstufe	Auslastung	Verhältnis t_{akt}/t_0
frei	< 80%	< 1,5
stockend	80% bis 99%	1,5 bis 2,0
gestaut	> 99 %	> 2,0

Tabelle 3: Ableitung des LOS

Um die Qualität der Tagesprognose abschätzen zu können, wurden Tagesprognosen für 68 Tage berechnet. Die daraus auf den Streckenabschnitten ermittelten Verkehrsstärken und Auslastungen werden dann mit den tatsächlich an diesem Tag gemesse-

nen Verkehrsbelastungen verglichen. Dies erfolgt für jede Stunde der 68 Testtage. Abbildung 11 zeigt die Häufigkeit der Abweichung zwischen der in der Tagesprognose vorhergesagten und der tatsächlich eingetretenen Qualitätsstufe.

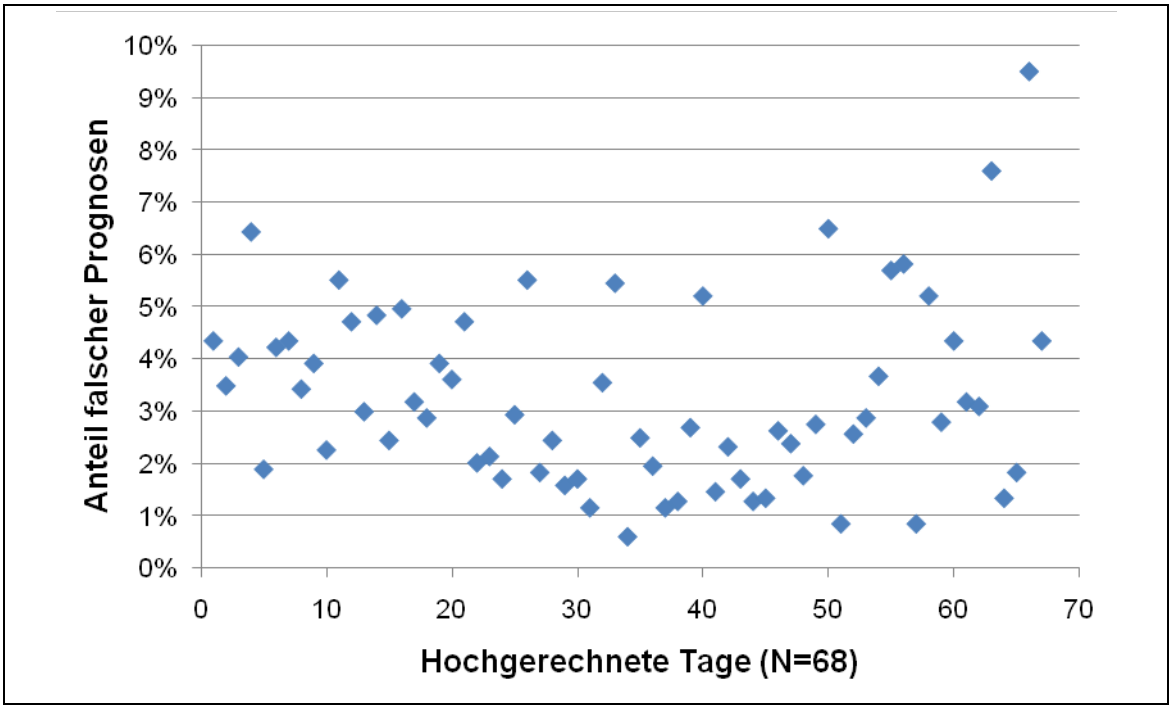


Abbildung 11: : Anteil der nicht korrekt prognostizierten Qualitätsstufen.

Der durchschnittliche Wert der Abweichung über alle untersuchten Tage liegt bei ca. 3 %. Das bedeutet, dass durchschnittlich je Strecke seltener als einmal am Tag für eine Stunde eine Qualitätsstufe falsch vorhergesagt wird.

Abweichungen zwischen Prognose und Realität sind unter anderem mit unvorhergesehenen Ereignissen wie Unfällen zu begründen. Aufgrund der Unberechenbarkeit dieser Situationen lässt sich deren Einfluss auf den Verkehrsablauf nicht im Voraus in die Tagesprognose einbeziehen.

Monitoring der Angebotsqualität (Kapitel 7)

Die Angebotsqualität in einem Verkehrsnetz beschreibt die Qualität des Verkehrsablaufs aus der Sicht des Nutzers. Das aus Nutzersicht entscheidende Kriterium ist der Zeitaufwand, der durch unterschiedliche Kenngrößen beschrieben werden kann. Im Folgenden werden die Kenngrößen Geschwindigkeit, Verlustzeit und Fahrzeitindex verwendet, zudem werden Verkehrsleistung und Verkehrszeitaufwand angegeben:

- Geschwindigkeit: mittlere Geschwindigkeit in einem Netzabschnitt [km/h]
- Verlustzeit: Summe der Verlustzeiten [Kfzh] in einem Netzabschnitt
- Fahrzeitindex: Quotient aus der mittleren Fahrzeit und der Soll-Fahrzeit

- Verkehrsleistung: Summe der gefahrenen Strecke aller Fahrzeuge in einem Netzabschnitt [Kfzkm]
- Verkehrszeitaufwand: Summe der Fahrzeiten aller Fahrzeuge in einem Netzabschnitt [Kfzh]

Unter Monitoring ist die kontinuierliche Erfassung der Angebotsqualität zu verstehen. Das Monitoring ermöglicht eine Analyse der Angebotsqualität über einen längeren Zeitraum. Es ermöglicht Trends zu erkennen und die Wirksamkeit verkehrlicher Maßnahmen zu quantifizieren. Für das Monitoring stehen sowohl Daten aus stationären Detektoren als auch Mobilfunktrajektorien zur Verfügung.

Dabei werden die Kenngrößen mittlere Geschwindigkeit, Fahrzeitindex, mittlerer Verkehrszeitaufwand sowie Verlustzeitaufwand auf einzelnen Streckenzügen sowie im gesamten Untersuchungsgebiet verwendet. Der Fahrzeitindex ist der Quotient aus der mittleren Fahrzeit und der Soll-Fahrzeit. Für einzelne Tage, an denen Messdaten der stationären Detektoren bzw. den FPD-Daten vorliegen, lassen sich die Kenngrößen Fahrzeit und mittlere Geschwindigkeit aus beiden Datenquellen ableiten.

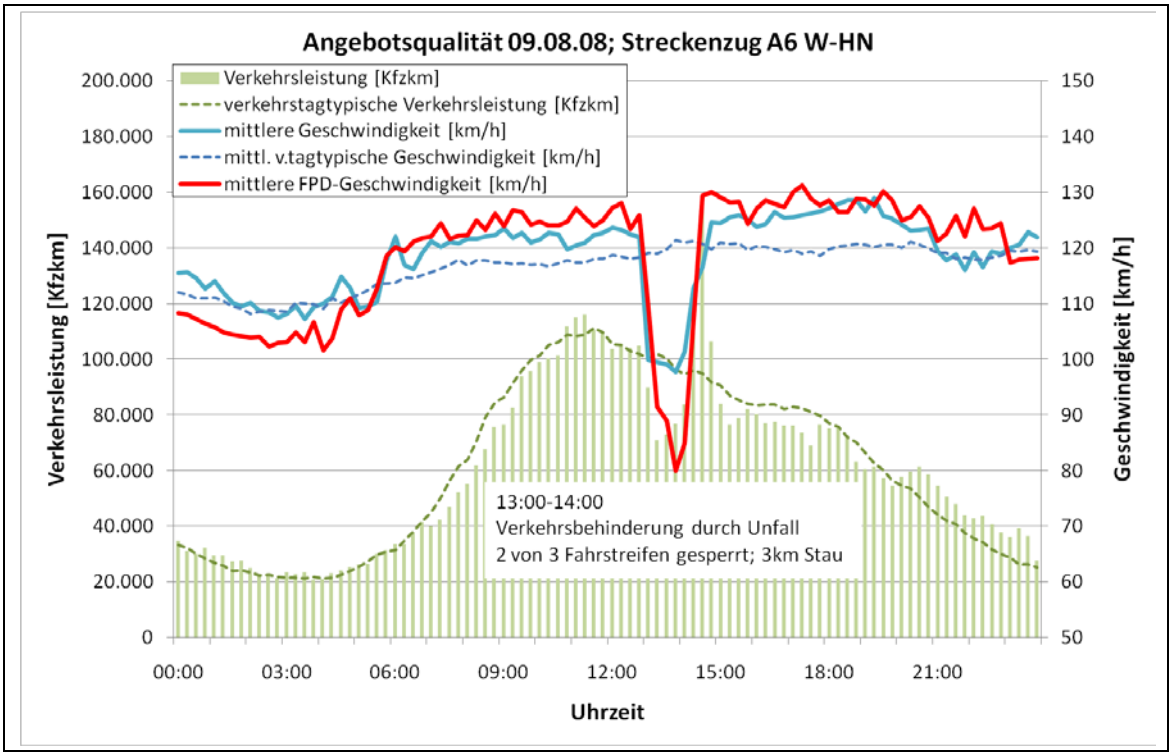


Abbildung 12: Tagesübersicht vom 09.08.2008 mit Verkehrsstörung durch Unfall.

In Abbildung 12 sind die Kenngrößen Verkehrsleistung sowie die Geschwindigkeitsverläufe für beide Datenquellen dargestellt. In diesem Tagesbericht der Angebotsqualität für die A6 von Walldorf nach Heilbronn ist deutlich zu sehen, wie die interpolierten Geschwindigkeiten der Zähldaten und die mittleren FPD-Geschwindigkeiten einen ähnlichen Verlauf über den Tag besitzen. Der Geschwindigkeitseinbruch

zwischen 13 und 14 Uhr von über 40 km/h ist sowohl bei den FPD-Geschwindigkeiten als auch bei den Zähldaten stationärer Detektoren zu erkennen, jedoch in unterschiedlich stark ausgeprägter Form. Ein Abgleich mit den Verkehrsmeldungen zeigt, dass die Störung auf einen Unfall zurückzuführen ist, der zu einer Sperrung von zwei Fahrstreifen auf der dreistreifigen Autobahn führte. Die mittleren verkehrstagtypische Geschwindigkeit repräsentiert die für diesen Tag typischerweise zu erwartende Geschwindigkeit. Der Vergleich zwischen dieser und der Ist-Geschwindigkeit zeigt, dass es sich hierbei um ein einmaliges Ereignis handelt.

Andere Streckenzüge weisen dagegen stark unterschiedliche Verläufe zwischen den Geschwindigkeiten von FPD- und Zähldaten auf. Abbildung 13 zeigt eine Tagesübersicht für den Streckenzug A8. An diesem Tag fanden Instandhaltungsarbeiten statt, die zu einem Stau führten. Die über den Streckenzug laufenden Mobilfunktrajektorien zeigen den aus dem Stau resultierenden Geschwindigkeitseinbruch. Die Geschwindigkeiten aus den stationären Messstellen bleiben jedoch von diesem Ereignis unbeeinflusst. Diese Tatsache lässt sich durch eine schlechte Messstellenversorgung der A8 auf dem Streckenzug der A8 erklären, auf dem es zwischen Karlsbad und Stuttgart einen 35 km langen Streckenabschnitt gibt, für den keine Messdaten vorliegen. Störungen in diesem Abschnitt können dort nicht erkannt werden, wenn sich Verkehrsstärke und Geschwindigkeit ausschließlich in diesem Abschnitt verändern und nicht auf benachbarte Strecken mit Detektoren auswirken. Bei der Erzeugung des Verkehrslageberichts wird dieser Streckenzug mit Messdaten der Detektoren bei Karlsbad und Stuttgart versorgt.

© VuV 2009

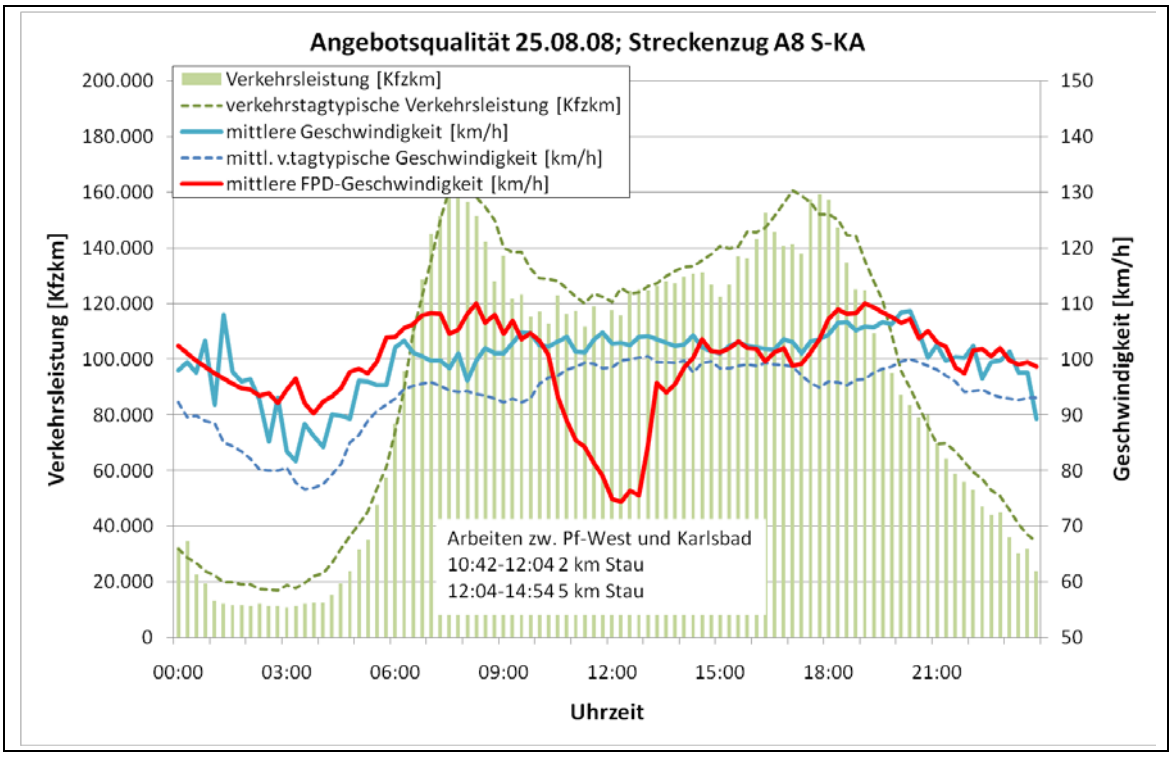


Abbildung 13: Tagesübersicht vom Streckenzug A8 S-KA am 25.08.2008 mit den Kenngrößen Verkehrsleistung und den Geschwindigkeiten aus FPD-Trajektorien und Zähldaten

Eine Monatsübersicht des Verkehrslagereports wird durch eine räumliche Aggregierung der Kenngrößen auf das gesamte Untersuchungsgebiet sowie eine zeitliche Aggregierung auf ganze Tage erzeugt. Am Beispiel des März 2008 Abbildung 14 ist deutlich der niedrigere Fahrzeitindex an Wochenenden zu erkennen, der sich durch die geringere Verkehrsnachfrage ergibt. Auch der letzte Arbeitstag vor den Osterferien, Freitag, der 14. März ist deutlich als Tag zu erkennen, der im Vergleich zu den sonstigen Freitagen eine um etwa 8% längere mittlere Fahrzeit besitzt.

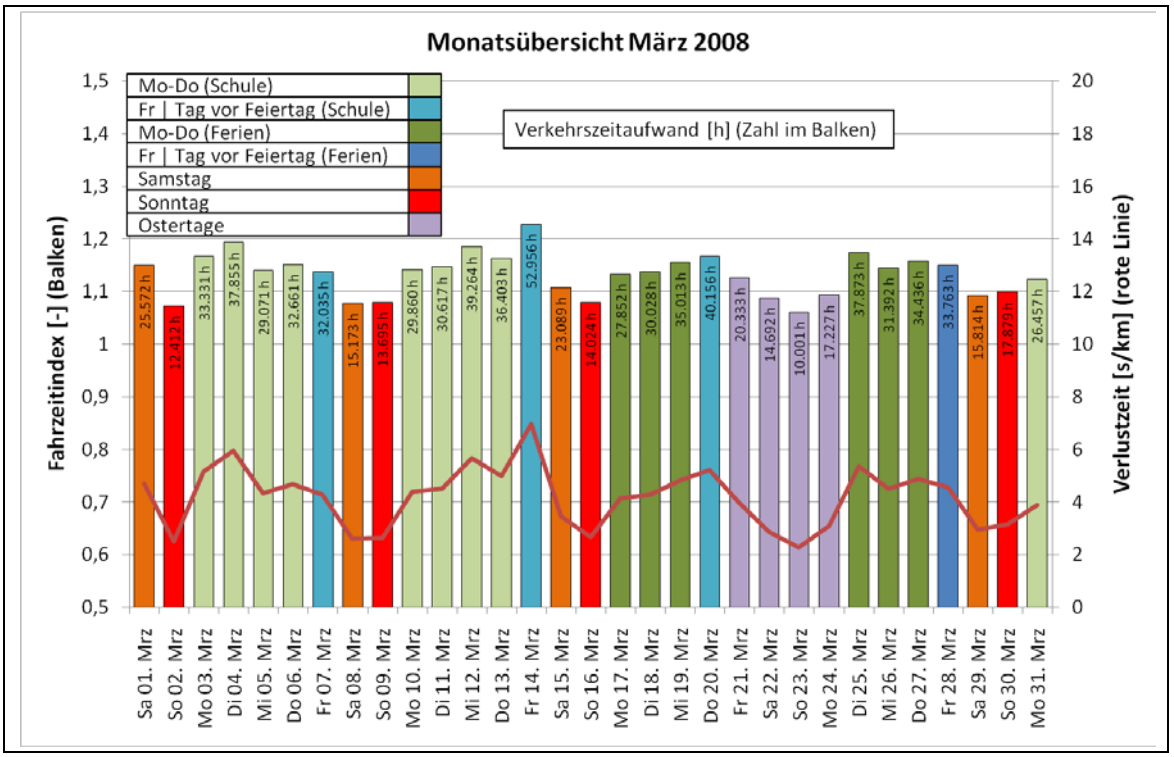


Abbildung 14: Übersicht der Kenngrößen des Monats März 2008 aus Messdaten stationärer Detektoren

Liegen Messdaten über einen längeren Zeitraum vor, sind Zeitreihenanalysen möglich. Durch einen Vergleich gemittelter Kenngrößen über ganze Monate hinweg, kann man z.B. eine langzeitbedingte Veränderung in der Angebotsqualität erkennen. Eine entsprechende Übersicht über den Fahrzeitindex und die und ein Jahresvergleich des Erfassungszeitraumes 2007 und 2008 ist in Abbildung 15 dargestellt. Tendenzen der Angebotsqualität lassen sich dadurch sowohl das gesamte Netz betreffend, als auch streckenzugfein darstellen.

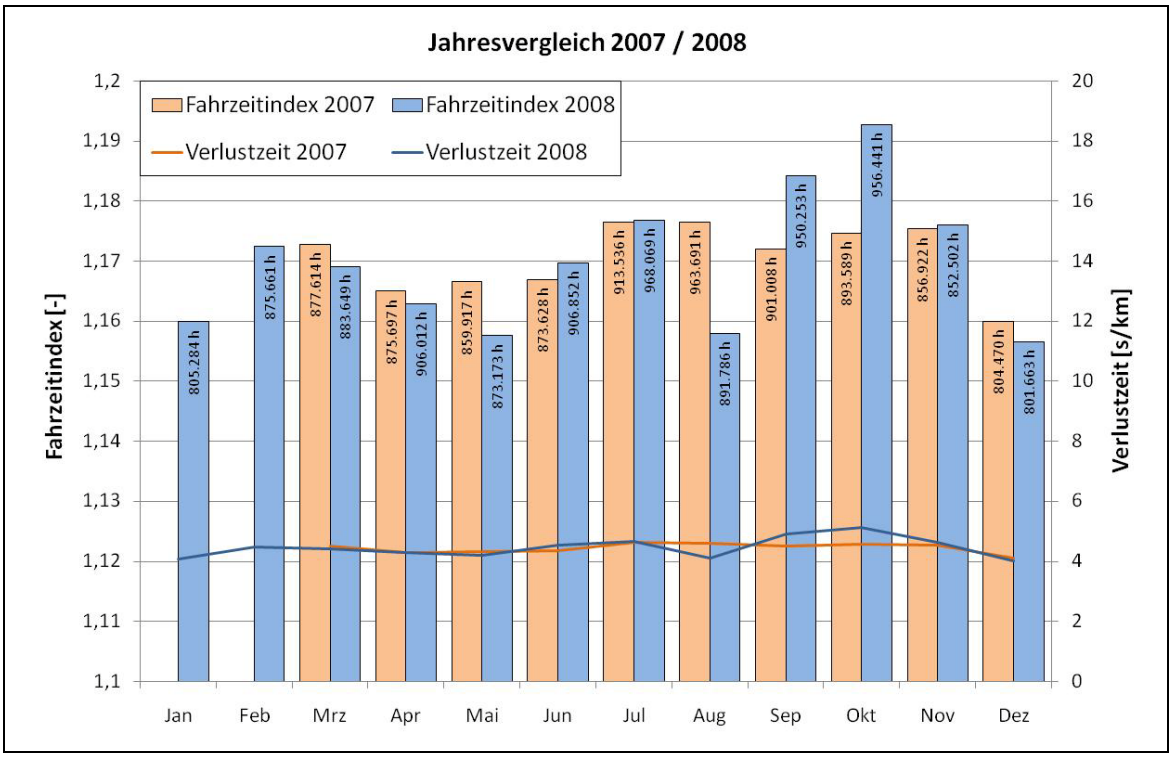


Abbildung 15: Vergleich der Fahrzeitindizes, Verkehrszeitaufwände und Verlustzeit der Jahre 2007 und 2008 im Untersuchungsgebiet.

1 Einleitung

In diesem Bericht zur Trajektoriengenerierung und zu den Arbeitspakten 5-7 werden die Beiträge des Lehrstuhls für Verkehrsplanung und Verkehrsleittechnik zum Schlussbericht zusammengefasst. Der Bericht beschreibt die fünf im Rahmen von Do-iT entwickelten Verfahren:

- Location-Area-basierte Trajektoriengenerierung (Kapitel 3):
  Hier wird erläutert das Verfahren erläutert, mit dem aus den Mobilfunkdaten der A-Ebene Trajektorien im Untersuchungsgebiet erzeugt werden. Dieses Kapitel enthält zudem die Ergebnisse der Trajektoriengenerierung über einen Zeitraum von ungefähr 80 Tagen.
- Schätzung der Parameter des Routenwahlverhaltens (Kapitel 4):
  Mit den Mobilfunk-Trajektorien wird das tatsächlich realisierte Routenwahlverhalten analysiert. Dabei werden die Einflüsse von Verkehrsmeldungen und von Wechselwegweisungen auf das Routenwahlverhalten untersucht.
- Generierung von Quelle-Ziel-Matrizen (Kapitel 5):
  Die Mobilfunktrajektorien werden mit einem dynamischen Verfahren anhand der Werte von stationären Detektoren hochgerechnet. Ergebnis dieser Hochrechnung sind Quelle-Ziel-Matrizen für das Untersuchungsgebiet. Durch Zusammenfassen der Matrizen für typische Verkehrstage können typische Quelle-Ziel-Matrizen für diese Verkehrstage berechnet werden.
- Durchführung einer Tagesprognose (Kapitel 6):
  Die typischen Quelle-Ziel-Matrizen werden dazu verwendet, um Tagesprognosen durchzuführen. Dafür wird die zu erwartende Verkehrsnachfrage auf das aktuelle, durch Baustellen veränderte Verkehrsnetz umgelegt.
- Monitoring der Angebotsqualität (Kapitel 7):
  Diese Kapitel zeigt eine Methode zum kontinuierlichen Monitoring der Angebotsqualität in Straßennetzen. In verschiedener zeitlicher Auflösung werden dabei Verkehrslagereports mit relevanten Kenngrößen entweder aus stationären Detektoren oder aus Mobilfunktrajektorien erstellt.

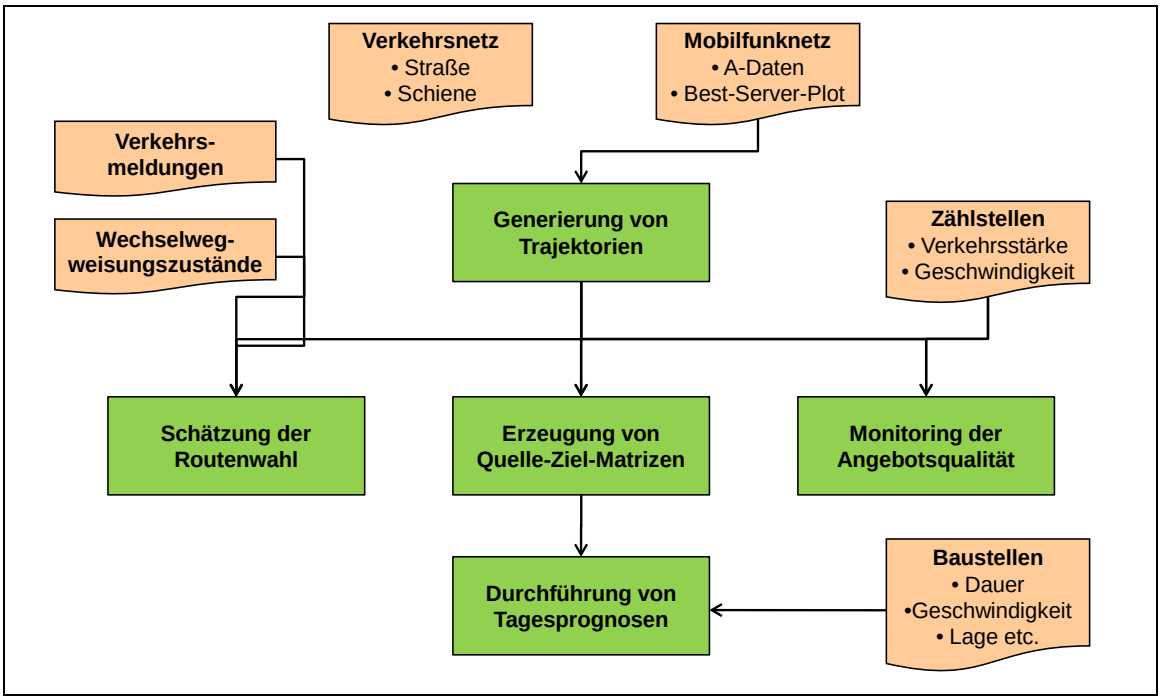


Abbildung 1: Vereinfachter Datenfluss und Struktur des Berichtes.

Die Abbildung 1 zeigt den vereinfachten Datenfluss. Aus den A-Daten des Mobilfunknetzes werden Trajektorien generiert, die anschließend in den verschiedenen Prozessen weiterverarbeitet werden. Für einige Prozesse sind weitere externe Informationen notwendig.

In den einzelnen Kapiteln werden zuerst die Verfahren ausführlich erläutert und analysiert. Danach werden die Ergebnisse für den Untersuchungszeitraum von rund drei Monaten dargestellt.

2 Eingangsdaten und Grundlagen

2.1 Vereinfachtes Projektnetz (VPN)

Die im AP6 entwickelten Verfahren verwenden Verkehrsstärken aus Richtungsmessquerschnitten und FPD-Trajektorien. Zur Verwendung dieser Eingangsdaten ist es notwendig, dass das zugrundeliegende Verkehrsnetz auf einem Knoten-Kanten-Modell basiert, dessen Kanten physikalischen Strecken entsprechen.

Ein auf Projektnetzstrecken basierendes Teilnetz des Lokalisierungsnetzes würde diese Bedingungen erfüllen. Der hohe Detailgrad der Roadelemente im Lokalisierungsnetz führt allerdings zu einer unverhältnismäßig hohen Rechenzeit bei der Durchführung der Verfahren. Aus diesem Grund wird auf Basis des Projektnetzes ein eigenständiges vereinfachtes Projektnetz (VPN) mit den geforderten Eigenschaften entwickelt. Im VPN werden komplexe Verkehrsknoten die aus mehreren Teilknoten bestehen zu einem Knoten zusammengefasst.

Die Knoten des VPN repräsentieren TMC-Locations, die auf den Strecken des Projektnetzes liegen. Dabei werden TMC-Locations mit denselben Koordinaten zu einem Knotenpunkt verschmolzen. Die Kanten zwischen den TMC-Locations entsprechen jeweils einer Strecke, deren Eigenschaften aus den dazwischen liegenden Roadelementen des Lokalisierungsnetzes bestimmt werden.

Es gibt einzelne Roadelemente im Projektnetz, die nicht eindeutig zwischen zwei TMC-Locations liegen und sich deshalb nicht eindeutig einer VPN-Kante zuordnen lassen. Abbildung 2 zeigt ein Beispiel für Roadelemente die in einen VPN-Knoten hineinfallen. Das Autobahnkreuz Walldorf wird im VPN durch genau einen Knoten repräsentiert (siehe Abbildung 3). Die im Lokalisierungsnetz blau und orange hervorgehobenen Roadelemente können keiner Von-Nach-TMC-Location-Strecke eindeutig zugeordnet werden. Im VPN liegen diese Roadelemente auf den Abbiegebeziehungen, deshalb können Zählstellen, die auf diesen Roadelementen liegen, nicht ins VPN übernommen werden.

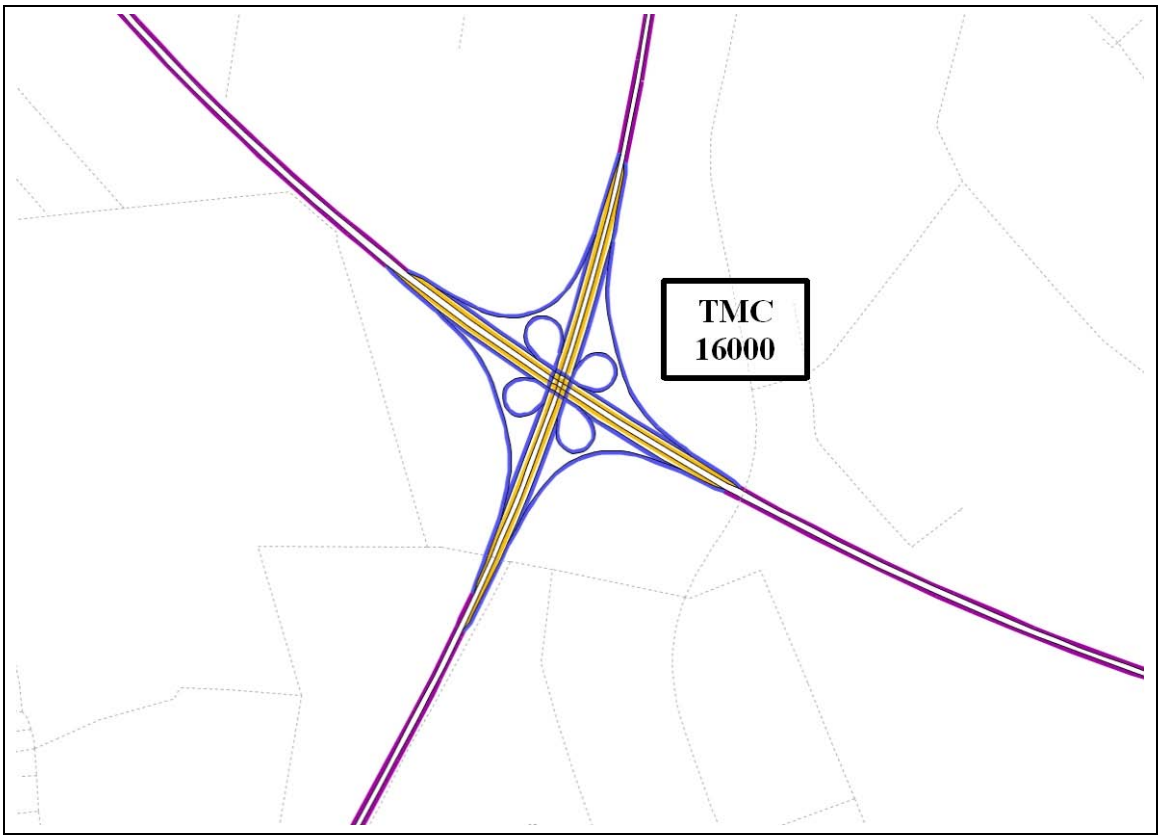


Abbildung 2: Autobahnkreuz Walldorf im LN.

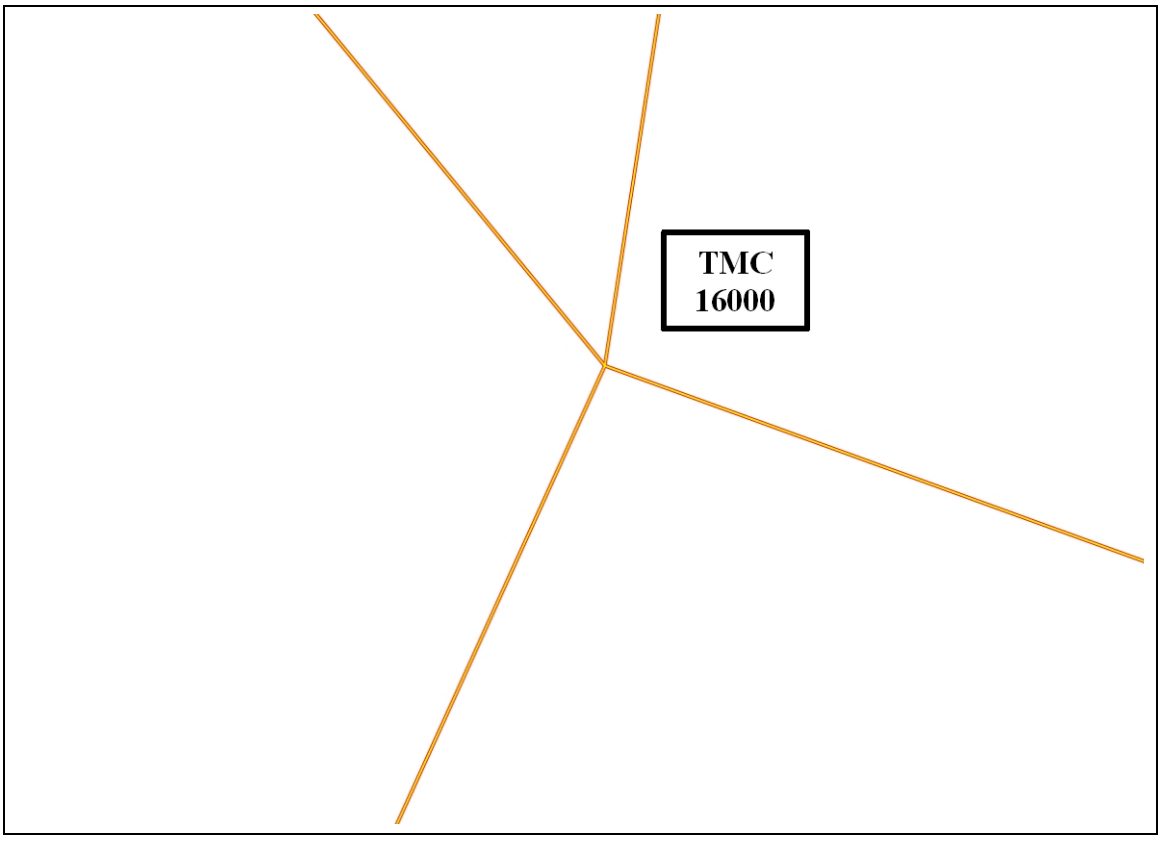


Abbildung 3: Autobahnkreuz Walldorf im VPN.

Die blau hervorgehobenen Roadelemente an der TMC-Location 20304 in Abbildung 4 können ebenfalls keiner eindeutigen Von-Nach-TMC-Location-Strecke zugeordnet werden und führen deshalb zu einer Verkürzung der Strecke zwischen den TMC-Locations 20304 und 30535. Die durch den Abstraktionsprozess auf diese Weise entfallenden Roadelemente des Projektnetzes führen zu einer geringen Abweichung (siehe auch Tabelle 1) der Gesamtstreckenlängen.

Strecke	VPN [km]	Googlemaps [km]	VPN/Googlemaps [%]
Leonberg-Heilbronn-Walldorf	101,4	103	98,45
Leonberg-Karlsruhe-Walldorf	89,8	90,1	99,67
Walldorf-Heilbronn-Leonberg	102,0	103	99,03
Walldorf-Karlsruhe-Leonberg	89,2	90,3	98,78

Tabelle 1: Vergleich VPN-Streckenlängen und Googlemaps.

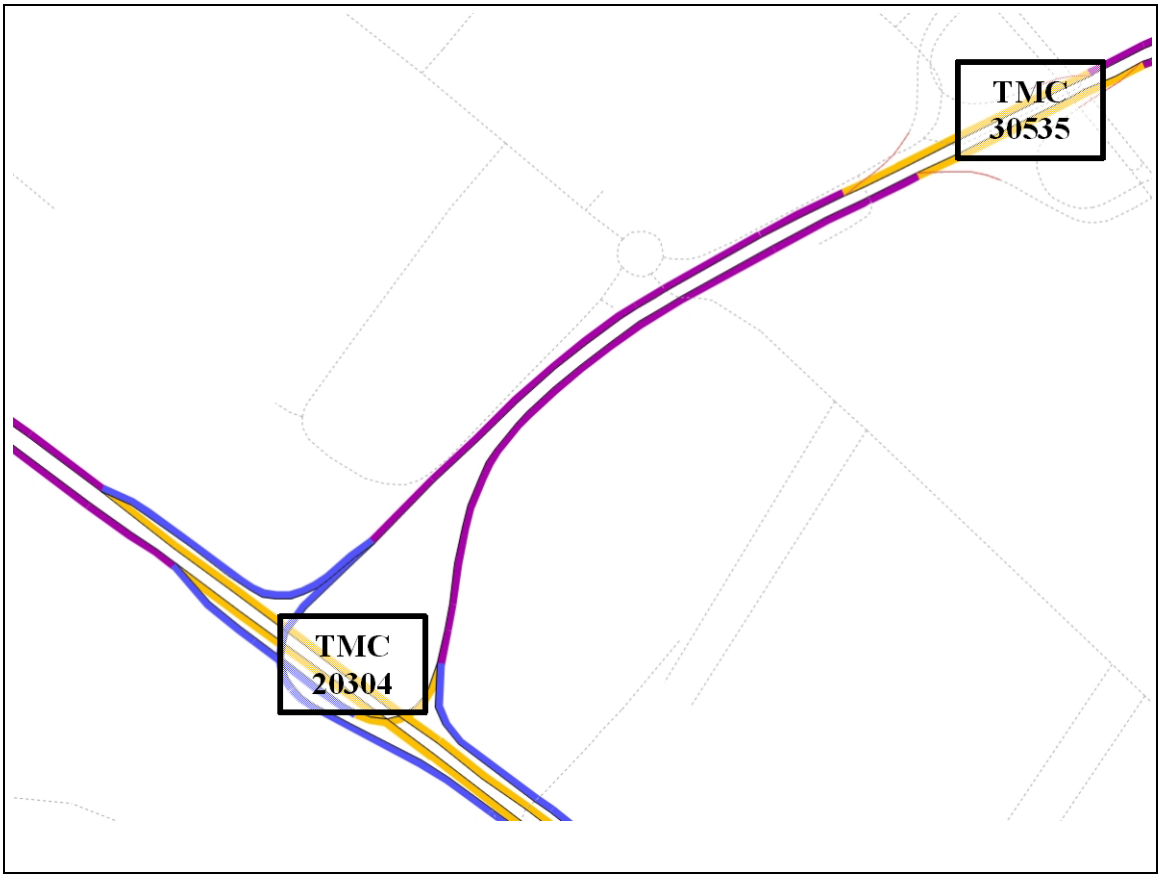


Abbildung 4: Nicht in das VPN referenzierbare Roadelemente am Beispiel des Knotens B10/B14-Uferstraße bei Stuttgart.

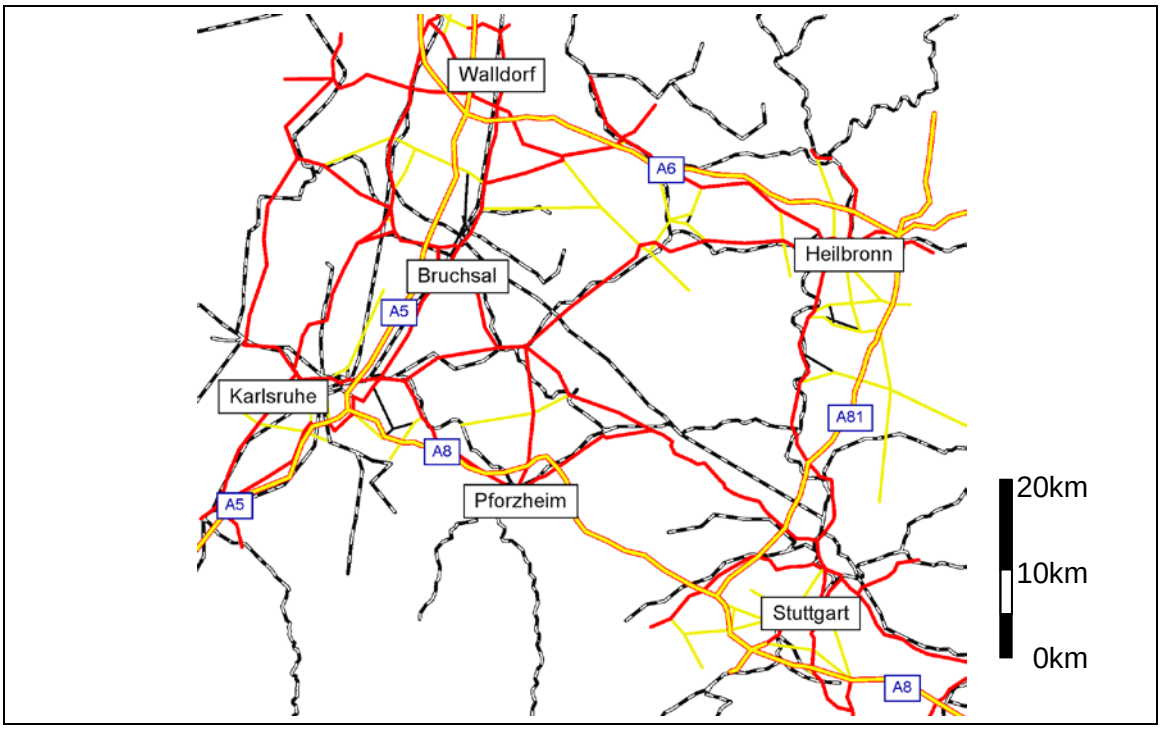


Abbildung 5: Vereinfachtes Projektnetz.

Das VPN (Abbildung 5) besteht aus 456 Knoten, zwischen denen sich 1128 Strecken mit einer Gesamtlänge von 2.815 km aufspannen. Es bildet dabei alle Strecken des Projektnetzes ab und bietet somit für die in diesem Bericht vorgestellten Verfahren eine geeignete Grundlage. Das ebenfalls in der Abbildung gezeigte Schienennetz ergibt sich aus den im Untersuchungszeitraum durch Züge des Fernverkehrs und Nahverkehrs (inklusive NE-Bahnen) befahrenen Strecken.

2.2 Zähldatenaufbereitung

Quellen von Zähldaten

An den Richtungsmessquerschnitten (RMQ) stehen Daten aus stationären Detektoren in Do-iT von insgesamt 5 Quellen zur Verfügung:

- Induktionsschleifen auf Autobahnen (Quelle: VRZ Ludwigsburg)
- Traffic-Eye-Daten der linken Fahrstreifen auf Autobahnen hochgerechnet auf den RMQ (Quelle: DDG)
- Induktionsschleifen auf Straßen der Stadt Stuttgart (Quelle: Landeshauptstadt Stuttgart)
- Induktionsschleifen auf Straßen der Stadt Karlsruhe (Quelle: Stadt Karlsruhe)
- Traffic-Eye-Daten des Richtungsquerschnitts auf Stadtstraßen in Karlsruhe (Quelle: Stadt Karlsruhe)

RMQ-Ganglinie $G(RMQ, Datum)$

Für das Projekt liegen die Messdaten je nach Quelle in unterschiedlichen Minutenintervallen aggregiert vor. Zur Vereinfachung wurden alle Messdaten zu 15-Minuten Intervallen Aggregiert.

Ein Datensatz einer RMQ-Ganglinie beinhaltet folgende Daten:

- RMQ-ID: Eindeutige ID der Zählstelle
- Datum: Datum im Format tt.mm.yyyy
- Q_{Kfz}: 96 Zählwerte für die aggregierte Verkehrsstärke je 15-Minuten-Intervall in [Kfz/h]
- V_{Kfz}: 96 Zählwerte für die gemittelte Durchschnittsgeschwindigkeit je 15-Minuten-Intervall in [km/h]
- Q_{Pkw}: 96 Zählwerte für die Verkehrsstärke von PKW je 15-Minuten-Intervall in [Kfz/h]
- Q_{Lkw}: 96 Zählwerte für die Verkehrsstärke von LKW je 15-Minuten-Intervall in [Kfz/h]
- V_{Pkw}: 96 Zählwerte für die Durchschnittsgeschwindigkeit von PKW je 15-Minuten-Intervall in [km/h]
- V_{Lkw}: 96 Zählwerte für die Durchschnittsgeschwindigkeit von LKW je 15-Minuten-Intervall in [km/h]

Weil bei einigen Quellen keine Unterscheidung zwischen LKW und PKW möglich ist und nur Zählwerte für Kfz übertragen werden, wird der LKW-Anteil mit statistischen Anteilen ermittelt, so dass alle RMQ-Ganglinien denselben Datenumfang besitzen.

Da die stationären Detektoren ihre Daten in unterschiedlichen Intervallen zusammenfassen und einige Systeme von Zeit zu Zeit ausfallen bzw. falsche oder inkonsistente Daten liefern (siehe Abbildung 6), wird der Dateneingang über ein eigenes Verfahren gefiltert und auf Verwendbarkeit überprüft. Nicht verwendbare Datenfragmente werden entfernt. Dies sind Messwerte bei denen eines der folgenden Kriterien zutrifft:

- durchschnittliche Geschwindigkeit > 200 km/h
- durchschnittliche Geschwindigkeit = 0 km/h
- durchschnittliche LKW-Geschwindigkeit > durchschnittliche PKW-Geschwindigkeit
- Verkehrsstärke = 0 Kfz/h
- Verkehrsstärke > 2.500 Kfz/h je Fahrstreifen

Im nächsten Schritt werden aus den Messdaten RMQ-Ganglinien erzeugt, die je Tag und RMQ genau einem Wert für je eines der 96 15-Minuten-Intervall des Tages besitzen. Dabei ist die Vollständigkeit ein ausschlaggebendes Kriterium. Fehlen in den RMQ-Ganglinien einzelne Intervallwerte, werden diese aus ihren Vorgängern und

Nachfolgern interpoliert, fehlen hingegen mehrere aufeinander folgende Intervallwerte, wird die gesamte RMQ-Ganglinie für den einen Tag und RMQ als nicht verwertbar markiert und anschließend gelöscht.

In Abbildung 6 ist die Verkehrsstärke einer RMQ-Ganglinie dargestellt, für die in einem etwa 3-stündigen Intervall keine Daten vorliegen hat. Eine Interpolation der Lücke ist nicht sinnvoll, die RMQ-Ganglinie ist nicht verwertbar und wird nicht in die Datenbank zu weiteren Auswertungen übernommen

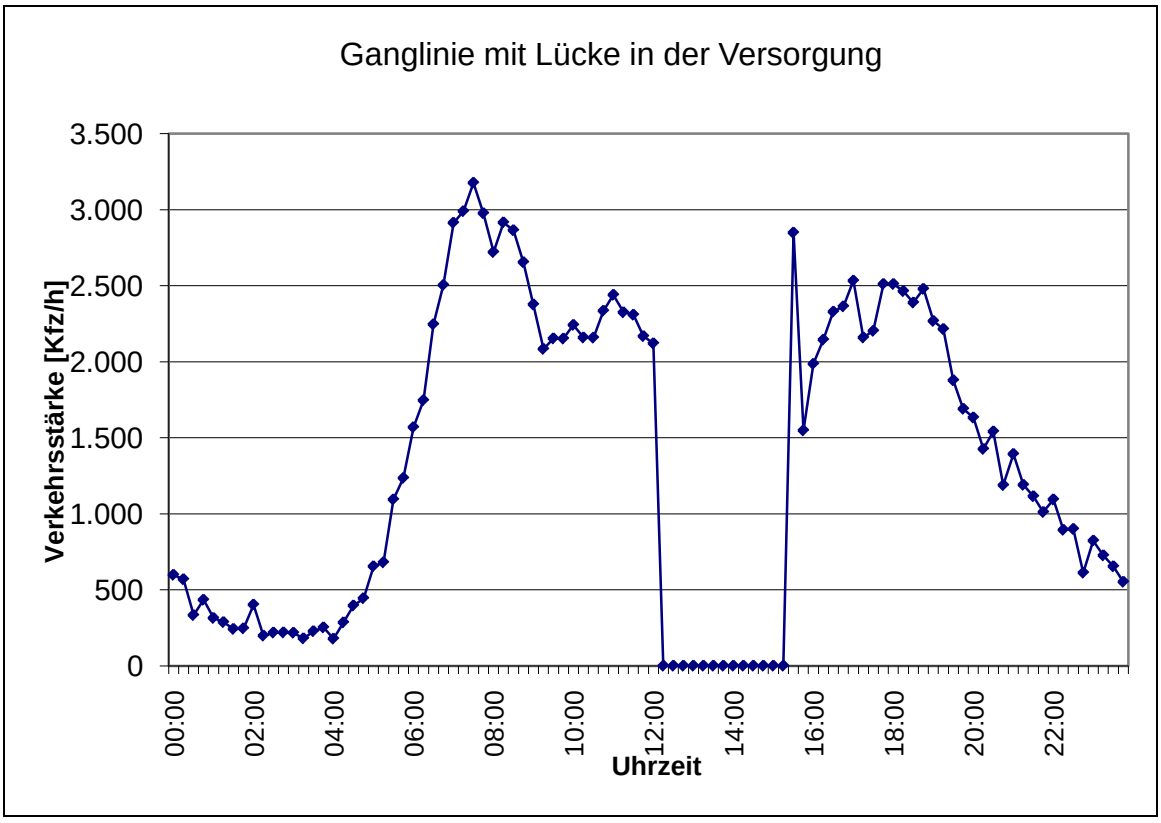


Abbildung 6: Inkonsistente Ganglinie durch Versorgungslücke von mehr als einem 15-Minuten-Intervall.

Datumsspezifische Attribute

Jede verwertbare – also vollständige – RMQ-Ganglinie bekommt entsprechend ihrem Datum folgende Eigenschaften zugeordnet:

- Wochentag (Montag =1, Sonntag = 7)
- Feiertag
- Brückentag
- Ferientag
- Ferienanfang und –ende in Baden-Württemberg
- Jeweils Ferienanfang und –ende in den Bundesländern Saarland, Rheinland-Pfalz, Nordrhein-Westfalen, Hessen und Bayern.

Das Attribut „Ferienanfang“ ist bei allen Tagen gesetzt, die zwischen einschließlich dem letzten Werktag ohne Ferien bis zum einschließlich darauf folgenden Sonntag liegen. Ist der Ferienanfang z.B. am Donnerstag, den 24. Juli 2008, so sind alle Tage vom Mittwoch, den 23. Juli 2008 bis zum Sonntag, den 27. Juli 2008 als Ferienanfang gezählt. Das Attribut „Ferienende“ bekommen die letzten 3 Tage (Freitag, Samstag, Sonntag) vor dem ersten Werktag ohne Ferien. Enden z.B. die Ferien am Samstag, den 6. September 2008, ist der Montag, der 8. September der erste Werktag ohne Ferien. Mit „Ferienende“ werden somit die 3 Tage Freitag, den 5.September bis Sonntag, den 7. September markiert.

Anhand einer Teilmenge dieser Attribute kann jedem Kalendertag ein TagTyp zugeordnet werden. Der TagTyp ist eine eindeutige Nummer, der die kalendarischen Eigenschaften in Form von einer TagTyp-ID speichert. Kalendertage mit gleicher TagTyp-ID haben den gleichen vorher festgelegten Eigenschaftssatz, wie z.B. <„Montag“;„Feiertag“;„Ferien“>.

Netzganglinie $G\left(\overrightarrow{RMQ}, Datum\right)$

Eine Zusammenfassung von mehreren RMQ-Ganglinien an genau einem Tag wird als Netzganglinie abgelegt. Eine Netzganglinie kann man sich als eine Aneinanderreihung der Zählwerte aller RMQ-Ganglinien vorstellen, wobei die Werte der nicht versorgten RMQ nicht gesetzt werden.

2.3 Binnenverkehrsproblematik

Mit Floating Phone Daten ist es nicht möglich, Fahrten jeder beliebigen Länge zu erfassen. Je kürzer eine Fahrt ist, desto unwahrscheinlicher ist eine erfolgreiche Ortung. Folgende Fahrten, die in diesem Bericht als „Binnenverkehr“ bezeichnet werden, können nicht erfasst werden:

- A-bis-Ebene:
  1. Fahrten innerhalb ein Funkzelle oder eine Location Area, während denen kein Telefonat geführt wird.
  2. Fahrten mit Telefonat, bei denen die Länge nicht ausreicht, um eine zuverlässige Ortung durchzuführen.
- A-Ebene:
  3. Fahrten ohne Telefonat innerhalb einer Location Area bzw. mit nur einem Location Area Update.
  4. Fahrten mit Telefonat innerhalb einer Funkzelle bzw. mit nur einem Handover.
  5. Fahrten mit zwei oder mehr Ortungspunkten (Handover, Location Area Update), die aber nicht ausreichen, um eine zuverlässige Ortung durchzuführen.

Die folgende Abbildung 7 zeigt drei der fünf oben genannten Fälle, bei denen die Ortung nicht erfolgreich durchgeführt werden kann:

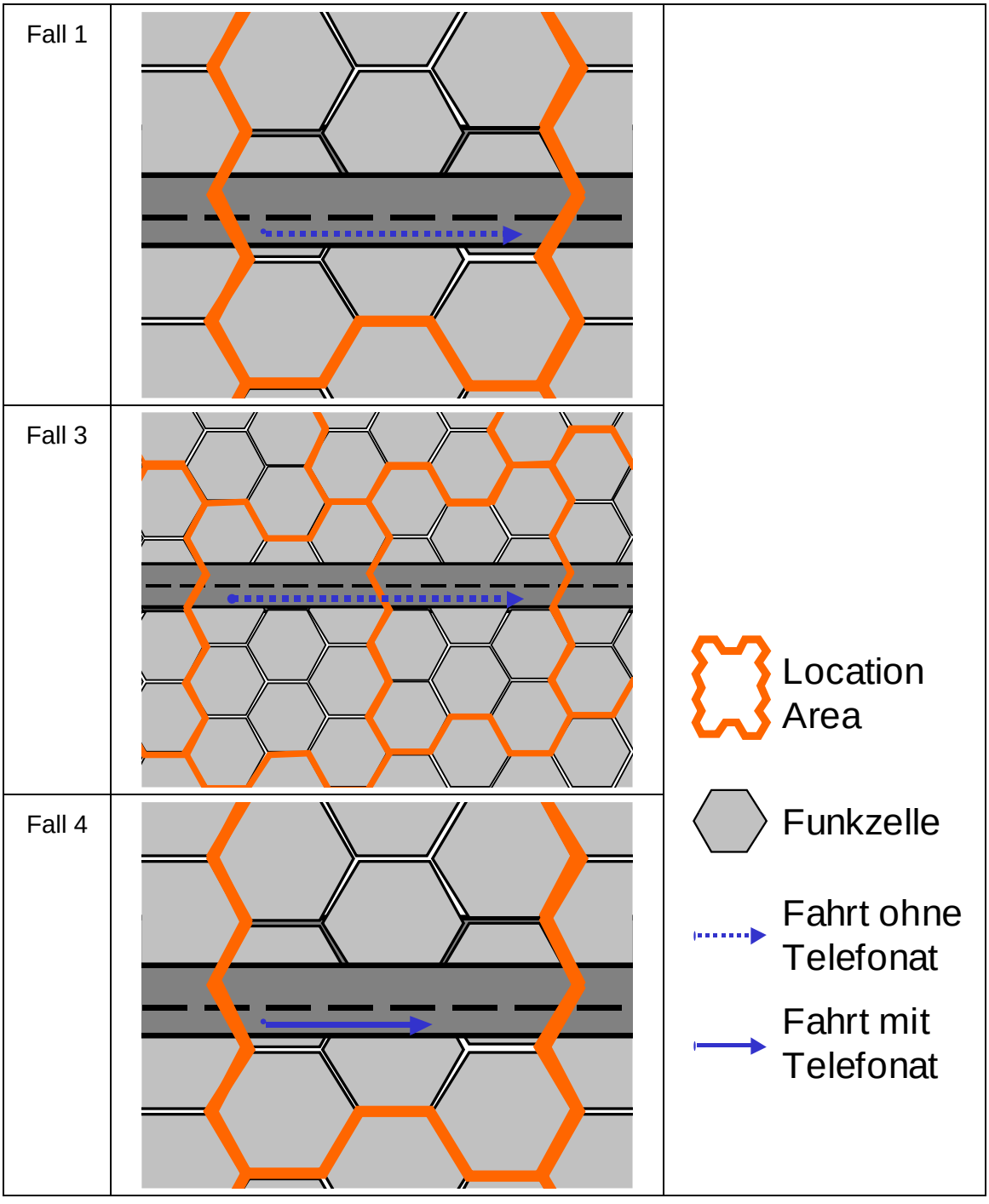


Abbildung 7: Nicht erfassbare Fahrten (Fälle 1, 3 und 4).

Mobiiltätsuntersuchungen wie „Mobilität in Deutschen“ (MID 2002) können dabei als Referenzgröße dienen. Solche Untersuchungen geben unter anderem Werte für die Fahrtweitenverteilung an (vgl. Abbildung 8).

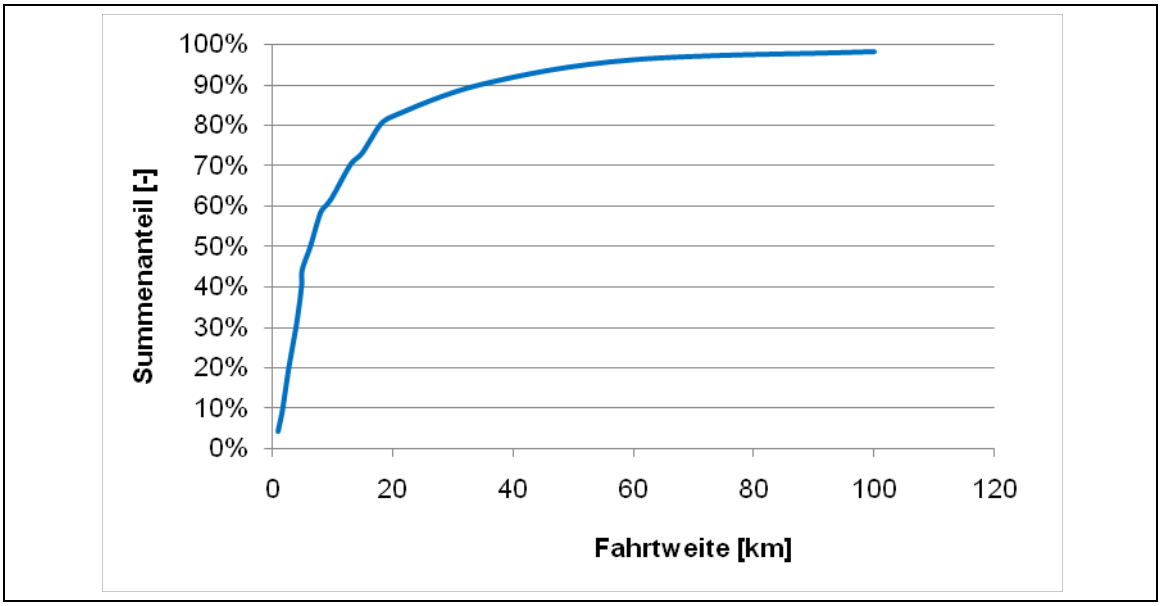


Abbildung 8: Mittlere Fahrtweiten im MIV (MID 2002).

Allerdings gilt diese Fahrtweitenverteilung global für das gesamte Verkehrsnetz, Werte für einzelne Streckentypen (z.B. Autobahn, Bundesstraße in Ballungsräumen) liegen nicht vor und müssen daher im Projekt Do-iT abgeschätzt werden. Auf den Strecken des VPN und den zugehörigen Zählstellen kann von höheren Fahrtweiten ausgegangen werden, da es sich bei diesen Strecken um höherwertige Strecken handelt.

Der konkrete Umgang mit den nicht durch FPD-Trajektorien erfassbaren Fahrten muss in Abhängigkeit der Anwendung erfolgen (vgl. Kapitel 5.2.2.4)..

2.4 Routenbaumgenerator

Um aus den auf der A-Ebene erfassten Mobilfunkdaten Trajektorien zu erstellen, werden die Location Area Folgen konkreten Routen im Verkehrsnetz zugeordnet. Dazu ist es erforderlich ein Choice Set aus Routen im Verkehrsnetz zu generieren. Für diese Aufgaben gibt es bereits verschiedene Ansätze, die im Folgenden kurz beschrieben. Anschließend wird der für dieses Projekt entwickelte Routenbaumgenerator beschrieben.

2.4.1 Übersicht über bestehende Ansätze

Labeling Approach

Dieses Verfahren basiert im Wesentlichen auf einer Kurzwegsuche. Um eine Vielzahl von Routen für eine Quelle-Ziel-Relation zu generieren, wird dabei die Funktion für den Streckenwiderstand gemäß verschiedener „labels“ variiert:

Label	Streckenwiderstandsfunktion
Minimale Zeit	Zeit
Minimale Entfernung	Entfernung
Landschaftlich schönste Route	$Zeit \cdot (1 + ß_1 \cdot Anteil_{\text{unattratkive Strecken}})$
Minimale Anzahl Lichtsignalanlagen	$Zeit + ß_2 \cdot N_{\text{Lichtsignalanlagen}})$
Minimale Staulänge	$Zeit \cdot (1 + ß_3 \cdot Dummy_{\text{hohe Auslastung}})$
Minimale Nutzung von Schnellstraßen	$Zeit \cdot (1 + ß_4 \cdot Dummy_{\text{Schnellstraße}})$
Maximale Nutzung von Strecken mit hoher Kapazität	$Zeit \cdot (1 + ß_5 \cdot Dummy_{\text{hohe Kapazität}})$
...	...

Tabelle 2: Beispiele für Labels (Ben-Akiva et al, 1984, Tabelle 2).

Je nach Label ergeben sich unterschiedliche Routen, die dann das Choice Set bilden. Für die Generierung eines vollständigen Choice Set für die Routenwahl ist dieser Ansatz ungeeignet, da keine ausreichende Anzahl von Routen für einen Quelle-Ziel-Relation generiert wird.

K-verschiedene Routen

Bei diesen Ansätzen soll eine Anzahl k von Wegen gefunden werden, die bestimmten Kriterien genügen. Häufig sollen dabei die k (zeit-)kürzesten Wege gefunden werden. Da es in Verkehrsnetzen zwischen einer Quelle und einem Ziel häufig sehr viele verschiedenen Routen gibt, die nur schwer alle zu berechnen sind, gibt es für diesen Ansatz vor allem heuristische Verfahren:

- Link Elimination Approach (z.B. Azevedo et. al.,1993):
  Nach einer Kurzwegsuche wird eine Route gespeichert, anschließend wird nach vordefinierten Regeln eine Strecke der Routen vor der nächsten Kurzwegsuche entfernt.
- Link Penalty Approach (z.B. de la Barra et. al., 1993):
  Dieses Verfahren ist ähnlich dem Link Elimination Approach, nur dass hier die Strecken nicht komplett entfernt wird, sondern nur ausgewählten Strecken (z.B. alle Strecken der letztgefunden Route) ein höherer Widerstand zugewiesen wird.

- Simulation Approach (z.B. Sheffi &Powell, 1982):
  Bei diesem Ansatz werden die Widerstände der Strecken bei jeder Kurzwegsuche stochastisch verändert, so dass sich verschiedene Routen ergeben.

Darüberhinaus gibt es auch mathematische Verfahren, die k kürzesten Wege zu finden. Diese Ansätze führen allerdings häufig zu Wegen, die sich nur minimal voneinander unterscheiden. Einige Ansätze sowie weiterführende Literatur ist unter www.optiv.de (Bastian et al., 2008) abgedruckt.

2.4.2 „Vollständige“ Enumeration

Um ein umfassendes Bild über mögliche Routen zwischen einer Quelle und einem Ziel zu haben, wäre es wünschenswert, alle Routen zu finden. Je nach Zielsetzung der jeweiligen Untersuchung könnten anschließend die geeigneten Routen in das Choice Set übernommen werden. Die Methode der nvollständige Enumeration soll anhand eines einfachen Verkehrsnetzes in Abbildung 9 verdeutlicht werden.

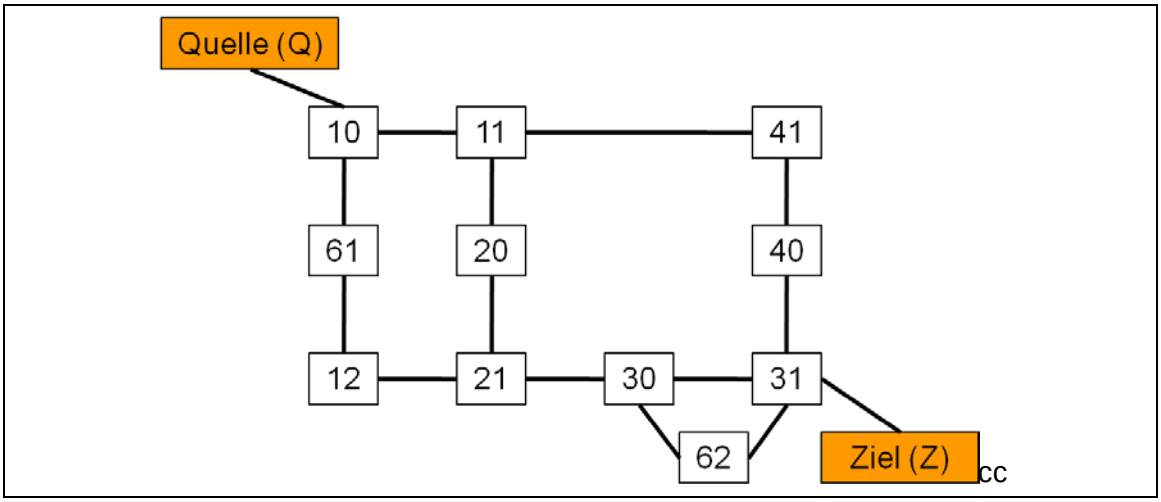


Abbildung 9: Beispielnetz.

Wie bei einer Kurzwegsuche wird auch bei der vollständigen Enumeration in einem ersten Schrift für jeden Knoten eine Nachfolgliste erstellt, die alle Knoten enthält, die direkt von k aus erreicht werden können. Anschließend wird für jede Quelle ein Routenbaum aufgespannt. Abbildung 10 zeigt den entstehenden Routenbaum unter der Bedingung dass innerhalb von einer Route ein Knoten nur einmal vorkommen darf.

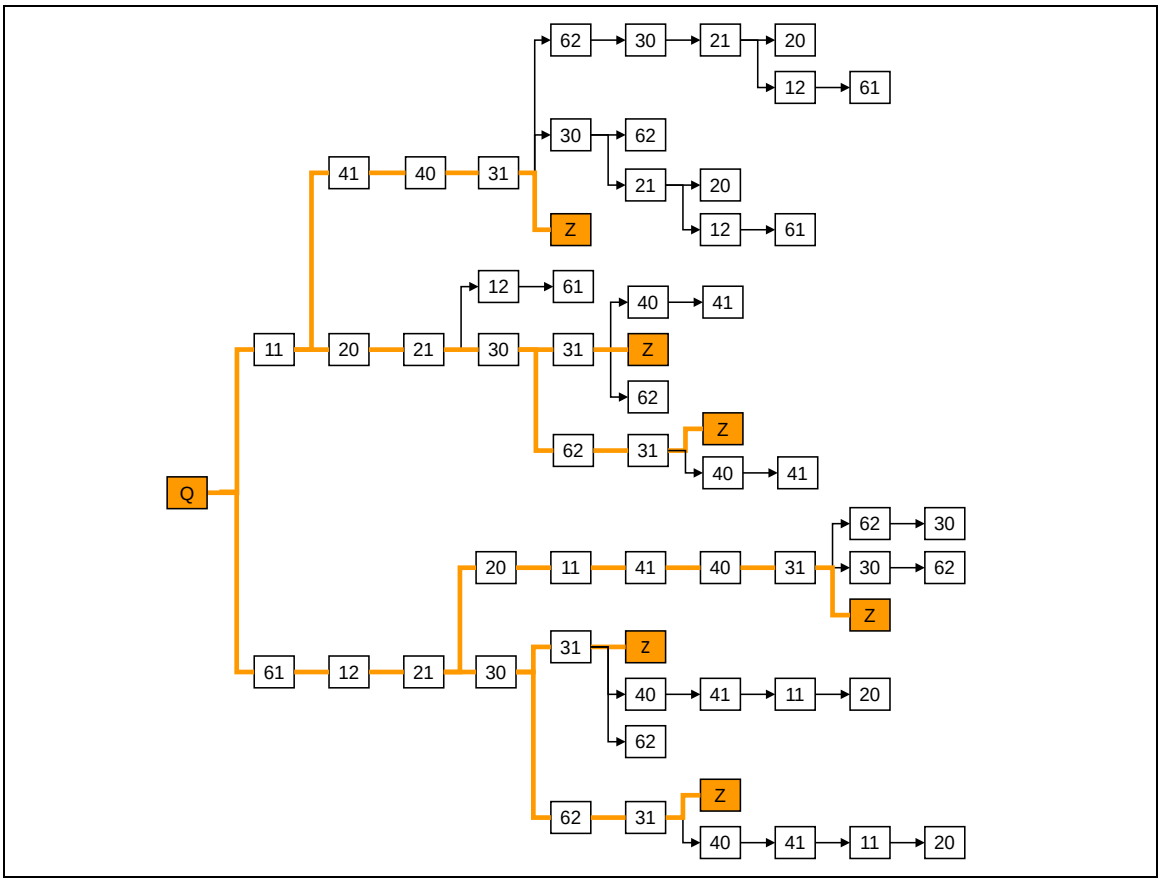


Abbildung 10: Routenbaum ausgehend von der Zelle 1.

Farblich hervorgehoben sind die Routen vom Zelle 1 zur Zelle 99, die auch noch einmal auf der nächsten Seite im Verkehrsnetz visualisiert sind.

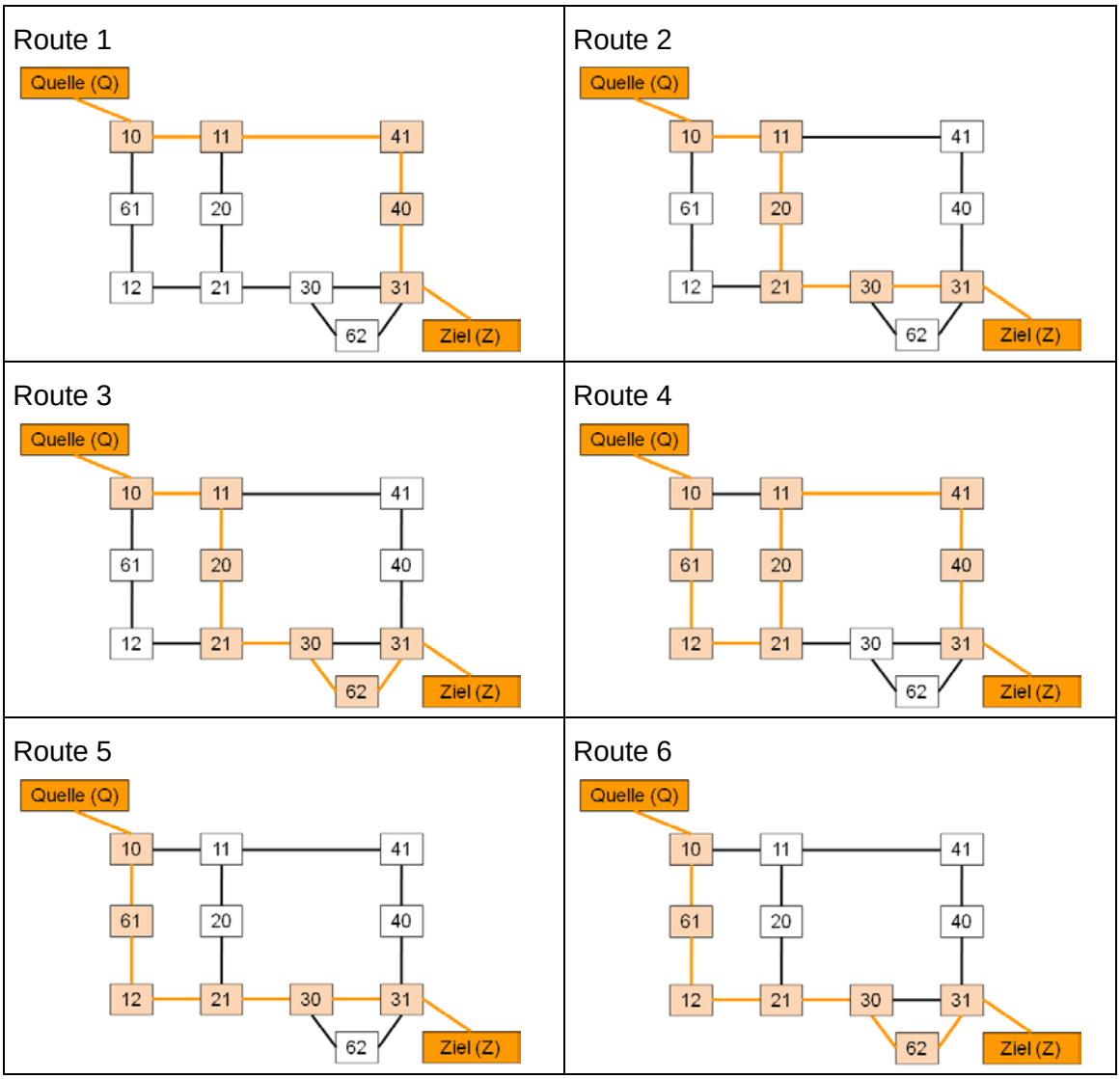


Abbildung 11: Sechs Routen zwischen Zelle 1 und Zelle 99.

Dieses Beispiel lässt bereits ahnen, dass eine vollständige Enumeration aller Routen in einem großen Verkehrsnetz nicht sinnvoll bzw. machbar ist, da es zu viele mögliche Routen zwischen zwei Knoten eines komplexeren Verkehrsnetzes gibt.

Daher wird bei der Generierung des Routenbaums beim Hinzufügen jeden Knotens geprüft, ob er den folgenden Bedingungen genügt:

Keine Kreisfahrten

Der hinzuzufügende Knoten darf auf der aktuellen Route nicht vorkommen. Dafür wird der Routenbaum von der aktuellen Position rückwärts bis zur Quelle durchlaufen und die alle auf dem Routenbaumast vorhandenen Knotennummern mit hinzuzufügenden Knoten verglichen.

Maximaler Umweg vom Start zum aktuellen Knoten

Der Widerstand bis zum hinzuzufügenden Knoten auf der jeweils betrachteten Route wird mit dem vorab berechneten Widerstand $w_{min}(n)$ auf dem kürzesten Weg verglichen. Dabei wird ein Knoten n nur dann hinzugefügt, wenn der aktuelle Widerstand $w_{akt}(n)$ kleiner oder gleich einem berechneten maximalen Widerstand $w_{max}(n)$ ist. Die Berechnung von $w_{max}(n)$ erfolgt dabei typischerweise in Abhängigkeit von $w_{min}(n)$, z.B.:

$$w_{akt}(n) \leq w_{max}(n) = c + f \cdot \left(w_{min}(n) - g \right)$$

mit

$w_{akt}(n)$	Widerstand auf der aktuellen Route von der Quelle bis zum Knoten n
$w_{max}(n)$	Zulässiger Widerstand von der Quelle bis zum Knoten n
$w_{min}(n)$	Widerstand einer Kurzwegsuche zwischen der Quelle bis zum Knoten n
c, f, g	Parameter

Maximaler Umweg auf dem hinteren Streckenabschnitt

Zusätzlich zum zweiten Abbruchkriterium wird ab einer bestimmten Länge auch der hintere Teil der Route auf einen maximalen Umweg geprüft. Dies soll verhindern, dass durch den Parameter f z.B. auf einer langen Autobahnroute ein zu großer Puffer für Umwege entsteht, der dann im hinteren Teil für große Umwege genutzt wird.

3 Trajektoriengenerierung

In diesem Kapitel wird ein Verfahren vorgestellt, das aus ausgewählten Daten auf der A-Ebene des Mobilfunknetzes Trajektorien einzelner Mobilfunkgeräte erzeugt. Dabei werden vorwiegend die Informationen über die aktuelle Location Area verwendet, funkzellenfeine Daten werden nur im nachgeschalteten Schritt berücksichtigt. Dies ermöglicht es, mit dem Routenbaumgenerator (vgl. Kapitel 2.4) erzeugte LAC-Folgen (Location-Area-Code-Folgen) mit den in den A-Daten aufgezeichneten LAC-Folgen zu vergleichen und daraus anschließend Trajektorien im VPN zu erzeugen. Dafür werden die in Abbildung 12 dargestellten und im Folgenden detailliert beschriebenen Schritte durchgeführt:

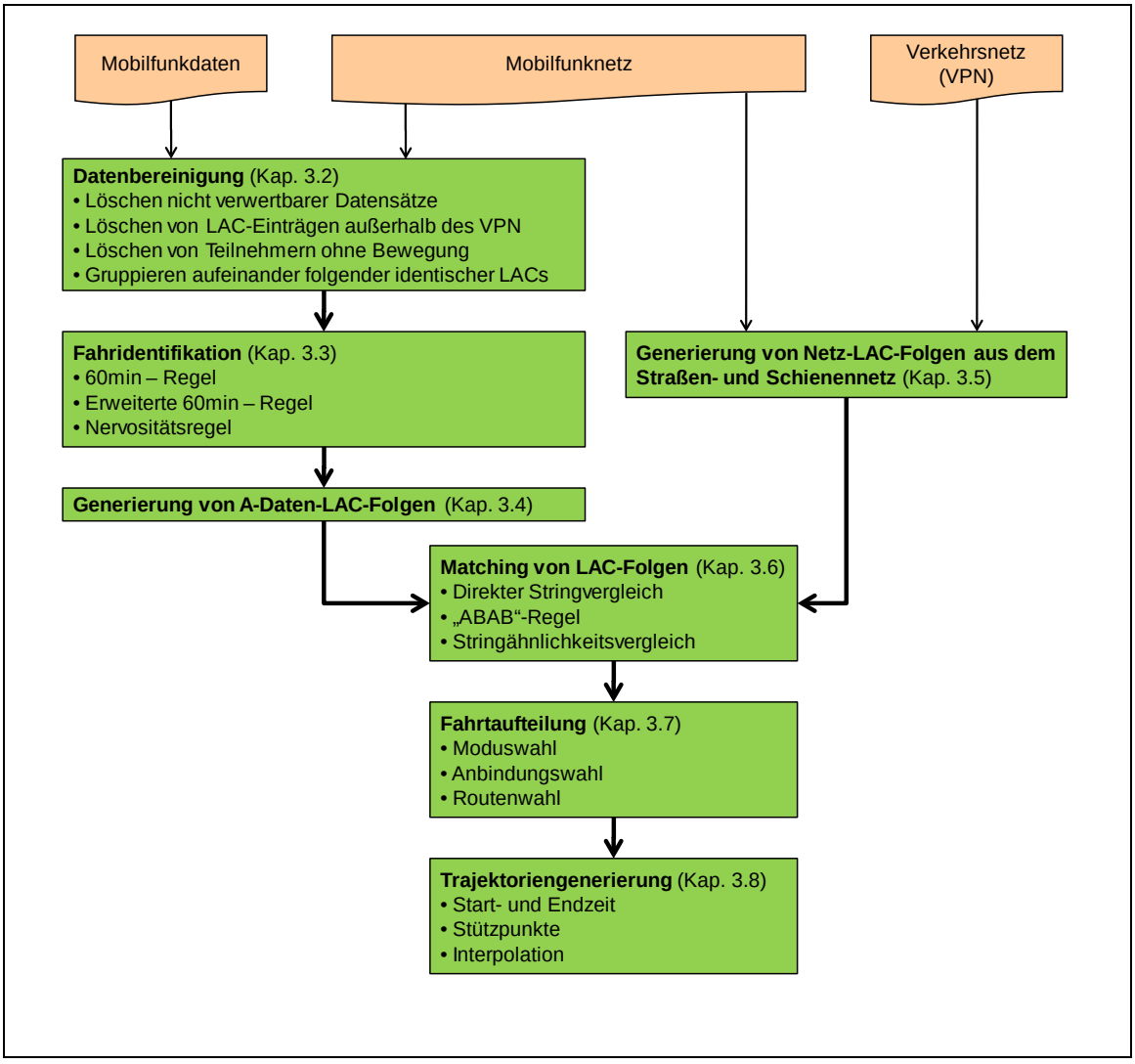


Abbildung 12: Ablauf der Generierung von Trajektorien aus A-Daten.

3.1 Beschreibung der A-Daten

Die A-Daten werden durch sogenannte Network Probes aus dem normalen Datenverkehr des Mobilfunknetzes auf der A-Ebene abgegriffen. Die Daten werden minutenweise in Text-Dateien gespeichert und enthalten die folgenden Attribute:

Name	Format	Beispiel	Beschreibung	Verwendet
UsrID	Long	12345	Tagesgültige UsrID eines Mobilfunkteilnehmers	Ja
t.s	Long	1207914695	Sekunden seit UNIX (GMT)	Ja
t.m	Int	65	Millisekunden seit „t.s“	Nein
type	Int	2	• 0 = Location Area Update (LUP) • 1..n-1 = Cell-ID-Wechsel während eines Telefonates • Nmax = 1024 = Gesprächsende	Nein
cLAC	Long	98765	Location Area Code der aktuellen Funkzelle	Ja
cCID	Long	43210	Cell ID der aktuellen Funkzelle	Ja
pLAC	Long	12345	Location Area Code der vorherigen Funkzelle	Nein
pCID	Long	65432	Cell ID der vorherigen Funkzelle	Nein

Tabelle 3: Vollständige Beschreibung der A-Daten.

Wie bereits erwähnt werden für die Generierung der Trajektorien vor allem die LACs verwendet, da das verwendete Verfahren kontinuierliche Übergänge zwischen den Location Areas benötigt und Sprünge zwischen den feiner aufgelösten Cell-IDs nicht verarbeiten kann. Darüberhinaus besteht die Gefahr, dass die Verwendung der Cell-IDs zu einer Scheingenauigkeit führt (siehe dazu Kapitel 3.10).

Da die Einträge in den Spalten pLAC bzw. pCID beim ersten Auftreten einer UsrID meistens Leereinträge bzw. Einträge von LACs außerhalb des Untersuchungsgebietes sind und bei den folgenden Einträgen redundante Informationen (pLAC/pCID entsprechen cLAC/pCID des vorherigen Eintrags) enthalten, werden diese Daten nicht weiter berücksichtigt. Auch die Angabe der Millisekunden wird zu Beginn der Datenverarbeitung gelöscht, da eine solche Genauigkeit das Ergebnis nicht verbessert, dagegen aber den erforderlichen Speicherplatz erhöht.

3.2 A-Daten-Datenbereinigung

Neben einer Reduzierung der verwendeten Eigenschaften je Datensatz werden nun in vier Schritten nicht benötigte Datensätze gelöscht:

- Löschen von nicht verwertbaren Datensätzen,
- Löschen von Mobilfunkteilnehmern ohne ausreichende Bewegung,
- Gruppieren aufeinander folgender identischer LACs und
- Verlassen des relevanten Netzes.

3.2.1 Löschen von nicht verwertbaren Datensätzen

Zum einen werden einzelne Datensätze – unabhängig von den anderen Datensätzen der UsrID gelöscht, die einem der beiden folgenden Kriterien genügen:

- Ungültige Einträge bei den verwendeten Spalten der A-Daten
  Dies betrifft im Wesentlichen den UNIX-Zeitstempel. Dieser ist für die Datensätze während aktiver Verbindungen, dabei insbesondere am Ende einer Verbindung öfters mit der Wert 0 („ungültig“) angegeben.
- Zeitstempel außerhalb des Untersuchungstages:
  Vereinzelt laufen in den A-Daten Einträge auf, die einige Stunden bis wenige Tage veraltet sind. Sofern diese außerhalb des Untersuchungstages liegen, werden sie gelöscht.

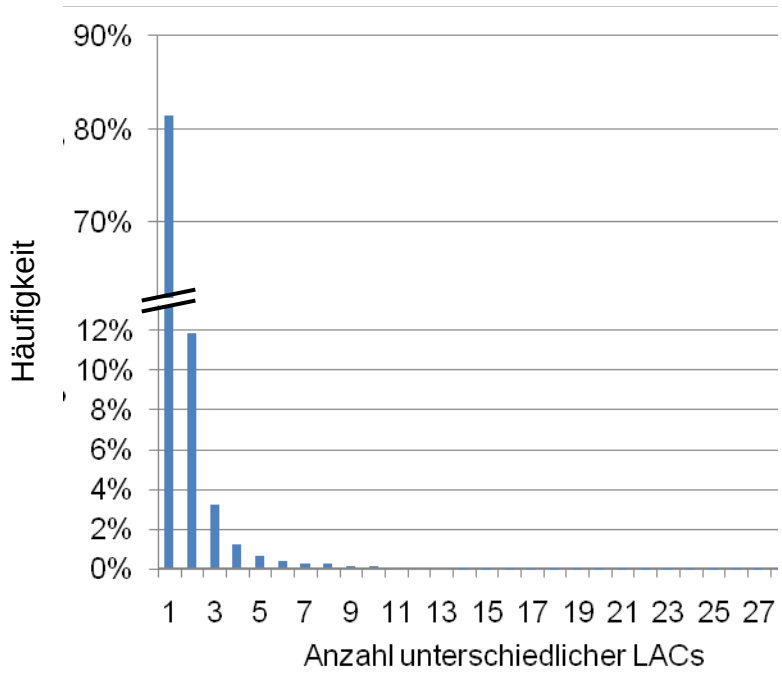
Darüberhinaus werden aus Gründen einer effizienten Datenverarbeitung alle Einträge einer UsrID gelöscht, wenn

- die UsrID mehr als 255 Location Area Wechsel hat, oder
- zwei Location Area-Wechsel einer UsrID innerhalb von einer Sekunde stattfinden.

Die Anzahl der durch diese beiden Kriterien gelöschter UsrID ist vernachlässigbar klein (i.d.R. 0-10 UsrIDs pro Tag), dafür vereinfacht die Löschung dieser UsrID die Datenstruktur erheblich.

3.2.2 Löschen von Mobilfunkteilnehmern ohne ausreichende Bewegung

Ein Großteil der Mobilfunkteilnehmer halten sich während des gesamten Beobachtungszeitraums in genau einer oder zwei unterschiedlichen LACs auf:



Anzahl untersch. LACs	Häufigkeit
1	81,52%
2	11,87%
3	3,24%
4	1,24%
5	0,63%
6	0,38%
7	0,26%
8	0,24%
9	0,16%
10	0,16%
11	0,09%
12	0,06%
>12	0,16%

Abbildung 13: Häufigkeiten der UsrID über die Anzahl unterschiedlicher LACs.

Wie in Kapitel 2.3 verdeutlicht, sind mindestens drei unterschiedliche LACs für eine Generierung einer Trajektorie nötig. Daher werden in diesem Schritt alle Datensätze von Mobilfunkteilnehmern mit ein oder zwei unterschiedlichen LACs gelöscht.

3.2.3 Gruppieren aufeinander folgender identischer LACs

Während eines Telefonates werden auf der A-Ebene mehrere Einträge in den A-Daten erzeugt, auch wenn dabei kein LAC-Wechsel stattgefunden hat. Diese Informationen sind bei dem hier vorgestellten Verfahren, das vor allem auf der Betrachtung von LAC-Wechseln basiert, nicht relevant. Daher werden diese Datensätze unter Beibehaltung des ersten Eintrags jeder Location Area gelöscht.

3.2.4 Verlassen des relevanten Netzes

Die A-Daten werden für 53 Location Areas geliefert. Davon haben 41 einen räumlichen Bezug zum VPN („VPN-LAC“). Die Datensätze mit LACs ohne Bezug zum VPN („Fremd-LAC“) werden gelöscht. Dabei werden die im Folgenden beschriebenen Fälle unterschieden.

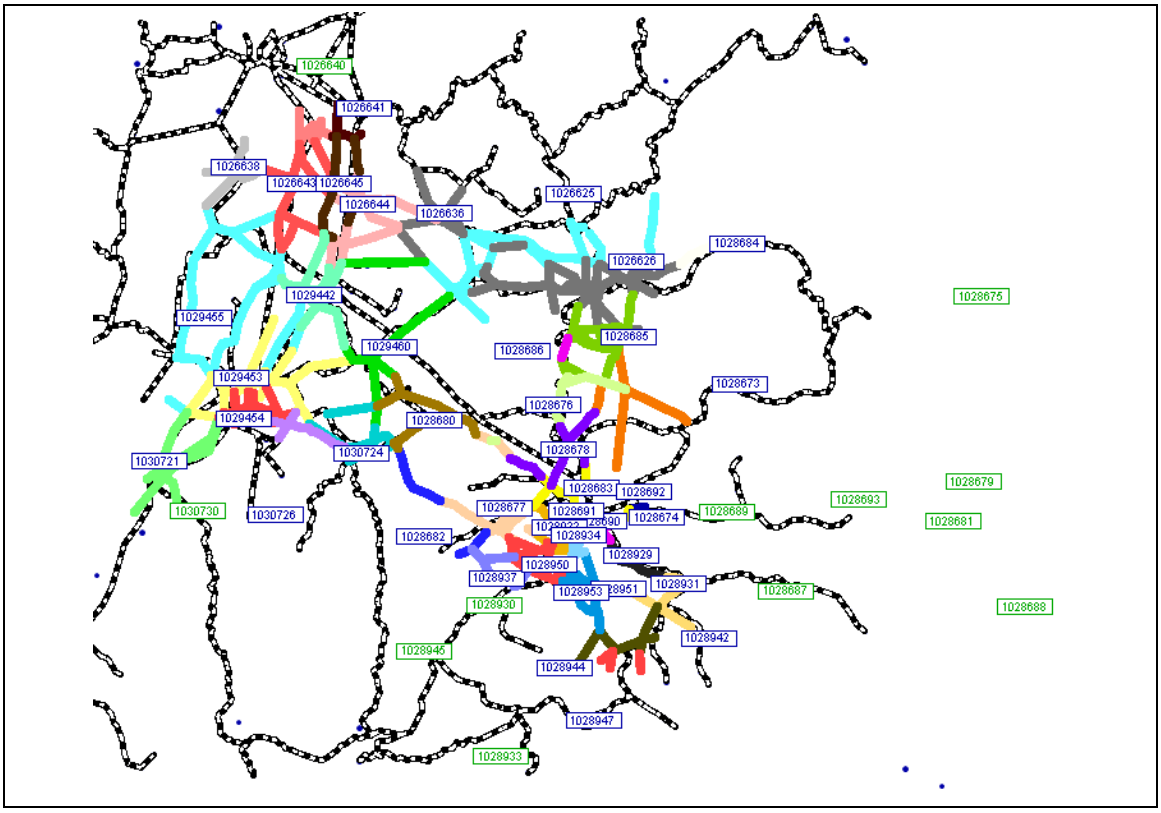


Abbildung 14: VPN und Schienennetz mit VPN-LACs (blau) und Fremd-LACs (grün), Einfärbung der VPN-Strecken nach Location Area.

Fremd-LACs zu Beginn bzw. am Ende einer UsrID

In den in Abbildung 15 verdeutlichten Fällen, in denen Mobilfunkgeräte zu Beginn bzw. am Ende ihren Bewegung Fremd-LACs haben, werden diese Einträge gelöscht. Die weiteren Einträge der betroffenen UsrID werden unverändert beibehalten, falls mindestens drei unterschiedlichen LACs in den restlichen Einträgen vorhanden sind.

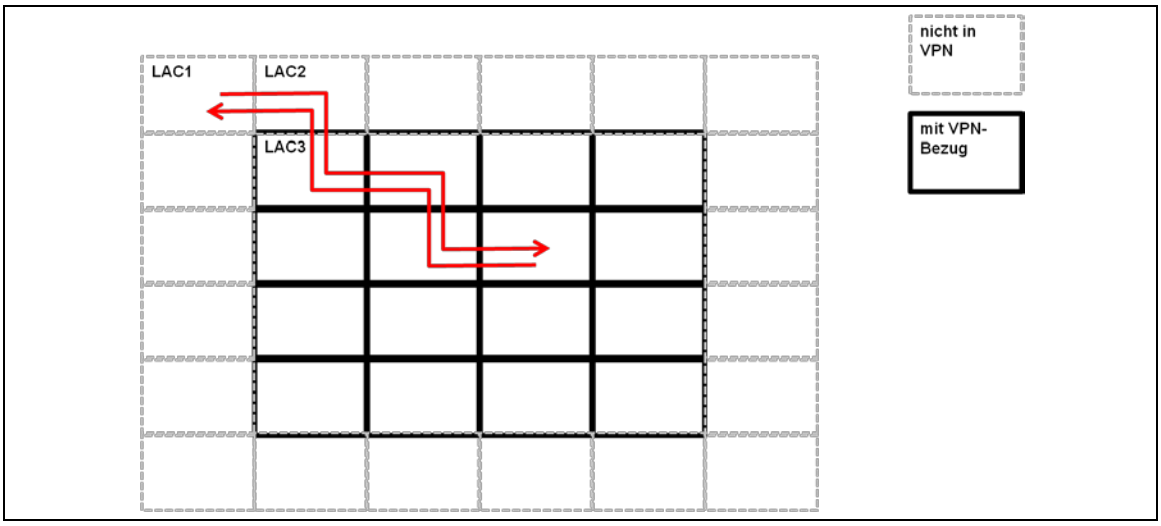


Abbildung 15: In das VPN hineingehende bzw. herausgehende Bewegung eines Mobilfunkteilnehmers.

Fremd-LACs zwischen VPN-LACs

Wenn ein Mobilfunkteilnehmer während seiner Bewegung das VPN verlässt, kann dies entweder eine Bewegung sein, die tatsächlich außerhalb des VPN stattfindet oder es handelt sich um ein kurzes Wechseln in eine Fremd-LAC aufgrund der Besonderheiten des Mobilfunknetzes.

Abbildung 16 zeigt ein Beispiel, bei dem die A-Daten eines Mobilfunkteilnehmers neben der LAC-Folge 6 – 5 – 4 zusätzlich einen oder mehrere Datensätze der Fremd-LAC 3 beinhalten.

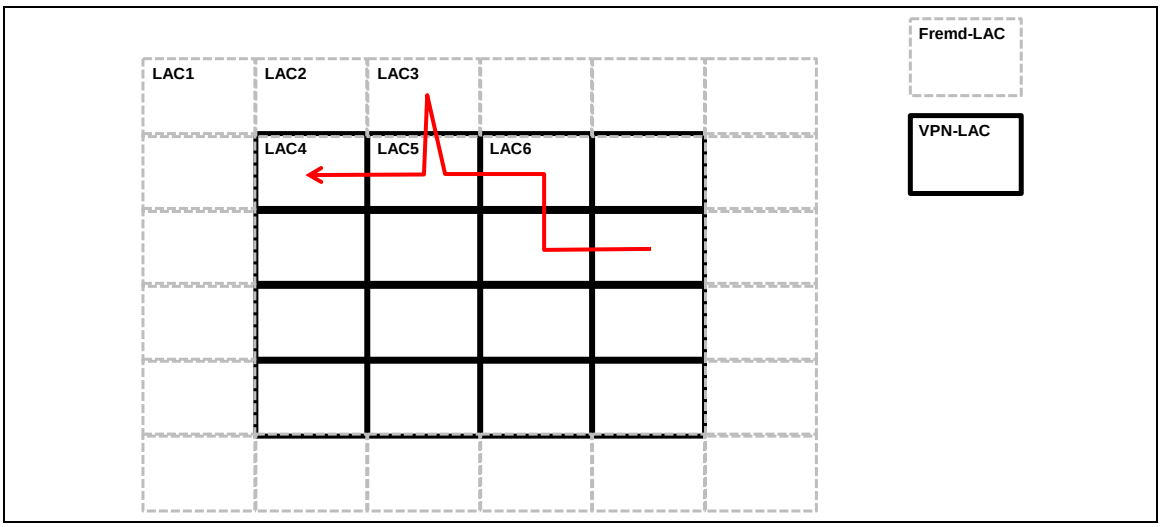


Abbildung 16: Eine Fremd-LAC zwischen VPN-LACs.

Die Einträge der Fremd-LAC werden gelöscht, dabei werden zwei Fälle unterschieden:

- Falls die Aufenthaltsdauer in der Fremd-LAC weniger als 10min beträgt, wird der Rest der Einträge nicht weiter verändert.
- Falls die Aufenthaltsdauer 10min oder mehr beträgt, wird davon ausgegangen, dass der Bereich des VPN tatsächlich verlassen worden ist. Deshalb werden in solchen Fällen zwei Fahrten für den Mobilfunkteilnehmer erzeugt. Die erste Fahrt von Beginn bis zum letzten VPN-LAC, d.h. vor Verlassen des VPN-Bereichs, sowie die zweite Fahrt vom ersten VPN-LAC nach Verlassen der Fremd-LAC also dem Wiedereintritt in den VPN-Bereich bis zum Ende.

Die Aufteilung auf zwei Fahrten wird unabhängig von der Aufenthaltsdauer auch dann durchgeführt, wenn zwei oder mehr unterschiedliche Fremd-LACs hintereinander in den A-Daten eines Mobilfunkteilnehmers sind (vgl. Abbildung 17).

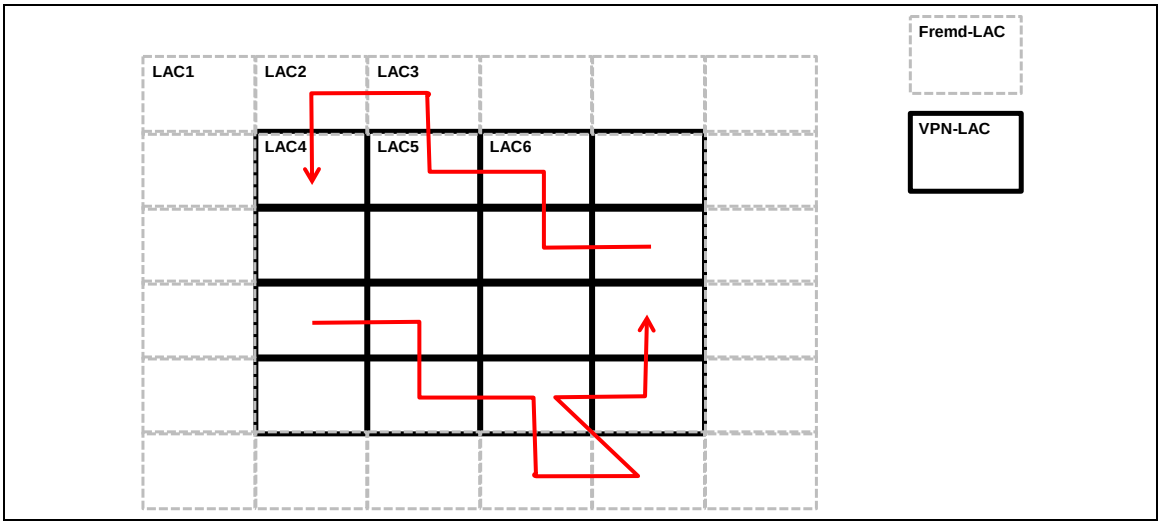


Abbildung 17: Zwei Beispiele mit zwei aufeinanderfolgenden Fremd-LACs.

3.3 Fahrtidentifikation

Mobilfunkteilnehmer können während der Aufzeichnung der A-Daten mehrere Fahrten im Straßen- oder Schienennetz durchführen, oder aber auch keine Fahrten machen. In diesem Kapitel werden daher drei Regeln beschrieben, nach denen aus den bereinigten A-Daten potentielle Fahrten identifiziert werden:

- 60min-Regel
- Erweiterte 60min-Regel
- Nervositätsregel

3.3.1 60min-Regel

Ein Mobilfunkteilnehmer, der sich deutlich länger in einer Location Area aufhält, als die direkte Durchfahrt unter Berücksichtigung von Stauereignissen dauert, hat in der jeweiligen Location Area einen potentiellen Fahrtbeginn bzw. –ende. Die Festlegung der Grenze, bei der eine Fahrt aufgetrennt wird, beeinflusst die Anzahl und Qualität der Trajektorien:

- Eine zu niedrige Grenze führt zu einer zu häufigen Trennung von Fahrten in zwei kürzere Fahrten.
- Eine zu hohe Grenze führt zu einer zu seltenen Trennung von Fahrten. Die daraus entstehenden langen Bewegungen können dann nicht als Trajektorie erkannt werden, wenn sie in ihrer Gesamtheit eine nicht identifizierbare Bewegung (z.B. Hin- und Rückfahrt) ergeben.

Abbildung 18 zeigt eine globale Betrachtung der Aufenthaltsdauern in den Location Areas. Knapp 25 % aller Aufenthaltsdauern sind unter 60 Sekunden, ca. 90 % liegen unter 20 Minuten. Danach nimmt der Anteil stetig ab, aber trotzdem sind noch knapp 5 % der Aufenthaltsdauern über 60 Minuten.

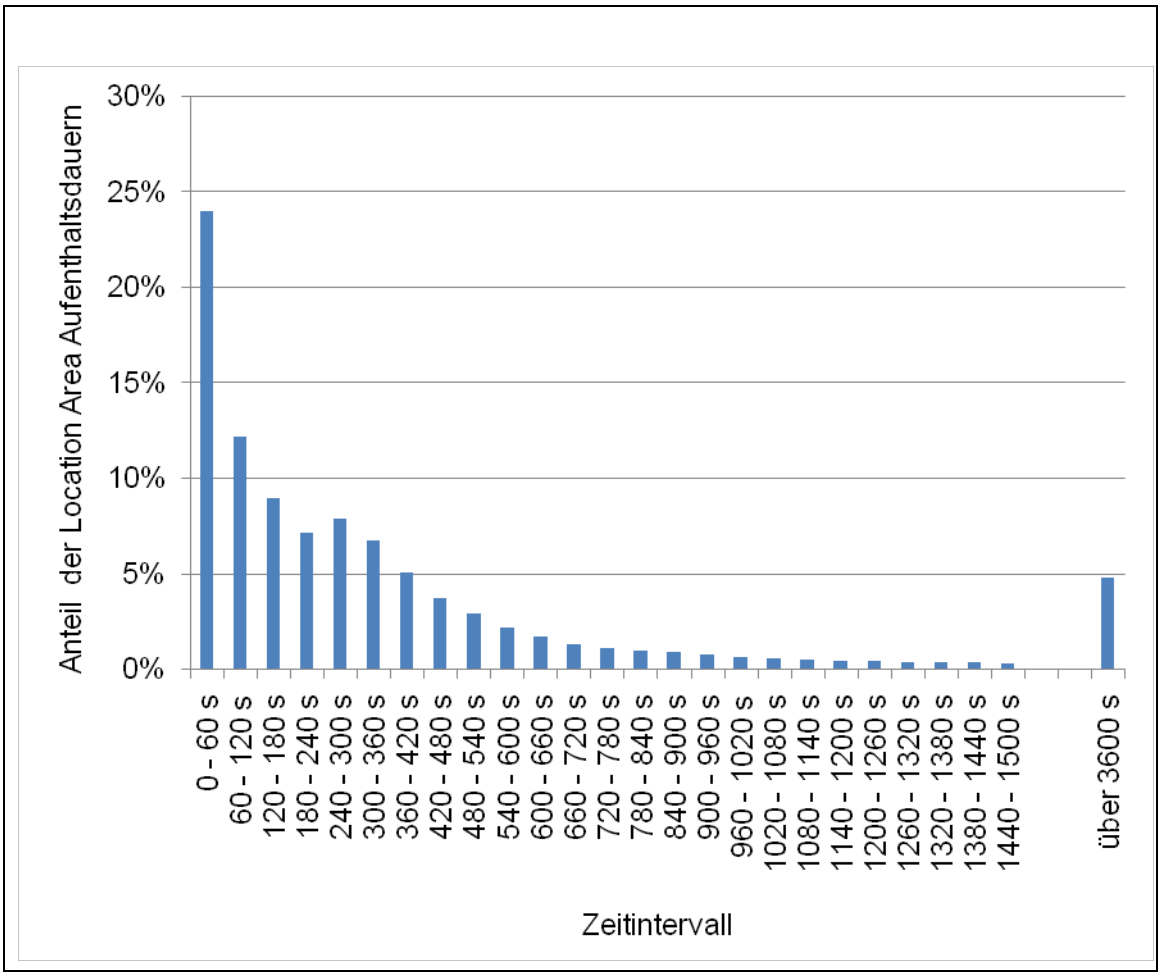


Abbildung 18: Aufenthaltsdauern in allen Location Areas.

Eine klare Grenze, ab der eine Bewegung eines Mobilfunkteilnehmers aufgeteilt werden sollte, ist dabei nicht zu erkennen. Auch eine LAC-weise Betrachtung in Abbildung 19 gibt keine weiteren Aufschlüsse über eine eventuelle LAC-feine Grenze, was sich auch aus der Vielzahl von möglichen Durchfahrtswegen durch eine Location Area ergibt.

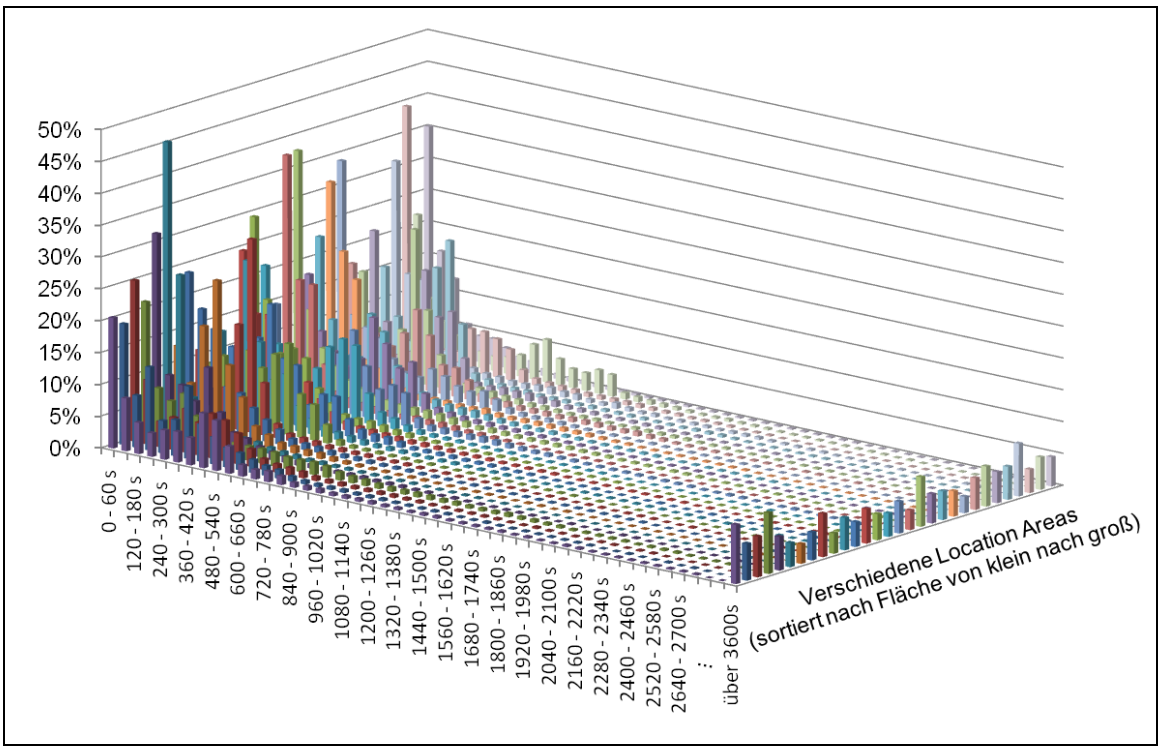


Abbildung 19: Aufenthaltsdauer je Location Area.

Daher muss die Grenze unter Berücksichtigung der LAC-Ausdehnungen und darauf resultierender Durchfahrtszeiten manuell festgelegt werden. Ein relativ hoher Wert von 60 Minuten ist dabei sinnvoll, da dadurch auch größere Stauereignisse abgedeckt sind und Fahrten nicht unnötig aufgetrennt werden. Im Gegenzug wird aber akzeptiert, dass fälschlicherweise nicht aufgeteilte Fahrten im späteren Verlauf nicht zu Trajektorien weiterverarbeitet werden können und somit die Anzahl der Trajektorien sinkt. Letzteres ist aufgrund der zu erwartenden hohen Anzahl von Trajektorien akzeptabel.

Die folgende Abbildung 20 zeigt schematisch die Anwendung der 60min-Regel anhand der LA-Aufenthaltsdauern entlang einer Bewegung eines Mobilfunkteilnehmers. In LAC 2 ist die Aufenthaltsdauer über der festgelegten Grenze von 60 Minuten. Daher wird hier davon ausgegangen, dass in LAC 2 eine Fahrt beendet wird. Somit entstehen zwei neue Fahrten, eine mit der LAC-Folge 4 – 3 – 2 und eine weitere mit der LAC-Folge 2 – 1 – 5 – 6 – 7. Die neu entstehenden LAC-Folgen müssen der Anforderung aus Kapitel 3.2.2 genügen und mindestens drei verschiedene LACs enthalten.

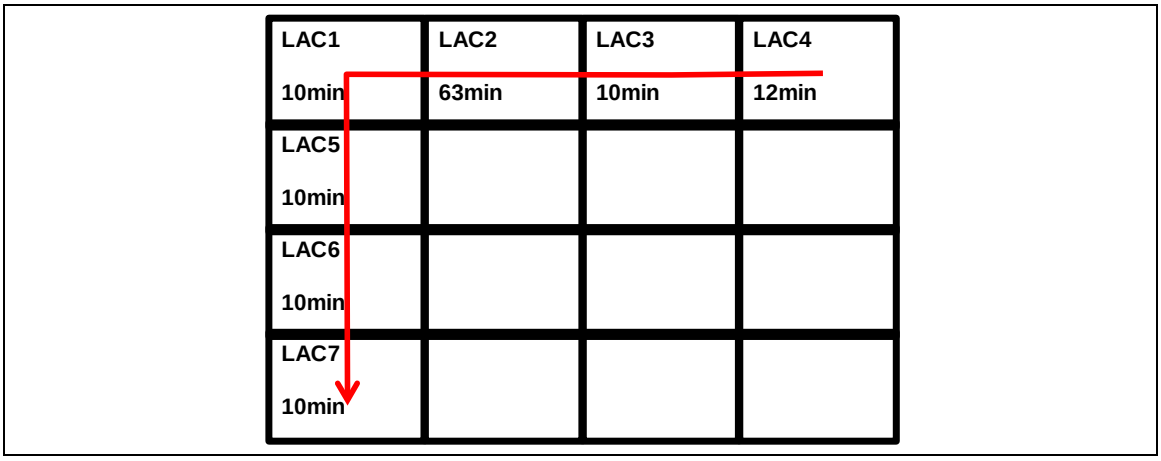


Abbildung 20: Bewegung eines Mobilfunkteilnehmers mit LAC-Aufenthaltsdauern.

Die Anwendung der 60min-Regel soll nun an zwei realen Beispielen verdeutlicht werden. Das erste Beispiel ist in Tabelle 4 gezeigt, in der links die ursprünglichen Einträge für den User 306394488 dargestellt sind. Dieser hat in der Location Area 29453 eine Aufenthaltsdauer von ca. 1,5 Stunden. Daher wird der User an dieser Stelle aufgeteilt. Die erste Fahrt behält die alte UsrID 306394488 und geht von Zeile 1 bis 6. Für den zweiten Teil wird eine neue UsrID (hier: 309496459) generiert und der Eintrag der Location Area 29453 als neue Zeile hinzugefügt. Der neu hinzugefügten Location Area wird eine Aufenthaltsdauer von einer Minute zugeordnet. Dies entspricht dem häufigsten Wert einer Aufenthaltsdauer in einer Location Area.

Vor Anwendung der 60min-Regel			
Zeile	**UsrID**	**Uhrzeit**	**cLAC**
1	306394488	14:12:44	30724
2	306394488	14:38:18	30726
3	306394488	14:43:16	29454
4	306394488	14:49:19	29453
5	306394488	16:32:07	29454
6	306394488	16:52:25	30726
7	306394488	16:57:58	30724

Nach Anwendung der 60min-Regel			
Zeile	**UsrID**	**Uhrzeit**	**cLAC**
1	306394488	14:12:44	30724
2	306394488	14:38:18	30726
3	306394488	14:43:16	29454
4	306394488	14:49:19	29453
1	*309496459*	*16:31:07*	*29453*
2	*309496459*	16:32:07	29454
3	*309496459*	16:52:25	30726
4	*309496459*	16:57:58	30724

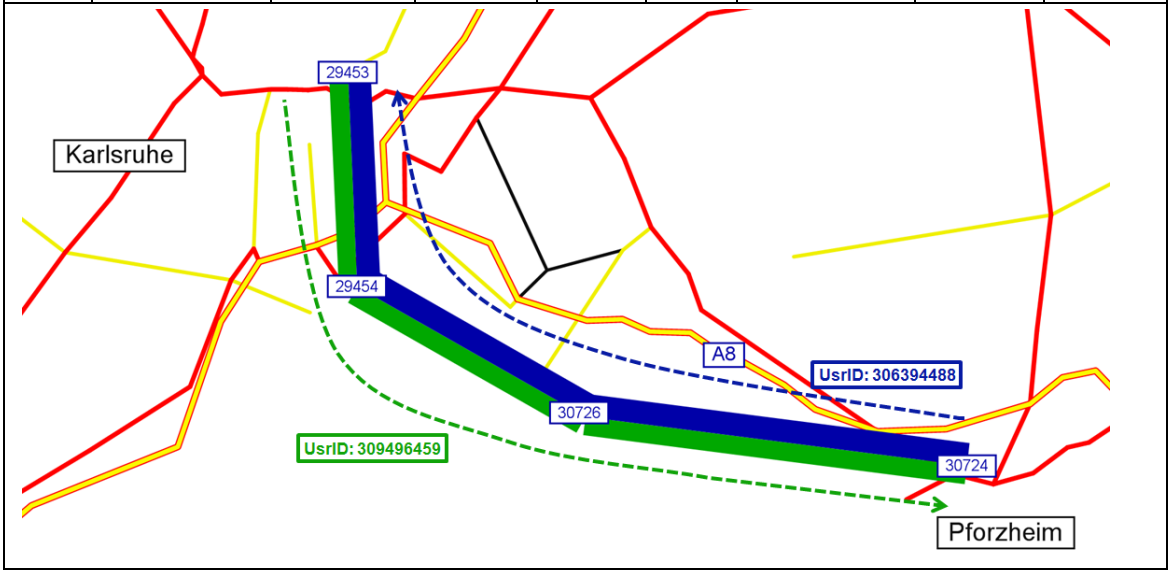


Tabelle 4: Aufteilung der UsrID 306394488 aufgrund der 60min-Regel in Zeile 5.

Das zweite Beispiel ist der Mobilfunkteilnehmer mit der UsrID 307744434 (vgl. Tabelle 5), der sich in der ersten Location Area mehr als 60 Minuten aufhält. Es ist davon auszugehen, dass sich der Mobilfunkteilnehmer in dieser relativ kleinen Location Area östlich von Walldorf (LAC 26638) aufgehalten hat, bevor er seine Fahrt auf der A6 Richtung Heilbronn begonnen hat. Dort wird ebenfalls eine zusätzliche UsrID erzeugt, die ursprüngliche UsrID wird allerdings aufgrund der Forderung, dass jede UsrID mindestens drei unterschiedliche LACs haben muss, gelöscht. Für die neue UsrID wird die erste Uhrzeit auf eine Aufenthaltsdauer von einer Minute angepasst.

Vor Anwendung der 60min-Regel					Nach Anwendung der 60min-Regel			
Zeile	**UsrID**	**Uhrzeit**	**cLAC**		**Zeile**	**UsrID**	**Uhrzeit**	**cLAC**
1	307744434	18:57:25	26638		~~1~~	~~307744434~~	~~18:57:25~~	~~26638~~
					1	*309655461*	*20:30:29*	26638
2	307744434	20:31:29	26643		2	*309655461*	20:31:29	26643
3	307744434	20:40:03	26645		3	*309655461*	20:40:03	26645
4	307744434	21:02:07	26644		4	*309655461*	21:02:07	26644
5	307744434	21:06:08	26636		5	*309655461*	21:06:08	26636
6	307744434	21:06:36	26644		6	*309655461*	21:06:36	26644
7	307744434	21:08:24	26636		7	*309655461*	21:08:24	26636

Tabelle 5: Änderung der UsrID 307744434 aufgrund der 60min-Regel in Zeile 1.

3.3.2 Erweiterte 60min-Regel

Es ist eine Besonderheit des Mobilfunknetzes, dass Mobilfunkteilnehmer in kurzen zeitlichen Abständen zwischen verschiedenen Cell-IDs und damit u.U. LACs wechseln (z.B. wenn sich diese im Grenzbereich zwischen mehreren Cell-IDs aufhalten). Damit entstehen Einträge in den A-Daten, ohne dass eine tatsächliche Bewegung stattfindet. Daher wird die vorher eingeführte „60min-Regel“ erweitert, so dass sie ein Wechseln zwischen verschiedenen LACs ohne tatsächliche Bewegung erkennt.

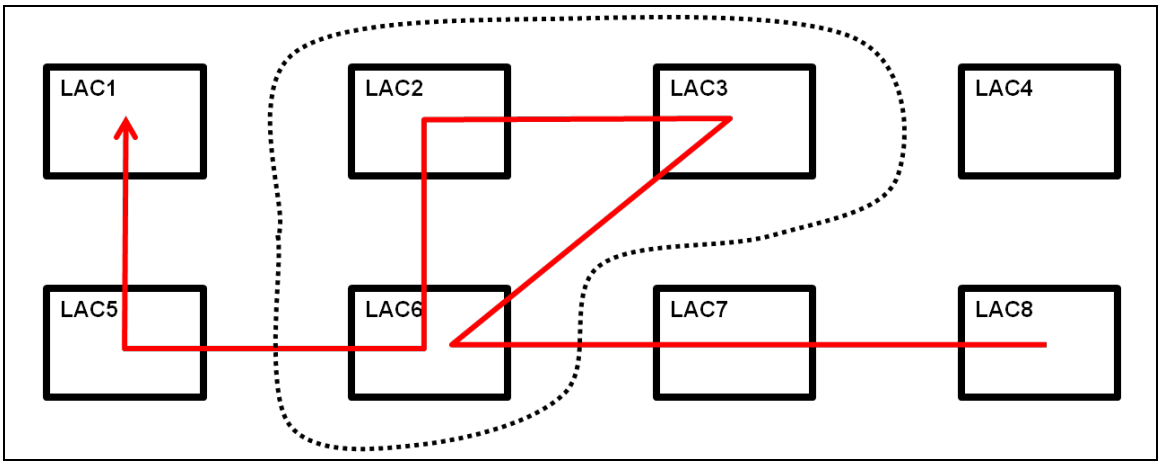


Abbildung 21: Bewegung eines Mobilfunkteilnehmers.

In der obigen Abbildung 21 meldet sich ein Mobilfunkteilnehmer mehrfach in der LAC 6 an. Dazwischen werden die zwei verschiedenen LACs 2 und 3 besucht. Unter den folgenden Bedingungen wird davon ausgegangen, dass es sich in dem gekennzeichneten Bereich um das Ende einer Fahrt handelt:

- Anzahl unterschiedlicher LACs zwischen der mehrfach besuchten LAC ≤ 2
- Zeitraum zwischen dem erstem Eintritt und dem letztem Austritt der mehrfach besuchten LAC ≥ 60min

Wenn beide Bedingungen eingehalten werden, werden zwei Fahrten für den Mobilfunkteilnehmer gebildet. Die erste (hier: 8 – 7 – 6) endet an der mehrfach besuchten LAC und die zweite beginnt dort (hier: 6 – 5 – 1). Die neu entstehenden Fahrten müssen wiederum mindestens drei verschiedenen LACs enthalten (vgl. Kapitel 3.2.2).

Die erste der beiden Bedingungen bezieht sich ausdrücklich auf die Anzahl unterschiedlicher LACs, um ein häufig auftretendes „Zittern“, also ein häufigeres Wechseln zwischen zwei LACs, auch als Fahrtende zu erkennen. Somit würde auch die in Abbildung 22 dargestellte Bewegung in zwei Fahrten aufgetrennt werden, wenn auch die zweite Bedingung eingehalten wird.

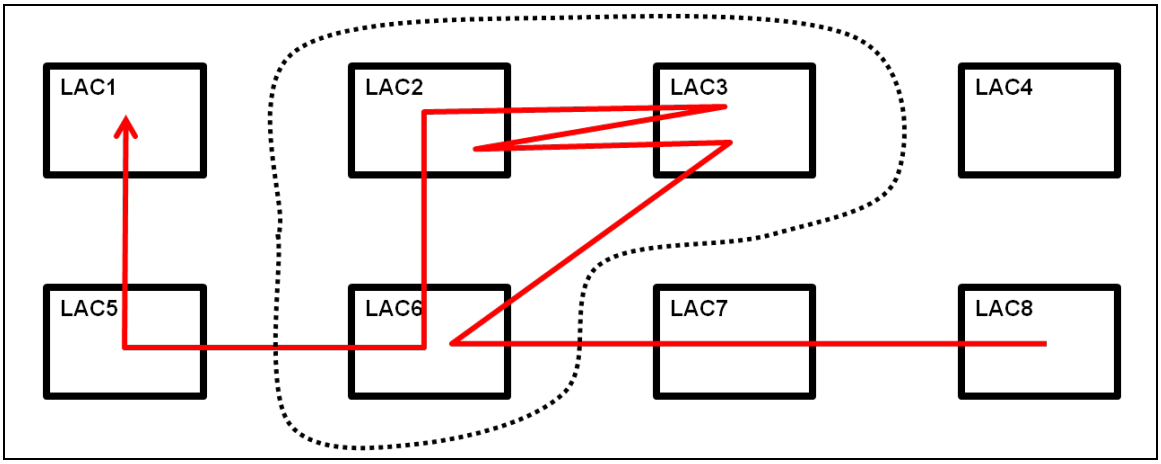


Abbildung 22: Bewegung eines Mobilfunkteilnehmers.

3.3.3 Nervositätsregel

Wie bereits im vorherigen Kapitel erwähnt, ist es im Mobilfunknetz üblich, dass Mobilfunkteilnehmer zwischen verschiedenen Cell-IDs bzw. LACs wechseln, ohne dass eine Bewegung stattgefunden hat. Um solche Teilnehmer zu erkennen wird ein Nervositätsfaktor für jede Fahrt berechnet:

$$Nervositätsfaktor = \frac{AnzLACWechsel}{AnzVerschiedeneLACs}$$

Die Berechnung soll anhand eines Beispiels in Abbildung 23 verdeutlicht werden.

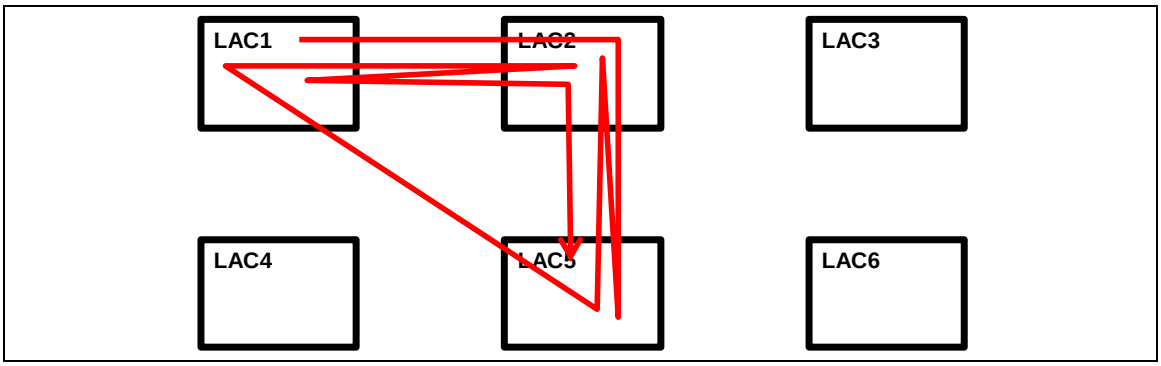


Abbildung 23: Bewegung eines Mobilfunkteilnehmers.

Der Mobilfunkteilnehmer, bei dem sich keine Fahrt eindeutig erkennen lässt, besucht drei unterschiedliche LACs, dabei werden 9 LAC-Wechsel gemacht. Damit berechnet sich der Nervositätsfaktor zu 3:

$$Nervositätsfaktor = \frac{AnzLACWechsel}{AnzVerschiedeneLACs} = \frac{9}{3} = 3$$

Mobilfunkteilnehmer, die die folgenden zwei Bedingungen erfüllen, werden gelöscht:

- $Nervositätsfaktor > 2$
- $AnzVerschiedeneLACs \leq 3$

Die angegebenen Grenzen sind großzügig gehalten, d.h. es werden tendenziell mehr Fahrten beibehalten als nötig. Da diese aber in der weiteren Bearbeitung im Regelfall keine falschen Trajektorien hervorrufen, wird diese vorsichtige Löschung angewendet.

3.4 Generierung von LAC-Folgen aus A-Daten

Aus den in Kapitel 3.3 ermittelten Fahrten können direkt LAC-Folgen gebildet werden. Dafür wird jeder LAC ein Code zugeordnet, der aus genau einem eindeutigen Zeichen (Buchstaben, Zahlen, Sonderzeichen) besteht. Abbildung 24 zeigt das Ergebnis dieser Umwandlung.

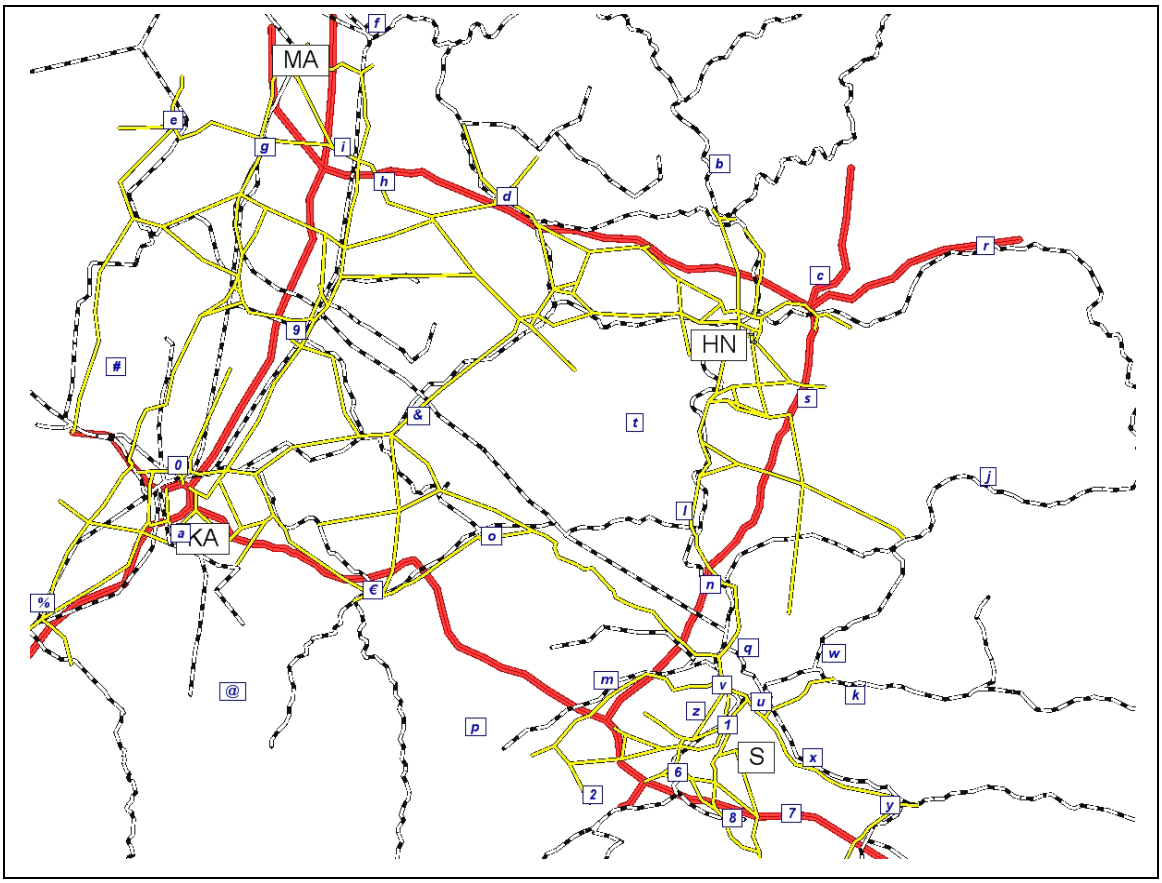


Abbildung 24: Untersuchungsgebiet mit VPN und Schienennetz sowie LA-Code.

Diese Vereinfachung ermöglicht es, LAC-Folgen als *String* (Zeichenfolge) darzustellen. Die unten dargestellte LAC-Folge zwischen Stuttgart („S“) und Karlsruhe („KA)“ kann damit als String „6mp€o€0“ dargestellt werden. Der wesentliche Grund für diese Umwandlung ist neben einer kompakteren Darstellung, dass auf diese Weise die Funktionalitäten von Strings für die weitere Verarbeitung genutzt werden können, also unter anderem ein Stringvergleich möglich ist (vgl. Kap. 3.6).

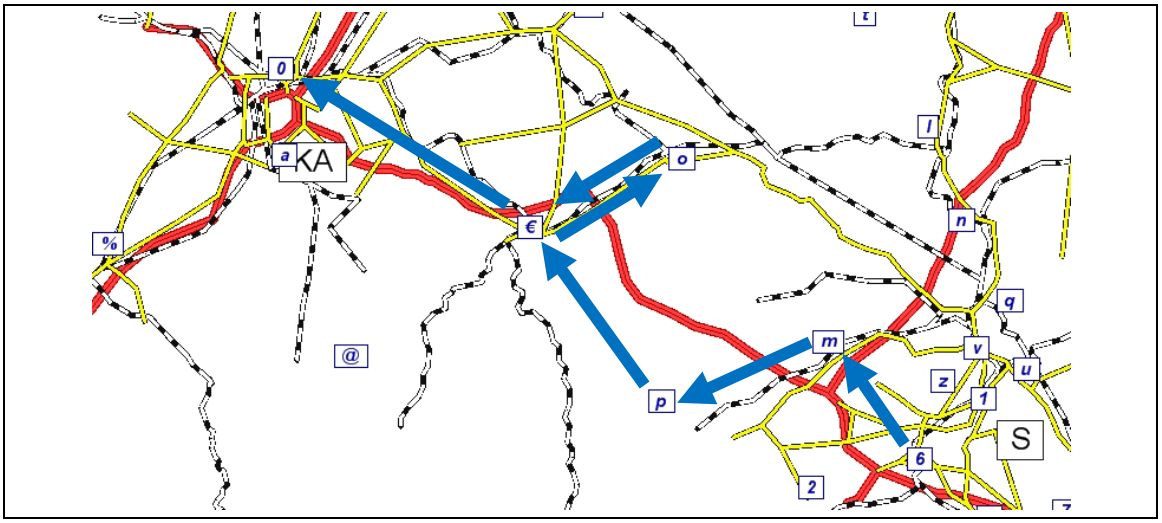


Abbildung 25: Darstellung der LAC-Folge „6mp€o€0“.

3.5 Netz-LAC-Folgen aus Routenbaumgenerator

Um den generierten A-Daten-LAC-Folgen Routen im Projektnetz zuordnen zu können, werden Netz-LAC-Folgen aus dem Projektnetz erzeugt. Dazu wird in einem ersten Schritt das Mobilfunknetz mit den folgenden Elementen in das Verkehrsnetz integriert:

- Die Location Areas werden als Verkehrszellen ins Verkehrsmodell des Projektnetzes aufgenommen. Diese Verkehrszellen sind Quelle und Ziel von Routen und haben eine flächenhafte Ausdehnung gemäß dem Best-Server-Plot.
- Die Verknüpfung zwischen den Location Areas als Verkehrszelle und Knoten des Verkehrsnetzes erfolgt über Anbindungen.
- Den Knoten und Strecken des Verkehrsnetzes wird ein Attribut hinzugefügt, das Auskunft über die zugehörige Location Area gibt. Für die Strecken wird dabei die zum Mittelpunkt gehörende Location Area zugeordnet.

Generierung von Anbindungen

Für die 41 LAC-Verkehrszellen werden nun manuell Anbindungen an das VPN und Schienennetz erzeugt. Dabei sollen alle voneinander getrennt liegenden Streckennetze einer Location Area angebunden werden. So müssen z.B. zwei VPN-Anbindungen und eine Eisenbahn-Anbindung für den Location Area „p“ (blau) erzeugt werden, um alle möglichen LAC-Folgen generieren zu können:

Abbildung 26: Anbindungen der LA „j“ (VPN: durchgezogene Linie / Eisenbahn: gestrichelt).

Desweiteren müssen auch Knoten angebunden werden, die selber nicht in der jeweiligen LAC liegen. Dies tritt vor allem am Rand des Testfeldes auf, wo Mobilfunkteilnehmer auf nicht-VPN-Strecken, aber in einer VPN-LAC fahren. Als Beispiel dient in Abbildung 27 die LAC „j“. Aus Nürnberg kommende Fahrzeuge fahren auf der A6, deren erste VPN-Strecke im LAC „r“ liegt. Da die A6 aber außerhalb des VPNs auch noch im LAC „j“ liegt, wird die LAC-Verkehrszelle „j“ auch an das Autobahnende angebunden.

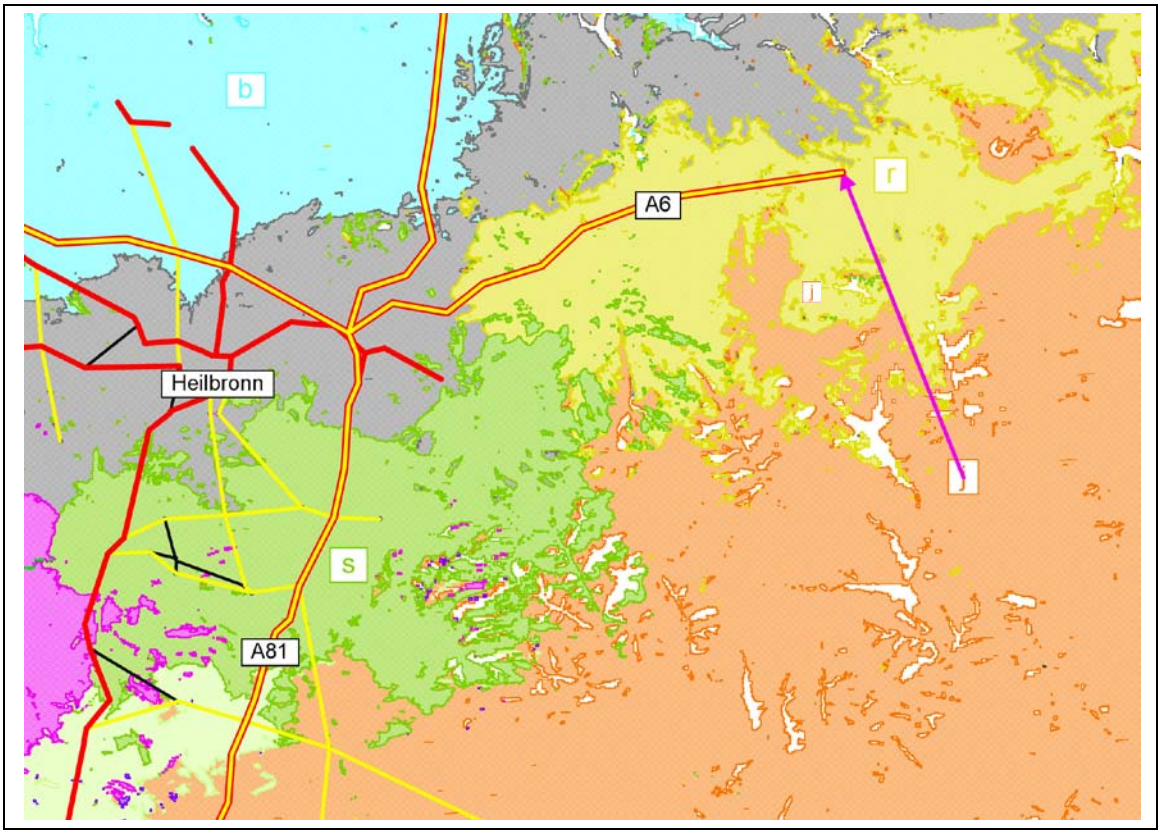


Abbildung 27: Zusätzliche Anbindung des LA „j“ (LA-Flächen aus Best-Server Plot).

Darüberhinaus werden große Location Areas mehrfach angebunden, auch wenn sich dies nicht aus den bisher genannten Kriterien ergibt. Das ermöglicht eine bessere Verteilung des Verkehrs innerhalb einer Location Area (vgl. Abschnitt 3.7.2).

Routengenerierung

Bei der Routengenerierung mit dem Routenbaumgenerator (vgl. Kap. 2.4) soll ein Parametersatz verwendet werden, der alle sinnvollen LAC-Folgen als Referenz für die LAC-Folgen aus den A-Daten generiert. Die Parameter *c*, *f* und *g* werden dabei in Abhängigkeit vom Widerstand $w_{min}(n)$ gewählt und ergeben den in Abbildung 28 dargestellten Verlauf:

- Zu Beginn einer Route bis zu einem Widerstand von 900 s werden nur geringe Toleranzen zugelassen, da diese in der Regel nur zu innerstädtisch minimal abweichenden Routen führen, die für die Anwendung in dieser Arbeit keinen Nutzen

darstellen. Mit den Parameter c=0, f=1,1 und g=0 ergibt sich die folgende Funktion für $w_{max}(n)$:

$$w_{max}(n) = 1{,}1 \cdot w_{min}(n)$$

- Bei Widerständen zwischen 900 s und 5.100 s werden die Parameter großzügiger gewählt, so dass auch grundsätzlich unterschiedliche Routen mit größeren Abweichungen gefunden werden:

$$w_{max}(n) = 990 + 1{,}5 \cdot \left(w_{min}(n) - 900\right)$$

- Ab 5.100 s wird die Differenz zwischen $w_{min}(n)$ und $w_{max}(n)$ auf einen Absolutwert von 2.190 s begrenzt, um bei langen Fahrten den Einfluss des Parameters *f*, der den erlaubten relativen Umweg beschreibt, zu begrenzen. Aufgrund der begrenzten Netzgröße greift diese Bedingung im Projektnetz nur in Ausnahmefällen.

$$w_{max}(n) = 2.190 + 1{,}0 \cdot w_{min}(n)$$

Als Quelle wird für alle Berechnungen der Widerstände der erste angebundene Knoten und nicht die LAC-Verkehrszelle verwendet, damit die Routen unabhängig von der Lage der anderen Anbindungen gefunden werden.

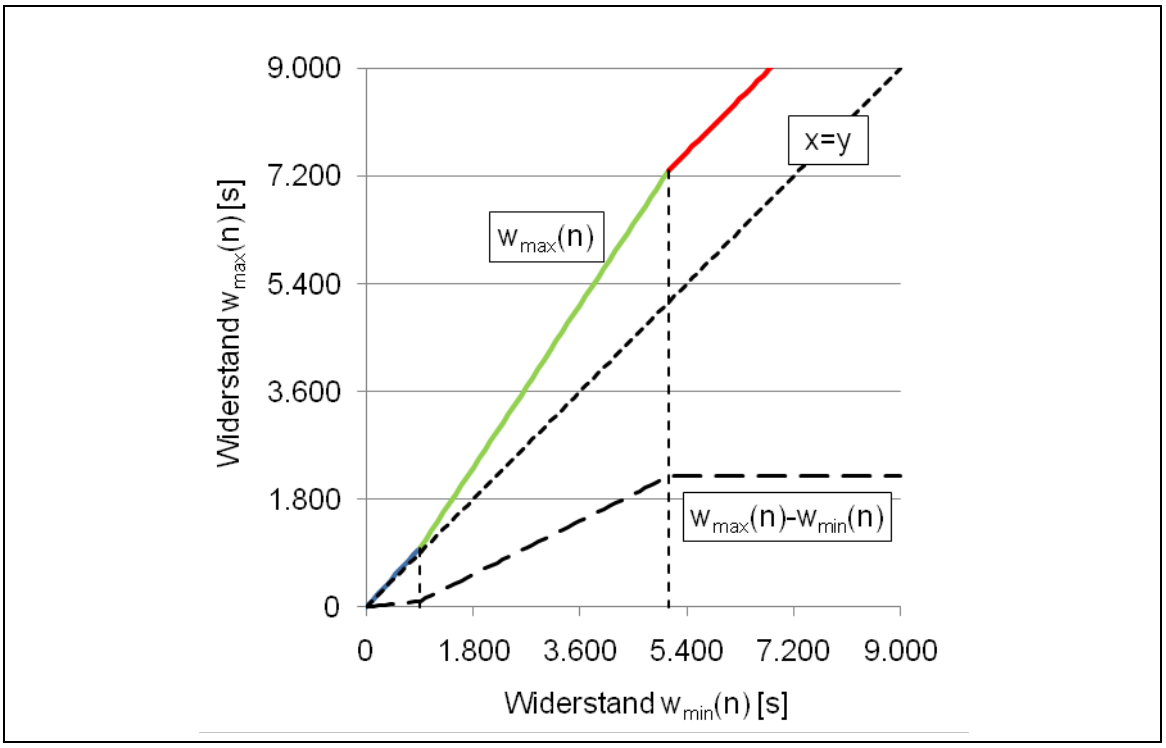


Abbildung 28: Visualisierung der Funktion $w_{max}(n)$ mit Angabe der Winkelhalbierende (x=y) und der Differenz zwischen $w_{max}(n)$ und $w_{min}(n)$.

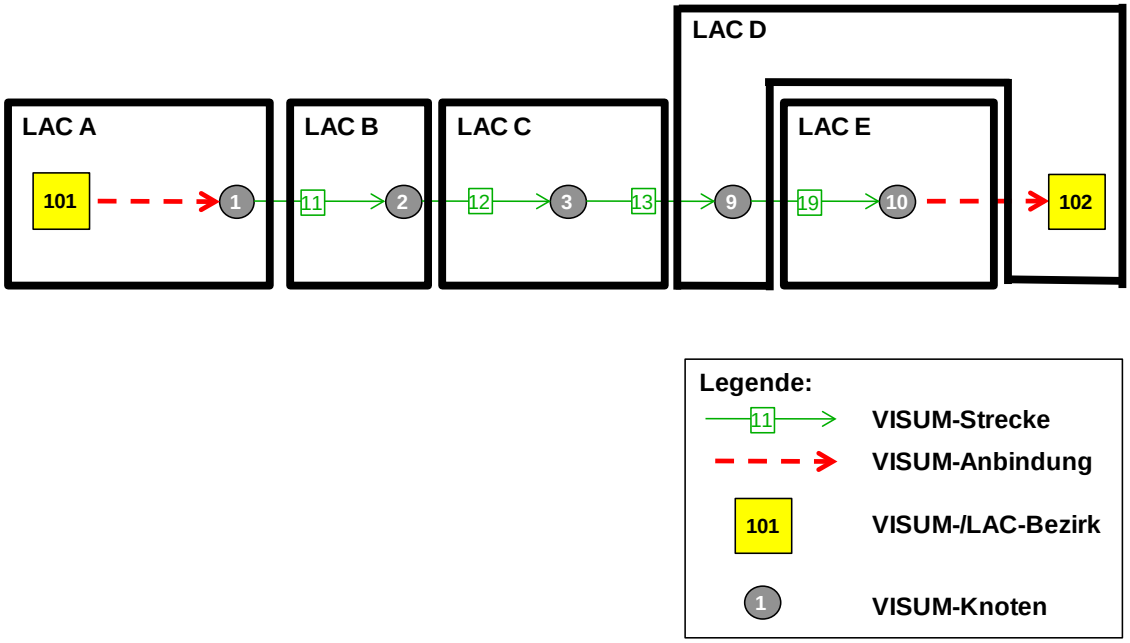
Die Prüfung des hinteren Teils der Route auf Umwegigkeit erfolgt für Routen mit einem Widerstand von 1200 Sekunden für die jeweils letzten 900 Sekunden.

Da bei dieser Routengenerierung den verschiedenen Anbindungen einer Verkehrszelle keine Widerstände zugeordnet werden können und eine Bevorzugung bestimmter Anbindungen verhindert werden soll, wird der Routenbaumgenerator so angepasst, dass dieser anbindungsfein sucht. Das bedeutet, dass die Vergleichswiderstände nicht wie in Kapitel 2.4 beschrieben von Quelle berechnet werden, sondern vom jeweils angebundenen Knoten. Auf diese Weise wird berücksichtigt, dass eine LAC-Verkehrszelle ein großes Gebiet abdeckt, das mehrere Ortschaften und damit Quellen enthalten kann. Die Anbindungen einer LAC-Verkehrszelle repräsentieren damit Zufahrtstraßen ins VPN (vgl. Abbildung 26).

Auf diese Weise werden 260.829 Routen gefunden, was bei 41 LAC-Verkehrszellen ca. 155 Routen je Quelle-Ziel-Relation entspricht. Mit angepassten Parametern werden zusätzlich ca. 11.000 Routen für das Schienennetz generiert. Die Zahl der Routen im Schienennetz ist vor allem deshalb deutlich geringer, da das Schienennetz grobmaschiger als das VPN ist.

Umwandlung von VPN-Routen in LAC-Folgen

Im letzten Schritt werden die im Routenbaumgenerator erzeugten VPN-Routen in LAC-Folgen umgewandelt. Dabei werden wie in Abbildung 29 verdeutlicht alle LACs in der auftretenden Reihenfolge (Verkehrszelle, Knoten, Streckenmitte) aneinander geschrieben, wobei doppelt hintereinander auftretende Einträge derselben LAC gelöscht werden (analog zu den LAC-Folgen aus A-Daten, vgl. Kapitel 3.2.3).



	Von Zelle	1. Knoten	2. Knoten	3. Knoten	...	n-1. Knoten	n. Knoten	Nach Zelle
VPN	101	1	2	3	...	9	10	102

	Von Zelle		1. Strecke		2. Strecke		...		...		n. Strecke			Nach Zelle
			Von Knoten	Mitte	Von Knoten	Mitte	Von Knoten	Mitte	Von Knoten	Mitte	Von Knoten	Mitte	Nach. Knoten	
VPN	101		11		12		...		...		19			102
LAC	A		A	B	B	C	...	...	...	...	D	E	E	D
Resultierende LAC-Folge: A B C ... D E D														

Abbildung 29: Umwandlung einer VPN-Route in eine LAC-Folge.

Da das VPN deutlich feinmaschiger als das Mobilfunk-LAC-Netz ist, ergeben etliche im Verkehrsnetz unterschiedliche Routen eine identische LAC-Folge. Bei den VPN-Routen ergeben sich aus den 260.829 Routen 40.875 unterschiedliche LAC-Folgen und im Schienennetz reduzieren sich die 10.956 gefundenen Routen auf 5.135 LAC-Folgen. 457 vorwiegend kürzere LAC-Folgen kommen sowohl im VPN als auch im Schienennetz vor. Die fehlende Eindeutigkeit erfordert nach dem Matching der LAC-Folgen eine Aufteilung, sowohl in Bezug auf die Moduswahl als auch in Bezug auf die Routenwahl innerhalb einer LAC. Dies wird in Kapitel 3.7 beschrieben.

3.6 Matching mit LAC-Folgen aus Routenbaum

3.6.1 Direkter Stringvergleich

Nachdem nun sowohl die A-Daten-LAC-Folgen als auch die Netz-LAC-Folgen aus dem Routenbaumgenerator vorliegen, können diese miteinander verglichen werden. In einem ersten Schritt wird dabei auf komplette Übereinstimmung der beiden Strings

geprüft. Alle LAC-Folgen aus A-Daten, die in diesem Schritt identifiziert worden sind, werden als bearbeitet gekennzeichnet und in den weiteren Schritten nicht berücksichtigt.

3.6.2 Berücksichtigung häufiger LAC-Wechsel („ABAB-Regel“)

Eine Analyse der nicht identifizierten LAC-Folgen ergibt als häufigen Grund für eine erfolglose Zuordnung einen häufigen Wechsel zwischen bestimmten LACs. Ein Beispiel ist die LAC-Folge „CABABABD“, wo einige Male zwischen den LACs A und B gewechselt wird, die erweiterte 60min-Regel aber aufgrund der kurzen Gesamtaufenthaltsdauer in den LACs A und B nicht gegriffen hat. Um solche LAC-Folgen zu vereinfachen und besser mit Vergleichsdaten matchen zu können, wird das Auftreten von „ABAB“ zu „A~~BAB~~“ reduziert. Die folgende Tabelle 6 zeigt den Vereinfachungsprozess für das bereits genannte Beispiel „CABABABD“.

Aktuelle LAC-Folge	Beschreibung
CABABABD	Erkennen des ersten Auftretens „ABAB“
CA BABD	Löschen der mittleren zwei LACs „BA“
CA BABD	Erkennen des zweiten Auftretens „ABAB“
CA BD	Löschen der mittleren zwei LACs „BA“, Ende der Bearbeitung

Tabelle 6: Zweifache Anwendung der „ABAB-Regel“.

Die mit der der „ABAB-Regel“ gekürzten Strings werden einem erneuten direkten Stringvergleich zugeführt. Alle A-Daten-LAC-Folgen, die in diesem Schritt nicht identifiziert worden sind, werden in dem folgenden Schritt in der ggf. gekürzten Version berücksichtigt.

In den Netz-LAC-Folgen des Routenbaums gibt es zwar vereinzelte Routen, die tatsächlich eine Folge „ABAB“ beinhalten, deren Identifizierung durch das Kürzen unwahrscheinlicher wird. Dabei handelt sich allerdings in allen Fällen um Routenteile im untergeordneten Straßennetz, für die es darüberhinaus immer eine Alternative der Form „A~~BAB~~“ gibt. Somit ist eine Routenidentifizierung in allen Fällen gleich wahrscheinlich, es kann aber durch die Weiterverwendung der kürzeren Strings im Einzelfall zu einer im untergeordneten Straßennetz leicht veränderten Routenführung kommen.

3.6.3 Stringähnlichkeitsvergleich

Es gibt eine Vielzahl von Ursachen, warum sich A-Daten-LAC-Folgen in den bisher durchgeführten zwei Schritten keiner LAC-Folge aus VPN- bzw. Schienenrouten zuordnen lassen. Diese Ursachen lassen sich in zwei Kategorien unterteilen:

1. LAC-Folge entspricht einer einzelnen Fahrt, es gibt aber keine passende LAC-Folge aus dem Routenbaumgenerator.
2. LAC-Folge entspricht keiner einzelnen Fahrt bzw. entspricht mehr als einer Fahrt.

Für LAC-Folgen der Kategorie 1 soll die bestmögliche LAC-Folge ermittelt werden, für die LAC-Folgen der Kategorie 2 soll dagegen keine LAC-Folge gefunden werden. Tabelle 7 fasst einige Gründe für das Auftreten in beiden Kategorien zusammen:

Nr.	**Beschreibung**	**Beispiel-LAC-Folge aus...**	
Kategorie 1: A-Daten-LAC-Folge entspricht einer Fahrt im VPN oder Schienennetz		**... A-Daten**	**... Routenbaum**
1	Anmeldung eines Mobilfunkteilnehmers an der nicht-nächsten LA (z.B. aufgrund Überlastung im Mobilfunknetz, Abschattung etc.), ebenso fehlende Anmeldung an einer nächsten LAC	ABCDXDF	ABCD F Mobilfunkteilnehmer wechselt vorübergehend in „X“, obwohl dies nicht die nächste LA ist.
2	Unschärfe bei der Generierung der Netz-LAC-Folgen aus dem VPN (vgl. Kap. 3.5)	ABCDEDF	ABCDE F Kurzer Rückwechsel in „D“ fehlt im Routenbaum, wenn dieser zwischen Knoten und Streckenmitte erfolgt.
3	Mobilfunkteilnehmer befährt Strecken, die nicht im VPN enthalten sind (z.B. untergeordnetes Straßennetz) und zu neuen LAC-Folgen führen.	---	
Kategorie 2: A-Daten-LAC-Folge ist keine einzelne Fahrt		**... A-Daten**	**... Routenbaum**
4	Mobilfunkteilnehmer nutzt Schienennetz und das VPN auf einer Quelle-Ziel-Relation	ABCDEFGH	ABCD (VPN) & DEFGH (Schiene) Multimodale Fahrt wird nicht erkannt.
5	Unzureichende Fahrtidentifikation, z.B. fehlende Aufteilung zwei verschiedenen Fahrten (vgl. Kap. 3.3).	ABCDEDCB	ABCDE & EDCB Ziel in „E“ wurde nicht erkannt, Gesamtfahrt nicht in Routenbaum vorhanden (zu umwegig)

Tabelle 7: Ursachen für das Nicht-Wiederfinden von LAC-Folgen aus A-Daten in LAC-Folgen aus dem Routenbaumgenerator.

Wie an den Beispielen der Kategorie 1 zu erkennen ist, folgen die Abweichungen gewissen Regeln (z.B. „ABA“ statt „A“). Für eine erfolgreiche Zuordnung einer A-Daten-LAC-Folge zu der wahrscheinlichsten Netz-LAC-Folge wird deshalb ein Stringvergleich angewendet, der auf dem aus der Biologie bekannten Needleman-Wunsch-Algorithmus (Needleman & Wunsch, 1970) basiert, der für die Zwecke des LAC-Folgen-Vergleichs angepasst wird.

Algorithmus des Stringähnlichkeitsvergleichs

Der Algorithmus vergleicht zwei Strings miteinander und bewertet deren Ähnlichkeit. Dabei wird eine Matrix $M(I,J)$ erzeugt und diese anhand der im Folgenden etwas vereinfacht dargestellten Regeln gefüllt:

$$M(0,0) = 0$$

$$M(i,0) = M(i-1,0) + w(Z_{1,i}, -)$$

$$M(0,j) = M(i-1,0) + w(-, Z_{2,j})$$

$$M(i,j) = \max \begin{pmatrix} M(i-1, j-1) + w(Z_{1,i}, Z_{2,j}); \\ M(i-1, j) + w(Z_{1,i}, -); \\ M(i, j-1) + w(0, Z_{2,j}) \end{pmatrix}$$

mit

$i = 1...I$	Index für Zeichen der A-Daten-LAC-Folge
$j = 1...J$	Index für Zeichen der Netz-LAC-Folge
$Z_{1,i}$	i. Zeichen des Strings Z_1 (A-Daten-LAC-Folge)
$Z_{2,j}$	j. Zeichen des Strings Z_2 (Netz-LAC-Folge)
$w(a,b)$	Bewertungsfunktion für die Zeichen a und b

Die Bewertungsfunktion $w(a,b)$ ergibt sich dabei wie folgt:

$$w(a,b) = \begin{cases} 1,0 & \text{für } a = b \\ 0,4 & \text{im Fall } aba \\ 0,2 & \text{im Fall } abca \\ -1,0 & \text{ansonsten} \end{cases}$$

„Fall aba“ bedeutet dabei, dass der Mobilfunkteilnehmer von einer LAC in eine andere wechselt und anschließend wieder zurück. Dies entspricht der häufig auftretenden Ursachennummer 1 aus der Tabelle 7. „Fall abca“ ist vergleichbar, nur dass hier zwei andere LACs zwischen der wiederholt auftretenden LAC liegen. Somit wird also im Gegensatz zum normalen Needleman-Wunsch-Algorithmus nicht nur ein Vergleich zwischen zwei Strings durchgeführt, sondern auch die Struktur der einzelnen Strings bei der Bewertung berücksichtigt.

Damit der Struktureinfluss bei kurzen A-Daten-LAC-Folgen mit weniger als 4 unterschiedlichen Location Areas nicht überschätzt wird, werden diese beiden Fälle dort mit einer Bewertung von 0 belegt.

Zur abschließenden Bewertung der Ähnlichkeit („$SeqAlignment$“) wird der Matrixwert $M(I,J)$ durch die mittlere Länge der beiden Strings geteilt:

$$SeqAlignment = \frac{M(I,J)}{0,5 \cdot (I+J)}$$

Bei den hier gewählten Parametern ergibt sich bei einer vollständigen Übereinstimmung von zwei Strings eine Ähnlichkeit von 100 %. Stark voneinander abweichende Strings können auch negative Bewertungen enthalten.

Die folgende Tabelle 8 zeigt einen String-Vergleich der A-Daten-LAC-Folge „abcdxdf“ mit der Netz-LAC-Folge „abcdf“. Fett gedruckt ist die Folge, die zur höchsten Punktezahl von 4,4 in der rechten unteren Ecke führt. Aus dieser Punktzahl ergibt sich eine Ähnlichkeit von 73,3 %.

		a	**b**	**c**	**d**	**x**	**d**	**f**
	0,0	*-1,0*	*-2,0*	*-3,0*	*-4,0*	*-5,0*	*-4,6*	*-5,6*
a	*-1,0*	**1,0**	0,0	-1,0	-2,0	-3,0	-2,6	-3,6
b	*-2,0*	0,0	**2,0**	1,0	0,0	-1,0	-0,6	-1,6
c	*-3,0*	-1,0	1,0	**3,0**	2,0	1,0	1,4	0,4
d	*-4,0*	-2,0	0,0	2,0	**4,0**	**3,0**	**3,4**	2,4
f	*-5,0*	-3,0	-1,0	1,0	3,0	3,0	3,4	**4,4**
$SeqAlignment = \frac{M(I,J)}{0,5 \cdot (I+J)} = \frac{4,4}{0,5 \cdot (7+5)} = 73,3\%$								

Tabelle 8: Beispielberechnung einer String-Ähnlichkeit.

Die Ähnlichkeit von 73,3 % kann auch anhand der folgenden Berechnung nachvollzogen werden, die allerdings bereits eine optimale Vorsortierung der LAC-Folgen benötigt. Dabei wird pro Übereinstimmung (*match*) von zwei Zeichen ein Punkt vergeben, eine Lücke (*gap*) führt zu einem Punkt Abzug und der „Fall aba“ wird mit 0,4 Punkten bewertet.

```
A-Daten-LAC-Folge: abcdxdf (Länge: 7)
Netz-LAC-Folge:    abcd  f (Länge: 5)
       match       ++++  +          → 5,0 Punkte
       gap             -            → -1,0 Punkte
       „Fall aba“       o           → 0,4 Punkte

Ähnlichkeit = Punkte / (0,5 * Gesamtlänge)
            = (5,0 – 1,0 + 0,4) / (0,5 * (5 + 7))
            = 0,733 = 73,3 %
```

Auswahl einer Ähnlichkeitsgrenze

Die Netz-LAC-Folge mit der höchsten Ähnlichkeit wird der A-Daten-LAC-Folge zugeordnet. Wie zu Beginn dieses Abschnitts in Tabelle 7 dargestellt, sind nicht alle A-Daten-LAC-Folgen Fahrten im VPN bzw. Eisenbahnnetz. Daher wird nur bei hoher Ähnlichkeit eine Fahrt generiert.

Die optimale Grenze der Ähnlichkeit kann nicht mathematisch bestimmt werden. Daher erfolgt die Festlegung anhand der Analyse einer Vielzahl von Stringvergleichsergebnissen. Dafür werden die A-Daten-LAC-Folgen, die ähnlichste Netz-LAC-Folge aus dem Routenbaumgenerator und die daraus resultierenden Routen im VPN bzw. Eisenbahnnetz für verschiedene Ähnlichkeiten und LAC-Folgen-Längen visualisiert. Im Folgenden werden drei Beispiele visualisiert:

1. Fahrt von Stuttgart nach Heidelberg (ca. 100 km, Ähnlichkeit 80,0 %)
2. Fahrt in der Region Stuttgart (ca. 20 km, Ähnlichkeit 57,1 %)
3. Fahrt von Stuttgart nach Heilbronn (ca. 100 km, Ähnlichkeit 59,5 %)

Abbildung 30 zeigt für das erste Beispiel links die in den A-Daten beoachtete LAC-Folge und rechts drei von 271.000 Routen in Straßen- und Schienennetz. Die beiden oberen Routen, die sich nur bei Heilbronn leicht unterscheiden, haben dieselbe Netz-LAC-Folge, die die höchste Ähnlichkeit zur A-Daten-LAC-Folge hat.

```
A-Daten-LAC-Folge:  z qnlnjlstscbdhif (Länge: 15)
Netz-LAC-Folge:     zmqnlnjls  cbdhif (Länge: 16)
        match       + +++++++  ++++++      → 14,0 Punkte
        gap          -       -             → -2,0 Punkte
        „Fall aba“            o            →  0,4 Punkte

Ähnlichkeit = Punkte / (0,5 * Gesamtlänge)
            = (14,0 – 2,0 + 0,4) / (0,5 * (15 + 16))
            = 0,800 = 80,0 %
```

Aufgrund der längeren Ubereinstimmungen der beiden LAC-Folgen und der Anwendung des Falls „aba“ (hier ... „sts“ ...) ergibt sich eine Ähnlichkeit von 80 %. Die gefundenen Routen im Projektnetz erscheinen passend zu der A-Daten-LAC-Folge, passenden Routen im Schienennetz gibt es nicht.

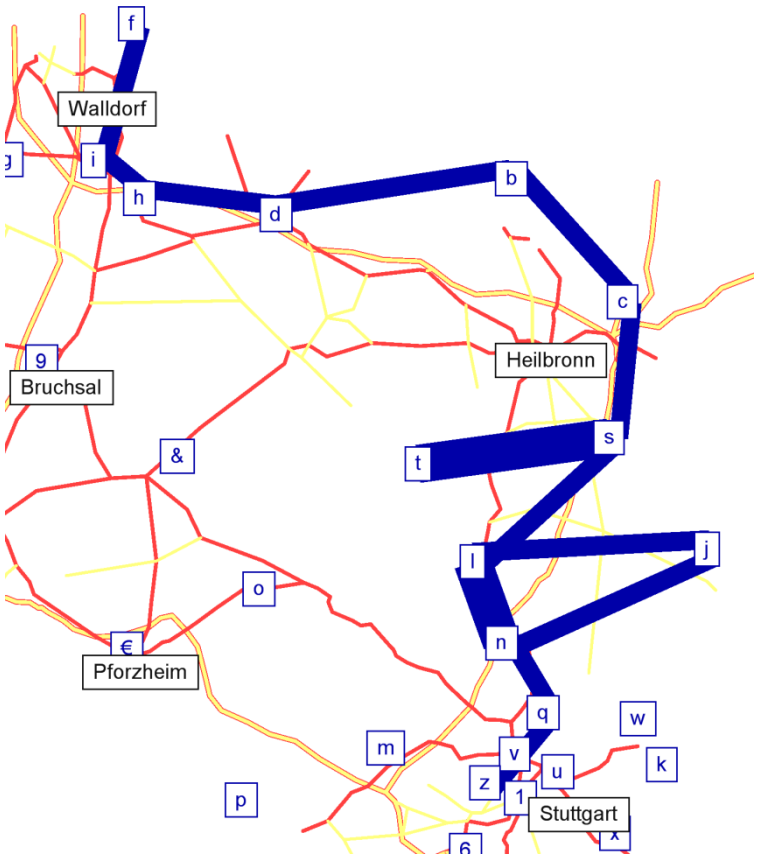
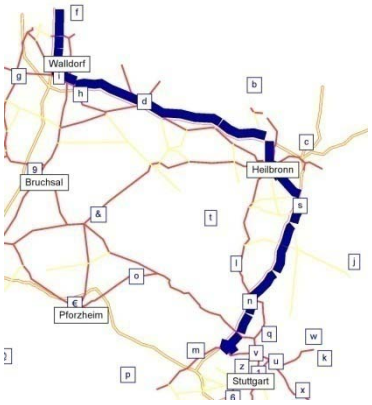
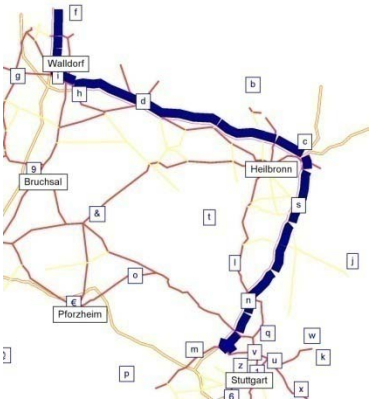
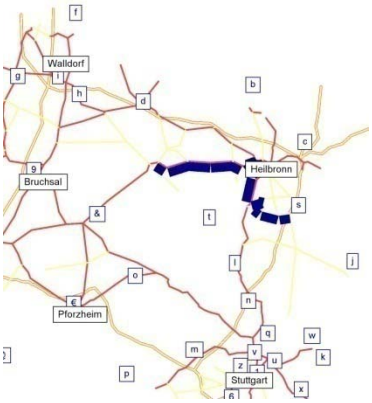
A-Daten-LAC-Folge “zqnlnjlstscbdhif”	**Routen im Projektnetz und resultierende Netz-LAC-Folgen**
	 → “zmqnlnjlscbdhif” → Ähnlichkeit: 80 %  → “zmqnlnjlscbdhif” → Ähnlichkeit: 80 %  → “bcs” → Ähnlichkeit: -105 % 271,000 weiteren Routen → Ähnlichkeiten < 80 %

Abbildung 30: Beispiel 1: Ergebnis des Stringvergleichs (Ähnlichkeit: 80,0 %).

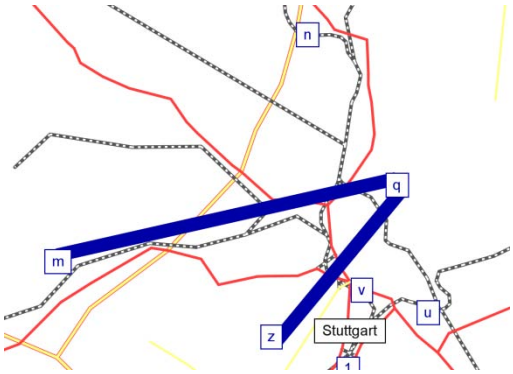
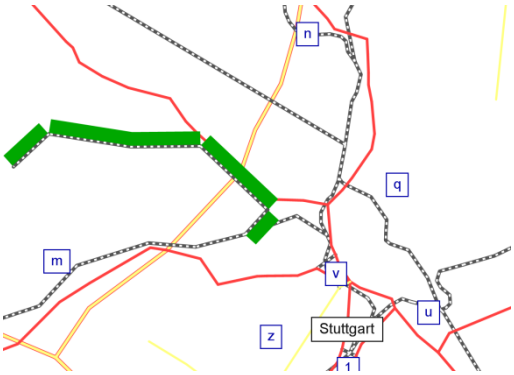
A-Daten-LAC-Folge “zqm”	**Ähnlichste Route im Projektnetz** „zqnm“
	

Abbildung 31: Beispiel 2: Ergebnis des Stringvergleichs (Ähnlichkeit: 57,1 %).

Das zweite Beispiel in Abbildung 31 zeigt links die A-Daten-LAC-Folge und rechts die aufgrund der Stringvergleiches als ähnlichste Route ermittelte Schienenroute. Es ergibt sich dabei eine Ähnlichkeit von 57,1 %:

```
A-Daten-LAC-Folge:  zq m (Länge: 3)
Netz-LAC-Folge:     zqnm (Länge: 4)
        match       ++ +        → 3,0 Punkte
        gap           -         → -1,0 Punkte

Ähnlichkeit = Punkte / (0,5 * Gesamtlänge)
            = (3,0 – 1,0) / (0,5 * (3 + 4))
            = 0,571 = 57,1 %
```

Bei dieser LAC-Folge lässt es sich, auch unter Hinzunahme des detaillierten Best-Server-Plots, der Cell-IDs, der Zeiten und Busfahrplänen nur erahnen, dass hier eine Kombination aus S-Bahn (S6) und Bus (Nr. 500 oder 501) stattgefunden hat. Eine Zuordnung zu einer Fahrt im Eisenbahnnetz ist deshalb nicht wünschenswert.

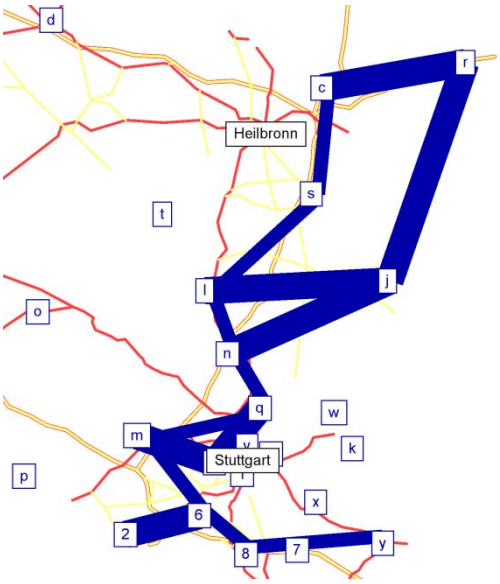
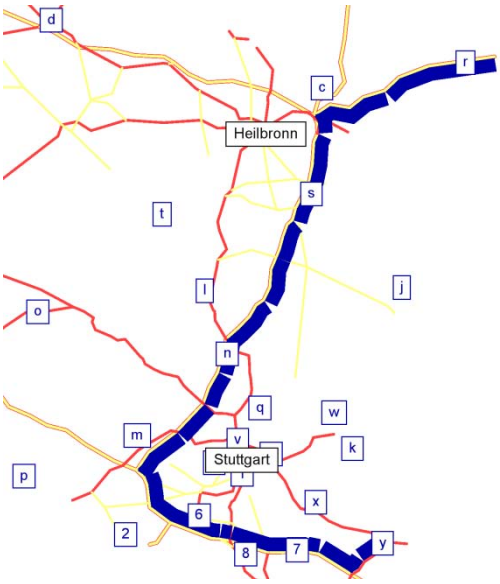
A-Daten-LAC-Folge “ y78626mzmqzqnljnjlscrjrcrj ”	**Ähnlichste Route im Projektnetz** „y786mqnlnjlscrj“
	

Abbildung 32: Beispiel 3: Ergebnisse des Stringvergleichs (Ähnlichkeit: 59,5 %).

Beim dritten Beispiel in Abbildung 32 ergibt sich bei Stringvergleich eine Ähnlichkeit von 59,5 %:

```
A-Daten-LAC-Folge: y78626mzmqzqnljnjlscrjrcrj (Länge: 26)
Netz-LAC-Folge:    y786  m  q  nl njlscrj     (Länge: 15)
       match       ++++  +  +  ++ +++++++     → 15,0 Punkte
       gap             -  -     -        - -  → -5,0 Punkte
       „Fall aba“      o  o  o          o o  →  2,0 Punkte
       „Fall abca“        o                  →  0,2 Punkte

Ähnlichkeit = Punkte / (0,5 * Gesamtlänge)
            = (15,0 – 5,0 + 2,0 + 0,2) / (0,5 * (26 + 15))
            = 0,595 = 59,5 %
```

Die A-Daten-LAC-Folge enthält deutlich mehr LACs als vergleichbare LAC-Folgen auf dieser Route, die häufigen Wechsel zwischen den LACs könnten durch spezielle Abschattungseffekte o.ä. zustande gekommen sein. Aufgrund der Berücksichtigung der aba- und abca-Fälle (z.B. „...zmqz...rcr...“) kann dennoch eine Ähnlichkeit von fast 60 % erreicht werden. Der gefundene Fahrverlauf entspricht dem Fahrtverlauf, der sich bei einer manuellen Analyse des Mobilfunkteilnehmers abschätzen lässt.

Aufgrund einer detaillierten Analyse einer Vielzahl von Stringvergleichen wird die Akzeptanzgrenze auf 60% festgelegt. Von den oben gezeigten Beispielen wird daher nur aus dem ersten eine Fahrt generiert werden. Kurze A-Daten-LAC-Folgen mit nur drei unterschiedlichen LACs wie im zweiten Beispiel können dann Ähnlichkeit über

60% erreichen, wenn A-Daten-LAC-Folge genau vier Zeichen beinhaltet und für das vierte Zeichen der aba-Fall zutrifft. Das dritte Beispiel ist durch die Wahl der 60 %-Grenze eine LAC-Folge, die knapp nicht weiter berücksichtigt wird.

Die gewählte Grenze und der Algorithmus des Stringvergleichs führen dazu, dass bei LAC-Folgen mit einer geringen Anzahl unterschiedlicher LACs relativ wenig LAC-Folgen die gewählte Grenze erreichen. Bei zunehmender Anzahl unterschiedlicher LACs, bei denen man von einer tatsächlichen Bewegung ausgehen kann, werden deutlich mehr Stringvergleiche mit einem erfolgreichen Matching (Ähnlichkeit ≥ 60 %) erreicht.

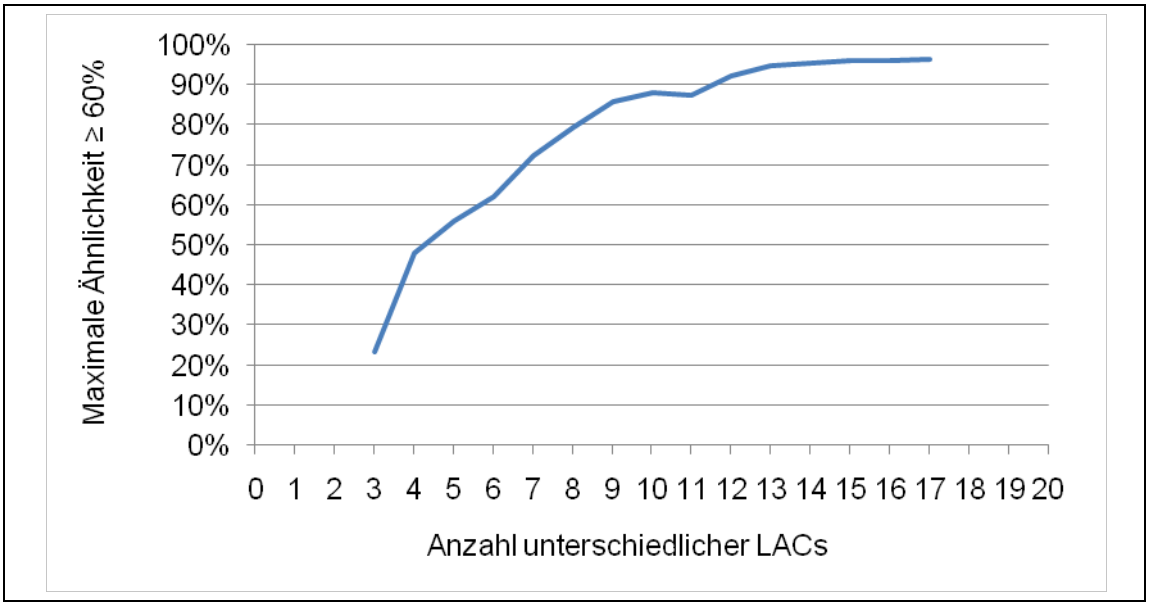


Abbildung 33: Stringvergleiche mit einem erfolgreichen Matching (Ähnlichkeit ≥ 60 %).

Performanceanalyse und -optimierung

Das Beispiel in Tabelle 8 zeigt, dass für einen Vergleich zweier Strings mit den Längen 7 und 5 die Bewertungsfunktion $w(a,b)$ insgesamt 105mal (3mal je Matrixeintrag) berechnet werden muss.

Untenstehende Tabelle 9 zeigt die notwendige Anzahl Funktionsaufrufe der Bewertungsfunktion zur Bestimmung des ähnlichsten Strings aus dem Routenbaum-generator für einen String aus den A-Daten mit der Länge 7. In der ersten Zeile lässt sich ablesen, dass der Routenbaumgenerator 304 unterschiedliche Strings mit der Länge 3 produziert hat. Für jeden dieser Strings muss die Bewertungsfunktion 63mal (7x3 Matrixeinträge und 3mal je Matrixeintrag) aufgerufen werden bzw. 19.152 mal für den Vergleich mit allen Strings der Länge 3. Für die 45.394 unterschiedlichen Netz-LAC-Folgen mit drei oder mehr LACs aus dem Routenbaumgenerator ergeben sich insgesamt über 13 Millionen Aufrufe dieser Funktion. Dieser Wert nimmt linear mit der

Länge der A-Daten-LAC-Folge zu, bei der maximal zulässigen LAC-Länge von 255 ergeben sich fast 500 Millionen Aufrufe für die Vergleiche eines Strings aus den A-Daten. Für einen Tag ergeben sich ca. 100.000 verschiedene A-Daten-LAC-Folgen mit einer mittleren Länge ca. 8 Zeichen.

LAC-Folgen aus dem Routenbaumgenerator		**Aufrufe der Bewertungsfunktion je ...**	
Anzahl LACs	**Anzahl je Klasse**	**... LAC-Folge**	**... LAC-Folgen-Klasse**
3	304	63	19.152
4	588	84	49.392
5	910	105	95.550
6	1.250	126	157.500
7	1.557	147	228.879
8	1.845	168	309.960
9	2.058	189	388.962
10	2.315	210	486.150
11	2.640	231	609.840
12	2.932	252	738.864
13	3.214	273	877.422
14	3.440	294	1.011.360
15	3.573	315	1.125.495
16	3.592	336	1.206.912
17	3.433	357	1.225.581
18	3.121	378	1.179.738
19	2.640	399	1.053.360
20	2.121	420	890.820
21	1.524	441	672.084
22	989	462	456.918
23	638	483	308.154
24	357	504	179.928
25	186	525	97.650
26	100	546	54.600
27	33	567	18.711
28	27	588	15.876
29	3	609	1.827
30	3	630	1.890
31	1	651	651
Summe	**45.394**		**13.463.226**

Tabelle 9: Statistik eines Stringvergleichs mit der LAC-Folge der Länge 7.

Dies führt dazu, dass aus Rechenzeitgründen der Stringvergleich nicht für alle Kombinationen von LAC-Folgen durchgeführt wird. Daher werden aus Vorab-Analysen verschiedene Regeln abgeleitet, wie die Rechenzeit des Stringvergleichs auf ein sinnvolles Maß reduziert werden kann:

- Reduzierung des Vergleichssets der LAC-Folgen aus dem Routenbaumgenerator
- Reduzierung der Anzahl der zu vergleichenden LAC-Folgen aus A-Daten

Die wichtigste Regel zur Perfomanceverbesserung ergibt sich aus einer Analyse des Zusammenhangs zwischen der Anzahl verschiedener LACs in der A-Daten-LAC-Folge und der ähnlichsten Routenbaum-LAC-Folge (vgl. Abbildung 34). Dabei zeigt sich, dass gerade bei geringer Anzahl unterschiedlicher LACs der ähnlichste String meistens dieselbe Anzahl unterschiedlicher LACs hat. Bei einer höheren Anzahl unterschiedlicher LACs in der A-Daten-LAC-Folge nimmt die Streuung erwartungsgemäß zu. Daher müssen die Grenzen bei längeren LAC-Folgen etwas großzügiger gewählt werden.

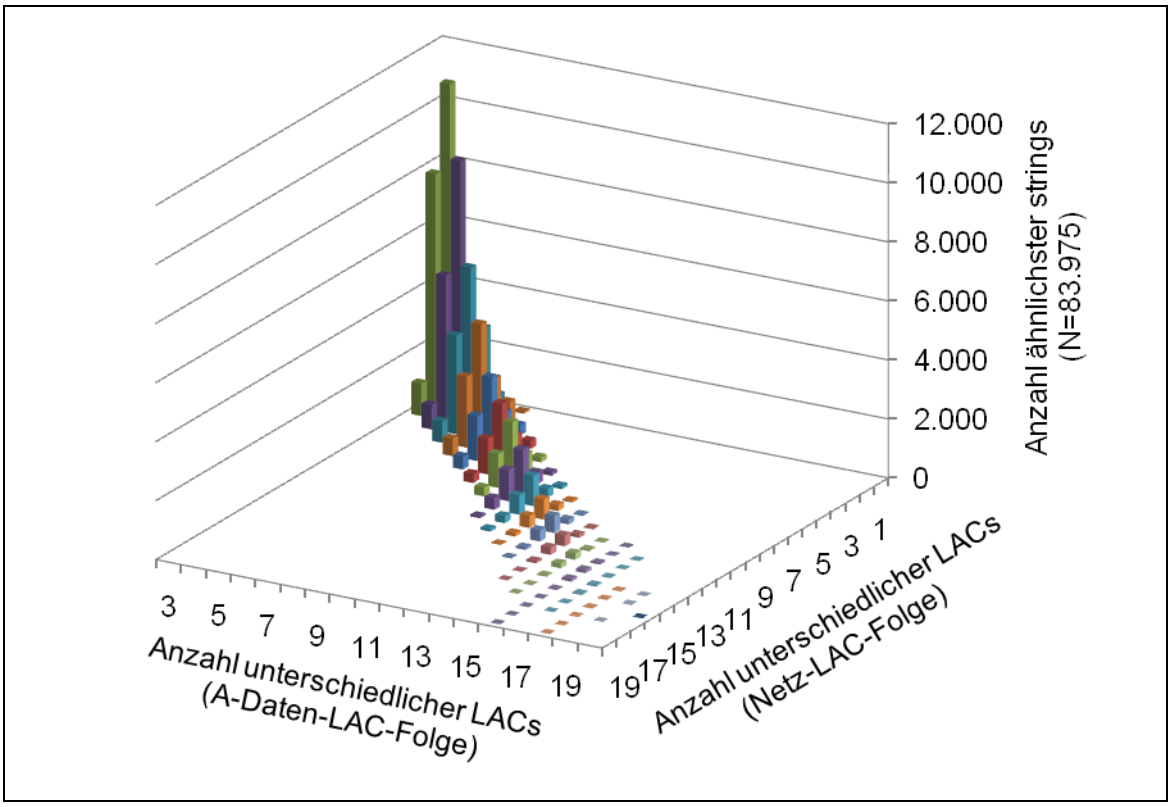


Abbildung 34: Zusammenhang zwischen der Anzahl verschiedener LACs in der A-Daten-LAC-Folge und der ähnlichsten Routenbaum-LAC-Folge (unabhängig von der Ähnlichkeit).

Für die Festlegung der Grenzen sind vor allem die Stringvergleiche mit einer Ähnlichkeit über 60 % entscheidend, da nur diese als Trajektorien weiterverarbeitet werden. Bei hoher Ähnlichkeit nimmt die Streuung gegenüber der in Abbildung 34 dargestellten Verteilung signifikant ab, so dass für die Auswahl des Vergleichssets der LAC-Folgen aus dem Routenbaumgenerator sehr enge Grenzen gewählt werden, die in der folgenden Tabelle dargestellt sind.

Verschiedene LACs in A-Daten-LAC-Folge	Verschiedene LACs im Vergleichsstring (Routenbaum)
3	-0 bis +1
4 - 8	-1 bis +1
> 9	-2 bis +1

Tabelle 10: Auswahl der zu vergleichenden LAC-Folgen aus dem Routenbaumgenerator für die Berechnung der Ähnlichkeit.

Weitere Einschränkungen, die sich aus einer vertieften Datenanalyse ergeben, sind:

- LAC-Folgen der Länge 3 werden nicht untersucht, da diese keine Ähnlichkeit über 60 % erreichen können. Ein Matching z.B. einer LAC-Folge ABC mit einer LAC-Folge ABX (Ähnlichkeit 33 %) ist nicht gewünscht.
- LAC-Folgen mit einem Nervositätsfaktor ≥ 2 werden unabhängig von der Anzahl der unterschiedlichen LACs nicht untersucht. Zum einen werden dort sehr selten akzeptable Ähnlichkeiten erreicht, zum anderen sind diese LAC-Folgen aufgrund ihrer Länge sehr rechenintensiv.

3.7 Aufteilung der identifizierten LAC-Folgen aus VPN-Routen

Es ist möglich, dass mehrere VPN-Routen und/oder Schienenrouten identische Netz-LAC-Folgen haben. Die Aufteilung der A-Daten-Fahrten mit mehrdeutigen LAC-Folgen erfolgt in den drei Schritten Moduswahl, Anbindungswahl und Routenwahl.

3.7.1 Moduswahl

457 der insgesamt 45.553 Netz-LAC-Folgen repräsentieren sowohl VPN-Routen (IV) als auch Schienenrouten (ÖV). Es gibt verschiedene Ansätze, die eine Aufteilung der Mobilfunkteilnehmer ermöglichen sollen.

In Fastenrath (2009) wird anhand eines Beispiels gezeigt, dass sich in den Mobilfunkdaten Züge durch Häufungen von Mobilfunkteilnehmern, die mehrere vorab definierte Location Area Updates an bestimmten Funkzellen durchführen, erkennen lassen (vgl. Abbildung 35). Ein Abgleich mit Fahrplänen ermöglicht es, die Punktwolken den verschiedenen Zügen auf dem jeweiligen Abschnitt zuzuordnen.

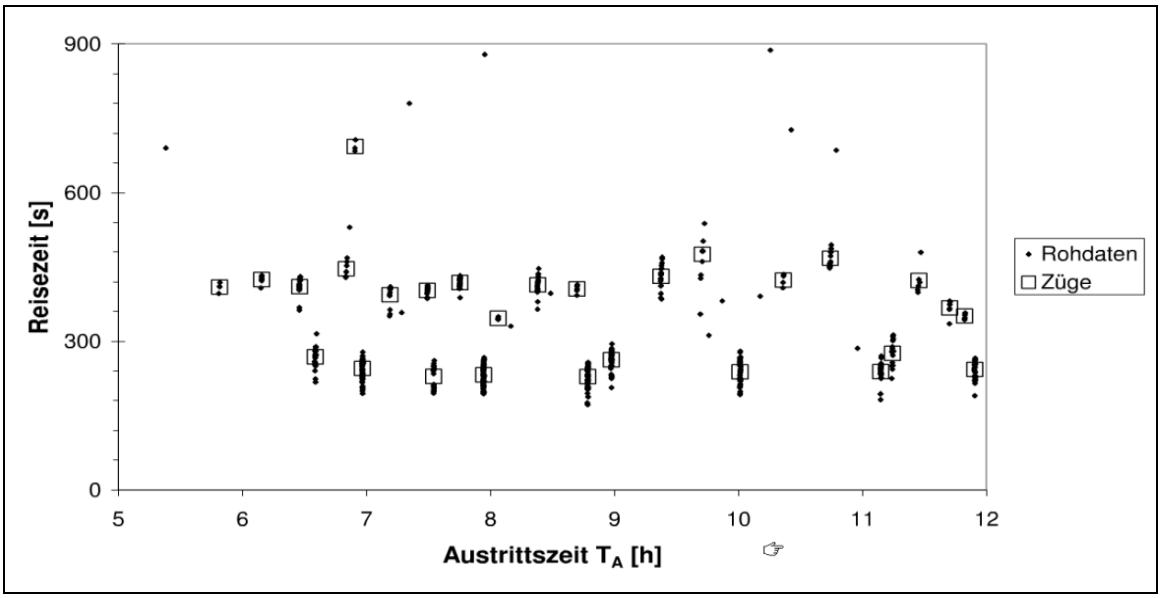


Abbildung 35: Zugerkennung aus Mobilfunkdaten (Quelle: Fastenrath 2009, S. 52).

Allerdings ist auch hier davon auszugehen, dass sich bei gebündelter Führung von Straßen und Schienenverkehrsstrecken Überlappungen ergeben, die eine eindeutige Zuordnung erschweren.

Einen anderer Ansatz (Schollmeyer & Wiltschko, 2007) basiert auf Untersuchungen einzelner Mobilfunkteilnehmer, der daher auch für schwach besetzte Busverkehre angewendet werden kann. Dazu wird das Verkehrsnetz mit Kacheln überzogen und anschließend für jede Kachel und jede Fahrt im Öffentlichen Verkehr eine zeitabhängige Zugehörigkeit ermittelt. Um eventuelle Verspätungen und Verfrühungen zu berücksichtigen, nimmt die Zugehörigkeit rund um die fahrplanmäßige Aufenthaltsdauer in einer Kachel nur langsam ab. Für eine Trajektorie wird für alle relevanten ÖV-Fahrten ein mittlerer Zugehörigkeitswert bestimmt und bei ausreichender Übereinstimmung die Fahrt als ÖV-Fahrt bewertet. Je nach Festlegung einer Mindestzugehörigkeit nimmt die Wahrscheinlichkeit für einen der folgenden Fehler zu:

- Bei einer geringen Mindestzugehörigkeit wird ein Kfz, das über einen längeren Zeitraum hinter einem Bus fährt, fälschlicherweise als ÖV erkannt.
- Bei einer hohen Mindestzugehörigkeit werden ÖV-Nutzer, die z.B. in verspäteten ÖV-Fahrzeugen fahren, als IV identifiziert.

Dieser Ansatz erfordert eine umfassende Datenbasis über den Öffentlichen Verkehr im Untersuchungsgebiet sowie sehr genaue Trajektorien.

Für das hier vorgestellte Trajektoriengenerierungsverfahren wäre es denkbar, die Fahrten mit vergleichbaren Methoden zwischen IV und ÖV aufzuteilen. Es wird aber an dieser Stelle bewusst auf eine solche Aufteilung verzichtet, da der zusätzliche Nutzen den Aufwand nicht rechtfertigt. Dies ergibt sich vor allem aus einer Analyse der nicht eindeutigen Fahrten. Diese haben, wie Abbildung 36 zeigt, zu über 50 % nur drei unterschiedliche LACs und sind somit Fahrten, deren Genauigkeit ohnehin etwas geringer

ist. Dies gilt insbesondere, da es sich zu einem großen Teil um Fahrten in der Region Stuttgart handelt, wo die S-Bahn identische LAC-Folgen mit dem Straßennetz hat.

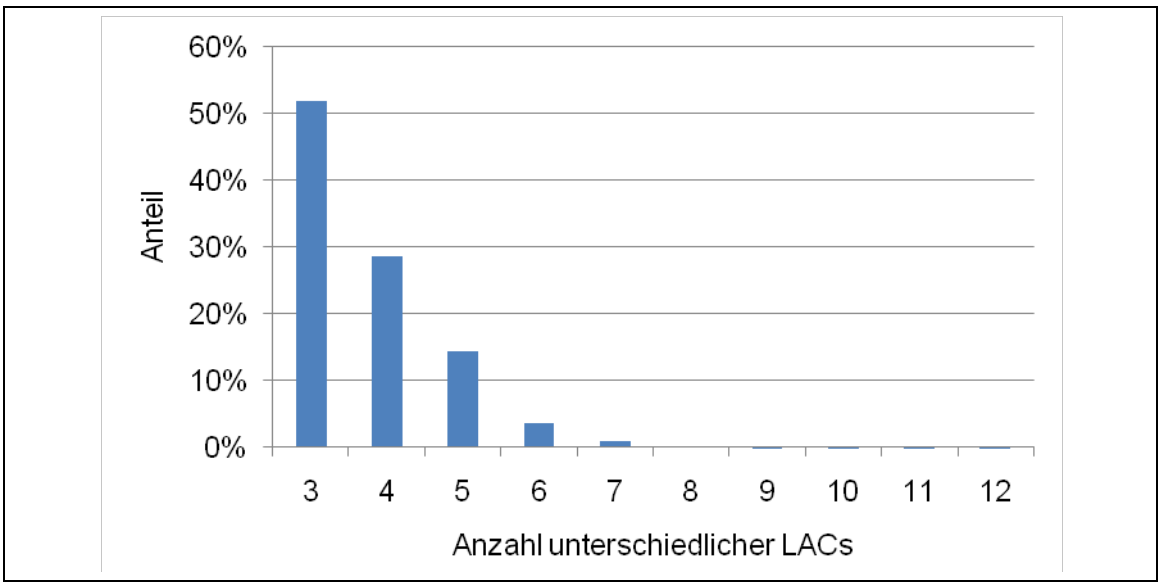


Abbildung 36: Anteil der LAC-Folgen, die nicht eindeutig einer Route zugeordnet werden können.

Daher wird für diese Fahrten eine pauschale Aufteilung zwischen IV und ÖV im Verhältnis von 80% zu 20% durchgeführt. Dies entspricht in etwa der Aufteilung der Fahrten bei eindeutigen LAC-Folgen und liegt im Bereich von Mobilitätsanalysen wie MiD (2002).

3.7.2 Anbindungswahl

Über verschiedene Anbindungen einer LAC-Verkehrszelle können identische LAC-Folgen entstehen, die in einem zweistufigen Verfahren aufgeteilt werden.

Stufe 1: Anbindungswahl über Start- und End-Cell-ID

Die Start- und End-Cell-ID stellen eine zusätzliche Information über Start und Ende der jeweiligen Fahrt dar. Auch wenn diese Cell-ID nicht zwingend Start oder Ende einer Fahrt sein muss, ist sie insbesondere in großen Location Areas (z.B. #) oder langgezogenen Locations Areas eine gute Näherung. Daher wird allen Cell-IDs eine Anbindung ins Projektnetz zugeordnet, indem für jede Cell-ID die jeweils nächste Anbindung gesucht worden ist. An wenigen Stellen, vor allem im Randbereich, erfolgt eine manuelle Anpassung der Zuordnung.

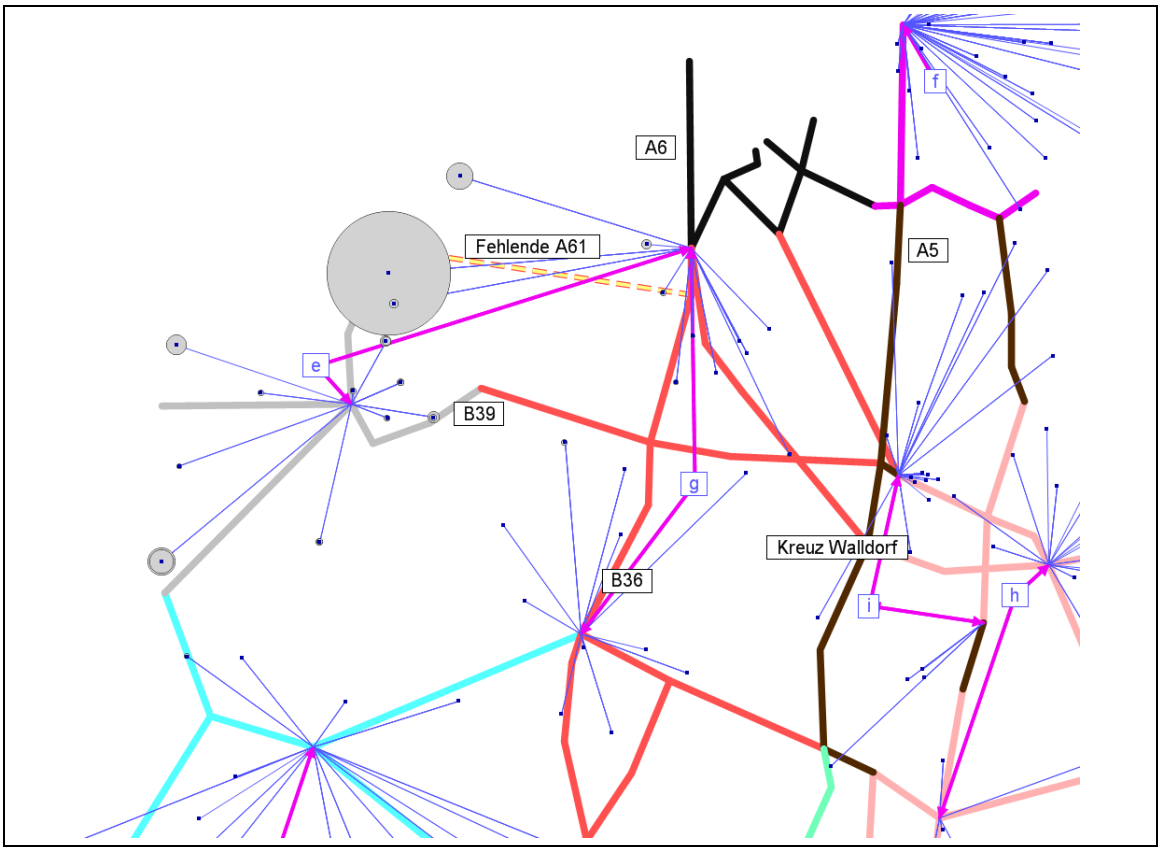


Abbildung 37: Zuordnung der Cell-IDs (blau) zu LAC-Anbindungen (lila) und das Cell-ID-Quellverkehrsaufkommen für die LAC „e“ (graue Kreise).

Abbildung 37 zeigt das Ergebnis der Zuordnung für den nordwestlichen Teil des Untersuchungsraums. Dort sind bei der LAC „e“ manuelle Korrekturen notwendig, da aufgrund der Lage am Rand des Projektnetzes die Autobahn A61 fehlt und somit eine LAC-Anbindung außerhalb der eigentlichen LAC an die A6 gesetzt worden ist. Die nördlichen Cell-IDs, deren Quellaufkommen in der LAC „e“ dominiert, werden hier manuell der Anbindung an der A6 zugeordnet, um zu verhindern, dass der Verkehr auf die B39 gelegt wird.

Zielvorstellung ist es nun, dass bei allen LAC-Folgen mit mehreren möglichen Anbindungen ein Verkehrsteilnehmer über Start- und Ziel-Cell-ID genau einer Start-und einer Zielanbindung zugeordnet werden kann. Dies kann allerdings nicht immer erreicht werden, da es für die zugeordneten Quell- und Zielanbindungen die gesuchte LAC-Folge nicht zwingend geben muss. Meistens kann aber zumindest eine Anbindung über Start- oder Ziel-Cell-ID festgelegt werden, in wenigen Fällen ist keine Zuordnung erfolgreich. Für diese Fahrten findet die Aufteilung auf die Anbindungen in Stufe 2 statt.

Stufe 2: Anbindungswahl nach manuell vorgebebener Gewichtung

Für die verbleibenden Routen mit unklarer Start- oder/und Ziel-LAC-Anbindung erfolgt die Aufteilung anhand einer manuellen Gewichtung der einzelnen Anbindungen. Eine Aufteilung nach den Widerständen der einzelnen Routen würde nicht das tatsächliche Verkehrsaufkommen widerspiegeln, sondern die jeweils am günstigsten gelegene Anbindung bevorzugen. Bei der Gewichtung werden die folgenden Größen berücksichtigt:

- Wichtigkeit der angebundenen Strecke
- Größe des Einflussbereichs des angebundenen LAC-Teilbereichs
- Einwohnerverteilung innerhalb der LAC

3.7.3 Routenwahl

Die Aufteilung von VPN-Fahrten erfolgt dagegen differenziert mit einem C-Logit-Modell (Cascetta, 2001). Dieses Modell berücksichtigt neben einer Nutzenfunktion auch noch die Eigenständigkeit E_i der betrachteten Route.

$$P_i = \frac{e^{V_i} \cdot E_i}{\sum_{j=1}^{N} \left(e^{V_j} \cdot E_j \right)}$$

mit

P_i	Auswahlwahrscheinlichkeit für Alternative j
V_i	Nutzen der Alternative i
E_i	Eigenständigkeit einer Route [0...1]
N	Zahl der Alternativen

In den Nutzen gehen die Fahrzeit im unbelasteten Netz t_0 sowie ein Indikator für ein mehrfaches Verlassen der Autobahn sowie zugehörige Skalierungsparameter β_i ein. Die Berücksichtigung von N_{BAB} soll das selten durchgeführte, kurzzeitige Verlassen von Autobahnen neben der zusätzlichen Fahrzeit zusätzlich „bestrafen“.

$$V_i = \beta_{Wid} \cdot t_0 + \beta_{BAB} \cdot N_{BAB}$$

mit

V_i Nutzen der Alternative i

t_0 Fahrzeit im unbelasteten Netz [s]

N_{BAB} Über 1 hinausgehende Anzahl des Verlassens von Autobahnen [-]

β_{Wid} Skalierungsparameter für die Größe t_0 (hier: -0,01) [1/s]

β_{BAB} Skalierungsparameter für die Größe N_{BAB} (hier: -1,00) [-]

Die Eigenständigkeit E_i berechnet sich mit Hilfe des Commonality-Faktors $C_{i,j}$ nach den folgenden Formeln:

$$C_{i,j} = \frac{L_{i,j}}{\sqrt{L_i \cdot L_j}}$$

$$E_i = \frac{1}{\sum_j C_{ij}} = \frac{1}{1 + \sum_{j, j \neq i} C_{ij}}$$

mit

E_i Eigenständigkeit der Route i [-]

$C_{i,j}$ Commonality-Faktor der Routen i und j [-]

$L_{i,j}$ Fahrzeit t_0 der gemeinsamen Strecken der Routen i und j

L_i Fahrzeit t_0 aller Strecken der Route i

Je größer die Überlappung zweier Routen i und j ist, desto größer wird der Commonality-Faktor $C_{i,j}$. Bei identischen Routen erreicht er einen Wert von 1, bei komplett getrennten Routen liegt er bei 0. Die Eigenständigkeit E_i jeder Route berechnet sich anschließend aus dem Kehrwert der Summe der Commonality-Faktoren.

Dieses Verfahren soll anhand eines Beispiels in Abbildung 38 verdeutlicht werden. Die drei gezeigten VPN-Routen nutzen identische Anbindungen und haben eine identische LAC-Folge.

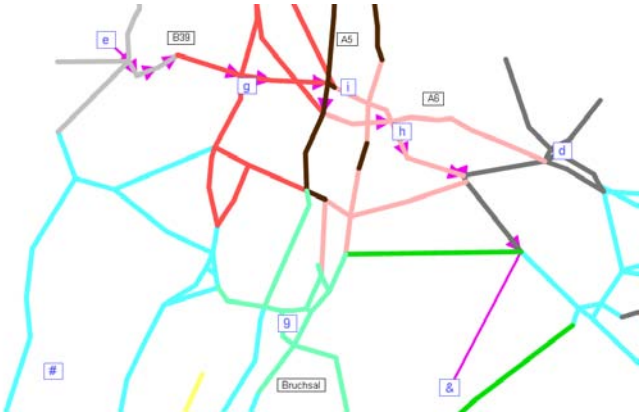
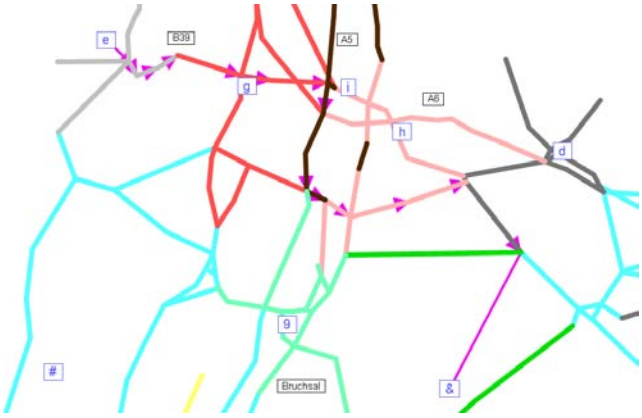
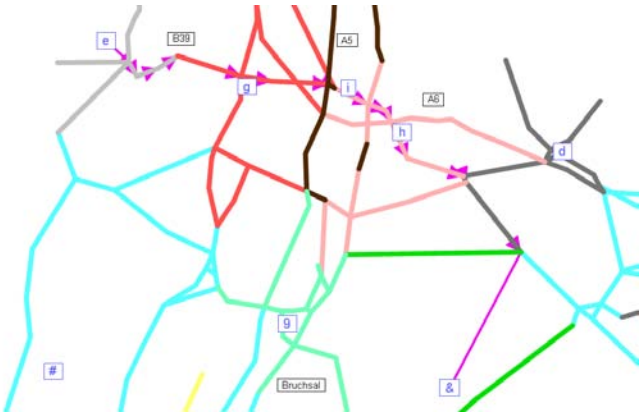
	Routennummer: 18962 Widerstand: 1970 [s] t_0 = 32,8 min $t_{Routenplaner}$ = 40 min
	Routennummer: 18966 Widerstand: 2159 [s] t_0 = 36,0 min $t_{Routenplaner}$ = 46 min
	Routennummer: 18982 Widerstand: 2069 [s] t_0 = 34,5 min $t_{Routenplaner}$ = 42 min

Abbildung 38: Drei VPN-Routen mit identischer LAC-Folge „egihd&“ mit Angabe der Fahrzeit t_0 im unbelasteten Netz und der Fahrzeit nach Google Maps (2008).

Die drei Routen unterscheiden sind im zweiten Teil der Route, wobei auch dort die Routen 18962 und 18982 relativ ähnlich sind. Deshalb sind die nach den oben genannten Formeln berechneten Eigenständigkeiten dieser beiden Routen geringer als die der Route 18966.

Nr.	$t_{0,i}$	$N_{BAB,1}$	V_i	E_i	$e^{V_i} \cdot E_i$	P_i
18962	1.970	0	-19,70	0,399	1,11E-09	64,7%
18966	2.159	0	-21,59	0,444	1,86E-10	10,9%
18982	2.069	0	-20,69	0,405	4,19E-10	24,4%

Tabelle 11: Beispielrechnung für die Routenwahl mit C-Logit.

Allerdings wird, wie die Berechnung in Tabelle 11 zeigt, der Einfluss der höheren Eigenständigkeit der Route 18966 von der höheren Fahrzeit überlagert, so dass diese Route nur einen Anteil von 10,9 % erhält. Die schnellste Route 18962 erhält 64,7 % einer LAC-Folge aus A-Daten und die Route 18982 24,4 %.

Für die ÖV-Fahrten erfolgt keine weitere Aufteilung, da diese nicht weiter verwendet werden.

3.8 Fahrzeiten einer Route

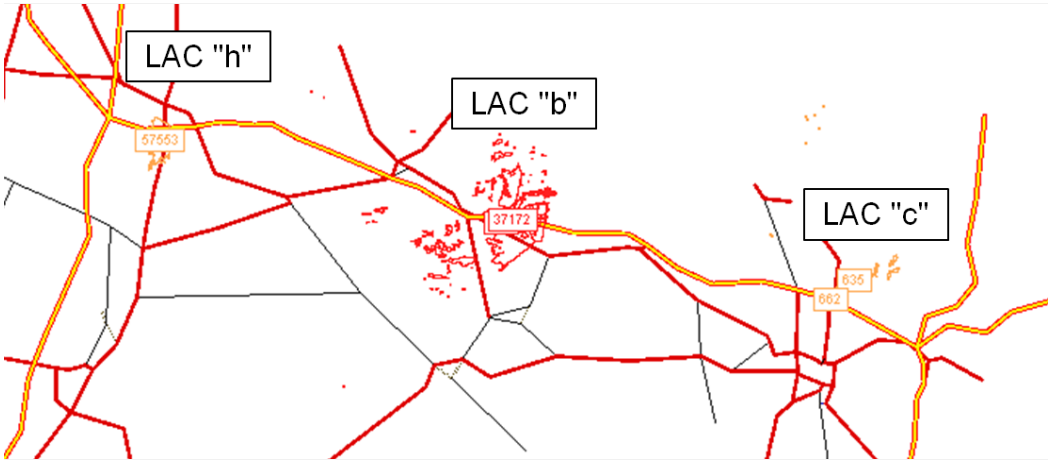
Die Umwandlung der aus A-Daten erzeugten VPN-Routen in Trajektorien erfolgt durch das Hinzufügen von Zeitinformationen für jeden VPN-Knoten der Routen. Als Minimalinformation liegen mit der ersten und letzten Meldung eines Users für jede Trajektorie zwei Zeitinformationen vor. Zwischen diesen Zeitinformationen können die Zeiten an weiteren Knoten der Route entweder längengewichtet oder gewichtet nach der Fahrzeit t_0 im unbelasteten Netz bestimmt werden. Letzteres ist dabei zu bevorzugen, da es die unterschiedlichen Geschwindigkeiten entlang einer Route berücksichtigt.

Theoretisch wäre es besser, nicht nach t_0, sondern nach der aktuellen Fahrzeit t_{akt} zu gewichten, um aktuelle Störungen entlang eine Route zu berücksichtigen. Da sich ein t_{akt} aber erst aus der Generierung der Trajektorien ergibt, ist dies nicht möglich. Um dennoch realitätsnahe Trajektorien zu erzeugen, wird im Folgenden eine Methode vorgestellt, wie neben Start- und Endzeitpunkt weitere zeitliche Stützpunkte generiert werden können.

Diese Methode basiert darauf, ausgewählte Location Area Updates, die in vordefinierten Cell-IDs stattfinden, bestimmten Stellen einer Route zuzuordnen. Da für diese Location Area Updates Zeitinformationen vorliegen, kann somit jeweils ein weiterer zeitlicher Stützpunkt gewonnen werden. Ziel ist es, für jeden der Autobahnstreckenzüge im Autobahnviereck (A5, A6, A8 und A81) pro Richtung zwei oder drei geeignete Stützstellen zu identifizieren, die eine hohe Qualität haben. Dafür werden für jede Richtung die Location Area Updates hinsichtlich der Cell-ID analysiert, wobei dort nur die Routen, die die komplette Achse mit der am häufigsten vorkommenden LAC-Folge befahren, berücksichtigt werden. Aus den fünf bis acht Location Area Updates werden dann nach den folgenden Kriterien zwei bis drei herausgesucht:

- Eine Cell-ID oder eine Gruppe von Cell-IDs lässt sich möglichst gut einem Punkt auf der Autobahn zuordnen. Wesentliches Beurteilungskriterium ist hierbei der Best-Server-Plot von T-Mobile, der erkennen lässt, ob es in der jeweiligen Fahrtrichtung eine klare Grenze zwischen den Location Areas gibt. Dies ist immer dann der Fall, wenn der Funkmast in der Nähe der Autobahn steht und eine Abstrahlung in Fahrtrichtung hat. Ebenso sind Cell-IDs mit kleiner Flächenausdehnung gut geeignet.
- Diese Cell-ID bzw. Gruppe von Cell-IDs wird von möglichst vielen der untersuchten Mobilfunkteilnehmer genutzt, so dass möglichst viele der Trajektorien tatsächlich eine Stützstelle an diesem Punkt erhalten.
- Die gefundenen Stützstellen liegen gleichmäßig auf dem Streckenabschnitt verteilt.

Die folgende Abbildung zeigt das Ergebnis des Auswahlprozesses für die A6 Fahrtrichtung Osten. Aus den Analysen ergeben sich drei Stützstellen bei den Location Area Updates in den LACs „h“, „b“ und „c“. Bezüglich der o.g. Kriterien zur Auswahl der Stützstellen kann bei den LACs „h“ und „c“ positiv festgehalten werden, dass durch die sehr kleinen Flächen der Cell-IDs eine klare Zuordnung zu einem Punkt auf der Autobahn möglich ist. Bei den Cell-IDs der LAC „c“ ist dies nicht gegeben, da diese entgegen der Fahrtrichtung abstrahlen und dadurch das Location Area Update im Best-Server-Plot in einem nicht klar abgrenzbaren Übergangsbereich liegt. Allerdings stellt diese Stützstelle sicher, dass auch in der Mitte der A6 eine zusätzliche Zeitinformation vorliegt. Die Abdeckung ist mit 84 – 94 % bei allen Stützstellen so hoch, dass ca. 96 % der untersuchten Fahrten mindestens zwei Stützstellen auf diesem Streckenabschnitt haben.



LAC "h"			LAC "b"			LAC "c"	
Cell-ID	**Anzahl**		**Cell-ID**	**Anzahl**		**Cell-ID**	**Anzahl**
57553	10277		37172	6834		635	7555
56230	149		948	2316		662	1706
15962	73		88555	623		16962	268
12342	45		52331	220		12667	152
32965	43		88206	95		29788	140
32983	34		51982	87		13287	124
...	...		...	...		...	...
• Gewählte Cell-IDs: 57553 • Abdeckung: 94 %			• Gewählte Cell-IDs: 37172 und 948 • Abdeckung: 84 %			• Gewählte Cell-IDs: 635 und 662 • Abdeckung: 85 %	

Abbildung 39: Cell-IDs der Stützstellen auf der A6 Richtung Osten (10.890 untersuchte Mobilfunkteilnehmer mit der LAC-Folge „...ihdbc...“).

Abbildung 40 zeigt schematisch die drei unterschiedlichen Methoden. Die Überschätzung der Geschwindigkeiten auf den Bundestraßen bei der längengewichteten Interpolation wird durch eine t_0-Gewichtung korrigiert. Durch die Stützstellen wird zusätzlich erreicht, dass die Geschwindigkeiten noch differenzierter berechnet werden. In diesem Fall ergibt sich auf der Autobahn zwischen den Stützstellen eine höhere Geschwindigkeit und entsprechend in den anderen Bereichen niedrigere Geschwindigkeiten.

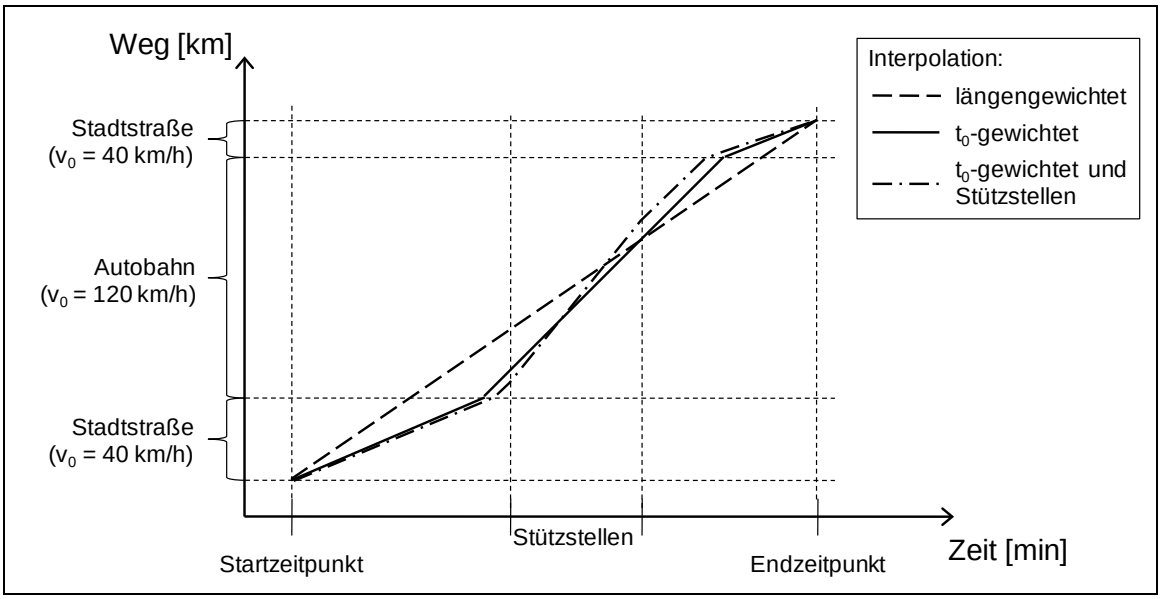


Abbildung 40: Drei Methoden zur Ermittlung des Zeit-Weg-Verlaufs einer Route.

3.9 Ergebnisse für das Testfeld

3.9.1 Ergebnisse für einen Tag

Die Ergebnisse des oben beschriebenen Verfahrens werden im Folgenden detailliert anhand des 23. Juni (Montag) beschrieben. Für diesen Tag sind ca. 37,7 Millionen Datensätze auf der A-Ebene vorhanden.

A-Daten-Datenbereinigung

Tabelle 12 zeigt die Auswirkungen der Datenbereinigungsschritte auf die Anzahl der Datensätze sowie die Anzahl der UsrIDs. Vor allem die Reduktion der Daten auf bewegte Teilnehmer mit drei oder mehr unterschiedlichen LACs führt zu einer deutlichen Reduktion der Anzahl der Datensätze und der Anzahl der verbleibenden UsrIDs. Die danach folgende Abbildung 41 zeigt die Belastungen auf virtuellen Strecken zwischen den Mittelpunkten der Location Areas (inkl. der Fremd-LACs), die bereits die Belastung entlang der Autobahnen erahnen lassen.

Kap.	Beschreibung		Datensätze		Anzahl UsrIDs
			Anzahl	Differenz	
	Rohdaten		37.665.000		11.787.948
3.2.1	Nicht verwertbare Einträge	cLac ≤ 0	37.490.097	- 174.903	
		UnixZeit ≤ 0	31.319.768	- 6.170.329	
		Falscher Tag	31.234.506	- 85.262	
		User-ID ≤ 0	31.234.504	- 2	
		Sonstiges	31.233.087	- 1.417	
3.2.2	Bewegte Teilnehmer		7.622.846	- 23.611.658	
3.2.3	Gruppieren		5.955.614	- 1.667.232	772.841
3.2.4	Fremd-LACs		5.164.155	- 791.459	660.528

Tabelle 12: Statistik der Datenbereinigung.

Die Anzahl gelöschter Einträge aufgrund einer ungültigen Unix-Zeit ist an anderen Tagen deutlich geringer, da die A-Daten-Aufzeichnung in der network probes durch ein Update der dort installierten Software verbessert werden konnte. Auswirkungen auf die sonstigen Ergebnisse hat dieser Verbesserung allerdings nicht, da es sich bei den Einträgen fast ausschließlich um Einträge während aktiver Verbindungen ohne Wechsel der Location Area handelt, die ohnehin im Schritte „Gruppieren“ gelöscht werden.

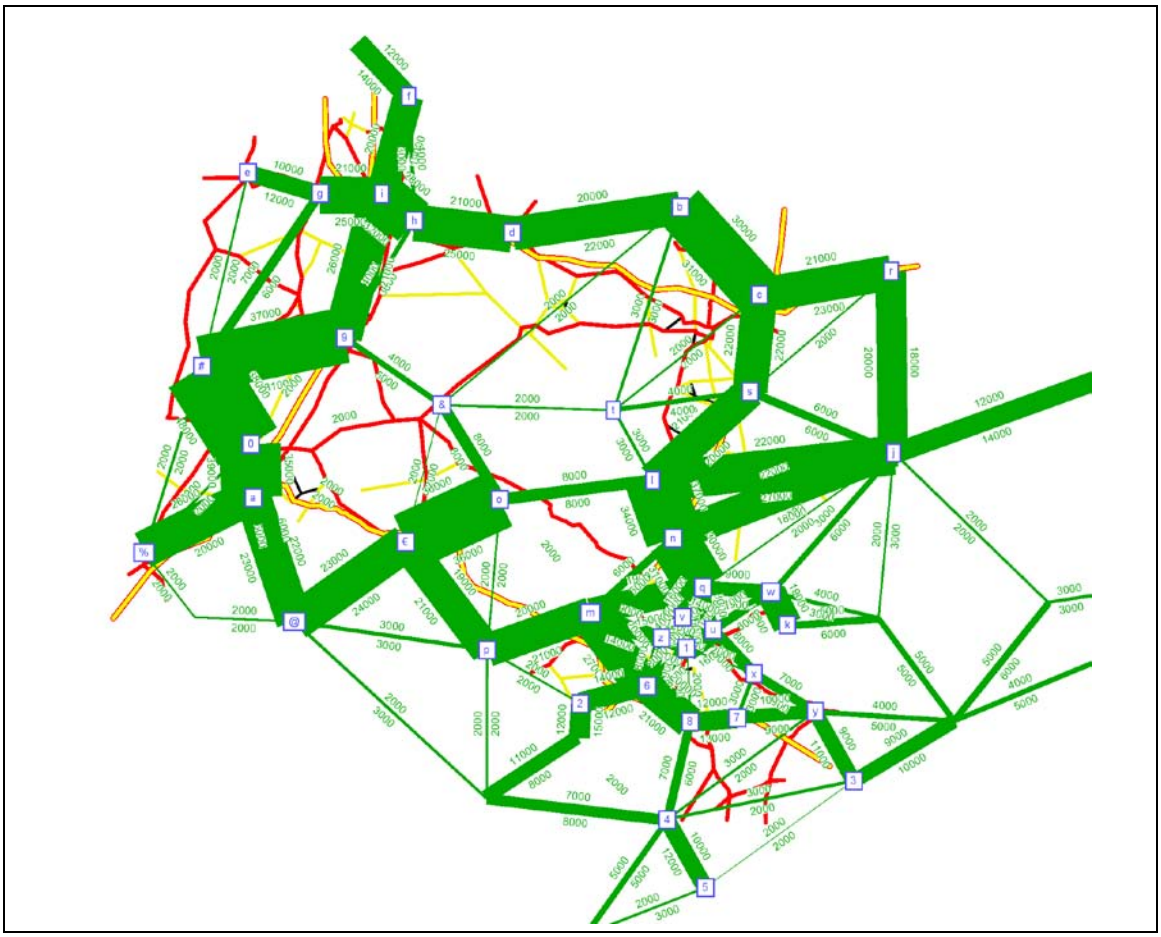


Abbildung 41: Belastungen auf virtuellen Strecken zwischen den Location Areas (ohne Buchstaben = Fremd-LAC).

Fahrtidentifikation

Bei der Fahrtidentifikation werden teilweise mehrere Fahrten für eine UsrID erzeugt, manche UsrIDs entfallen dagegen komplett. Die angegebenen Werte beziehen sich dabei jeweils auf das Ende des Schrittes, beinhaltet also nach dem Anwenden der eigentlichen Regel auch das Gruppieren von LAC-Einträgen und Löschen von Fahrten mit weniger als drei unterschiedlichen LACs.

Kap.	Beschreibung	Datensätze		Anzahl UsrIDs
		Anzahl	Differenz	
	Datenbereinigung	5.164.155		660.528
3.3.1	60min-Regel	4.985.046	- 179.109	683.739
3.3.2	Erweiterte 60min-Regel	4.355.308	- 629.738	658.902
3.3.3	Nervositätsregel	3.977.691	- 377.617	617.638

Tabelle 13: Statistik für die Fahrtidentifikation.

Matching mit LAC-Folgen aus Routenbaum

In der folgenden Tabelle sind die Ergebnisse der drei verschiedenen Stufen des Matching der LAC-Folgen aus A-Daten mit den LAC-Folgen aus dem Routenbaum-generator abgedruckt.

Kap.	Beschreibung	Identifizierte Fahrten			Anteil identifzierter Fahrten	
		Straße	Schiene	Gesamt	je Schritt	kumuliert
3.6.1	Direkter Stringvergleich	209.938	36.494	246.432	39,9%	39,9%
3.6.2	„ABAB“-Regel	73.221	7.021	80.242	13,0%	52,9%
3.6.3	Stringähnlichkeitsvergleich	136.575	38.395	174.970	28,3%	81,2%
	Gesamt	419.733	81.911	501.644		

Tabelle 14: Statistik für das Matching der LAC-Folgen (gerundet).

Die folgenden Abbildungen zeigen die Belastungen im VPN, die sich aus einer pauschalen Aufteilung bei der Modus-, Anbindungs- und Routenwahl ergeben. Die Aufteilung der mehrdeutigen LACs erfolgt hier gleichmäßig auf alle VPN- und Schienenrouten. Die Ergebnisse sind daher nur qualitativ zu bewerten, was allerdings für die folgenden Analysen ausreichend ist.

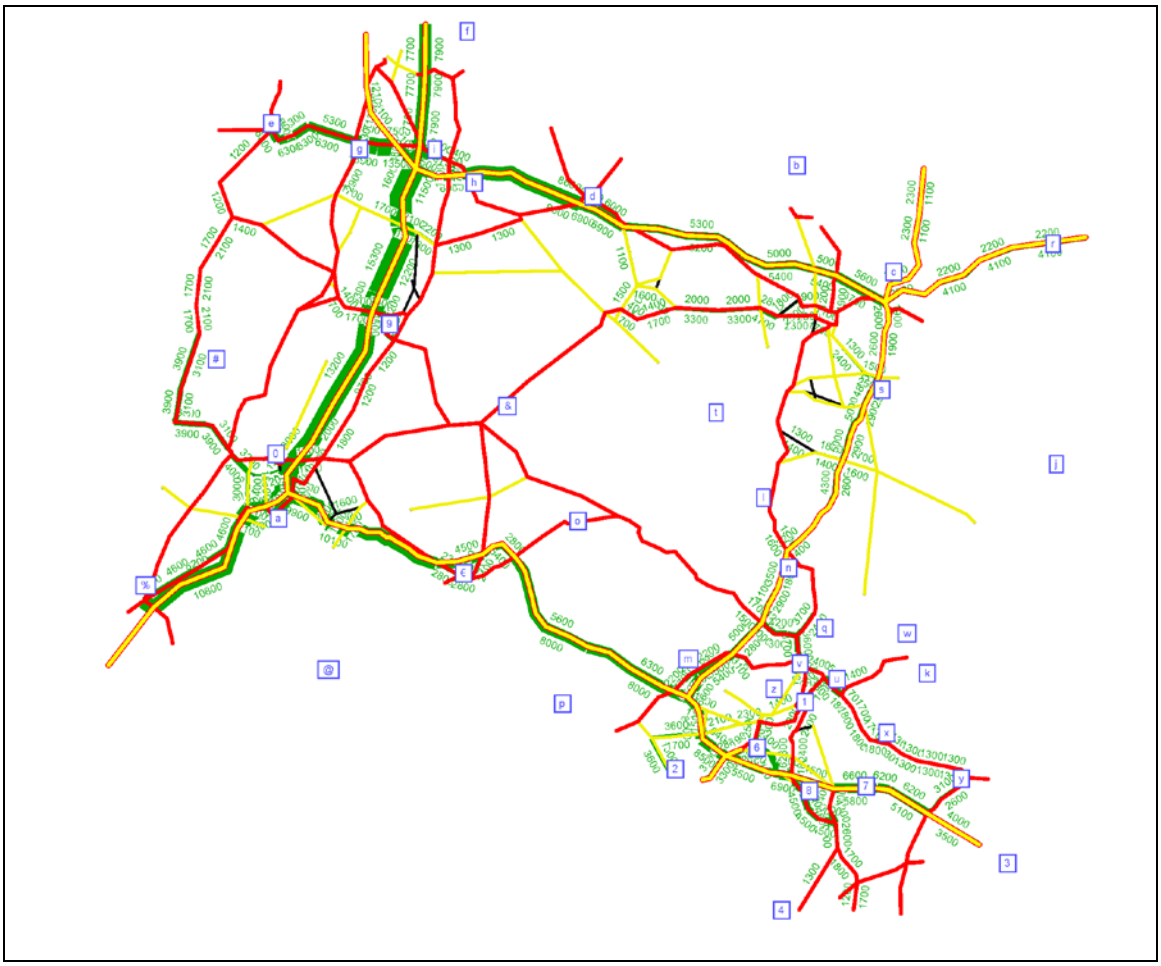


Abbildung 42: Belastungen aus den Fahrten des direkten Stringvergleichs.

Nach dem ersten Schritt zeigt sich deutlich eine räumliche Schiefe der Ergebnisse. Während die Belastungen zwischen Karlsruhe und Walldorf bei 10.000 – 15.000 Fahrzeugen liegen, sind es auf der Achse zwischen Stuttgart und Heilbronn meist weniger als 5.000 Fahrzeuge.

Diese räumliche Schiefe nimmt durch das Matching mit der ABAB-Regel ab, wie die folgende Abbildung 43 verdeutlicht, in der die durch die Anwendung der ABAB-Regel hinzugekommenen Fahrten zu sehen sind. Dies liegt vor allem an dem Systemrand hinter Heilbronn Richtung Nürnberg. Dort wechseln Mobilfunkteilnehmer auf nicht-VPN-Strecken der A6 zwischen den LACs „r“ und „j“ hin und her. Daher ist bei den in diesem Schritt identifizierten Fahrten der IV-Anteil auch besonders hoch. Die folgende Abbildung 43, in der die durch die Anwendung der „ABAB-Regel“ hinzugekommenen Fahrten zu sehen sind, verdeutlicht dies.

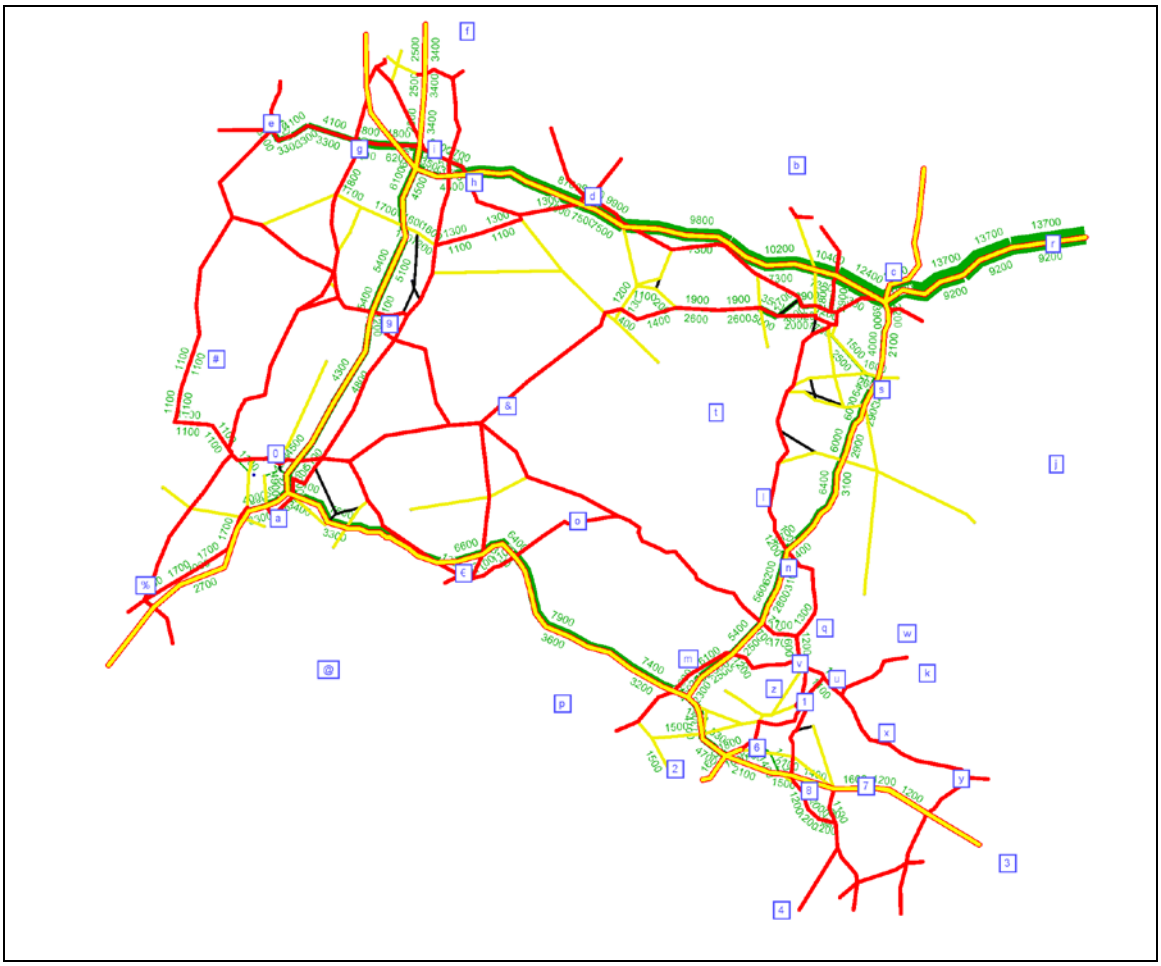


Abbildung 43: Belastungen aus den Fahrten der „ABAB-Regel“.

Durch den Stringähnlichkeitsvergleich können noch einmal 28,3 % der bislang nicht identifizierten Fahrten einer LAC-Folge aus dem Routenbaumgenerator zugeordnet werden. Die folgende Abbildung zeigt die Belastungen aus den Fahrten durch den Stringähnlichkeitsvergleich.

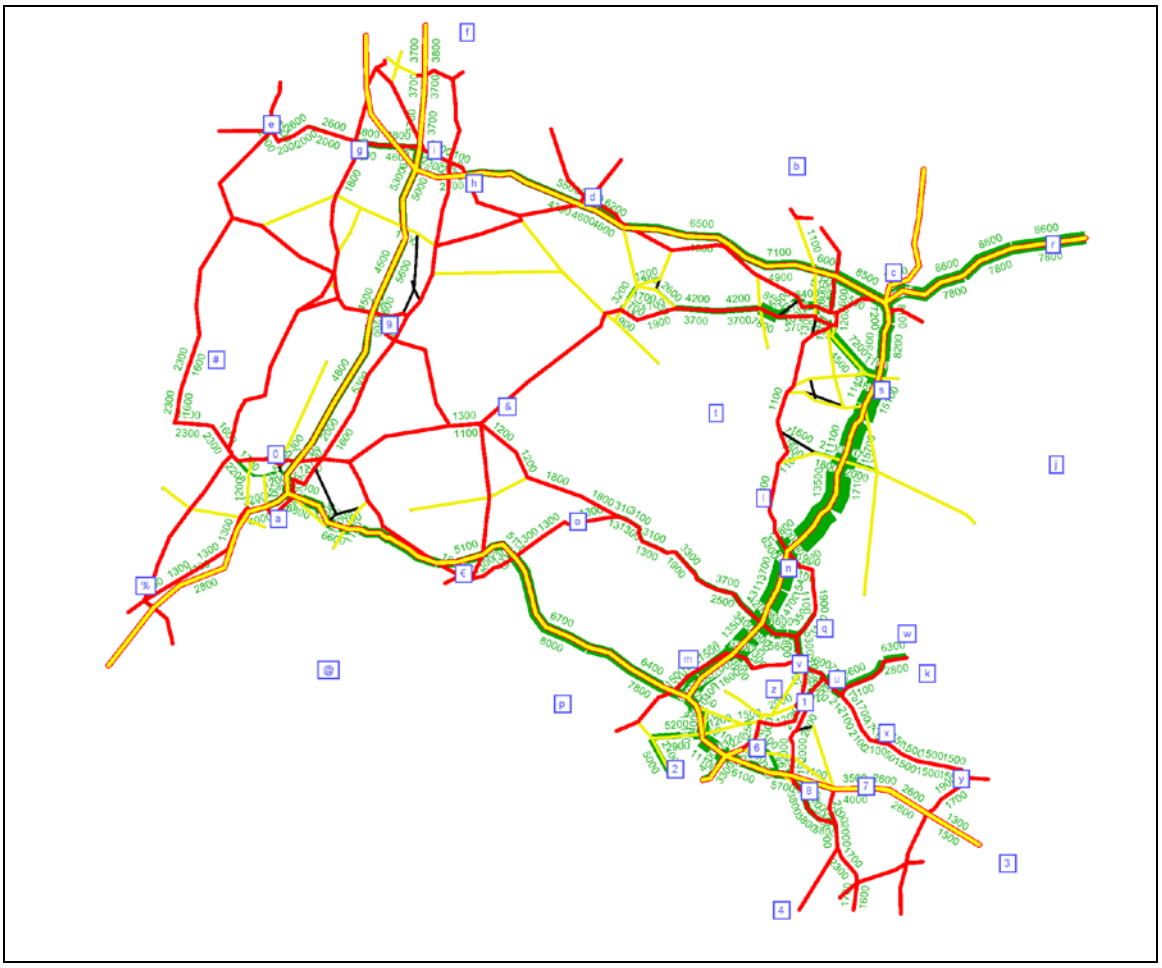


Abbildung 44: Belastungen aus den Fahrten des Stringähnlichkeitsvergleich.

Auch hier wird die bisher noch bestehende räumliche Schiefe weiter ausgeglichen. Insbesondere auf der Achse Stuttgart – Heilbronn gibt es viele zusätzliche Fahrten. Dies lässt sich vor allem auf eine Besonderheit des Mobilfunknetzes an einem Streckenabschnitt bei Ludwigsburg-Nord zurückführen. Hier wechseln ca. 70 % der Mobilfunkteilnehmer aus dem LAC „n“ vorübergehend in den LAC „l“, während die anderen diesen Wechsel nicht ausführen (vgl. Abbildung 45).

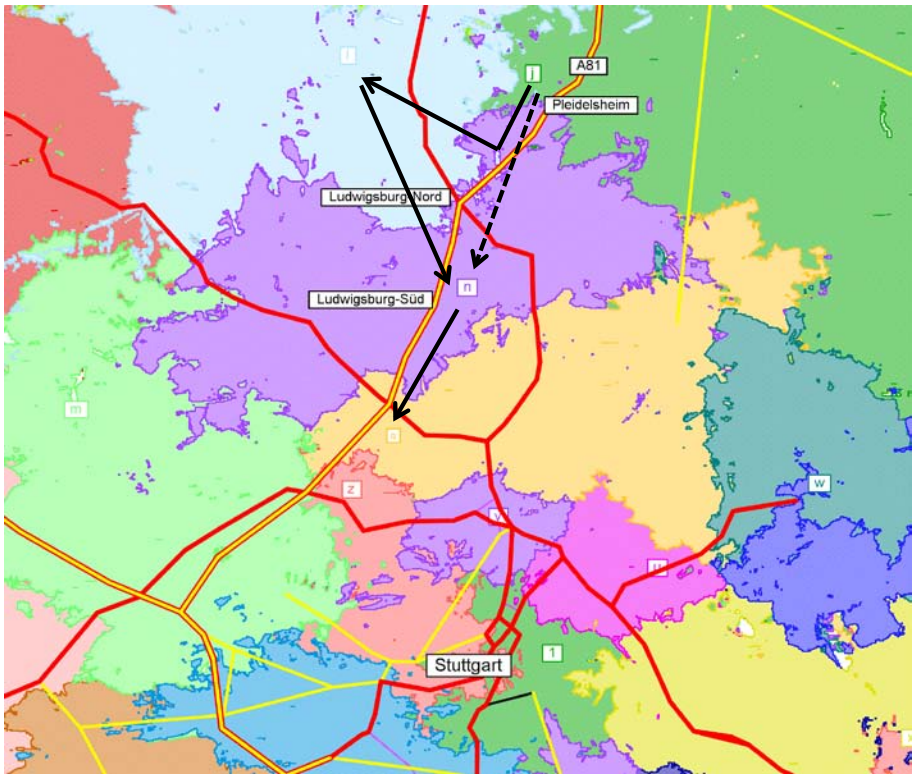


Abbildung 45: LAC-Wechsel in der Nähe von Ludwigsburg-Nord (durchgezogene Linien = Standard-LAC-Folge / gestrichelt: alternative LAC-Folge).

Aufteilung der identifizierten LAC-Folgen aus VPN-Routen

Während bislang die Aufteilung mehrdeutiger LAC-Folgen ohne Gewichtung erfolgte, soll nun das oben ausführlich beschriebene Aufteilungsverfahren angewendet werden. Dadurch ergibt sich eine Aufteilung der Fahrten auf 419.733 Fahrten im VPN sowie 81.911 Fahrten im Schienennetz sowie auf alternative Routen innerhalb der zwei Verkehrssysteme. Die folgende Abbildung 46 zeigt die Belastungen im VPN nach der Aufteilung.

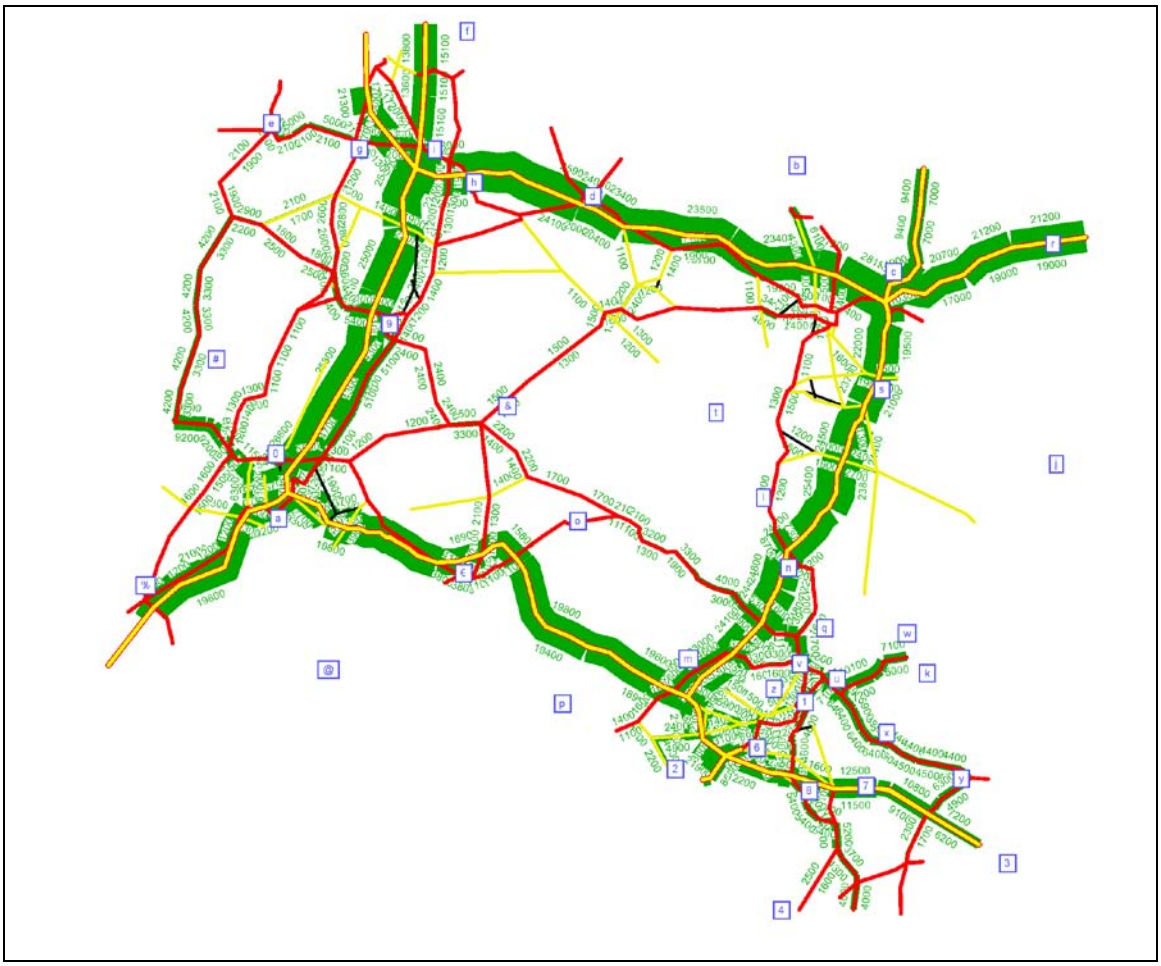


Abbildung 46: Belastungen im VPN nach Modus-, Anbindungs- und Routenwahl.

Das Verhältnis zwischen den Belastungen aus A-Trajektorien und stationären Detektoren liegt bei den Messstellen in der BAB-Netzmasche im Mittel bei 40-50%. Eine zeitlich differenzierte Analyse der Belastungen zeigt, dass diese den bekannten Mustern von Ganglinien auf Fernstraßen folgen. Die folgende Abbildung 47 zeigt für die A8 zwischen Karlsruhe-Nord und Bruchsal die Belastungswerte einer stationären Zählstelle und der Trajektorien.

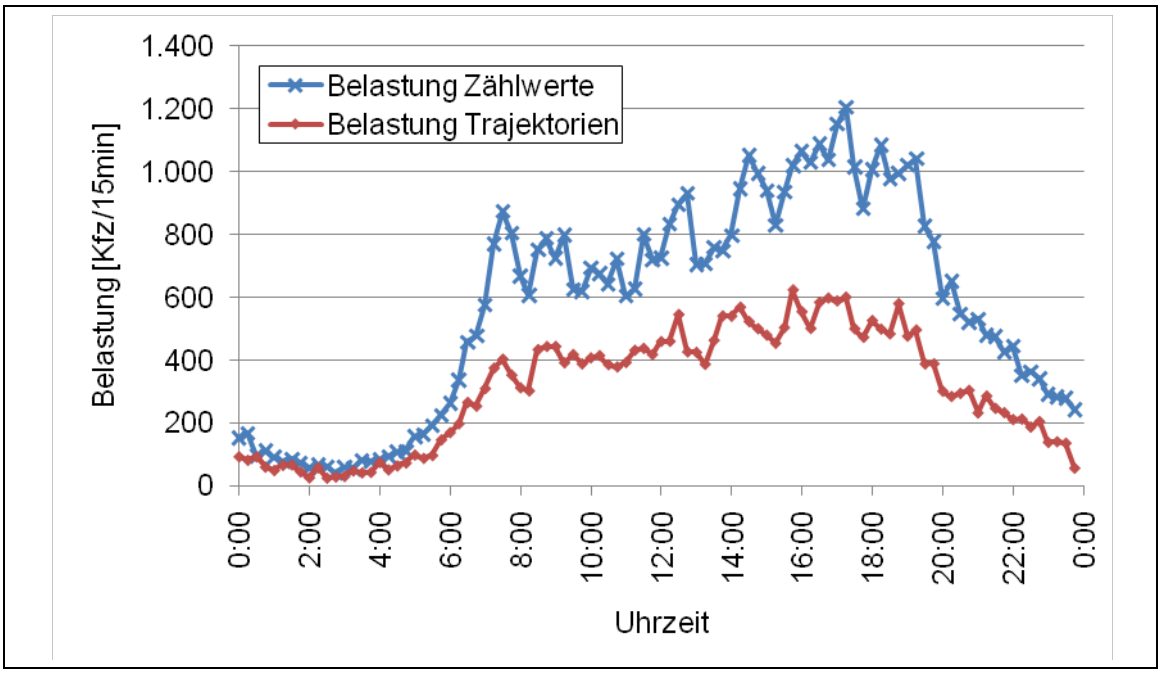


Abbildung 47: Streckenbelastungen aus stationärer Zählstelle und FPD-Trajektorien auf der A81 zwischen Karlsruhe-Nord und Bruchsal.

Am Untersuchungstag gab es abends eine Vollsperrung der A8 zwischen Pforzheim-Ost und Pforzheim-Nord. Der Verkehr wurde dort örtlich ausgeleitet. Die getrennte Detektion einer Umleitungsstrecke durch Pforzheim ist nicht möglich, da sie in derselben Location Area liegt. Die Abbildung 48 zeigt deutlich den Rückgang der Belastungen in beiden Richtungen während der Vollsperrung. Nach Aufhebung der Vollsperrung gegen 20 Uhr steigt die Belastung noch einmal kurz an.

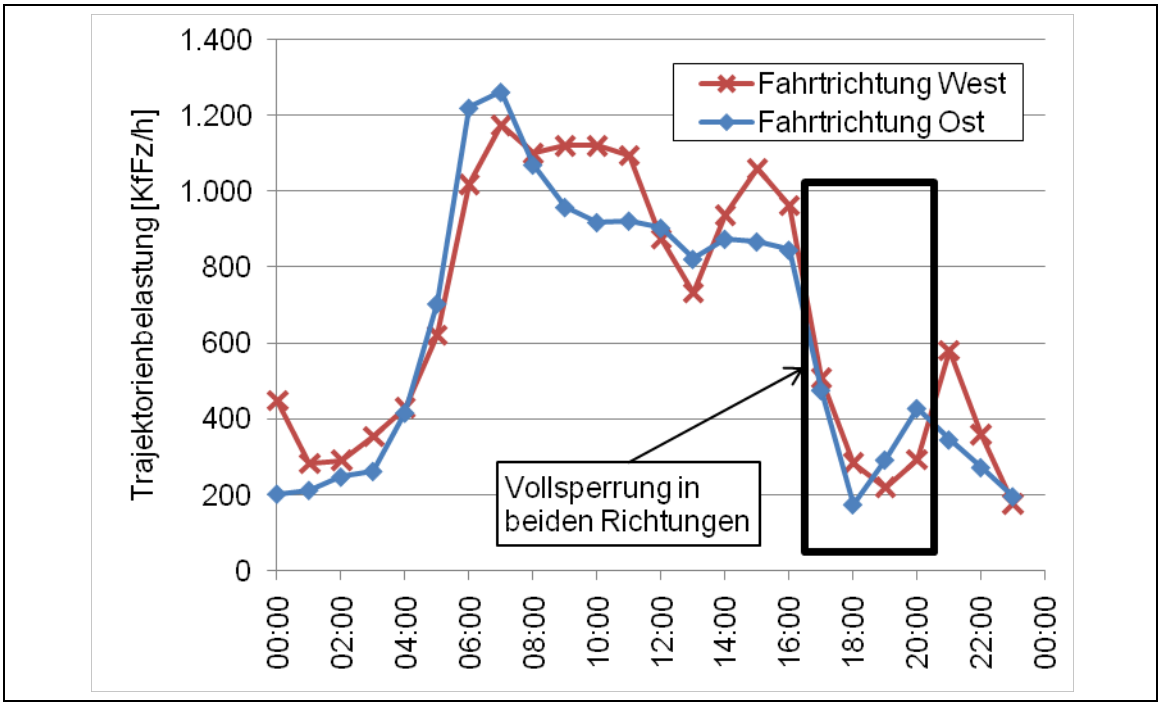


Abbildung 48: Streckenbelastungen aus Trajektorien auf der A8 zwischen Pforzheim-Ost und Pforzheim-Nord.

Die folgende Abbildung 49 zeigt als abschließende Auswertung die Fahrtweitenverteilung der Fahrten im VPN. Als Vergleich ist die Fahrtweitenverteilung nach der Erhebung Mobilität in Deutschland (MID, 2002) in der Abbildung eingetragen. Die Fahrtweiten im VPN-Netz haben eine deutlich größere Länge, da viele kurze Fahrten durch das Verfahren und die Beschränkung auf Strecken des Projektnetzes nicht erfasst werden.

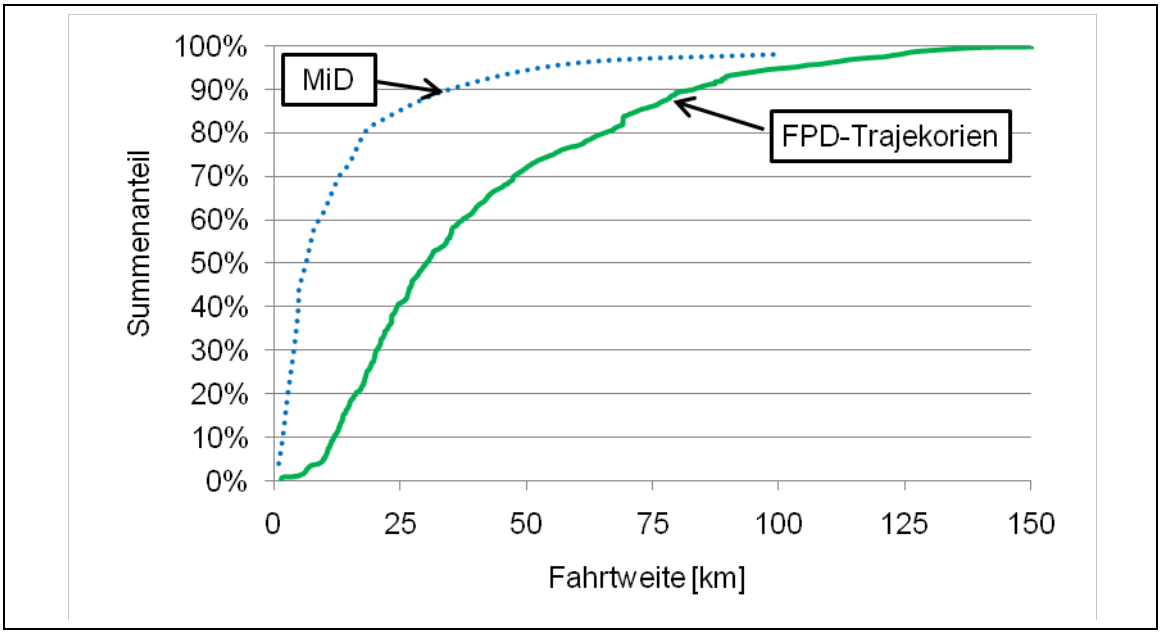


Abbildung 49: Fahrtweitenverteilung aus FPD-Trajektorien und der Erhebung „Mobilität in Deutschland“ (MID, 2002).

Auswertung Fahrzeiten

Die Auswertung der Fahrzeiten stellt eine Validierung der generierten Trajektorien dar und ermöglicht – wie die folgenden Beispiele zeigen – darüber hinaus ermöglichen, die Fahrzeugklassen Lkw und Pkw zu erkennen.

Das erste Beispiel in Abbildung 50 zeigt die Fahrzeiten auf einem ca. 10 km langen Autobahnabschnitt entlang der A81 zwischen den Anschlussstellen Ludwigsburg und Mundelsheim. Die meisten Messungen ergeben Geschwindigkeiten zwischen 86 km/h (Lkw) und 145 km/h (Pkw).

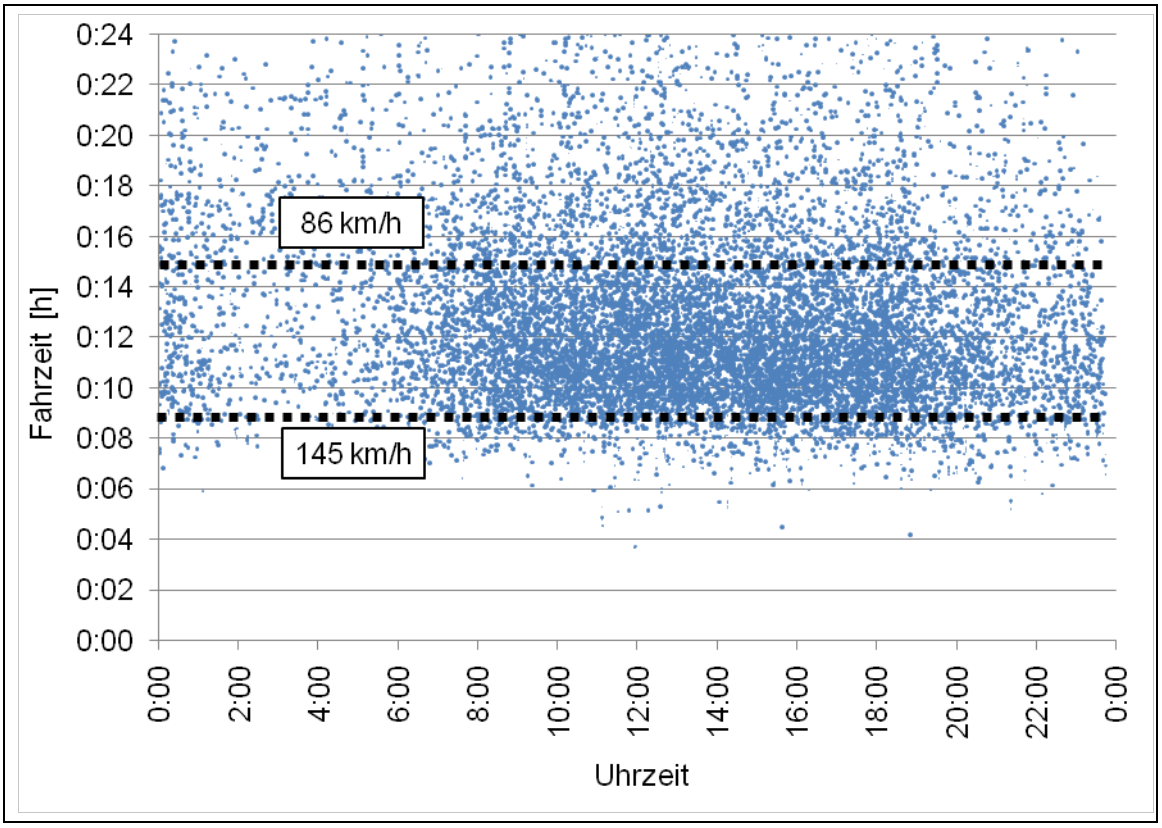


Abbildung 50: Fahrzeiten von der AS Ludwigsburg zur AS HN/Untergruppenbach (A81), ca. 22 km.

Eine weitere Analyse in Abbildung 51 für den ca. 90 km langen Abschnitt vom Autobahndreieck Leonberg über Karlsruhe zum Kreuz Walldorf zeigt, dass sich bei längeren Strecken Punktwolken für Pkw bei ca. 120 km/h und für Lkw (bzw. langsame Pkw) bei 85 km/h ergeben. Gegen 10 Uhr gibt es eine größere Störung, bei der die Punktwolken ineinander übergehen.

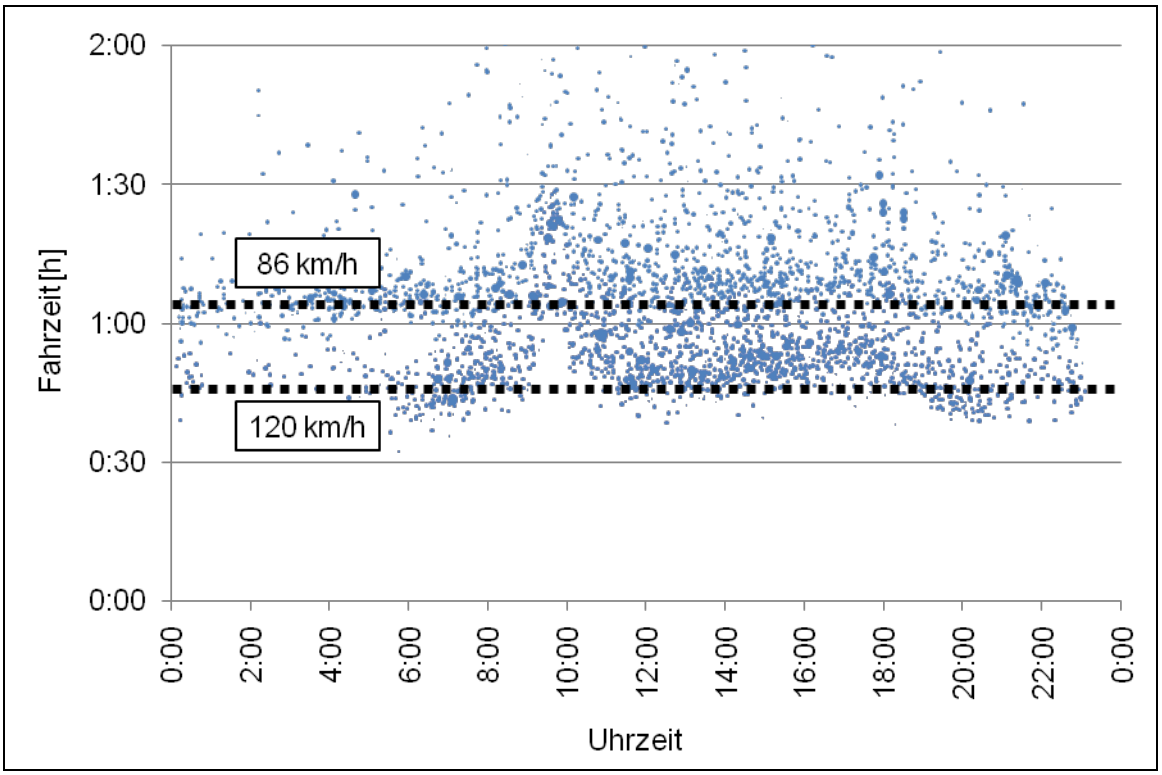


Abbildung 51: Fahrzeiten vom Dreieck Leonberg über Karlsruhe zum Kreuz Walldorf (A8 & A5), ca. 90 km.

Beim zweiten Beispiel ist ein hoher Lkw-Anteil von ungefähr 50 % erkennbar, der über den Werten von 10 - 15 % aus manuellen Zählungen (BASt, 2007) liegt. Dies lässt sich auf zwei Gründe zurückführen.

1. On-Board-Units GSM-Module im Netz von T-Mobile
   Nach Angaben der Betreiberfirma Toll Collect Gmbh (2008) sind fast 70 % der mautzahlenden Lkw in Deutschland mit On-Board-Units ausgestattet, der Anteil der Mautbuchungen mit On-Board-Units liegt bei über 90 %. Das GSM-Modul der On-Board-Units erscheint in den A-Daten wie ein normales Mobilfunkgerät, so dass nahezu jeder Lkw zusätzlich zu den ansonsten mitgeführten T-Mobile-Geräten ein weiteres Gerät mit sich führt.

2. Fahrtweitenverteilung von Lkw vs. Pkw
   Für den hohen Anteil an Lkw auf langen Routen spricht darüberhinaus die von Wermuth (2007) veröffentlichte Fahrtweitenverteilung von Pkw und Lkw. Danach haben Lkw > 3,5 t eine mittlere Fahrtweite von 55 km, während diese bei Pkw zwischen 16 und 26 km beträgt. Auch wenn sich diese Werte nicht direkt auf die untersuchten Autobahnen übertragen lassen, kann von einem erhöhten Lkw-Anteil bei längeren Autobahnrouten ausgegangen werden.

Weitere Gründe für den hohen Anteil der Lkw könnten in GSM-basierten Flottenmanagementsystemen, dem hohen Geschäftskundenanteil von T-Mobile und im Ausbauzustand des Netzes von T-Mobile, das allgemeinhin als sehr hochwertig bezeichnet

wird, liegen. Diese Aspekte lassen sich mangels verlässlicher Zahlen allerdings weder belegen noch quantifizieren.

Der hohe Lkw-Anteil verursacht insbesondere durch die On-Board-Units eine Schiefe der Stichprobe, ermöglicht aber auch eine relativ einfache Trennung von Pkw- und Lkw. Diese wird für die Routenwahl und das Monitoring der Angebotsqualität (Kap. 4 und 7) durchgeführt.

3.9.2 Auswertung im Längsschnitt

Die Auswertungen in den vorherigen Abschnitten beziehen sich jeweils auf einen Untersuchungstag. Im Folgenden sollen einige Analysen zeigen, dass das Verfahren auch über einen längeren Zeitraum konstant gute Ergebnisse liefert. In den folgenden zwei Abbildungen sind jeweils 78 aufeinander folgende Tage zwischen dem 08. Juli und dem 06. Oktober dargestellt.

Abbildung 52 zeigt auf der linken Achse die Anzahl der Fahrten (d.h. Ortsveränderungen von Mobilfunkgeräten) pro Tag. Auch wenn einige Tage teilweise oder komplett aufgrund von Problemen bei der A-Daten-Aufzeichnung fehlen, lässt sich ein Muster aus jeweils 5 Werktagen mit höherer Nachfrage und Samstag und Sonntag mit einer geringeren Nachfrage erkennen. Während der Schulferien in Baden-Württemberg (24.07. - 06.09.) nimmt die Nachfrage etwas ab.

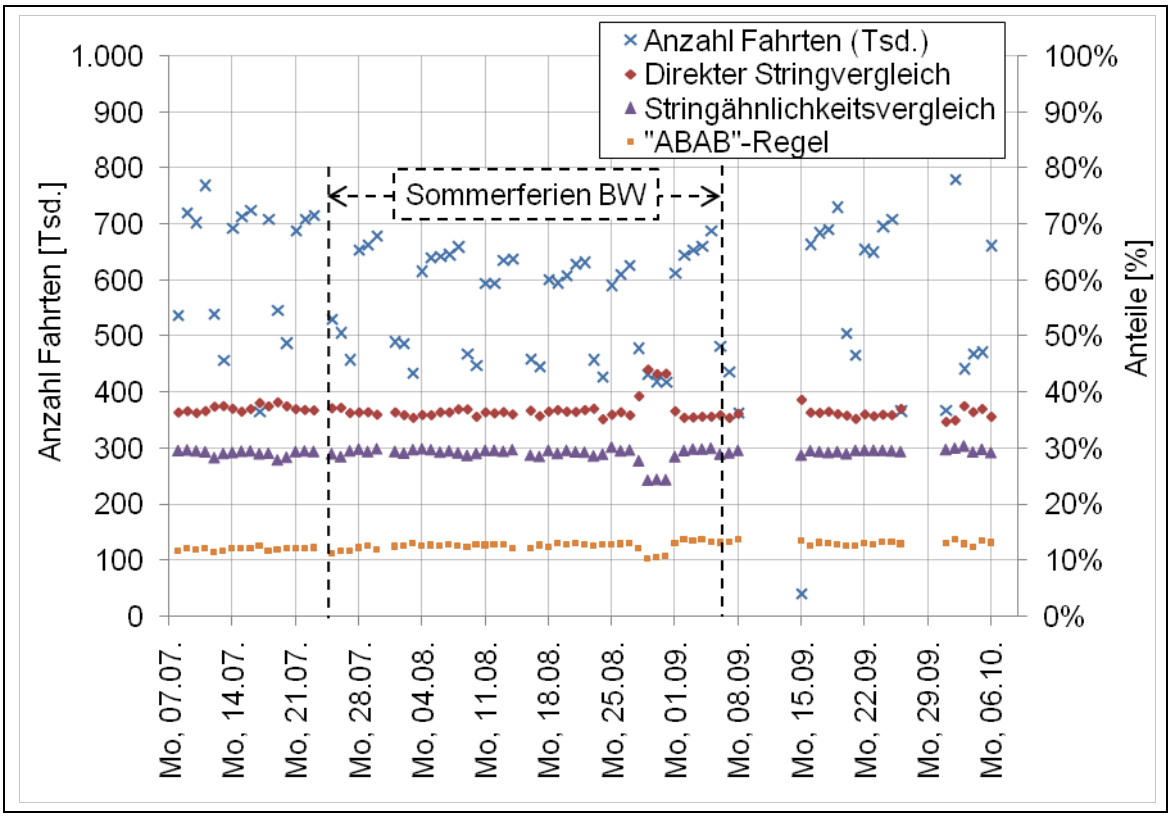


Abbildung 52: Analyse des Trajektoriengenerierungsverfahrens über 78 Tage.

Auf der rechten Achse sind die Anteile der drei Stufen des Matching von LAC-Folgen aus A-Daten und aus dem Routenbaumgenerator aufgetragen. Diese Anteile sind wie auch weitere globale Auswertungen, abgesehen von drei Tagen Ende August, relativ konstant.

An diesen Tagen mit abweichendem Verhalten ist der Anteil der Fahrten, die im direkten Stringvergleich erkannt werden, leicht erhöht, während die anderen Anteile entsprechend abnehmen. Eine weitere Auffälligkeit an diesen Tagen ist eine um ca. 30 % unter den Werten vergleichbarer Tage liegende mittlere Fahrtweite. Diese und weitere Kenngrößen (z.B. Statistiken der Datenreduktion) zeigen, dass an diesen Tagen die Teilnehmerverfolgung im Mobilfunknetz fehlerhaft war. Mobilfunkteilgeräte haben eine neue UsrID bekommen, obwohl sie regelmäßig Location Area Updates hatten. Dadurch erhöht sich der Anteil der UsrID, die als unbewegte Teilnehmer (weniger als 3 unterschiedliche LACs) erkannt werden und die verbleibenden Fahrten sind im Durchschnitt kürzer. Da eine nachträgliche Korrektur der Rohdaten nicht möglich ist, werden die als problematisch erkannten Tage nicht weiter verwendet.

3.10 Diskussion der Einbeziehung der Cell-IDs

Das vorgestellte Verfahren zur Generierung von Trajektorien verwendet die Information, in welcher Cell-ID ein Location Area Update stattgefunden hat, nur bei der Start- und End-Cell-ID zur Auswahl der am nächsten liegenden Anbindung der jeweiligen Location Area und als Stützpunktinformation für die Zeit. In diesem Abschnitt soll erläutert werden, warum diese Information nicht auch bei der Routenwahl bei mehrdeutigen LAC-Folgen verwendet wird.

Der wesentliche Grund für das Vernachlässigen dieser Information ist, dass die jeweils verwendete Cell-ID nicht zwingend diejenige ist, die nach dem Best-Server-Plot der T-Mobile die Optimale wäre. Dies hat zwei Gründe:

- Der Best-Server-Plot stellt nach Angaben der T-Mobile nur eine ungefähre Näherung dar (Felderhoff, 2005). Für eine exakte Berechnung der Ausbreitung der Funksignale wären sehr detailierte Gelände- und Bebauungsinformationen nötig.
- Je nach Auslastung der einzelnen Funkzellen und der Abschirmung des einzelnen Mobilfunkgerätes werden unterschiedliche Cell-IDs verwendet.

Diese Ungenauigkeiten bei der Ortung können zu falschen Routen führen. Dies soll anhand der A5 und der parallel führenden B3 nördlich von Karlsruhe verdeutlicht werden. Für die Analyse wurden ca. 500 Mobilfunkteilnehmer ausgewählt, die aufgrund der LAC-Folge, den Fahrzeiten sowie der Cell-IDs in den LACs „0“ und „i“ eindeutig der Süd-Nord-Route von Karlsruhe Richtung Walldorf zugeordnet werden können. Die folgende Abbildung zeigt die wichtigsten Cell-ID-Wechsel dieser User sowie die Ausdehnungen der Cell-IDs gemäß dem Best-Server-Plot der T-Mobile.

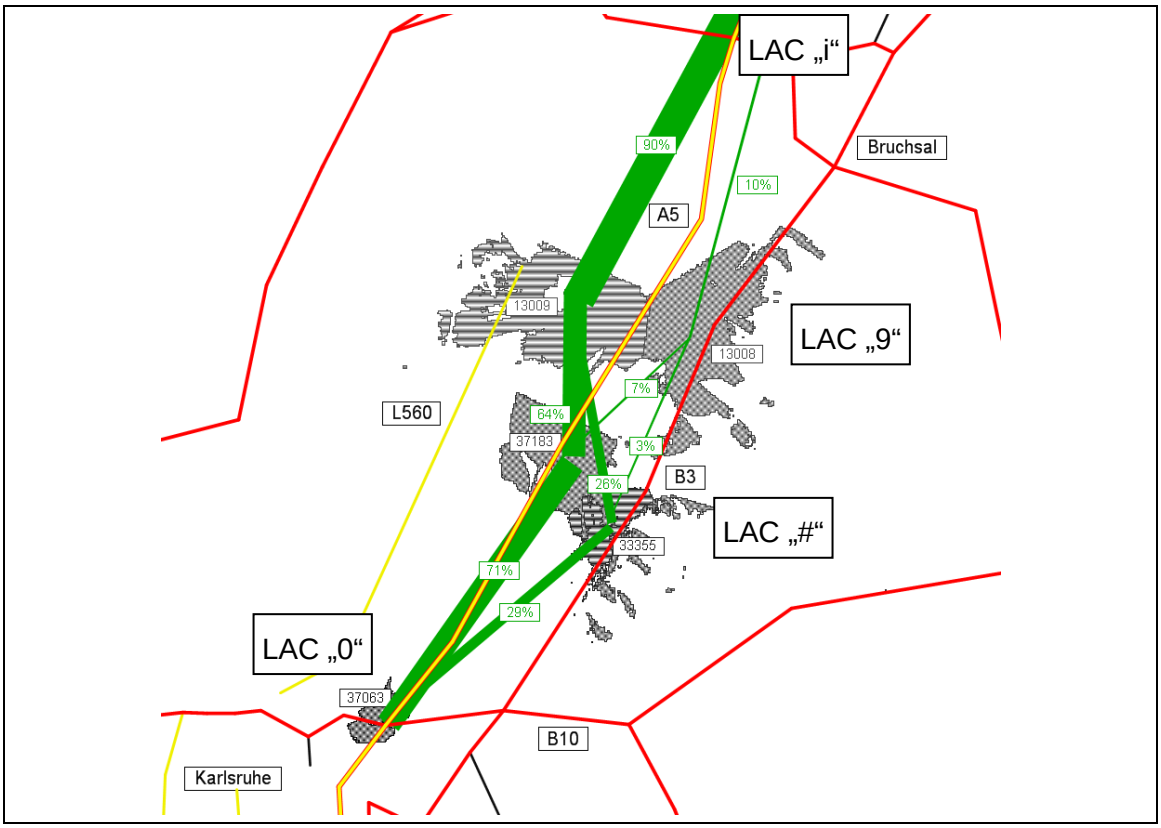


Abbildung 53: Cell-ID-Wechsel ausgewählter Mobilfunkteilnehmer (vereinfachte Darstellung).

Ca. 30 % der Mobilfunkteilnehmer melden sich in der LAC „#“ in der Cell-ID 33355 an, obwohl diese gemäß dem Best-Server-Plot der B3 zugeordnet werden müsste. Beim nächsten LAC-Wechsel wechseln davon wieder ca. 90 % in die Cell-ID 13009, die auch nach dem Best-Server-Plot als erstes erreicht wird. Die restlichen 10 % verwenden für diese Location Area Update die Cell-ID 13008, die der B3 zugeordnet werden müsste. Somit würden 3 % der Route über die B3 zugeordnet werden, weitere 33 % würden nur dann korrekt zugeordnet werden, wenn Algorithmen verwendet werden, die die Autobahn bevorzugen. Nur 64 % könnten zweifelsfrei der A5 zugeordnet werden.

Bereits die 3 % der Fahrzeuge, die in diesem Beispiel falsch der B3 zugeordnet werden, würden sich auf Routenwahlschätzungen erheblich auswirken. Die Route über die B3 hat einen Umwegfaktor von ca. 2,5 (bezogen auf die t_0-Fahrzeit), so dass sie abgesehen von großen Stauereignissen auf der Autobahn sehr selten genutzt wird und somit bereits 3 % des relevanten Verkehrs eine erhebliche Verzerrung der Stichprobe darstellt. Entsprechend würde sich das ebenfalls negativ auf die Generierung von Quelle-Ziel-Matrizen auswirken.

Trotz dieser Feststellung wird an der Nutzung der Start- und End-Cell-ID für die Anbindungswahl festgehalten, da die Anbindungswahl in der Regel großräumiger

Entscheidungen trifft, bei denen sich die genannten Probleme nur in Ausnahmefällen auf das Ergebnis auswirken.

3.11 Verfahrensanpassungen für die Netzbeeinflussungsanlage Stuttgart

Im Bereich der Netzbeeinflussungsanlage Stuttgart (A81, B10, B295, B27) können die Ergebnisse des Trajektoriengenerierungsverfahrens nicht ohne weiteres für Routenwahlanalysen übernommen werden. Dies hat im Wesentlichen drei Ursachen, die im Folgenden mit der gewählten Lösung erläutert werden:

1. *Problem:*
   In der Region Stuttgart gibt es durch die in der Nähe zu wichtigen Straßen geführten S-Bahn-Linien (z.B. B10/27 zwischen Hauptbahnhof und Pragsattel) häufig bereits den Fall, dass eine Aufteilung zwischen IV und ÖV nicht eindeutig anhand der LAC-Folgen möglich ist.

   Lösung:
   Um das Problem der möglichen Verwechslung mit dem Schienennetz zu vermeiden, werden nur Fahrten berücksichtigt, die über die parallel geführten Schienenstrecken hinaus auf der Autobahn bis Mundelsheim gehen und bei denen daher keine Verwechselungsgefahr besteht.

2. *Problem:*
   Bei einer kleinräumigen Analyse ist der Einfluss des Stringähnlichkeitsvergleichs sowie der Einfluss nicht im Projektnetz vorhandener Strecken nicht zu vernachlässigen.

   Lösung:
   Es erfolgt eine manuelle Zuordnung von LAC-Folgen zu den Routen. Dabei ergeben sich für den kurzen Abschnitt bereits 54 unterschiedliche LAC-Folgen.

3. *Problem:*
   Die Lage *der* Location Areas führt dazu, dass die Routen über die B10 und B27 im Regelfall identische LAC-Folgen haben.

   Lösung:
   Es erfolgt eine manuelle Zuordnung der Cell-IDs beim Location Area Update von der LAC „q“ zu der LAC „n“ zu den Routen der B10 und B27. Abbildung 54 zeigt die Cell-ID-Wechsel, die sich unter Berücksichtigung der Abstrahlwinkel zuordnen lassen.

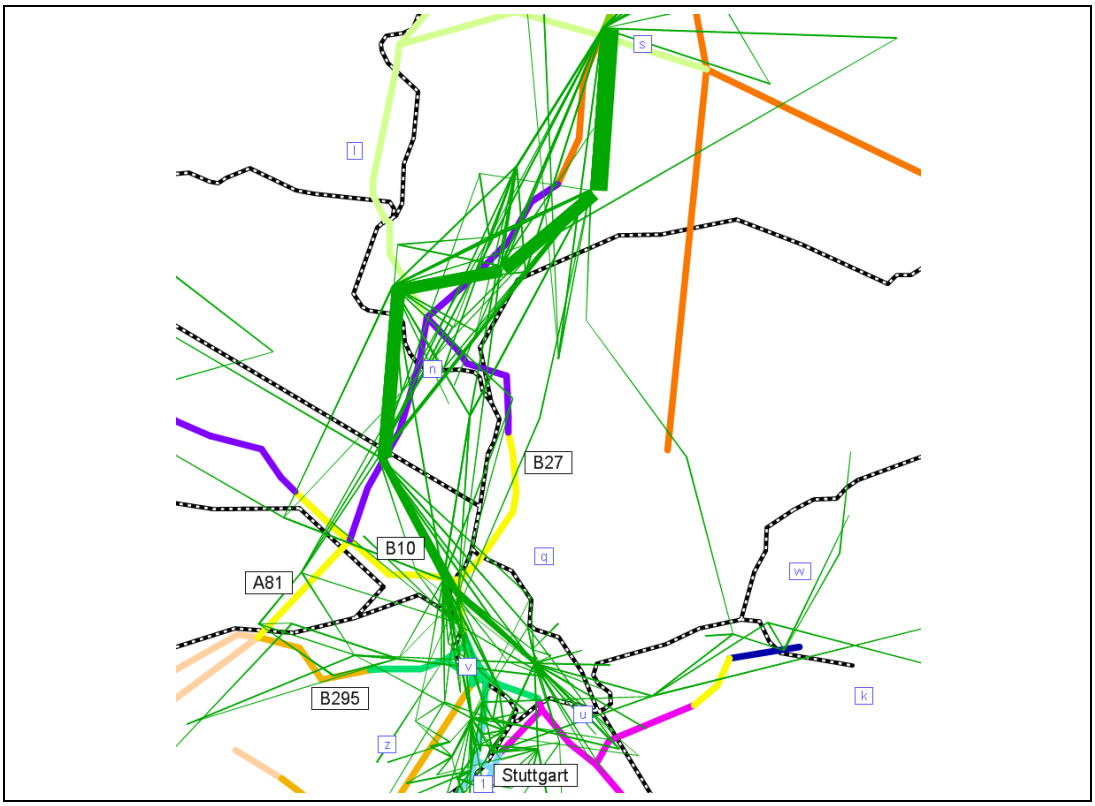


Abbildung 54: Cell-ID-Wechsel (grün) der Mobilfunkteilnehmer zwischen den LACs „v“ und „s“.

Dieses verfeinerte Verfahren ermöglicht die Analyse der Routenwahl im Bereich der Netzbeeinflussungslage. Es zeigt exemplarisch, wie es möglich ist, das vor allem LAC-basierte Verfahren für kleinräumige Analysen zu verbessern. Aufgrund des hohen manuellen Aufwands bei der Zuordnung von LAC-Folgen und Cell-IDs, der überproportional zur Größe des Untersuchungsgebietes steigt, lässt es sich nicht auf das gesamte Untersuchungsgebiet ausweiten und stößt zudem auch bei eng beieinander liegenden Strecken, wie im vorherigen Kapitel erläutert, an seine Grenzen.

3.12 Bewertung

Das hier vorgestellte Verfahren zur Generierung von Trajektorien erzeugt eine Vielzahl von Trajektorien, da alle Mobilfunkteilnehmer mit mindestens drei unterschiedlichen Location Areas in die Auswertung eingehen. In diesem Abschnitt sollen nun die Qualität und die Einsatzmöglichkeiten der erzeugten Trajektorien bewertet werden. Dabei werden die drei Einsatzbereiche

- Autobahnen (BAB),
- außerstädtische Bundes- und Landstraßen (U-Strecken), und
- Stadtverkehr (KA und LHS)

unterschieden.

Beschränkung des Verkehrsnetzes auf das Projektnetz

Längere Fahrten werden in der Regel auf BAB, Bundestraßen oder sonstigen U-Strecken sowie ausgewählten anderen Strecken durchgeführt. Zielsetzung bei der Auswahl des Projektnetzes im AP1 war es, genau dieses relevante Streckennetz abzubilden. Daher ist eine Beschränkung des untersuchten Verkehrsnetzes auf das VPN und damit die Vernachlässigung des nachgeordneten Straßennetzes für die Generierung von Trajektorien aus Mobilfunkteilnehmern mit mindestens drei unterschiedlichen LACs ausreichend. Einzig im Stadtverkehr Stuttgart ergeben sich dabei Einschränkungen, da das Netz der Location Areas sehr engmaschig ist und damit drei unterschiedliche Location Areas auch auf nicht im VPN abgebildeten Strecken durchfahren werden können.

Beschränkung der Ortungsgenauigkeit auf Location Areas

Eine wesentliche Beschränkung des Verfahrens ergibt sich aus der Tatsache, dass die kleinste verwendete räumliche Einheit für die Ortung die Location Areas sind. Diese haben teilweise sehr große Ausdehnungen, wie die folgende Abbildung 55 zeigt.

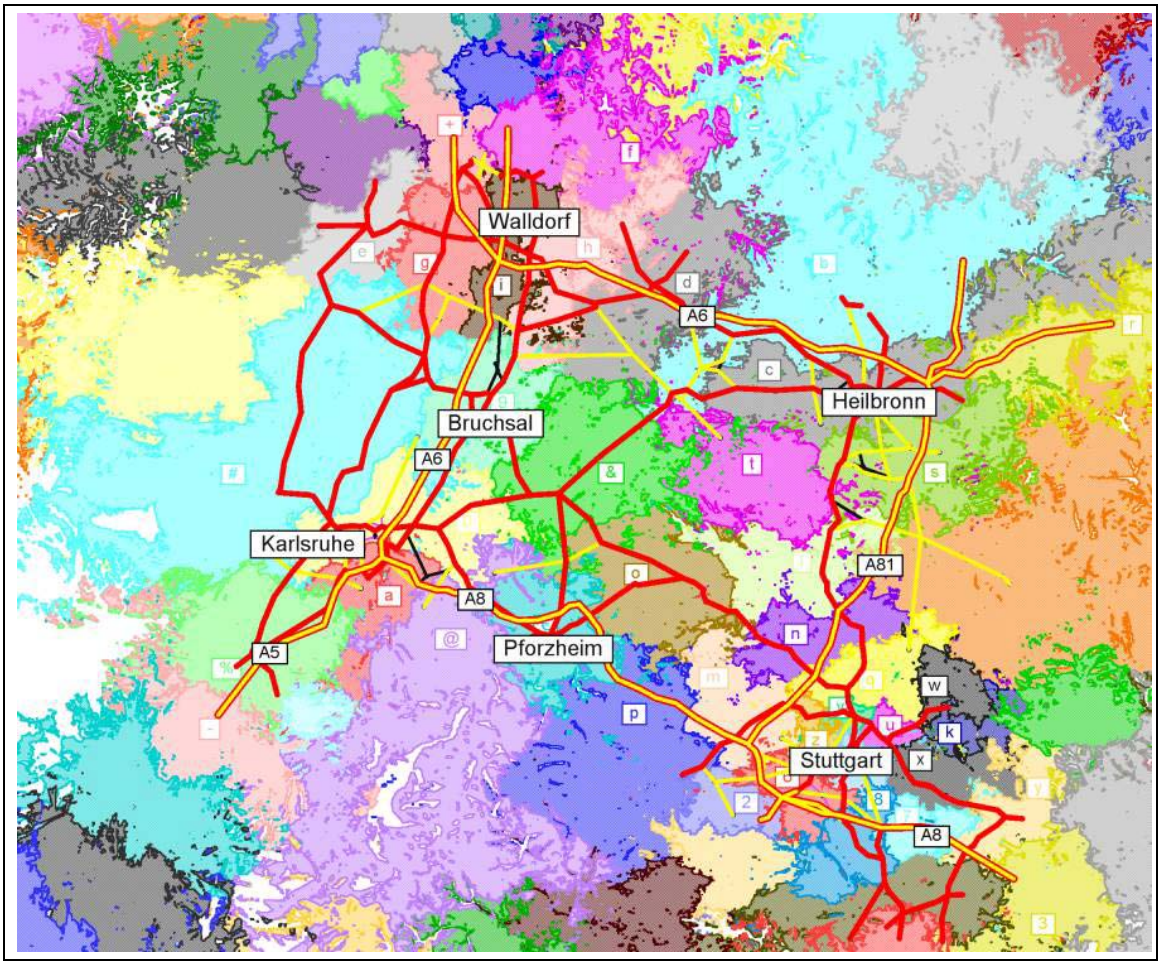


Abbildung 55: Projektnetz mit Location Areas.

Große LAs machen es teilweise unmöglich, bei einer dort beginnenden Fahrt die richtige Bundesstraße als Start der Fahrt zu bestimmen. Im Fall der LAC „#“ nordwestlich von Karlsruhe (türkis) sind z.B. die Bundesstraßen B9 und B36, die beide mit einer großen Länge in dieser LA liegen. Trotz allem ist auf ausgewählten Strecken aber auch innerhalb dieser großen LAs eine Ortung möglich, wenn längere Fahrtbeziehungen durch die jeweilige LA betrachtet werden.

Entlang der BAB ergeben sich zumeist eindeutige LAC-Folgen, da es dort keine auffallend großen LAs gibt. U-Strecken der BAB ergeben in einigen Fällen unterschiedliche LAC-Folgen als die jeweilige Hauptstrecke und können dann mit hoher Wahrscheinlichkeit korrekt detektiert werden. Hier muss allerdings für jede U-Strecke geprüft werden, ob eine entsprechend günstige Lage der Location Areas vorliegt.

Innerstädtisch ist die Einheit der Location Areas zu groß. Insbesondere in Karlsruhe, aber in Stuttgart mit kleineren Location Areas, gibt es eine Vielzahl von möglichen Routen beim Durchqueren der Location Areas. Innerstädtisch ist auch der Einfluss des nicht detektierbaren Binnenverkehrs (vgl. Kapitel 2.3) sehr groß.

Aufteilung mehrdeutiger LAC-Folgen

Bei der Aufteilung der mehrdeutigen LAC-Folgen werden bestimmte Annahmen getroffen. Diese beeinflussen das Ergebnis auf den drei Netzebenen unterschiedlich.

	Auswirkung auf Genauigkeit für ...		
	städtische Netze	**U-Strecken**	**BAB**
Aufteilungsverhältnis IV – ÖV	hoch	mittel	sehr gering
Gewichtung Anbindungen	hoch	gering	sehr niedrig
Aufteilungsverfahren für Routenwahl	sehr hoch	mittel	sehr niedrig

Tabelle 15: Bewertung der Schritte bei der Aufteilung mehrdeutiger LAC-Folgen.

Die Wahl des Aufteilungsverhältnisses zwischen IV und ÖV wirkt sich vor allem auf städtische Gegenden aus, in denen zwischen S-Bahn und dem Straßennetz auf vielen Relationen nicht anhand der LAC-Folge unterschieden werden kann. Je länger eine Route ist, desto unwahrscheinlicher ist es im Untersuchungsgebiet, dass sich für IV und ÖV identische LAC-Folgen ergeben, so dass sich diese pauschale Aufteilung auf BAB nur sehr gering auswirkt.

Die Gewichtung der Anbindungen wirkt sich vor allem auf die feineren Netzebenen aus. Je höher die betrachtete Netzebene ist, desto eher ist der Verkehr bis zum Erreichen dieser Ebene ohnehin wieder zusammengeflossen. Durch die Verfeinerung der Anbindungswahl mit der Start- bzw. Ziel Cell-ID sind die Auswirkungen für die BAB sehr gering.

Ähnliches gilt für das Aufteilungsverfahren bei der Routenwahl. Zumeist handelt es sich dort um Alternativen innerhalb von Städten. So liegen z.B. die B14, B27, B27a und der Arnulf-Klett-Platz im Stuttgarter Zentrum innerhalb einer LA. Fälle, wie das zur Verdeutlichung des Verfahrens gewählte Beispiel (vgl. Abbildung 38), bei denen es mehrere attraktive Alternativen im übergeordneten Straßennetz gibt, sind sehr selten.

Zusammenfassung

Die Bewertung des Verfahrens zeigt, dass es für den *Stadtverkehr* nicht geeignet ist. Neben den genannten Schwierigkeiten in allen Teilbereichen ist auch die Identifikation von Fahrgästen im öffentlichen Nahverkehr nur schwer mit diesem Verfahren möglich.

Für *außerörtliche Bundes- und Landstraßen* kann keine pauschale Aussage getroffen werden - die Einsetzbarkeit des Verfahrens hängt ganz wesentlich von der Größe und Lage der Location Areas ab. Hier kann nur im Einzelfall je nach Ziel der jeweiligen Anwendung eine Aussage getroffen werden, ggf. muss für das Verfahren, wie am Beispiel der Netzbeeinflussungsanlage Stuttgart gezeigt, ein verfeinertes Verfahren angewendet werden.

Für die *BAB* kann das Verfahren dagegen für die Ziele in den Arbeitspaketen 5 bis 7 angewendet werden. Die beobachtete Stichprobe kann innerhalb der Fahrzeugklassen Pkw und Lkw als zufällige Stichprobe des gesamten Verkehrs betrachtet werden. Dies ermöglicht eine Analyse des Routenwahlverhaltens und eine Hochrechnung zu stundenfeinen Quelle-Ziel-Matrizen.

Die Trajektorien aus dem LAC-basierten Trajektoriengenerierungsverfahren bieten somit eine sehr gute Grundlage für verkehrstechnische und verkehrsplanerische Anwendungen auf Autobahnen. Dabei ist anwendungsbezogen zu entscheiden, ob bestimmte Trajektorien gefiltert werden sollten:

- Für die Analyse der Routenwahl in der BAB-Netzmasche A5/A6/A8/A81 bieten die Trajektorien eine hervorragende Grundlage, da die Routen mit ca. 100 km sehr lang sind. Wie auch bei Untersuchungen mit Kennzeichenerfassungssystem sind hier die Trajektorien zu filtern, die aufgrund ihrer hohen Fahrzeit nicht als Durchgangsverkehr zu bezeichnen sind, weil sie z.B. eine Warenlieferung in Autobahnnähe gemacht haben.
- Für die Generierung von Quelle-Ziel-Matrizen muss beurteilt werden, wie es sich auswirkt, dass nur Fahrten ab einer gewissen Länge (≥ 3 unterschiedliche LACs) erkannt werden. Darüberhinaus muss analysiert werden, inwieweit Fahrten mit nur 3 unterschiedlichen LACs mit geringer Ortungsgenauigkeit berücksichtigt werden sollen.
- Für das Monitoring der Angebotsqualität muss bei der Bestimmung der Geschwindigkeiten aus den Trajektorien analysiert werden, ob sehr kurze

Trajektorien sinnvolle Geschwindigkeiten liefern, da dort insbesondere der Einfluss der Anbindungswahl sehr hoch ist.

4 Schätzung der Parameter des Routenwahlverhaltens (AP5)

Verkehrsteilnehmer müssen im Verlauf einer Ortsveränderung eine Vielzahl von Routenwahlentscheidungen treffen, da theoretisch an jedem Knotenpunkt eine oder mehrere Alternativrouten beginnen können. Viele diese Alternativrouten werden aufgrund der deutlich längeren Fahrzeit oder Fahrtweite vom Verkehrsteilnehmer nicht als solche wahrgenommen.

Mit den Detektoren, deren Anwendung heutzutage Standard ist, ist es nicht möglich, für große Stichproben das Routenwahlverhalten zu beobachten. Stationäre Detektoren ermöglichen in der Regel keine Fahrzeugwiedererkennung und bei im Verkehr mitfließenden Detektoren ist in der Regel die Stichprobenmenge sehr gering. Zur Beobachtung der Routenwahl muss daher in der Regel auf Untersuchungen mit kleinen Stichproben (z.B. Wermuth, Sommer & Wulff, 2004) zurückgegriffen werden. Alternativ können auch Interviews oder Laborversuche (z.B. Schreckenberg et al., 2001) durchgeführt werden. Die Erhebungsmethoden sind allerdings zeit- und kostenintensiv und lassen, insbesondere bei Interviews und Laborversuchen, oft einen realistischen Bezug vermissen. Vorteil dieser Erhebungsmethoden ist allerdings, dass zusätzliche Informationen über den Fahrer sowie die gesamte Wegekette, den Fahrtzweck, etc. erfasst werden können.

Zur Beobachtung der Routenwahl sind Detektoren notwendig, die entweder die komplette Route erfassen oder mindestens die Fahrzeugwiedererkennung an mehreren Stellen im Verkehrsnetz ermöglichen. Dazu eignen sich zum einen die Trajektorien aus Mobilfunkdaten, deren Erzeugung in Kapitel 3 beschrieben ist, und zum anderen Messungen mit Kennzeichenerfassungsgeräten.

In diesem Kapitel werden für die BAB-Netzmasche in Baden-Württemberg (A5, A6, A8 und A81) sowie die Netzbeeinflussungsanlage Stuttgart durch die Auswertung von beobachteten Routenwahlentscheidungen mögliche Einflussgrößen auf die Routenwahl systematisch identifiziert und deren Gewichtung in der Nutzenfunktion von Entscheidungsmodellen ermittelt. Bei Ben-Akiva et al. (1984) finden sich eine Vielzahl von möglichen Einflussgrößen auf die Routenwahl, darunter die Fahrtzeit, Fahrtweite, aber auch die Straßenqualität und Landschaft (vgl. Kapitel 2.4.1). In dieser Untersuchung, bei der nur ausgewählte Routenwahlentscheidungen betrachtet werden, können nur über die Untersuchungszeit dynamische Größen wie z.B. Verkehrsmeldungen als Einflussgrößen betrachtet werden. Statische Größen wie die Fahrtweite gehen in dem konstanten Nutzenanteil ein. Diese könnten nur dann statistisch gesichert ermittelt werden, wenn eine Vielzahl Routenwahlentscheidungen mit unterschiedlichen Fahrtweiten analysiert werden würden und dabei auf den konstanten Nutzenanteil verzichtet werden könnte.

Abbildung 56 zeigt für die Richtung Stuttgart → Walldorf die Routenwahlanteile eines ausgewählten Tages in Stundenintervallen. Im Mittel fahren in dem dargestellten Zeit-

raum ca. 67 % über Karlsruhe, der Anteil schwankt dabei zwischen 40 und 80 %. Die Untersuchungen werden zeigen, ob und wie sich die Schwankungen der Anteile auf den verschiedenen Alternativen mit Verkehrsmeldungen etc. erklären lassen.

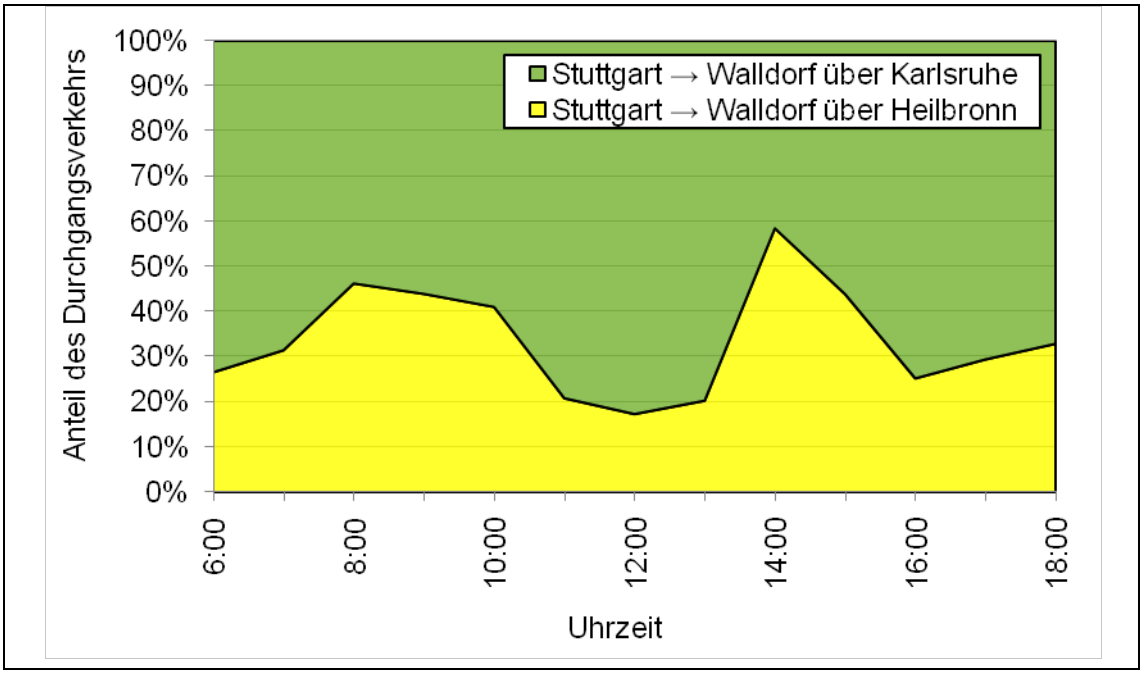


Abbildung 56: Tageszeitabhängige Routenwahl (Stundenintervalle, 10.07.2008).

In Kapitel 4.1 werden einige für die Routenwahlanalysen grundlegende Theorien und Daten erörtert. Die Kapitel 4.2 und 4.3 behandeln anschließend die Routenwahlanalysen in der BAB-Netzmasche, einmal aus Untersuchungen mit Kennzeichenerfassungssystemen und einmal mit FPD-Trajektorien aus Mobilfunkdaten. In Kapitel 4.4 werden dann die Analysen zur Netzbeeinflussungsanlage (NBA) Stuttgart vorgestellt. Abschließend werden in Kapitel 4.5 die Ergebnisse zusammengefasst und bewertet.

4.1 Grundlagen

4.1.1 Nutzenfunktion

Die meisten Routenwahlmodelle basieren auf multi-kriteriellen Entscheidungsmodellen (z.B. Logit-Modell), d.h. die verschiedenen Einflussfaktoren werden über eine Nutzenfunktion zu einem Wert zusammengefasst. Der berechnete Nutzen ist im Verkehrswesen häufig negativ und wird daher auch als Widerstand oder generalisierte Kosten bezeichnet. Eine lineare Nutzenfunktion hat die folgende Form:

$$V_j = \alpha_j + \sum_{n=1}^{N} \beta_n \cdot E_{n,j}$$

mit

V_j	Nutzen der Alternative *j*
$n...N$	Anzahl der betrachteten Einflussfaktoren
$E_{n,j}$	n-ter Einflussfaktor der Alternative *j*
α_j , β_n	Parameter der Nutzenfunktion

Mögliche Einflussfaktoren bei der Routenwahl sind die Länge der Route, die Fahrtzeit im unbelasteten Netz, typische Fahrzeiten zum Zeitpunkt der Routenwahlentscheidung, Verkehrsmeldungen und Routenwahlempfehlungen im Radio und durch Wechselwegweisungen. Die meisten Einflussfaktoren können für einzelne Strecken ermittelt und anschließend über alle Strecken der jeweiligen Alternativroute aufsummiert werden. Dagegen werden z.B. Routenwahlempfehlungen einer gesamten Route zugewiesen, da sie keinen Nutzen für einzelne Streckenabschnitte darstellen.

Die Werte der Nutzenfunktion verschiedener Alternativen bilden den Eingangswert für Entscheidungsmodelle. Eines dieser Modelle, das Logit-Modell, wird im folgenden Kapitel erläutert.

4.1.2 Entscheidungsmodell (Logit-Modell)

Entscheidungsmodelle ermitteln Auswahlwahrscheinlichkeiten für die Alternativen *j*. Bei Verkehrsteilnehmern werden unter anderem die Verkehrsziel-, Verkehrsmodus- und Routenwahl mit Entscheidungsmodellen modelliert. Ein im Verkehrswesen häufig verwendetes diskretes Entscheidungsmodell ist das Logit-Modell, bei dem die Auswahlwahrscheinlichkeiten wie folgt ermittelt werden:

$$P_j = \frac{e^{V_j}}{\sum_{j=1}^{J} e^{V_j}}$$

mit

Pj	Auswahlwahrscheinlichkeit für Alternative *j*
V_j	Nutzen der Alternative *j*
j...J	Alternativen der Entscheidung

Maßgebend für die Aufteilung nach der obigen Formel ist die Differenz der Nutzen V_j der betrachteten Alternativen.

Für eine ausführliche Diskussion des Logit-Modells sowie eine mathematische Herleitung wird an dieser Stelle auf McFadden (2003) und Maier & Weiss (1990) verwiesen.

4.1.3 Maximum-Likelihood-Schätzung

4.1.3.1 Durchführung der Schätzung

Die Maximum-Likelihood-Schätzung ist eine Methode zur Schätzung von Parametern in Nutzenfunktionen auf Basis von empirischen Daten (Cowan, 2003, Louviere et al., 2000). Dafür wird für jede Alternative der Nutzen mit der Nutzenfunktion berechnet und mit einer geeigneten Entscheidungsmodell (z.B. Logit-Modell) die Auswahlwahrscheinlichkeit für jede Alternative berechnet. Durch Aufsummieren der logarithmisierten Auswahlwahrscheinlichkeiten der gewählten Alternativen wird die Log-Likelihood-Funktion $L\left(\alpha_j,\beta_n\right)$ berechnet, die bei der Maximum-Likelihood-Schätzung maximiert wird. Variabel sind dabei die Parameter α_j und β_n der Nutzenfunktion:

$$L\left(\alpha_j,\beta_n\right)=\sum_{k=1}^{K}\sum_{j=1}^{J} g_{k,j}\cdot ln\,P\left(V\left(\alpha_n,\beta_j\right)\right)_{k,j}$$

mit

$L\left(\alpha_j,\beta_n\right)$	Log-Likelihood-Funktion für die Parameter α_j und β_n
k...K	Anzahl der empirisch erhobenen Datensätze
$j...J$	Anzahl der Alternativen
$P(V)_{n,j}$	Wahrscheinlichkeit, dass im Datensatz *k* die Alternative *j* gewählt wird
$V\left(\alpha_j,\beta_n\right)$	Nutzenfunktion für verschiedene Alternativen
$g_{k,j}$	Wahlvariable = 1, wenn Alternative *j* im Datensatz *k* gewählt wurde = 0, wenn Alternative *j* im Datensatz *k* nicht gewählt wurde

Ausführlich beschrieben ist die Maximum-Likelihood-Schätzung bei Ben-Akiva & Lerman (1985). Darüberhinaus wird die Berechnung der Log-Likelihood-Funktion anhand eines einfachen Beispiels in Kapitel 4.2.4 verdeutlicht.

In dieser Arbeit werden die Schätzungen mit der Software biogeme 1.8 (Bierlaire, 2003 & 2009) durchgeführt.

4.1.3.2 Bestimmung der Modellgüte

Zur Bestimmung der Modellgüte gibt es eine Vielzahl von Gütemaßen, die entweder die *Güte der Anpassung* direkt aus der Log-Likelihood-Funktion ermitteln oder die *Güte des Gesamtmodells* mit Hilfe eines Pseudo-R² beschreiben. Im Folgenden wird jeweils ein geeigneter Vertreter für beide Arten von Gütemaßen beschrieben. Eine umfassende Übersicht über weitere Gütemaße findet sich bei Backhaus et al. (2006).

Die *Güte der Anpassung* kann mit Hilfe des Likelihood Ratio Tests (LRT, vgl. u.a. Huelsenbeck & Crandall, 1997) bestimmt werden. Dieser Test untersucht, ob eine Schätzung mit zusätzlichen Parametern ein signifikant besseres Ergebnis erzeugt. Das Likelihood Ratio LR berechnet sich zu:

$$LR = 2 \cdot (L_2 - L_1)$$

mit

LR Likelihood Ratio

L_2 Log-Likelihood der Variante mit mehr Parametern

L_1 Log-Likelihood der Variante mit weniger Parametern oder ohne Parameter („Null-Log-Likelihood“)

Das LR folgt nach den Angaben in der Literatur in etwa der Chi-Quadrat-Verteilung. Aus Tabellen können mit der Anzahl der Freiheitsgrade (Differenz der Anzahl der Parameter) und dem Signifikanzniveau die Werte der Chi-Quadrat-Verteilung ermittelt und dem berechneten LR gegenübergestellt werden. In der Literatur wird ein Signifikanzniveau von 95,0 % meistens als „signifikant“ und 99,0 % als „hochsignifikant“ bewertet. Ein ebenfalls gelegentlich angegebenes Signifikanzniveau von nur 90 % bedeutet, dass mit einer Irrtumswahrscheinlichkeit *α* von bis zu 10 % die Nullhypothese H_0 (hier: zusätzliche Parameter sind geeignet, um die Güte der Anpassung zu steigern) angenommen wird, obwohl sie nicht korrekt ist.

Freiheitsgrade	Signifikanzniveau		
	90 %	95 %	99 %
1	2,71	3,84	6,63
2	4,61	5,99	9,21
3	6,25	7,81	11,34
4	7,78	9,49	13,28
5	9,24	11,07	15,09
6	10,64	12,59	16,81
7	12,02	14,07	18,48
8	13,36	15,51	20,09
9	14,68	16,92	21,67
10	15,99	18,31	23,21

Tabelle 16: Chi-Quadrat-Tabelle (aus Backhaus et al., 2006, S. 818).

Der Tabelle 16 kann entnommen werden, dass beim Hinzufügen eines Parameters β_i das Likelihood Ratio LR einen Wert von mindestens 3,84 annehmen muss, um als „signifikant“ bewertet zu werden, und 6,63, um als „hochsignifikant“ zu gelten. Da das LR die doppelte Differenz der beiden Log-Likelihoods ist, muss sich also das Log-Likelihood in diesen Fällen um ca. 1,9 bzw. 3,3 Log-Likelihood-Punkte erhöhen.

Die *Güte des Gesamtmodells* kann mit Hilfe von ρ^2 (*rho-square*) ermittelt werden. Dabei wird das Verhältnis zwischen dem Log-Likelihood des geschätzten Modells und dem Null-Log-Likelihood ermittelt und von 1 abgezogen. Eine Variante vom ρ^2 ist das *adjusted-* ρ^2, das zusätzlich die Anzahl der Parameter berücksichtigt.

$$\rho^2 = 1 - \frac{L_1}{L_0}$$

$$\textit{adjusted-}\ \rho^2 = 1 - \frac{L_1 - K}{L_0}$$

mit

ρ^2	rho-square (Wertebereich zwischen 0 und 1)
K	Anzahl Parameter
L_1	Log-Likelihood des geschätzten Modells
L_0	Null-Log-Likelihood

Der Unterschied zwischen diesen beiden Werten nimmt mit zunehmendem Stichprobenumfang ab und ist für die Schätzungen in dieser Arbeit vernachlässigbar.

ρ^2 kann Werte zwischen 0 und 1 annehmen, wobei die erreichten Werte in der Regel deutlich geringer als das Bestimmtheitsmaß R² der linearen Regression sind. In der Literatur wird allgemeinhin für ρ^2 ein Bereich von 0,2 bis 0,4 angegeben, bei dem man von einer hohen Güte ausgehen kann (vgl. u.a. Backhaus et. al, 2006). Dieser Bereich entspricht nach Domencich & McFadden (1975) einem R² der linearen Regression von 0,44 bis 0,72.

4.1.3.3 Bestimmung der Parametersignifikanz

Für die geschätzten Parameter wird bei der Maximum-Likelihood-Schätzung ein Signifikanztest durchgeführt. Dadurch kann für ein gewähltes Signifikanzniveau sichergestellt werden, dass der Parameter ungleich 0 ist und damit in das Modell über-

nommen werden sollte. Die Prüfgröße T_i ermittelt sich damit aus dem bei der Schätzung ermittelten Wert β_i eines Parameters und seiner Standardabweichung s_{b_i}:

$$T = \frac{\beta_i}{s_{b_i}}$$

mit

T_i Prüfgröße für den Parameter i

β_i Schätzwert des Parameters i

s_{b_i} Standardfehler des Parameters β_i

Unter Voraussetzung einer ausreichend großen Stichprobe ergeben sich die folgenden Werte aus der Student-t-Verteilung, die die Prüfgröße T_i im Betrag nicht unterschreiten sollte, um als signifikant (95 %) bzw. hochsignifikant (99 %) zu gelten.

Freiheitsgrade	Signifikanzniveau		
	90 %	95 %	99 %
>1.000	1,67	1,96	2,58

Tabelle 17: t-Tabelle für den zweiseitigen Test (aus Backhaus et al., 2006, S. 808).

Dieser Test mit der Student-t-Verteilung liefert bei der Maximum-Likelihood-Schätzung identische Ergebnisse wie der in der Literatur gelegentlich verwendete Wald-Test. Die Prüfgröße des Wald-Tests ist T_i^2 und wird mit Werten der Chi-Quadrat-Tabelle für einen Freiheitsgrad verglichen.

4.1.4 Verkehrsmeldungen

Für die Erhebungen mit Kennzeichenerfassungssystemen werden vom Südwestrundfunk (SWR) zur Verfügung gestellte Verkehrsmeldungen verwendet. Für die Auswertung der Trajektorien aus Mobilfunkdaten kann auf die Daten der Landesmeldestelle Baden-Württemberg, die vom Landeskriminalamt verwaltet werden, zurückgegriffen werden (vgl. Tabelle 18).

Spalte	Name	Typ	Beschreibung
1	Meldungsnummer	Long Integer	Eindeutige Meldungsnummer der Landesmeldestelle
2	Version	Integer	Index für verschiedene Versionen einer Meldung (mit identischer Meldungsnummer), beginnt bei 0
3	Eingabestelle	Integer	z.B. „LPD S“ für die Landespolizeidirektion Stuttgart. Der Wert „SYSTEM“ beschreibt automatische Systemabmeldungen, die 15min nach der Aufhebungsmeldung erfolgen.
4	Datum	Datum	TT.MM.JJJJ
5	Zeit	Uhrzeit	HH:MM
6	Strasse	Text	Angabe der Straße (nur bei Autobahn und Bundesstraße)
7	Richtung	Text	Angabe der Richtung im Klartext, z.B: „Basel – Heidelberg“
8	Meldungstext	Text	Meldungstext, voll- und teilstandardisiert
9	Stau	Integer	Angabe der Staulänge in km (optional)

Tabelle 18: Definition der Verkehrsmeldungen der Landesmeldestelle.

Bei den sehr seltenen Vollsperrungen (nur 0,04 % aller Meldungen) werden entweder die gemeldete Staulänge verwendet oder 10 km eingesetzt. Staumeldungen ohne Staulängenangabe werden mit 1 km weiterverarbeitet.

Die Daten werden so aufbereitet, dass für die späteren Analysen alle Informationen streckenbezogen zur Verfügung stehen. Dabei werden die folgenden Regeln verwendet:

- Bei Meldungen mit nur einer Ortsangabe (z.B. „A8 Höhe Pforzheim-Ost“) wird die stromabwärts liegende Strecke ausgewählt.
- Bei Staulängenangaben, die größer als die Strecken zwischen den angegebenen Knoten sind, wird die Staulänge längengewichtet auf die betroffenen Strecken verteilt.
- Zusätzliche Angaben zur Stauursache etc. werden ausgelesen und gespeichert.

Um für einige Auswertungen die Meldungen über Stau und stockenden Verkehr zusammenfassen und für die Verkehrslage einen einheitlichen Wert verwenden zu können, werden diese wie folgt zusammengefasst:

$$VL_i = l_{Stau,i} + f \cdot l_{Stockend,i}$$

mit

VL_i Verkehrslage eines Streckenabschnittes bzw. einer Route [km]

$l_{Stau,i}$ Gemeldete Länge gestauten Verkehrs auf einer Strecke bzw. Route i [km]

$l_{Stockend,i}$ Gemeldete Länge stockenden Verkehrs auf einer Strecke bzw. Route i [km]

f Faktor zu Abminderung von $l_{Stockend,i}$

Dabei sollte der Faktor f zwischen 0 und 1 liegen. Der Wert 0 bedeutet dabei, dass eine Meldung über stockenden Verkehr nicht in die Verkehrslage eingeht, also vom

Verkehrsteilnehmer als nicht relevant erachtet wird. Der Wert 1 bedeutet dagegen, dass Meldungen über Stau und stockenden Verkehr gleichwertig eingeschätzt werden.

Für den Verkehrsteilnehmer gibt es eine Vielzahl von Möglichkeiten, sich über die aktuelle Verkehrslage zu informieren. Zum einem gibt es die Möglichkeit der Pre-Trip-Informationen z.B. auf Webseiten von Radiosendern. Zum anderen kann der Verkehrsteilnehmer während der Fahrt (on-trip) die Verkehrsnachrichten verschiedener Radiosender hören oder über TMC ausgestrahlte Meldungen im Radio oder im Navigationsgerät sehen. Die Meldungen der meisten Quellen basieren auf den Daten der Landesmeldestelle, die auf ftp-Servern als xml-Dateien abgerufen werden können, werden aber insbesondere von den Radiostationen durch eigene Verkehrsredaktionen weiterverarbeitet (Voeske, 2008).

Datenquelle	**14:30**	**14:45**	**15:00**	**15:15**	**15:30**	**15:45**	**16:00**	**16:15**
Landesmeldestelle	-	5	5	5	5	5	5	-
SWR1 Radio	-	-	5	5	5	5	-	-
SWR1 Internet	-	-	-	5	5	5	5	-
Hitradio Radio	-	-	5	5	5	5	5	-
Hitradio Internet	-	5	5	5	5	5	5	-

Tabelle 19: Vergleich einer Störungsmeldung [Staulänge in km] auf der A6 im zeitlichen Verlauf.

Eine manuelle Analyse über einen Tag zeigt allerdings, dass die Staumeldungen bei ausgewählten Quellen sehr ähnlich sind. Tabelle 19 zeigt, dass die Meldung „A6 Heilbronn Richtung Nürnberg zwischen AS Heilbronn/Neckarsulm und Kreuz Weinsberg 5 km Stau“ im gleichen Zeitintervall bei der Landesmeldestelle und auf der Webseite von Hitradio Antenne 1 Baden-Württemberg erscheinen. Die Meldungen im Radio sind etwas verzögert, auch da diese nur halbstündig ausgestrahlt werden. Die Anzeige auf der Webseite von SWR1 erscheint erst um 15:15 Uhr. Das Stauende ist bei allen Medien nahezu identisch.

Bei weiteren Störungsmeldungen zeigen sich andere Muster, auch die angegebenen Staulängen variieren teilweise um geringe Werte. Es lässt sich allerdings feststellen, dass die Meldungen sehr ähnlich verlaufen und somit auf eine differenzierte Analyse der 12 landesweiten Radiosender sowie der über 30 lokalen Radiosender verzichtet werden kann. Dies gilt vor allem vor dem Hintergrund, dass es für Verkehrsmeldungen keine einheitliche Schnittstelle gibt und somit die Datenlieferungen – auch für die Radiosender – erheblichen Aufwand darstellen würden. So ist beispielsweise beim SWR die einzig vorhandene Export-Schnittstelle ein Papierausdruck der Dateien.

4.1.5 Wechselwegweisungen im Untersuchungsbiet

4.1.5.1 Wechselwegweisung BAB-Netzmasche (A5/A6/A8/A81)

In der BAB-Netzmasche Baden-Württemberg, die aus den Autobahnen A5, A6, A8 und A81 zwischen Walldorf, Heilbronn, Karlsruhe und Stuttgart besteht, gibt es in beiden Diagonalen zwei Alternativen mit vergleichbaren Fahrtweiten:

- Diagonale Stuttgart ↔ Walldorf:
  - Stuttgart (AD Leonberg) – Karlsruhe – AK Walldorf (S-KA-W): 91 km
  - Stuttgart (AD Leonberg) – Heilbronn – AK Walldorf (S-HN-W): 103 km
- Diagonale Heilbronn ↔ Karlsruhe:
  - Karlsruhe – Walldorf – Heilbronn (AK Weinsberg) (KA-W-HN): 92 km
  - Stuttgart (AD Leonberg) – Heilbronn – AK Walldorf (S-HN-W): 102 km

Für die Diagonale Stuttgart ↔ Walldorf gibt es eine additive Wechselwegweisung, deren Normalroute über Heilbronn verläuft, obwohl diese etwas länger ist. Das dWiSta-Anzeigesytem (vgl. Abbildung 57) czeigt bei Schaltung durch den Disponenten in der Verkehrsrechnerzentrale Ludwigsburg bei Störungen auf der Hauptroute eine Empfehlung für die Alternativroute über Karlsruhe.

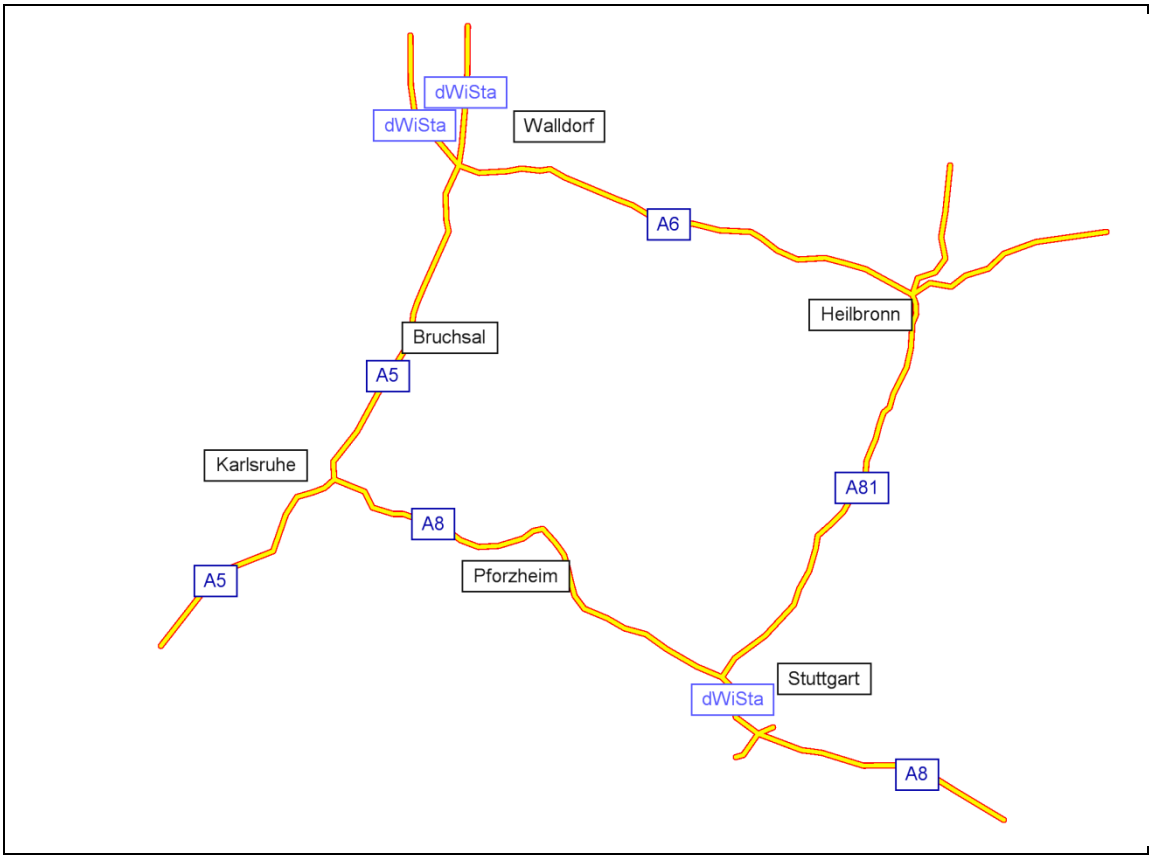


Abbildung 57: Darstellung der Wechselwegweisung in der BAB-Netzmasche Leonberg / Heilbronn / Walldorf / Karlsruhe).

Neben der Empfehlung wird auch eine Begründung für die Empfehlung angezeigt. Diese enthält neben dem Ort der Störung (Autobahnnummer und Anschlussstelle) auch die Staulänge. Neben der Angabe der Kilometer sind auch die Anzeige „Stau“ (ohne Längenangabe) und „Vollsperrung“ vorgesehen. Da die Vollsperrung nur sehr selten angezeigt wird, ist eine gesonderte Auswertung nicht möglich. In diesen Fällen wird eine Staulänge von 10 km für die Berechnungen angenommen.

Zeile	Format	Beispiel
1	TT.MM.JJJ HH:MM:SS:	21.06.2007 17:03:26:
2	Richtung | via AK | Fz.-art (optional)	Stuttgart | via AD Karlsruhe |
3	XX km Stau | auf A X hinter | (YY) Name Anschlussstelle	20 km Stau | auf A 6 hinter | (32) Wiesloch/Rauenberg
4	Ausgeschaltet: TT.MM.JJJ HH:MM:SS	Ausgeschaltet: 21.06.2007 20:44:53

Tabelle 20: Definition und Beispiel (A5) des Formats der Meldungen der Wechselwegweisung („dWiSta-Anzeigen“).

Tabelle 20 zeigt das Format der Meldungen sowie ein Beispiel mit Illustrierung der Anzeigentafel. Die dort vorgesehene Option, Empfehlungen nur für bestimmte Fahrzeugarten herauszugeben, wurde im Untersuchungszeitraum nicht genutzt. Neben den genannten Anzeigen ist noch vorgesehen, dass statt Stau eine Baustelleninformation angezeigt wird. Dies ist aber derzeit nicht in den Schaltstrategien vorgesehen.

Abbildung 58 zeigt, wie lange die Wechselwegweisung im Jahr 2008 geschaltet war. Die Wechselwegweisung von Walldorf Richtung Stuttgart ist im Jahr 2008 in 7,4 % der Zeit geschaltet, in der Gegenrichtung sind es nur 1,3 %.

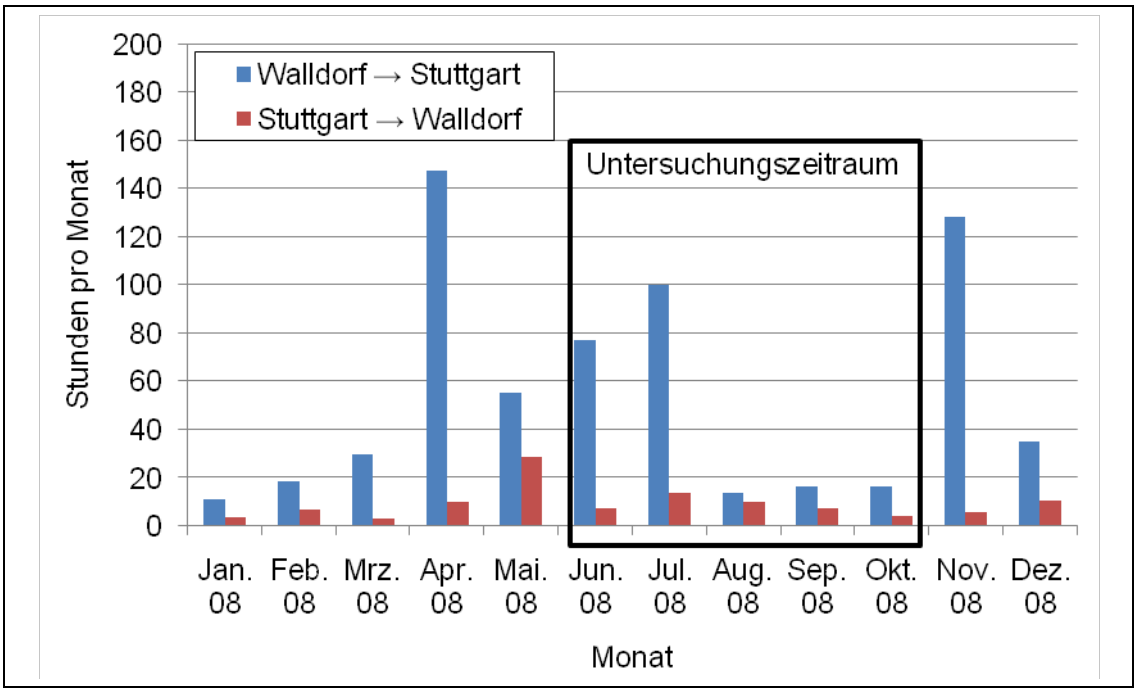


Abbildung 58: Monatliche Summe der aktiven Wechselwegweisung der BAB-Netzmasche.

4.1.5.2 Netzbeeinflussungsanlage Stuttgart (A81/B10/B27/B295)

Zur Fußballweltmeisterschaft 2006 wurde die Netzbeeinflussungsanlage Stuttgart in Betrieb genommen. Diese steuert den von der A81 aus Richtung Heilbronn kommenden Verkehr entweder über die Hauptroute B10 oder über die Alternativroute B295 Richtung Stuttgart (vgl. Abbildung 59). Neben der B10 und B295 gibt es mit der B27 noch eine dritte Möglichkeit, um von der A81 von Norden kommend nach Stuttgart zu fahren. Diese ist allerdings nicht in der Netzbeeinflussungsanlage integriert.

Die Nutzungsstatistik in Abbildung 60 zeigt, dass diese Netzbeeinflussungsanlage deutlich seltener als die der BAB-Netzmasche geschaltet wird. In der Stadteinwärtsrichtung, in der die Anlage zwischen Mitte Februar und Mitte August 2008 aufgrund technischer Schwierigkeiten nicht geschaltet werden konnte, ist die Netzbeeinflussungsanlage nur in 0,4 % der Zeit geschaltet. In der Gegenrichtung sind es 1,3 %.

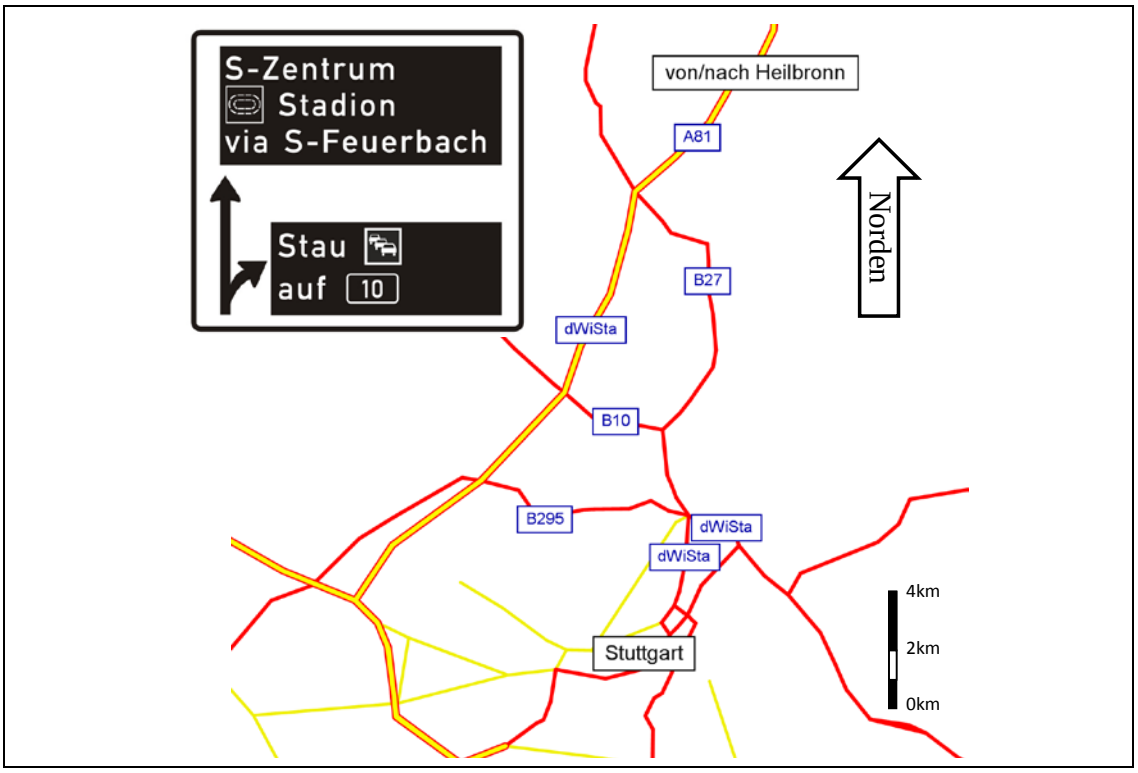


Abbildung 59: Netzbeeinflussungsanlage Stuttgart.

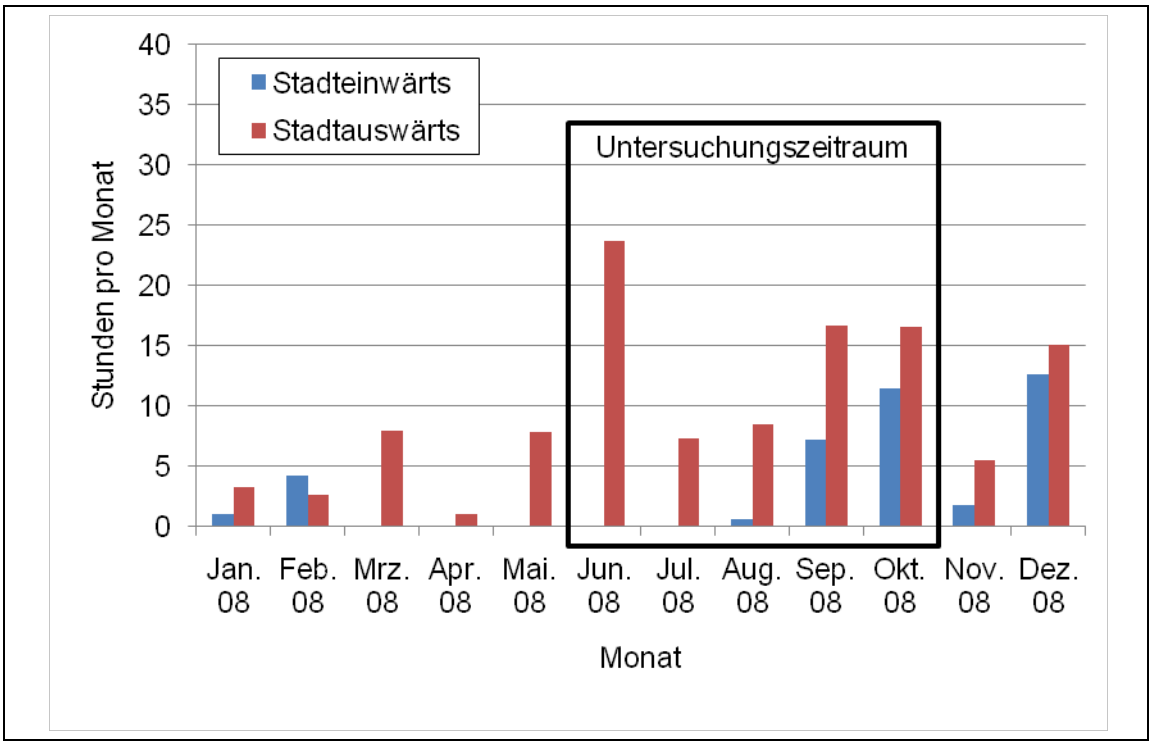


Abbildung 60: Monatliche Summe der geschalteten NBA Stuttgart.

4.2 Routenwahlanalysen in der BAB-Netzmasche mit ANPR-Messungen

4.2.1 Datenaufbereitung für Routenwahlanalysen

Untersuchungsmethodik

Die Erhebung wurde mit automatischen Kennzeichenerfassungssystemen (Automatic Number Plate Recognition, ANPR) durchgeführt. Eine Einheit dieses Systems besteht aus einer Infrarot-Kamera, die die Breite eines Fahrstreifens erfasst und die Videodaten an einen Rechner liefert. Auf dem Rechner erfolgt die Echtzeit-Erfassung der Kennzeichen mit einem Bildverarbeitungssystem, die anschließend mit Zeitstempel und weiteren Informationen in eine Datenbank geschrieben werden (vgl. Friedrich et al., 2009). Ergänzend wurden handelsübliche Camcorder verwendet, bei denen die Auswertung durch nachträgliches Auswerten der aufgezeichneten Videobilder erfolgt.

Es wurde an drei Messstellen jeweils die komplette Fahrbahnbreite des Verkehrs erfasst. Die Lage der Messstellen kann der Abbildung 61 entnommen werden.

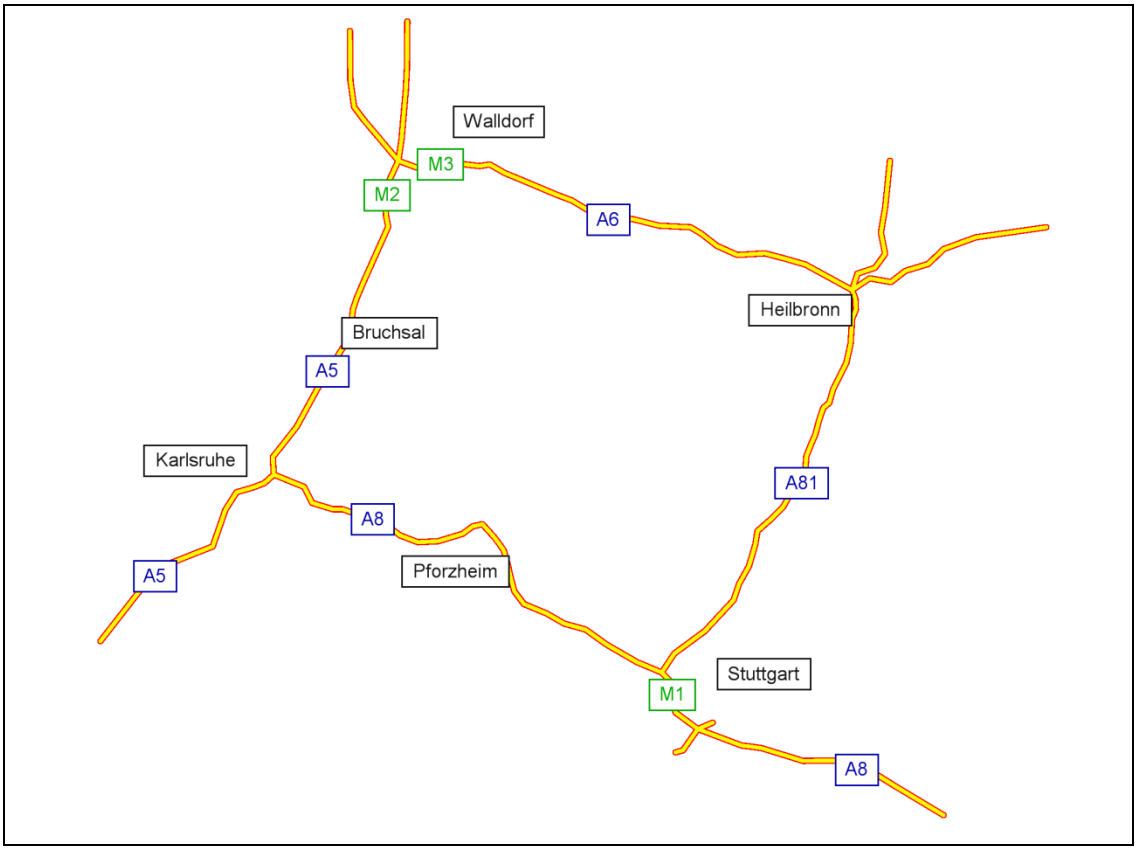


Abbildung 61: Lage der Messstellen M1 bis M3 der BAB-Netzmasche Leonberg / Heilbronn / Walldorf / Karlsruhe.

Dabei sind die Messstellen so angeordnet, dass der gesamte Durchgangsverkehr zwischen den Autobahnknoten Leonberg und Walldorf erfasst wird. Dies wird dadurch

sichergestellt, dass zwischen den Messstellen und den Autobahnknoten Walldorf bzw. Stuttgart keine Anschlussstellen liegen.

Für die Auswertung werden die Messungen in den folgenden Zeiträumen berücksichtigt:

- Fahrtrichtung Norden (München – Frankfurt):
  - Fr, 28.10.2005: A8: 9:30-14:00 Uhr / A5 & A6: 10:00-16:15 Uhr
  - Mo, 08.05.2006: A8: 7:30-17:00 Uhr / A5 & A6: 08:00-19:00 Uhr
- Fahrtrichtung Süden (Frankfurt – München):
  - Fr, 12.05.2006: A5 & A6: 08:00-17:00 / A8: 8:30-19:30 Uhr
  - Mo, 15.05.2006: A5 & A6: 08:00-11:15 / A8: 8:30-13:15 Uhr

Durchgangsverkehr

Als *Durchgangsverkehr* wird der Verkehr bezeichnet, der vom Kreuz Leonberg bis zum Kreuz Walldorf (oder umgekehrt) fährt, ohne dazwischen ein routenwahlbeeinflussendes Ziel zu haben. Verkehr, der sowohl am Dreieck Leonberg als auch am Kreuz Walldorf erfasst wird, aber vermutlich zwischenzeitlich ein Ziel (z.B. Beladen, Pause) angesteuert hat, wird als *gebrochener Durchgangsverkehr* bezeichnet.

Da es allein auf Basis der Kennzeichen keine Möglichkeit gibt, den echten Durchgangsverkehr vom gebrochenen Durchgangsverkehr zu unterscheiden, wird als Hilfsgröße die Durchfahrtszeit verwendet. Ist die Durchfahrtszeit eines Kfz deutlich höher als die zur gleichen Zeit fahrender Kfz, wird dieses Fahrzeug als gebrochener Durchfahrtsverkehr gekennzeichnet. Im Rahmen dieser Untersuchung wird der Durchgangsverkehr dabei wie folgt ermittelt.

1. Zuerst werden durch einen Kennzeichenvergleich alle Fahrzeuge mit der zugehörigen Durchfahrtszeit ermittelt, die sowohl an der A8 als auch an der A5 oder der A6 erfasst worden sind.
2. Dann wird das über 40 Minuten gleitende Mittel der Fahrzeit über die Fahrzeuge gebildet. Unberücksichtigt bleiben dabei die jeweils 5% schnellsten sowie 10% langsamsten Fahrzeuge.
3. Abschließend werden alle Fahrzeuge gelöscht, deren Durchfahrtszeit um mehr als 40 % größer als das über 40 Minuten gleitende Mittel ist.

Abbildung 62 veranschaulicht das Verfahren am Beispiel des Durchgangsverkehrs über Karlsruhe am 08.05.2006. Durch die blauen Punkte ist der Durchgangsverkehr dargestellt, die grünen Dreiecke repräsentieren den gebrochenen Durchgangsverkehr. Die gestrichelte Linie zeigt die Trennung zwischen Durchgangsverkehr und gebrochenem Durchgangsverkehr, die sich anhand der oben beschriebenen Ermittlung ergibt.

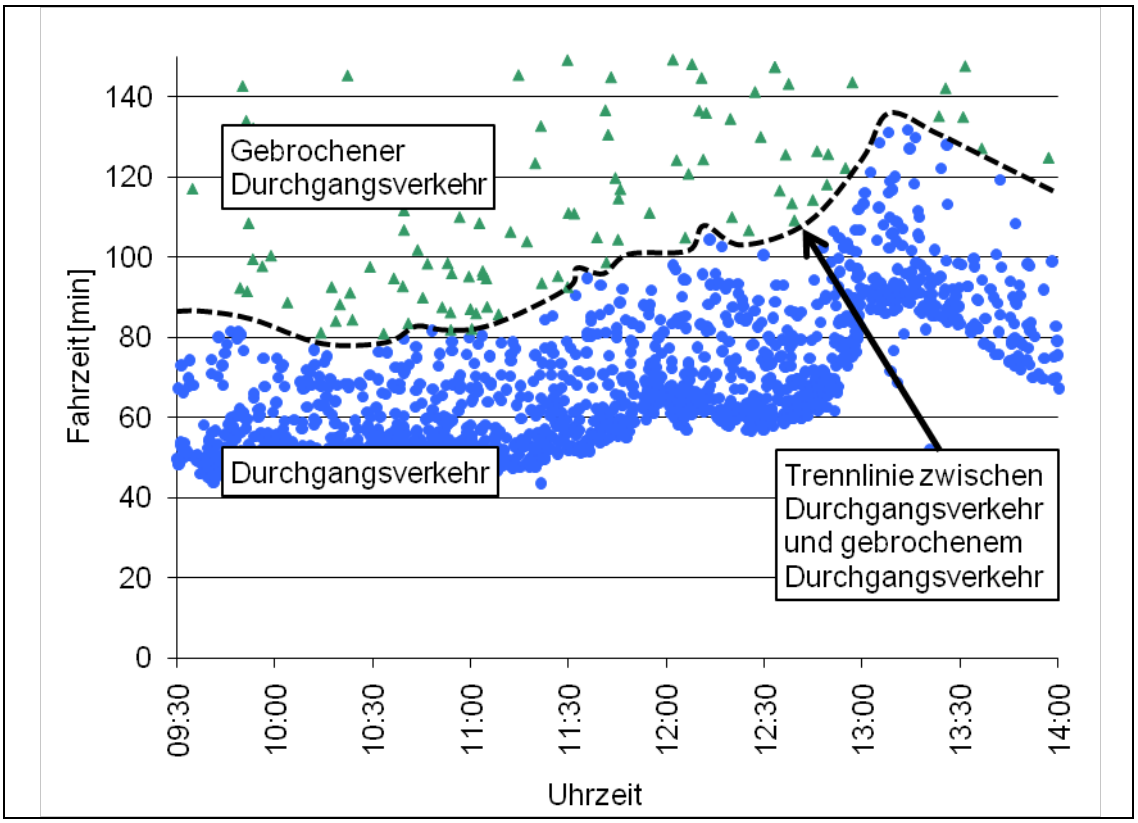


Abbildung 62: Unterscheidung von Durchgangsverkehr und gebrochenem Durchgangsverkehr am Beispiel der Route über Karlsruhe (08.05.2006).

4.2.2 Stichprobenbeschreibung

Insgesamt wurden an den vier Messtagen 9.841 Fahrzeuge als Durchgangsverkehr erkannt. Eine automatische Unterscheidung von Pkw und Lkw ist mit den eingesetzten Kennzeichenerfassungsgeräten nicht möglich. Die folgende Tabelle 21 gibt einige Kennwerte dieses Durchgangsverkehrs wieder.

Richtung	Route	Absolute Anzahl des Durchgangs-verkehrs	Anteil (je Richtung)	Relativer Anteil am Gesamtverkehr (bezogen auf die Verkehrsstärke am Messquerschnitt A8)	
Süd-Nord	S-KA-W	3.896	71,5 %	8,0 %	11,2 %
	S-HN-W	1.552	28,5 %	3,2 %	
Nord-Süd	W-KA-S	2.820	64,2 %	5,8 %	9,0 %
	W-HN-S	1.573	35,8 %	3,2 %	
Gesamt		**9.841**			**10,2 %**

Tabelle 21: Übersicht über den Durchgangsverkehr
(S = Stuttgart, KA = Karlsruhe, HN = Heilbronn, W = Walldorf).

Da die eingesetzten Kennzeichenerfassungssysteme in der Regel nur zwischen 85% und 90% der Fahrzeuge korrekt erfassen, muss der Durchgangsverkehrsanteil hochge-

rechnet werden. In einer vergleichbaren Untersuchung (Friedrich et al., 2007) wurden 86 % der Fahrzeuge korrekt erfasst. Da die nicht korrekt erkannten Fahrzeuge zufällig verteilt sind, ergibt sich ein durchschnittlicher Hochrechungsfaktor für die relativen Durchgangsverkehrsanteile von $1/0{,}86 = 1{,}16$ und für den absoluten Durchgangsverkehr von $1/0{,}86^2 = 1{,}35$.

Eine Auswertung der Zulassungsbezirke der Kennzeichen zeigt, dass 11 % des Durchgangsverkehrs aus dem Bereich der BAB-Netzmasche, 32 % aus Baden-Württemberg und 89 % aus Deutschland kommen. Die Zulassungsbezirke, die von den großen Autovermietungen verwendet werden und auffällig oft vorkommen (so ergibt sich z.B. für den Zulassungsbezirk HH, in dem viele Mietwagen zugelassen sind, der Rang 10 über alle erfassten Kennzeichen), liegen weder im Bereich der BAB-Netzmasche noch in Baden-Württemberg. Die folgende Abbildung zeigt die relative Verteilung des Durchgangsverkehrs auf die zwei Routen sowie die Staumeldungen auf der Route über Karlsruhe. Der gut erkennbare Zusammenhang zwischen den Staumeldungen und den Routenwahlanteilen wird in den nächsten Kapiteln weiter untersucht.

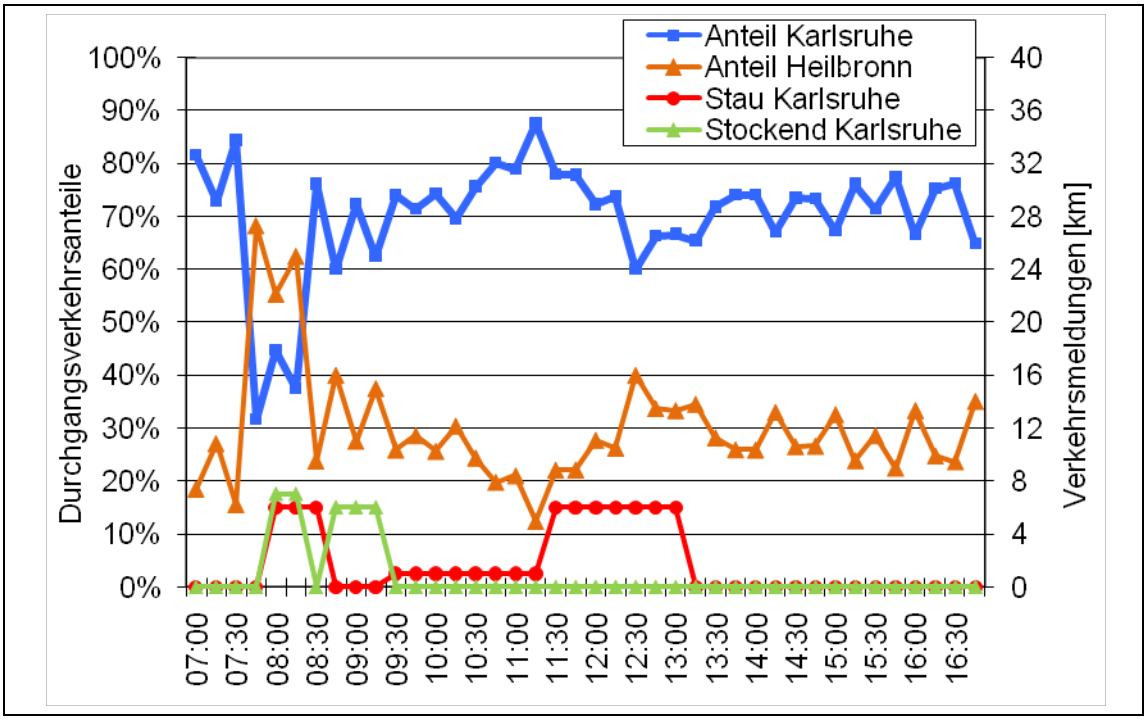


Abbildung 63: Staumeldungen S-KA-W und Routenwahlentscheidung aus der Kennzeichenerfassung (keine Verkehrsmeldungen auf der Route über Heilbronn, 08.05.2006).

4.2.3 Untersuchung einzelner Einflussgrößen mittels linearer Regression

In diesem Abschnitt wird untersucht, ob die Verkehrsmeldungen mit dem Anteil der über Karlsruhe fahrenden Fahrzeuge korrelieren. Dafür wird für jede der 9.841 beobachteten Routenwahlentscheidungen die Differenz der gemeldeten Längen des gestauten und stockenden Verkehrs auf den beiden Alternativen über Karlsruhe bzw. Heilbronn ermittelt.

Staulänge

Abbildung 64 zeigt den Anteil der Routenwahlentscheidungen über Karlsruhe in Abhängigkeit von der Differenz der gemeldeten Staulängen. Ein positiver Wert der horizontalen Achse bedeutet dabei, dass der gemeldete Stau über Karlsruhe länger als der über Heilbronn war. Zur besseren Übersichtlichkeit sind für die Grafiken die Werte der vertikalen Achse auf 5%-Werte gerundet und zusammengefasst. Die Kreisfläche entspricht dabei der Anzahl der Beobachtungen.

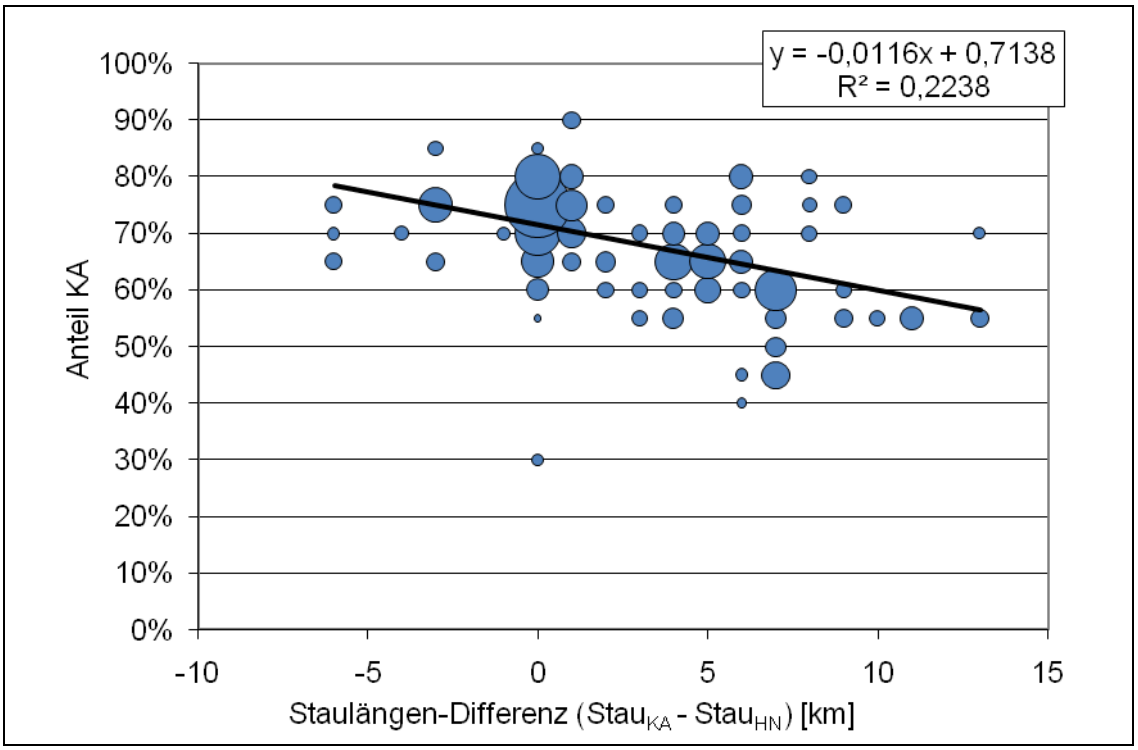


Abbildung 64: Routenwahlentscheidung in Abhängigkeit von der Differenz der gemeldeten Staulängen (beide Richtungen, Kreisfläche entspricht der Anzahl der Beobachtungen).

Die Regressionsgerade berechnet sich zu

$$A_{KA} = 0{,}7138 - 0{,}0116 \cdot \left(l_{Stau,KA} - l_{Stau,HN} \right)$$

mit

A_{KA}	Anteil der Routenwahlentscheidung über Karlsruhe [-]
$l_{Stau,KA}$	Gemeldete Länge gestauten Verkehrs auf der Route über Karlsruhe [km]
$l_{Stau,HN}$	Gemeldete Länge gestauten Verkehrs auf der Route über Heilbronn [km]

Dies bedeutet, dass ohne Staumeldungen der Anteil des Durchgangsverkehrs über Karlsruhe bei ca. 71 % liegt und mit jedem Kilometer zusätzlicher Staumeldung über Karlsruhe um 1,16 % abnimmt.

Das im Vergleich zu anderen Regressionsuntersuchungen relativ geringe Bestimmtheitsmaß R^2 bei dieser und den folgenden Regressionen kann vor allem damit erklärt werden, dass hier nur eine einzelne Einflussgröße untersucht wird, die alleine die Routenwahl nicht hinreichend erklärt.

Verkehrslage (Stau und Stockend)

Um neben den Staumeldungen auch die Meldungen über den stockenden Verkehr zu berücksichtigen, wird in der folgenden Abbildung 65 die gesamte gemeldete *Verkehrslage VL* (vgl. Abschnitt 4.2.1)

$$VL_i = l_{Stau,i} + f \cdot l_{Stockend,i}$$

dem Routenwahlverhältnis gegenübergestellt.

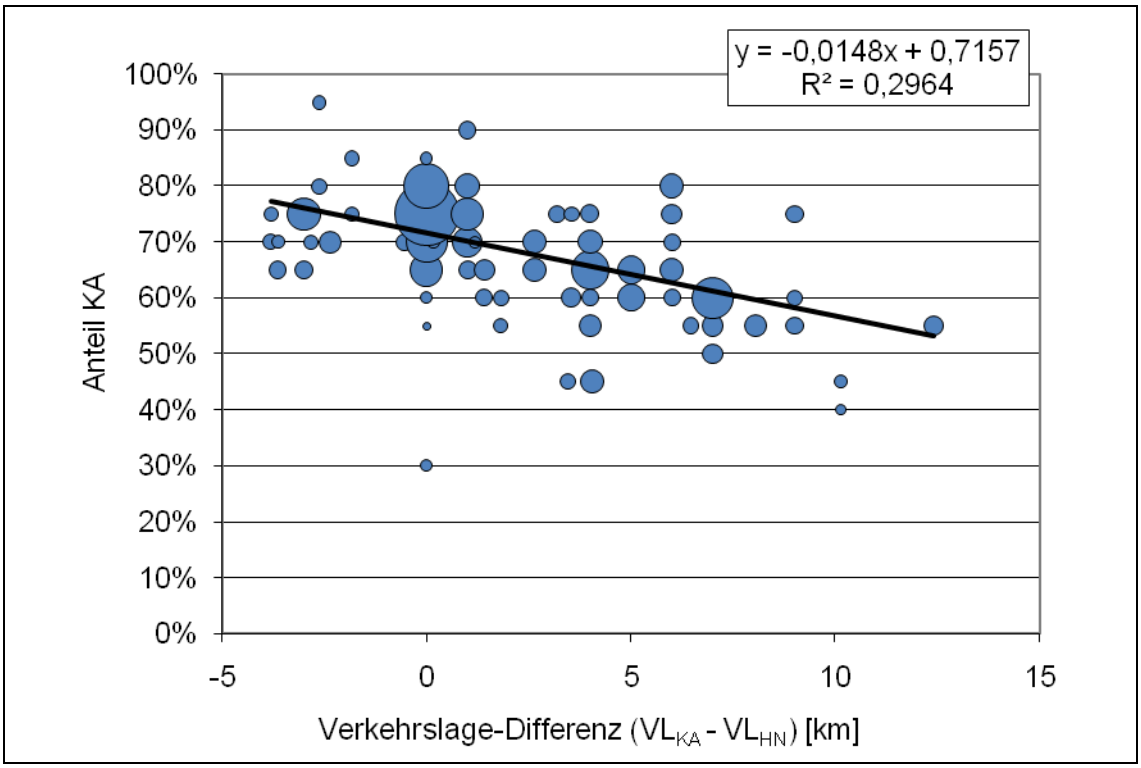


Abbildung 65: Routenwahlentscheidung in Abhängigkeit von der Differenz der Verkehrslage (beide Richtungen, Kreisfläche entspricht der Anzahl der Beobachtungen).

Die zugehörige Regressionsgerade berechnet sich zu

$$A_{KA} = 0{,}7157 - 0{,}0148 \cdot \left(VL_{KA} - VL_{HN}\right)$$

Das Bestimmtheitsmaß liegt mit R^2 = 0,296 dabei höher als bei der vorherigen Variante, bei der die Meldungen über stockenden Verkehr nicht berücksichtigt worden sind. Für den Faktor f wurde mit dem Excel Solver der Wert 0,59 ermittelt. Diesem Berechnungsergebnis folgend bedeutet dies, dass der Verkehrsteilnehmer eine Meldung über stockenden Verkehr gegenüber einer Staumeldung mit 59 % gewichtet. Dieser Wert für f weicht von dem Ergebnis 0,63 bzw. 63 % aus Schlaich & Friedrich (2008) ab, da dort die tatsächliche Auswirkung der Verkehrsmeldungen auf die Fahrzeit ermittelt worden ist, während hier die Auswirkung auf die Routenwahl untersucht wird, der eine subjektive Nutzenfunktion zugrunde liegt.

Verkehrsmeldung vorhanden: ja/nein

Die folgende Tabelle 22 zeigt eine ähnliche Untersuchung wie die obigen Korrelationsanalysen. Allerdings bleibt hier die Länge der Verkehrsmeldungen unberücksichtigt. Es wird ausschließlich berücksichtigt, ob eine Verkehrsmeldung (Stau oder stockend) vorliegt oder nicht.

	Verkehrsmeldung über …		Anteil über …		Anzahl Fahrzeuge
	Karlsruhe	Heilbronn	Karlsruhe	Heilbronn	
1	Nein	Ja	67,9%	32,1%	140
2	Nein	Nein	73,5%	26,5%	2.858
3	Ja	Ja	65,6%	34,4%	4.176
4	Ja	Nein	66,8%	33,2%	2.667

Tabelle 22: Einfluss von Staumeldungen (unabhängig von deren Art und Länge).

Diese Analyse ergibt im Gegensatz zu den bisherigen Untersuchungen keine eindeutige Tendenz. So wäre z.B. von Zeile 1 zu Zeile 2 eine Zunahme des Anteils über Heilbronn zu erwarten, da die Verkehrsmeldung über Heilbronn entfällt. Gemessen wurde dagegen eine Abnahme von 32,1 % auf 26,5 %.

Diese Beobachtung lässt vermuten, dass die Verkehrsteilnehmer in der Lage sind, Verkehrsmeldungen detailliert aufzunehmen und differenziert zu verarbeiten.

4.2.4 Entwicklung des Basismodells der Routenwahl

Nachdem die Voruntersuchungen im vorherigen Kapitel gezeigt haben, dass grundsätzlich ein Einfluss der Verkehrsmeldungen auf die Routenwahl besteht, soll nun mit Hilfe von Maximum-Likelihood-Schätzungen der Einfluss der verschiedenen dynamischen Größen untersucht und quantifiziert werden. Dafür wird – wie auch später bei den Untersuchungen mit Trajektorien aus Mobilfunkdaten – zuerst ein Basismodell entwickelt, bei dem die Logit-Funktion die folgenden Bestandteile hat:

- α_i: Konstante, die jeweils für N-1 Alternativen jeder Routenwahlentscheidung vergeben werden
- $\beta_{WWW} \cdot WWW_{aktiv}$: Term zur Berücksichtigung, ob eine Wechselwegweisung aktiv ist
- $\beta_{l_Stau} \cdot l_{Stau,j}$: Term für die Länge des gemeldeten Staus [km] sowie der zugehörige Parameter
- $\beta_{l_Stockend} \cdot l_{Stockend,j}$: Term für die Länge des gemeldeten stockenden Verkehrs [km] sowie der zugehörige Parameter

Zur Veranschaulichung des Verfahrens wird im Folgenden schrittweise jeweils einer der genannten Parameter zur Nutzenfunktion hinzugefügt und die Veränderung der Parameter sowie die sonstigen Ergebnisse erläutert. Da die Wechselwegweisung während der ANPR-Messungen nicht geschaltet worden ist, entfällt dieser Term. Jedes Modell erhält dabei eine eindeutige Nummer.

(0) Null-Log-Likelihood

Die Nutzenfunktionen sind bei beiden Alternativen konstant Null. Damit wird der sogenannte Null-Log-Likelihood berechnet werden, der sich zu –6.821 ergibt. An diesem Beispiel soll die Berechnung der Log-Likelihood erläutert werden. Tabelle 23 zeigt einen Auszug aus der Berechnungstabelle der Maximum-Likelihood-Schätzung.

Nr.	**Routenwahl**	**Wert der Nutzenfunktion**		**Wahrscheinlichkeit für die Wahl der Alternativen**		**Beitrag zur Log-Likelihood-Funktion**
		V_{KA}	V_{HN}	P_{KA}	P_{HN}	
1	KA	0	0	0,5	0,5	-0,693
2	HN	0	0	0,5	0,5	-0,693
3	KA	0	0	0,5	0,5	-0,693
...	...	...	...	...	...	...
9.839	KA	0	0	0,5	0,5	-0,693
9.840	KA	0	0	0,5	0,5	-0,693
9.841	HN	0	0	0,5	0,5	-0,693
Summe						**-6.821,261**

Tabelle 23: Modell (0): Auszug aus der Schätztabelle.

In der Tabelle sind die 9.841 beobachteten Routenwahlentscheidungen aufgelistet. Da die Nutzenfunktion beim Null-Log-Likelihood zu Null gesetzt wird, haben beide Alternativen eine Wahrscheinlichkeit von 50 %. Daraus ergibt sich über

$$\sum_{j=1}^{J} g_j \cdot \ln P\left(V\left(\alpha_j, \beta_n\right)\right)_j = 1 \cdot \ln(0{,}5) + 0 \cdot \ln(0{,}5) = -0{,}693$$

für jede Zeile ein konstanter Beitrag zur Log-Likelihood-Funktion von -0,693. Durch Aufsummieren über die 9.841 Beobachtungen ergibt sich ein Null-Log-Likelihood von -6.821.

(1) Konstante

Auch bei dieser Nutzenfunktion werden keine veränderlichen Einflussgrößen beachtet. Allerdings wird für jeweils eine Verbindung eine Konstante angesetzt:

- $V_{KA,r} = \alpha_{KA,r}$ mit *r* für Richtung (Walldorf-Stuttgart/W-S bzw. Stuttgart-Walldorf/S-W)
- $V_{HN} = 0$

Anzahl Beobachtungen	9.841
Null-Log-Likelihood	-6.821,3
Log-Likelihood	-6.120,7
Likelihood-Ratio (LR)	1.401,0 (hoch-signifikant)
Adjusted-ρ²	0,1024

Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	0,920	0,030	30,663	hoch-signifikant
2	$\alpha_{KA,W-S}$	0,584	0,031	18,550	hoch-signifikant

Tabelle 24: Modell (1): Ergebnis der Schätzung mit Konstanten (BAB, ANPR).

Die Schätzung ergibt einen Log-Likelihood von –6.121. Dies ist eine statistisch hochsignifikante Verbesserung der Erklärung des beobachteten Routenwahlverhaltens, da sich der Log-Likelihood gegenüber dem Null-Log-Likelihood um 700 Punkte verbessert hat. Grund für diese hohe Signifikanz ist, dass ca. 68 % der Verkehrsteilnehmer die Route über Karlsruhe gewählt haben und somit mit einem positiven Nutzenanteil für die Route über Karlsruhe die Güte der Schätzung deutlich verbessert wird.

Die folgenden Berechnungen der Nutzenfunktionen und der Wahrscheinlichkeiten für die Wahl veranschaulicht das Ergebnis (Richtung Stuttgart → Walldorf):

$$\begin{matrix} V_{KA,S\text{-}W} = 0{,}920 \\ V_{HN,S\text{-}W} = 0 \end{matrix} \rightarrow \begin{matrix} P\left(V_{KA,S\text{-}W}\right) = \dfrac{e^{0{,}920}}{e^{0{,}920} + e^{0}} = 71{,}5\% \\ P\left(V_{HN,S\text{-}W}\right) = \dfrac{e^{0}}{e^{0{,}920} + e^{0}} = 28{,}5\% \end{matrix}$$

Die Wahrscheinlichkeit für die Wahl der Route über Karlsruhe in der Fahrtrichtung Stuttgart → Walldorf beträgt 71,5 %. Dies entspricht exakt dem tatsächlichen Anteil der beobachteten Routenwahlentscheidungen. Da der Anteil der über Karlsruhe fahrenden Kfz in dieser Richtung etwas höher als in der Gegenrichtung ist, ist das $\alpha_{KA,S-W}$ höher als der Wert der Gegenrichtung.

(2) Konstante und Staulänge

Zusätzlich zu (1) werden nun die gemeldeten Staulängen der zwei Routen in den Nutzenfunktionen berücksichtigt. Der Parameter β_{l_Stau} sollte einen negativen Wert annehmen, da Staumeldungen einen negativen Nutzen darstellen:

- $V_{KA,r} = \alpha_{KA,r} + \beta_{l_Stau} \cdot l_{Stau,KA,r}$
- $V_{HN,r} = 0 \quad + \beta_{l_Stau} \cdot l_{Stau,HN,r}$

Anzahl Beobachtungen	9.841
Null-Log-Likelihood	-6.821,3
Log-Likelihood	- 6.087,7
Likelihood-Ratio (*LR*)	1.469,2 (hoch-signifikant)
Adjusted-ρ²	0,1072

Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	1,207	0,033	31,114	hoch-signifikant
2	$\alpha_{KA,W-S}$	0,749	0,038	19,780	hoch-signifikant
3	β_{l_Stau}	-0,047	0,006	-8,231	hoch-signifikant

Tabelle 25: Modell (2): Ergebnis der Schätzung mit Konstanten und der Meldung „Stau“ (BAB, ANPR).

Die Ergebnisse in Tabelle 25 zeigen, dass der Parameter β_{l_Stau} das geforderte negative Vorzeichen hat und die Parameter $\alpha_{KA,r}$ sich entsprechend etwas vergrößern. Die Verbesserung um 44 Log-Likelihood-Punkte auf –6.106 rechtfertigt die Einbeziehung dieses zusätzlichen Parameters.

(3) Konstante, Staulänge und Stockend

Zusätzlich zu (2) wird nun auch die Länge des gemeldeten stockenden Verkehrs berücksichtigt. Der Parameter $\beta_{l_Stockend}$ sollte wiederum negativ sein und analog zum Faktor f zwischen 0 und 1 des Parameters β_{l_Stau} liegen:

- $V_{KA,r} = \alpha_{KA,r} + \beta_{l_Stau} \cdot l_{Stau,KA,r} + \beta_{l_Stockend} \cdot l_{Stockend,KA,r}$
- $V_{HN,r} = 0 \quad + \beta_{l_Stau} \cdot l_{Stau,HN,r} + \beta_{l_Stockend} \cdot l_{Stockend,HN,r}$

Anzahl Beobachtungen		9.841			
Null-Log-Likelihood		-6.821,3			
Log-Likelihood		-6.078,6			
Likelihood-Ratio (*LR*)		1.485,3 (hoch-signifikant)			
Adjusted-ρ²		0,1083			
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	1,022	0,033	31,030	hoch-signifikant
2	$\alpha_{KA,W-S}$	0,791	0,038	20,670	hoch-signifikant
3	β_{l_Stau}	-0,060	0,006	-9,140	hoch-signifikant
4	$\beta_{l_Stockend}$	-0,027	0,006	-3,962	hoch-signifikant

Tabelle 26: Modell (3): Ergebnis der Schätzung des Basismodells (BAB, ANPR).

Die Parameter haben weiterhin das richtige Vorzeichen und auch das Verhältnis zwischen $\beta_{l_Stockend}$ und β_{l_Stau} liegt mit ca. 0,45 in dem sinnvollen Intervall zwischen 0 und 1. Der Unterschied zwischen den beiden Konstanten des Basismodells ist gegenüber dem Schätzmodell nur mit Konstanten deutlich geringer. Das bedeutet, dass ein Teil der Asymmetrie der Routenwahl in den beiden Fahrtrichtungen durch Staumeldungen erklärt werden kann. Der Log-Likelihood beträgt –6.078, so dass das Likelihood Ratio LR im Vergleich zur vorherigen Untersuchung einen Wert von 30 annimmt und somit hochsignifikant ist.

In Abbildung 66 ist eine grafische Darstellung des Basismodells dargestellt, Meldungen über stockenden Verkehr sind dabei mit dem ermittelten Faktor *f* von 0,45 zu gewichten.

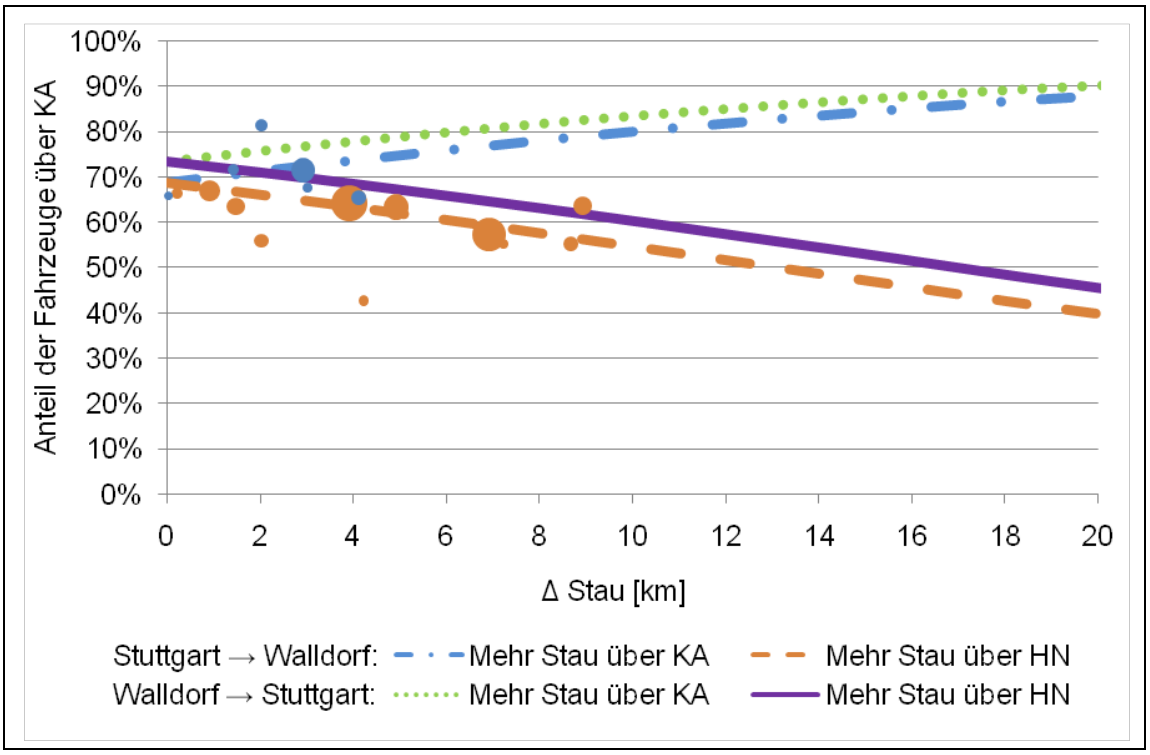


Abbildung 66: Grafische Darstellung des Basismodells (Linien) sowie Darstellung der beobachteten Werte der Richtung Stuttgart → Walldorf (Kreise).

Die Linien stellen teilweise eine Extrapolation der beobachteten Werte dar. Wie für die Richtung Stuttgart → Walldorf durch die tatsächlich gemessenen Werte als Kreise dargestellt, reichen die tatsächlichen Messwerte nur von ca. 5 km mehr Stau über Heilbronn bis ca. 10 km mehr Stau über Karlsruhe.

Interessant ist auch festzuhalten, dass eine Schätzung, die nur die Länge der Meldungen ohne Unterscheidung von „stockend“ und „Stau“ berücksichtigt, deutlich schlechter ist. Die Verkehrsteilnehmer reagieren also auf die Art der Meldung. Diese Reaktion ist – wie eine weitere Schätzung zeigt – nahezu unabhängig von der Richtung.

Das geschätzte Basismodell für die Untersuchungen aus Kennzeichenerfassungen liefert somit für die beta-Parameter die erwarteten Vorzeichen und Abstufungen zwischen den Werten. Sowohl der Likelihood-Ratio-Test als auch die Signifikanzanalysen für die einzelnen Parameter liefern hoch-signifikante Ergebnisse. Die Güte des Gesamtmodells, beschrieben durch das *adjusted-* ρ^2, ist dagegen relativ gering. Der Wert von ca. 0,11 entspricht nach Domencich & McFadden (1975) einem R^2 der linearen Regression von ca. 0,25. Weitere Untersuchungen in den folgenden Kapiteln sollen nun zeigen, ob weitere Einflussgrößen identifiziert werden können.

4.2.5 Einfluss der tatsächlichen Fahrzeit

In diesem Abschnitt soll nun geprüft werden, inwieweit der Verkehrsteilnehmer die Fahrzeiten auf den Alternativrouten abschätzen kann und daraus ableitend die schnellere Route wählt. Wie und aufgrund welcher Informationen der Verkehrsteilnehmer seine Abschätzung macht, kann nicht mit Hilfe der Schätzung ermittelt werden. Mögliche Gründe, warum diese Analyse zu einer im Vergleich zum Basismodell besseren Schätzgüte führen könnte, sind im Folgenden aufgelistet:

- Kenntnis der Verkehrsteilnehmer über tageszeitabhängige „typische“ Fahrzeiten
- Kenntnis der Verkehrsteilnehmer über potentielle Störungen aufgrund der Baustellensituation
- Kenntnis der Auswirkungen bestimmter Staumeldungen auf die Fahrzeit.

(4) Konstante, Stau, Stockend und Fahrzeitdifferenz

Aus den ANPR-Messungen können direkt mittlere Durchfahrtszeiten abgeleitet werden. Da die Messungen allerdings nur für wenige Tage durchgeführt worden sind, können daraus keine tageszeitabhängigen „typischen“ Fahrzeiten berechnet werden, so dass die tatsächliche Fahrzeitdifferenz $\left(t_{akt,KA,r} - t_{akt,HN,r}\right)$ in Minuten als typische Fahrzeit angenommen wird. Das führt dazu, dass die Fahrzeiten stark mit den Staumeldungen korrelieren (vgl. Schlaich & Friedrich, 2008).

Trotz dieser Schwäche soll die Analyse durchgeführt werden. Die Ergebnisse in Tabelle 27 sind allerdings mit großer Vorsicht zu interpretieren.

- $V_{KA,r} = \alpha_{KA,r} + \beta_{l_Stau} \cdot l_{Stau,KA,r} + \beta_{l_Stockend} \cdot l_{Stockend,KA,r} + \beta_{Fahrzeitvorteil_KA} \cdot \left(t_{akt,KA,r} - t_{akt,HN,r}\right)$
- $V_{HN,r} = 0 \quad + \beta_{l_Stau} \cdot l_{Stau,HN,r} + \beta_{l_Stockend} \cdot l_{Stockend,HN,r} + 0$

Anzahl Beobachtungen	9.841				
Null-Log-Likelihood	-6.821,3				
Log-Likelihood	-6.076,5				
Likelihood-Ratio (*LR*)	1.489,5	(hoch-signifikant)			
***LR* gegenüber Basismodell**	-4,2	(signifikant bei 96 %)			
Adjusted-ρ²	0,1084				
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	0,989	0,033	27,249	hoch-signifikant
2	$\alpha_{KA,W-S}$	0,832	0,038	18,744	hoch-signifikant
3	β_{l_Stau}	-0,058	0,006	-8,722	hoch-signifikant
4	$\beta_{l_Stockend}$	-0,024	0,006	-3,739	hoch-signifikant
5	$\beta_{Fahrzeitvorteil_KA}$	0,004	0,002	2,059	signifikant (α=4 %)

Tabelle 27: Modell (4): Ergebnis der Schätzung des Basismodells mit Fahrzeitdifferenz (BAB, ANPR).

Die Ergebnisse des nach dem Likelihood-Ratio-Test gegenüber dem Basismodell bei 96 % signifikant besseren Modells (4) zeigen, dass der Einfluss der Fahrzeitdifferenz $\beta_{Fahrzeitvorteil_KA}$ bei 0,004 liegt. Der neue Parameter führt zu einer Verringerung der Parameter β_{l_Stau} und $\beta_{l_Stockend}$, da die Einflussgrößen Verkehrsmeldung und Fahrzeitdifferenz wie oben erwähnt miteinander korrelieren. Die Verbesserung des *adjusted-ρ²* ist sehr gering.

(5) Konstante, Stau, Stockend und Fahrzeitdifferenz („Ortskenntnis“)

In einem weiteren Schritt soll nun geprüft werden, ob Fahrer mit Zulassungsbezirken aus dem Bereich der BAB-Netzmasche (11 % des Durchgangsverkehrs) bzw. aus Deutschland (89 %) sich besser als die jeweils anderen Fahrer auskennen und damit die unter (4) ermittelte Kenntnis über die Fahrzeitdifferenz stärker ausgebildet ist. Dafür werden verschiedene Varianten getestet:

- Nur Fahrzeuge aus der BAB-Netzmasche haben Kenntnis über Fahrzeitdifferenz.
- Nur Fahrzeuge aus Deutschland haben Kenntnis über Fahrzeitdifferenz.
- Bei Fahrzeugen aus der BAB-Netzmasche geht die Fahrzeitdifferenz doppelt in die Nutzenfunktion ein.

Auf die Darstellung der Ergebnisse wird an dieser Stelle verzichtet, da für keine der genannten Varianten eine signifikante Verbesserung des Log-Likelihood erreicht wird. Vielmehr ist bei den ersten beiden Varianten sogar eine Verschlechterung der Log-

Likelihood festzustellen, da nicht mehr alle Fahrzeuge die Fahrzeitdifferenz als Teil ihrer Nutzenfunktion haben.

(6) Fahrzeitdifferenz, Stau und Stockend

Umlegungsmodelle verteilen den Verkehr häufig anhand der Widerstände, die sich zu einem wesentlichen Teil aus der Fahrzeit t_{akt} ergeben, auf die Alternativrouten. Ein konstanter Nutzen, wie in den obigen Schätzungen der Alternative über Karlsruhe zugeordnet, ist in der Regel nicht vorgesehen. Im Rahmen einer detaillierten Modelleichung werden diese dann allerdings z.B. in Form von Abbiegezuschlägen und Abminderungsfaktoren für bestimmte Streckentypen eingearbeitet.

Daher wird in einer weiteren Schätzung analysiert, ob die tatsächliche Fahrzeitdifferenz mit dem zugehörigen Parameter $\beta_{Fahrzeitvorteil_KA}$ anstelle der Konstanten die beobachtete Routenwahl erklären kann.

- $V_{KA,r} = \beta_{Fahrzeit} \cdot t_{akt,KA,r} + \beta_{l_Stau} \cdot l_{Stau,KA,r} + \beta_{l_Stockend} \cdot l_{Stockend,KA,r}$
- $V_{HN,r} = \beta_{Fahrzeit} \cdot t_{akt,HN,r} + \beta_{l_Stau} \cdot l_{Stau,HN,r} + \beta_{l_Stockend} \cdot l_{Stockend,HN,r}$

Anzahl Beobachtungen		9.841			
Null-Log-Likelihood		-6.821,3			
Log-Likelihood		-6.698,6			
Likelihood-Ratio (*LR*)		245,3	(hoch-signifikant)		
***LR* gegenüber Basismodell**		-1240,0	Verschlechterung des Log-Likelihoods		
Adjusted-ρ²		0,0175			
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\beta_{Fahrzeit}$	-0,010	0,001	-7,350	hoch-signifikant
2	β_{l_Stau}	0,069	0,005	12,983	hoch-signifikant
3	$\beta_{l_Stockend}$	0,004	0,006	0,704	nicht signifikant

Tabelle 28: Modell (6): Ergebnis der Schätzung des Basismodells mit räumlicher Unterteilung der Verkehrsmeldungen (BAB, ANPR).

Die Modellergebnisse in Tabelle 28 zeigen, dass sich der Log-Likelihood gegenüber dem Basismodell bei gleicher Anzahl von Parametern deutlich um über 600 Log-Likelihood-Punkte verschlechtert hat. Die Güte des gesamten Modells *adjusted-ρ²* ist nahe bei Null und auch die Vorzeichen der Parameter entsprechen nicht immer den Erwartungen, da Verkehrsmeldungen nach den Schätzergebnissen einen positiven Nutzen haben.

Insgesamt ist damit festzuhalten, dass die tatsächliche Fahrzeit nur einen geringen Mehrwert bei der Erklärung des Gesamtmodells liefert. Da dies allerdings auch auf die angesprochenen methodischen Schwächen zurückgeführt werden kann, sollten aus den Untersuchungen dieses Kapitels keine Schlüsse gezogen werden. Bei den Untersuchungen mit Trajektorien aus Mobilfunkdaten stehen über einen längeren Zeitraum gemittelte Fahrzeiten zur Verfügung mit denen die hier gemachten Analysen wiederholt werden.

4.2.6 Einfluss der Lage von Störungsmeldungen

Im Basismodell werden die Verkehrsmeldungen in den zwei Kategorien „Stau“ und „stockend“ über die gesamte Länge der Route aufsummiert, ohne die Lage der Störung zu berücksichtigen. Mit einer Schätzung, die die Verkehrsmeldungen in zwei Abschnitte aufteilt, soll nun untersucht werden, ob der Verkehrsteilnehmer die Meldungen nach Lage differenziert aufnimmt und Meldungen in weiterer Entfernung geringer gewichtet. Die Aufteilung in zwei Abschnitte erfolgt dabei an Ecken der BAB-Netzmasche. Somit umfasst z.B. der erste Abschnitt eines Verkehrsteilnehmers von Stuttgart über Karlsruhe nach Walldorf die A8 (Stuttgart-Karlsruhe) und der 2. Abschnitt die A5 (Karlsruhe-Walldorf).

Anzahl Beobachtungen	9.841				
Null-Log-Likelihood	-6.821,3				
Log-Likelihood	-6.077,5				
Likelihood-Ratio (*LR*)	1.487,5	(hoch-signifikant)			
***LR* gegenüber Basismodell**	2,2	nicht signifikant (α=66 %)			
Adjusted-ρ²	0,1082				
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	1,017	0,033	30,517	hoch-signifikant
	$\alpha_{KA,W-S}$	0,794	0,048	16,508	hoch-signifikant
2	$\beta_{l_Stau_1.Abschnitt}$	-0,063	0,007	-8,501	hoch-signifikant
3	$\beta_{l_Stau_2.Abschnitt}$	-0,058	0,009	-6,703	hoch-signifikant
4	$\beta_{l_Stockend_1.Abschnitt}$	-0,029	0,007	-4,066	hoch-signifikant
5	$\beta_{l_Stockend_2.Abschnitt}$	-0,012	0,015	-0,815	nicht signifikant

Tabelle 29: Modell (7): Ergebnis der Schätzung mit Konstanten und den Meldungen „Stau“ und „stockend“ (BAB, ANPR).

Tabelle 29 zeigt, dass die Parameter des 1. Abschnittes jeweils größer als die des zweiten Abschnittes sind und damit die These unterstützen, dass Verkehrsteilenehmer in der Lage sind, Verkehrsmeldungen differenziert aufzunehmen. Allerdings gibt es wichtige Gründe, die gegen die Annahme dieser These sprechen:

- Das Likelihood-Ratio gegenüber dem Basismodell beträgt nur 2,2. Damit ist die Güte der Anpassung nicht signifikant besser als die des Basismodells.
- Die Güte des Gesamtmodells (*adjusted-ρ²)* der abschnittsfeinen Untersuchung ist aufgrund der höheren Anzahl an Parametern schlechter als die des Basismodells.
- Einer der geschätzten Parameter ($\beta_{l_Stockend_2.Abschnitt}$) ist nicht signifikant. Dies liegt vor allem daran, dass die Meldung „stockend" seltener als die Meldung „Stau" ist.

Die Beantwortung der Frage, ob Störungsmeldungen räumlich differenziert vom Verkehrsteilnehmer erfasst und bewertet werden, kann daher nur mit einer größeren Datenbasis beantwortet werden.

Ähnliches gilt auch für Schätzungen mit getrennten Parametern nach der genannten Stauursache (Baustelle, Unfall, etc.), da aufgrund zu seltenen Auftretens der einzelnen Ursachen keine verwertbaren Ergebnisse erreicht werden können.

4.3 Routenwahlanalysen in der BAB-Netzmasche mit Mobilfunktrajektorien

Aus den Trajektorien aus Mobilfunkdaten des LAC-basierten Trajektoriengenerierungsverfahrens (vgl. Kapitel 3) werden die Trajektorien ausgewählt, die eine Diagonale (Stuttgart ↔ Walldorf oder Heilbronn ↔ Karlsruhe) durchfahren. Für jede FPD-Trajektorie wird dabei die Ein- und Ausfahrtszeit an den Autobahnkreuzen/-dreiecken sowie in der Diagonale Stuttgart ↔ Walldorf den Zeitpunkt des Erreichens der dWiSta-Anzeigen der Wechselwegweisung ermittelt. Jede Trajektorie wird außerdem einer Alternative zugeordnet. Dabei ist es, wie bei den Kennzeichenerfassungen, nicht nötig, dass die Trajektorien komplett auf den Autobahnen der BAB-Netzmasche verlaufen. Die Anzahl der Trajektorien, die die Autobahn kurzzeitig verlassen, ist sehr gering und häufig nur darauf zurückzuführen, dass eine alternative Strecke entlang einer BAB innerhalb derselben LAC liegt (z.B. Pforzheim) und daher ein sehr kleiner Anteil der Belastung bei der automatischen Routenwahl auf die alternative Strecke verteilt worden ist. Darüber hinaus hat sich ein Verkehrsteilnehmer, auch wenn er z.B. bei Pforzheim kurzzeitig die BAB verlässt, bei der Zufahrt zur BAB-Netzmasche für eine der zwei Alternativen entschlossen, so dass auch dieser Verkehrsteilnehmer für die Routenwahlanalysen verwendet werden kann.

4.3.1 Datenaufbereitung für Routenwahlanalysen

Auch bei den Trajektorien aus dem LAC-basierten Trajektoriengenerierungsverfahrens ist es ähnlich wie bei den Daten aus den Kennzeichenerfassungen notwendig, sehr langsame Fahrzeuge herauszufiltern. Neben den in Kapitel 4.2.1 genannten Gründen ist dies bei den Trajektorien aus Mobilfunkdaten auch deshalb nötig, um Trajektorien niedriger Qualität herauszufiltern. Darüberhinaus bieten die Trajektorien bereits durch Anschauen der Fahrzeiten die Möglichkeit, zwischen Pkw und Lkw zu unterscheiden (vgl. Abbildung 67).

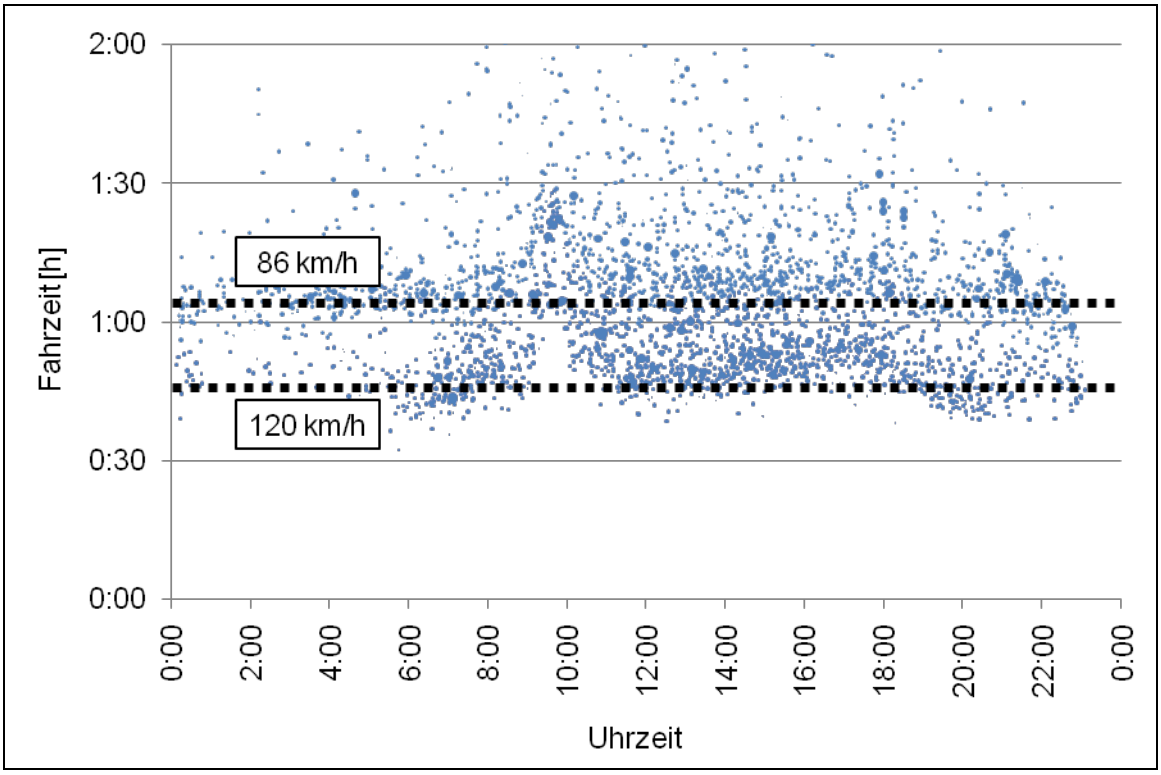


Abbildung 67: Fahrzeiten vom Dreieck Leonberg über Karlsruhe zum Kreuz Walldorf (A8 & A5), ca. 90 km.

Über drei einfache Zugehörigkeitsfunktionen (vgl. Abbildung 68) wird für jede der drei Klassen Pkw, Lkw (umfasst auch langsame Pkw), gebrochener Durchgangsverkehr eine Gesamtpunktzahl ermittelt, aus der dann eine Zuordnung zu einer der Klassen erfolgt.

1. Die erste Zugehörigkeitsfunktion orientiert sich an dem gleitenden Mittel der Fahrzeit aller Fahrzeuge. Bis zu 80 % der mittleren Fahrzeit bekommt nur die Klasse Pkw einen Punkt. Zwischen 80 und 120 % sinkt der Wert für Pkw und der Wert für Lkw steigt. Ab 150 % bekommt die Klasse gebrochener Durchgangsverkehr die höchste Punktzahl.
2. Da die erste Zugehörigkeitsfunktion zu Zeiten, in denen die mittlere Fahrzeit sehr schnell ist, dazu führen kann, dass Lkw unrealistisch hohe Geschwindigkeiten

haben, vergibt die zweite Zugehörigkeitsfunktion bis zur gewählten maximalen Lkw-Geschwindigkeit von 100 km/h einen Punkt an die Klasse Pkw und keinen Punkt an die Klasse Lkw. Die Grenze von 100 km/h ist aufgrund der Unschärfen der Trajektorien aus Mobilfunkdaten großzügig gewählt.

3. Die dritte Zugehörigkeitsfunktion berücksichtigt den erhöhten Lkw-Anteil zwischen 21 Uhr und 5 Uhr mit einer erhöhten Punktzahl für die Fahrzeugklasse Lkw.

Die Zuordnung der Fahrzeugklasse für eine Trajektorie erfolgt nun, indem die Klasse mit der höchsten Punktzahl ausgewählt wird. Dabei können ggf. Abminderungsfaktoren für einzelne Fahrzeugklassen vergeben werden, sofern die Ergebnisdarstellung nicht das gewünschte Ergebnis zeigt.

Zur Verringerung der Auswirkungen von Ausreißern empfiehlt es sich, das Verfahren zweimal durchzuführen und beim zweiten Mal die Fahrzeuge, die im ersten Durchgang als gebrochener Durchgangsverkehr identifiziert worden sind, nicht mehr bei der Ermittlung des gleitenden Mittels der Fahrzeit zu berücksichtigen.

© VuV 2009

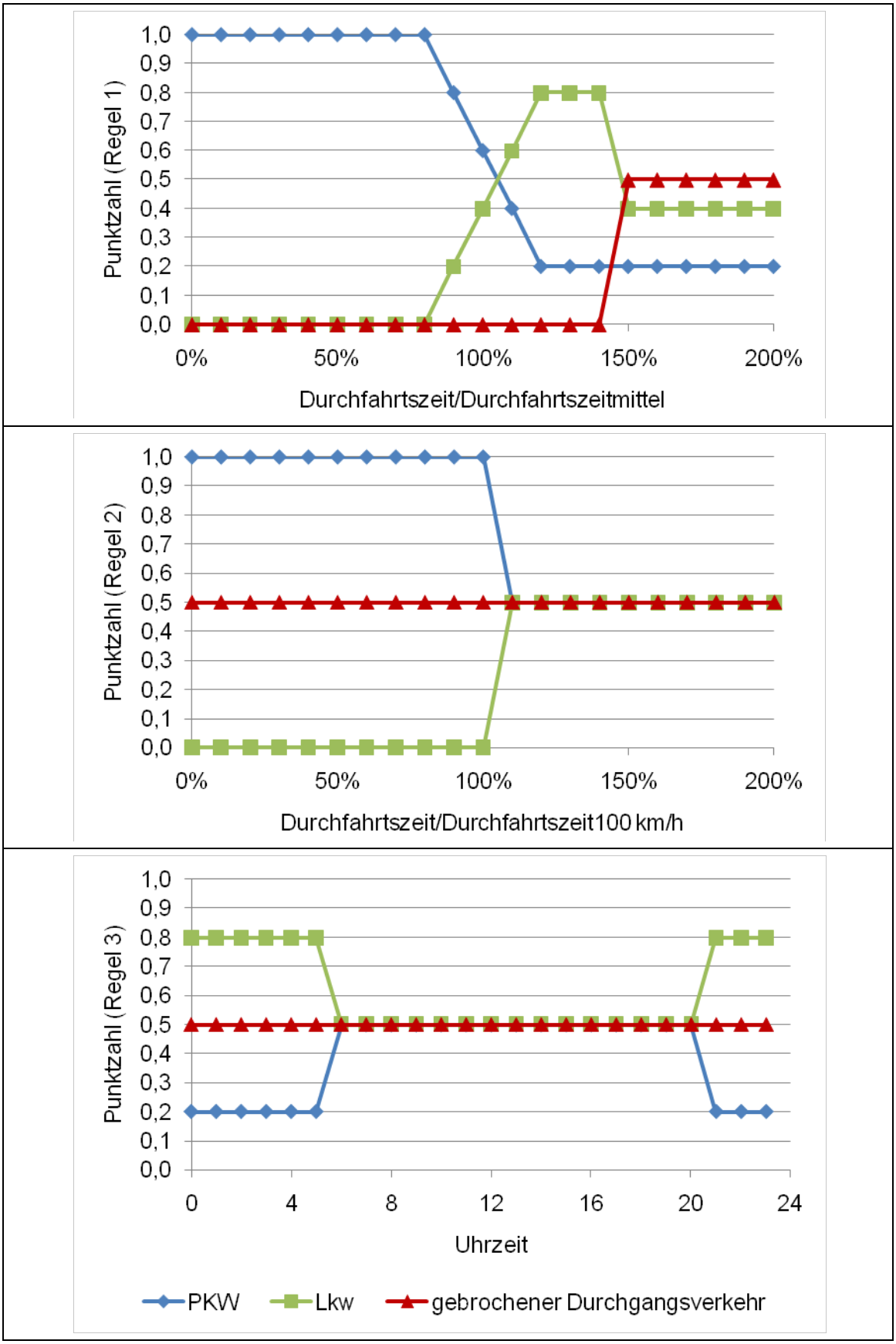


Abbildung 68: Regeln (1) - (3) zur Unterscheidung von Fahrzeugklassen.

Das Ergebnis der Aufteilung ist in Abbildung 69 für einen Tag abgedruckt. Es zeigt sich, dass auch bei allgemein hohen Fahrzeiten die Unterscheidung zwischen Pkw und Lkw zu erkennen ist, was bei Cluster-basierten Ansätzen nicht möglich wäre. Dafür gibt es im Gegenzug bei dem vor allem auf Fahrzeiten basierenden Verfahren einige Zuordnungen, die bei einer manuellen Zuordnung sicherlich anders gemacht würden.

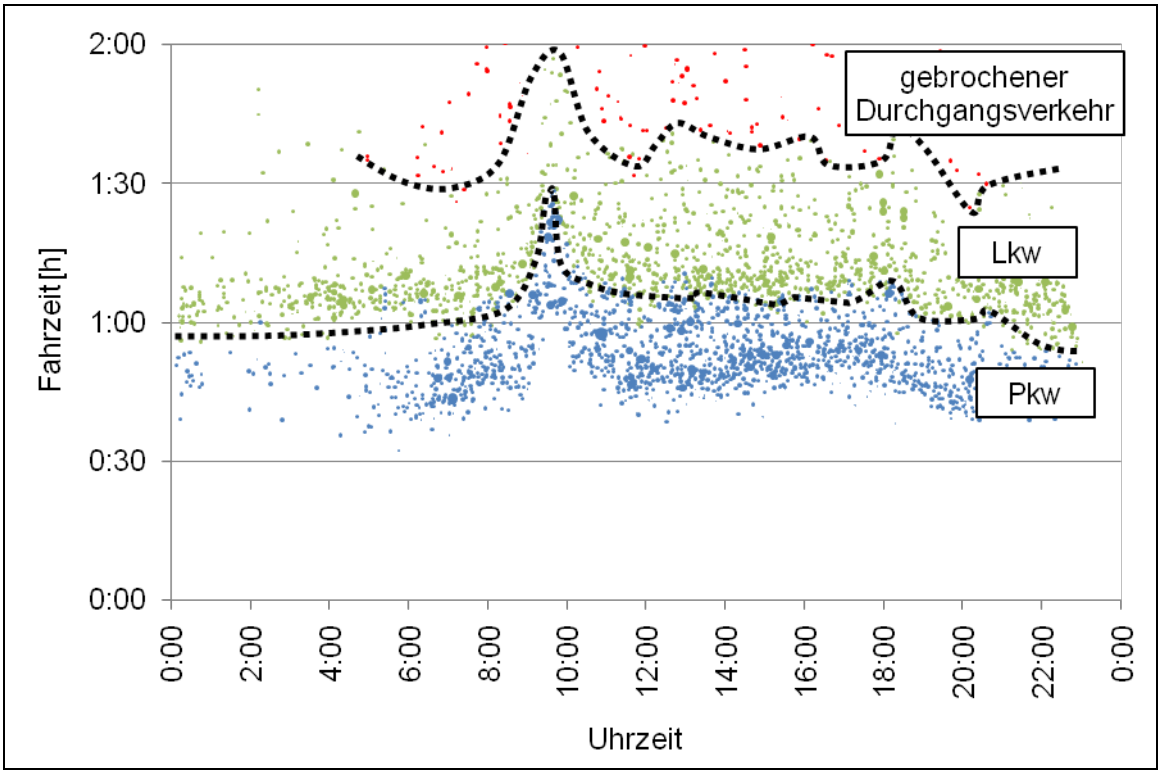


Abbildung 69: Aufteilung in die Fahrzeugklassen Pkw, Lkw und gebrochener Durchgangsverkehr.

Die Aufteilung auf Pkw und Lkw kann nur Montag bis Freitag erfolgreich durchgeführt werden, da am Wochenende der Lkw-Anteil zu gering ist, um Punktwolken von Pkw und Lkw zu trennen. Dies liegt unter anderem am Lkw-Ferienfahrverbot an Samstagen auf allen Autobahnen des Untersuchungsgebietes nach der Ferienreiseverordnung (BMJ, 2009a) sowie dem allgemeinem Lkw-Fahrverbot an Sonn- und Feiertagen gemäß §30 StVO (BMJ, 2009b).

4.3.2 Stichprobenbeschreibung

Die folgende Tabelle 30 zeigt den Stichprobenumfang für die Routenwahlschätzungen in der BAB-Netzmasche mit insgesamt über 1.000.000 Mobilfunktrajektorien. Für die Diagonale Stuttgart ↔ Walldorf ergibt sich in beiden Richtungen eine Aufteilung von ca. 70 zu 30 % zugunsten der Route über Karlsruhe. Abbildung 70 zeigt, dass der Anteil über Heilbronn während der 80 Untersuchungstage zwischen 20 und 40 % variiert. In der anderen Diagonale Karlsruhe ↔ Heilbronn dominiert die Route über

Walldorf mit über 95 % Anteil. Der mit ungefähr 50 % sehr hohe Lkw-Anteil wird am Ende des Kapitels 3.9.1 erläutert.

Richtung	Route	Anzahl Durchgangs-verkehr	Anteil (je Richtung)	Pkw-Anteil (nur Mo-Fr)	Lkw-Anteil (nur Mo-Fr)
S-W	S-KA-W	235.628	71,2%	50,9%	49,1%
	S-HN-W	95.262	28,8%		
W-S	W-KA-S	271.188	70,5%	50,0%	50,0%
	W-HN-S	113.372	29,5%		
Gesamt S ↔ W		**715.451**			
KA-HN	KA-W-HN	160.701	95,2%	48,1%	51,9%
	KA-S-HN	8.092	4,8%		
HN-KA	HN-W-KA	169.424	96,6%	50,9%	49,1%
	HN-S-KA	6.044	3,4%		
Gesamt KA ↔ HN		**344.261**			

Tabelle 30: Übersicht über den Durchgangsverkehr von Mobilfunkgeräten (S = Stuttgart, KA = Karlsruhe, HN = Heilbronn, W = Walldorf).

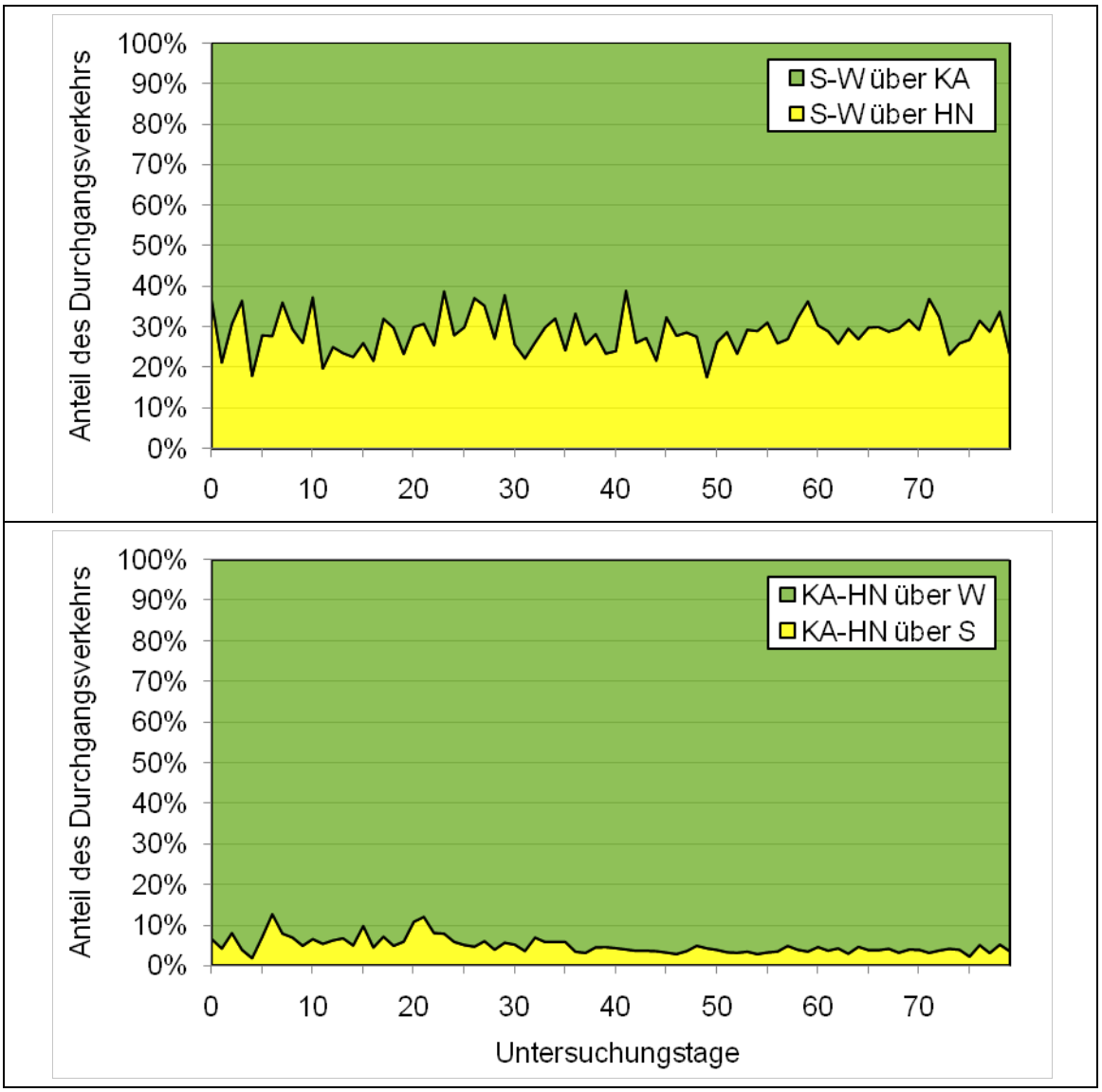


Abbildung 70: Routenwahl zwischen Stuttgart und Walldorf sowie zwischen Karlsruhe und Heilbronn.

4.3.3 Basismodell der Routenwahl in der BAB-Netzmasche

Analog zu den Maximum-Likelihood-Schätzungen der Daten aus Kennzeichenerfassungssystemen (Modell (3), vgl. Kapitel 4.2.4) wird auch für Trajektorien aus Mobilfunkdaten zuerst ein Basismodell geschätzt, das als Ausgangsbasis für weitere Analysen dienen wird. Das Modell wird für die beiden Diagonalen in der BAB-Netzmasche gemeinsam geschätzt, die Konstanten α werden dabei für die jeweils stärker genutzte Alternative vergeben. Die Nutzenfunktionen des Basismodells lauten daher wie folgt:

Diagonale Stuttgart ↔ Walldorf:

$$V_{KA,r} = \alpha_{KA,r} + \beta_{WWW} \cdot WWW_{aktiv,r} + \beta_{l_Stau} \cdot l_{Stau,KA,r} + \beta_{l_Stockend} \cdot l_{Stockend,KA,r}$$

$$V_{HN,r} = 0 \quad + \beta_{WWW} \cdot WWW_{aktiv,r} + \beta_{l_Stau} \cdot l_{Stau,HN,r} + \beta_{l_Stockend} \cdot l_{Stockend,HN,r}$$

Diagonale Karlsruhe ↔ Heilbronn:

$$V_{W,r} = \alpha_{W,r} + \beta_{l_Stau} \cdot l_{Stau,W,r} + \beta_{l_Stockend} \cdot l_{Stockend,W,r}$$

$$V_{S,r} = 0 \quad + \beta_{l_Stau} \cdot l_{Stau,S,r} + \beta_{l_Stockend} \cdot l_{Stockend,S,r}$$

mit

$V_{j,r}$	Nutzen der Alternative j in der Richtung r
$\alpha_{j,r}$	Konstanter Nutzen der Alternative j (j = KA, W) in der Richtung r
β_{WWW}	Parameter zur Gewichtung der Größe $WWW_{aktiv,r}$
$WWW_{aktiv,r}$	1, wenn Wechselwegweisung der Richtung r aktiv ist, ansonsten 0
β_{l_Stau}	Parameter zur Gewichtung der Größe $l_{Stau,j,r}$
$l_{Stau,j,r}$	Länge der gemeldeten Staus der Alternative j in der Richtung r
$\beta_{l_Stockend}$	Parameter zur Gewichtung der Größe $l_{Stockend,j,r}$
$l_{Stockend,j,r}$	Länge der Meldung stockend der Alternative j in der Richtung r

Auf die Darstellung des schrittweisen Hinzufügens der Einflussgrößen wird an dieser Stelle verzichtet, die Vorgehensweise kann in Kapitel 4.2.4 nachgelesen werden. Tabelle 31 zeigt die Ergebnisse für das Basismodell für die FPD-Trajektorien.

Anzahl Beobachtungen	1.059.712				
Null-Log-Likelihood	-734.431,6				
Log-Likelihood	-486.300,8				
Likelihood-Ratio (*LR*)	496.261,7 (hoch-signifikant)				
Adjusted-ρ²	0,3375				
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	0,936	0,021	45,188	hoch-signifikant
2	$\alpha_{KA,W-S}$	0,924	0,021	43,886	hoch-signifikant
3	$\alpha_{W,KA-HN}$	2,968	0,048	61,702	hoch-signifikant
4	$\alpha_{W,HN-KA}$	3,328	0,057	58,734	hoch-signifikant
5	β_{l_Stau}	-0,062	0,003	-19,929	hoch-signifikant
6	$\beta_{l_Stockend}$	-0,052	0,005	-11,624	hoch-signifikant
7	β_{www}	0,428	0,070	6,079	hoch-signifikant

Tabelle 31: Modell (8): Ergebnis der Schätzung des Basismodells (BAB, FPD).

Bei den Konstanten ergeben sich für beide Richtungen einer Diagonalen ähnliche Werte, die absoluten Werte sind jedoch bei der Diagonale Karlsruhe ↔ Heilbronn deutlich höher, da dort die Route über Walldorf sehr stark dominiert. Auf diesen hohen Anteil lässt sich auch das mit über 0,3 sehr hohe *adjusted-ρ²* zurückführen, da bei einer sehr stark genutzten Alternative die Gesamtgüte des Modells bereits durch die Diagonale stark verbessert wird.

Die *β*-Parameter haben die erwarteten Vorzeichen: Verkehrsmeldungen gehen negativ in die Nutzenfunktion ein, wobei die Meldung „Stau“ etwa 20 % höher ist. Eine Empfehlung der Route über Karlsruhe wirkt sich positiv auf die Nutzenfunktion dieser Route aus. Mathematisch ausgedrückt, ergibt eine Umrechnung der *β*-Parameter, dass die Wirkung einer aktiven Wechselwegweisung der Wirkung einer Staumeldung von ca. 7 km entspricht.

4.3.4 Einfluss der Fahrzeugklasse (Lkw-Pkw)

Wie in Kapitel 4.3.1 gezeigt, können montags bis freitags Pkw und Lkw aufgrund der Fahrzeiten unterschieden werden. Dies ermöglicht eine differenzierte Analyse der Routenwahl für die beiden Fahrzeugklassen. Dafür werden die 7 Parameter des Basismodells jeweils für Pkw und Lkw angelegt (vgl. Tabelle 32).

Anzahl Beobachtungen	737.183	
Null-Log-Likelihood	-510.976,1	
Log-Likelihood	-328.844,4	
Likelihood-Ratio (*LR*)	364.263,4	(hoch-signifikant)
***LR* gegenüber Basismodell**	446,5	(hoch-signifikant ggü. Basismodell Montag-Freitag)
Adjusted-ρ²	0,3564	(Basismodell Montag-Freitag: 0,3560)

Parameter			**Schätzergebnisse**			
Nr.	**Name**	**Pkw/Lkw**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	Lkw	1,004	0,034	29,730	hoch-signifikant
2		Pkw	0,914	0,032	28,933	hoch-signifikant
3	$\alpha_{KA,W-S}$	Lkw	0,912	0,034	27,101	hoch-signifikant
4		Pkw	0,946	0,032	29,232	hoch-signifikant
5	$\alpha_{W,KA-HN}$	Lkw	3,102	0,081	38,281	hoch-signifikant
6		Pkw	2,836	0,070	40,503	hoch-signifikant
7	$\alpha_{W,HN-KA}$	Lkw	3,647	0,106	34,208	hoch-signifikant
8		Pkw	3,159	0,080	39,274	hoch-signifikant
9	β_{l_Stau}	Lkw	-0,066	0,005	-13,492	hoch-signifikant
10		Pkw	-0,067	0,005	-14,550	hoch-signifikant
11	$\beta_{l_Stockend}$	Lkw	-0,050	0,008	-6,702	hoch-signifikant
12		Pkw	-0,054	0,007	-7,681	hoch-signifikant
13	β_{www}	Lkw	0,344	0,091	3,772	hoch-signifikant
14		Pkw	0,414	0,143	2,090	hoch-signifikant

Tabelle 32: Modell (9): Ergebnis der Schätzung des Basismodells mit Unterscheidung von Pkw und Lkw, Montag-Freitag (BAB, FPD).

Eine Analyse der einzelnen Parameter zeigt die folgenden Auffälligkeiten:

- Die Fahrzeugklasse Lkw hat bei den α–Konstanten, die die kürzere Route kennzeichnen, in drei von vier Fällen einen höheren Wert. Dies bedeutet, dass Lkw grundsätzlich eine stärkere Präferenz für die kürzere Route haben, was sich u.a. durch die Maut erklären lassen könnte. Das umgekehrte Verhältnis bei der Richtung Walldorf → Stuttgart lässt sich nicht erklären.
- Die *β*-Parameter für die Reaktionen auf die Verkehrsmeldungen zeigen etwas höhere Werte bei der Fahrzeugklasse Pkw. Auch die Reaktion der Pkw auf die Hinweise der Wechselwegweisung ist größer als bei den Lkw.

Der tatsächliche Einfluss der Unterschiede der *β*-Parameter auf die Routenwahl ist dabei deutlich geringer als die der α–Konstanten. Die Darstellung der Routenwahl der Relation Stuttgart → Walldorf in Abbildung 71 verdeutlicht dies. Dort unterscheidet sich vor allem der durch die α–Konstante beeinflusste Werte ohne Staumeldungsdifferenz, während die Verläufe der Kurven sehr ähnlich sind.

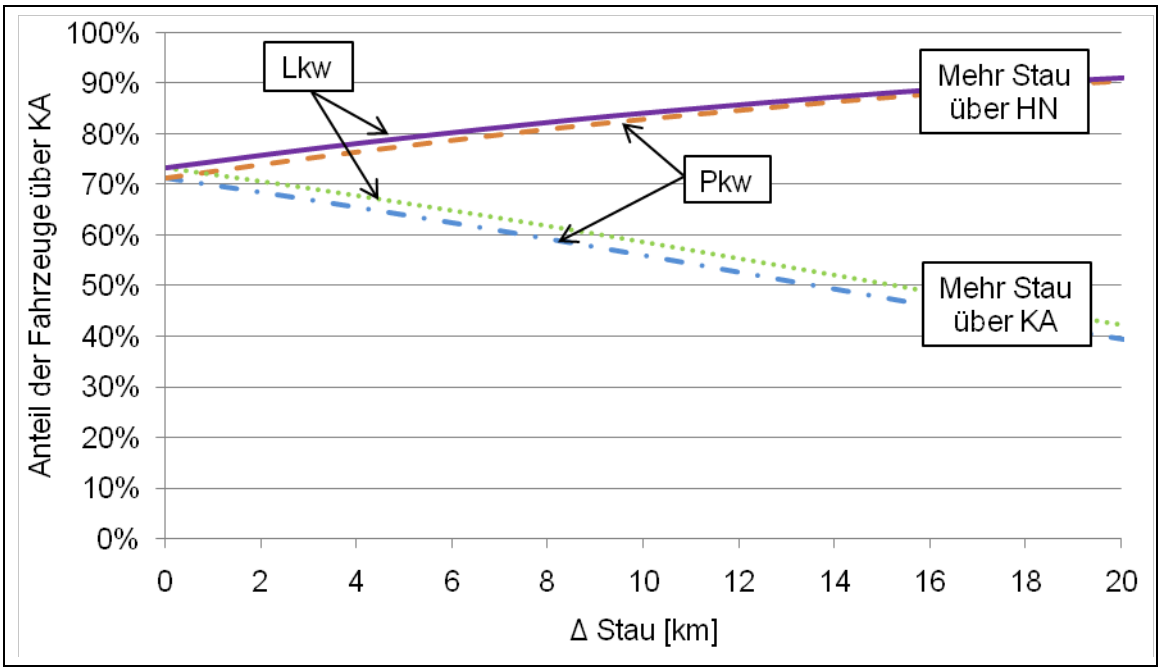


Abbildung 71: Nach Pkw und Lkw unterschiedene Routenwahl von Pkw und Lkw auf Relation Stuttgart → Walldorf.

Auch wenn sowohl die einzelnen Parameter als auch der Likelihood-Ratio-Test gegenüber dem Basismodell hoch-signifikant sind und die Ergebnisse sich weitergehend gut erklären lassen, ist der zusätzliche Nutzen der Unterscheidung von Pkw und Lkw zur Güte des Gesamtmodells gering. Das *adjusted-ρ²* ist nur ca. 0,1 % höher das des vergleichbaren Basismodells, obwohl es Einfluss auf alle Nutzenfunktionen des Schätzmodells hat und nicht nur auf die, bei denen sich eine bestimmte Einflussgröße (z.B. Anzeige der Wechselwegweisung) ändert.

4.3.5 Einfluss der untersuchten Relation

Im Basismodell gibt es zwar für jede Relation jeweils α–Konstante, die Einflüsse der Verkehrsmeldungen und der Wechselwegweisung, beschrieben durch die *β*-Parameter, werden dabei aber für alle Relationen gemeinsam berechnet.

Eine relationsfeine Untersuchung führt zu jeweils vier Parametern für β_{Stau} und $\beta_{Stockend}$ sowie zwei Parametern für β_{WWW}, also insgesamt zu einer Verdoppelung der Anzahl der Parameter gegenüber dem Basismodell. Die Ergebnisse der Schätzung zeigen, dass sich ähnlich wie bei der getrennten Betrachtung von Pkw und Lkw keine wesentliche Verbesserung der Güte des Gesamtmodells einstellt, sondern dass das *adjusted-ρ²* sich nur sehr wenig ändert. Innerhalb der α- und *β*-Parameter kommt es zu einigen Verschiebungen mit einhergehenden höheren Standardabweichungen, aber nicht zu grundsätzlich neuen Erkenntnissen. Auf eine tabellarische Darstellung der Ergebnisse wird daher verzichtet.

4.3.6 Einfluss der typischen, tageszeitabhängigen Fahrzeit

Mit den in den Kapiteln 5.1 und 7 beschriebenen Methoden ist es möglich, für jeden Zeitpunkt eine typischerweise zu erwartende Fahrzeit auf den Autobahnen in der BAB-Netzmasche zu bestimmen. Dafür werden die typischen Fahrzeiten in Stundenintervalle in den Hauptclustern C01 bis C06 aus Tabelle 54 ermittelt und anschließend den Trajektorien des Durchgangsverkehrs zugeordnet. Die in Abbildung 72 dargestellten Fahrzeiten und die sich daraus ergebende Differenz ist somit eine dynamische Größe, die in ein Schätzmodell eingehen kann.

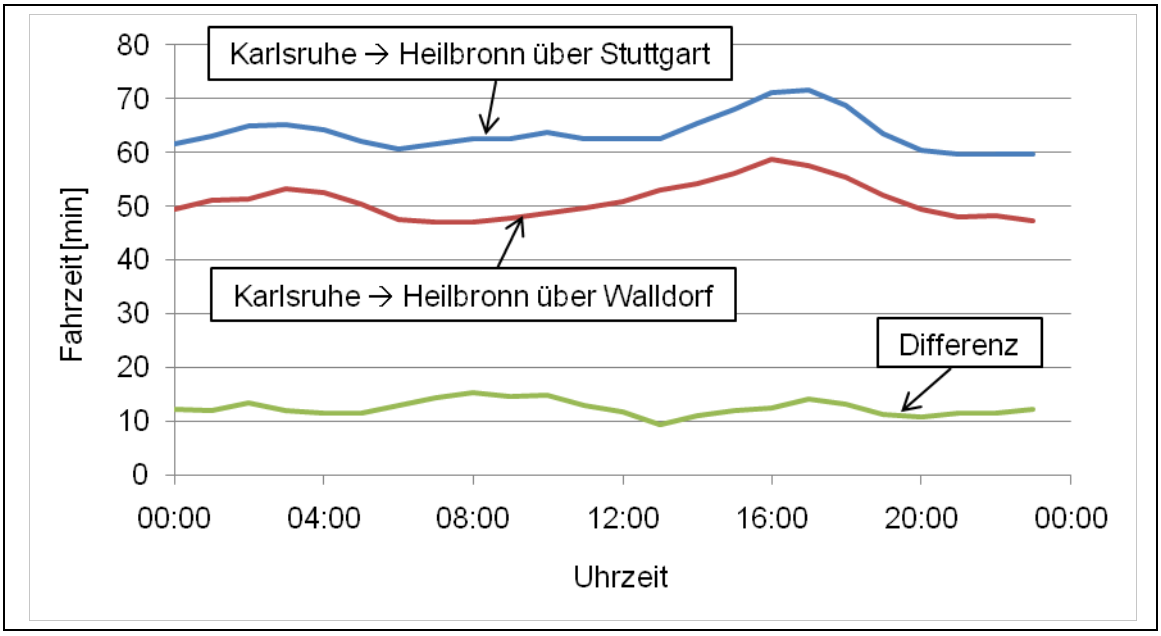


Abbildung 72: Typische Fahrzeiten zwischen Karlsruhe und Heilbronn an einem Freitag, Schule.

In den folgenden zwei Kapiteln wird nun untersucht, ob die Fahrzeiten die α–Konstanten komplett ersetzen können oder ob die Fahrzeiten zusätzlich zu den α–Konstanten die Modellgüte erhöhen können.

4.3.6.1 Substitution der Konstanten

In Verkehrsplanungsmodellen wird zur Aufteilung häufig eine Widerstandsfunktion verwendet, die sich komplett oder zu großen Teilen aus der Fahrzeit ergibt. Daher soll untersucht werden, ob die Konstante durch einen Term zur Berücksichtigung der typischen Fahrzeit in Minuten ersetzt werden kann. Daraus ergeben sich die folgenden Nutzenfunktionen:

Diagonale Stuttgart ↔ Walldorf:

$$V_{KA,r} = \beta_{Fahrzeit} \cdot t_{hist,r,j} + \beta_{WWW} \cdot WWW_{aktiv,r} + \beta_{l_Stau} \cdot l_{Stau,KA,r} + \beta_{l_Stockend} \cdot l_{Stockend,KA,r}$$

$$V_{HN,r} = \beta_{Fahrzeit} \cdot t_{hist,r,j} + \beta_{WWW} \cdot WWW_{aktiv,r} + \beta_{l_Stau} \cdot l_{Stau,HN,r} + \beta_{l_Stockend} \cdot l_{Stockend,HN,r}$$

Diagonale Karlsruhe ↔ Heilbronn:

$$V_{W,r} = \beta_{Fahrzeit} \cdot t_{hist,r,j} + \beta_{l_Stau} \cdot l_{Stau,W,r} + \beta_{l_Stockend} \cdot l_{Stockend,W,r}$$

$$V_{S,r} = \beta_{Fahrzeit} \cdot t_{hist,r,j} + \beta_{l_Stau} \cdot l_{Stau,S,r} + \beta_{l_Stockend} \cdot l_{Stockend,S,r}$$

mit

$V_{j,r}$ Nutzen der Alternative j in der Richtung r

$\beta_{Fahrzeit}$ Parameter zur Gewichtung der Größe $t_{hist,r,j}$

$t_{hist,r,j}$ Historische Fahrzeit der Alternative j in der Richtung r

β_{WWW} Parameter zur Gewichtung der Größe $WWW_{aktiv,r}$

$WWW_{aktiv,r}$ 1, wenn Wechselwegweisung der Richtung r aktiv ist, ansonsten 0

β_{l_Stau} Parameter zur Gewichtung der Größe $l_{Stau,j,r}$

$l_{Stau,j,r}$ Länge der gemeldeten Staus der Alternative j in der Richtung r

$\beta_{l_Stockend}$ Parameter zur Gewichtung der Größe $l_{Stockend,j,r}$

$l_{Stockend,j,r}$ Länge der Meldung stockend der Alternative j in der Richtung r

Anzahl Beobachtungen		1.059.712			
Null-Log-Likelihood		-734.431,6			
Log-Likelihood		-503.777,0			
Likelihood-Ratio (*LR*)		461.309,2	(hoch-signifikant)		
***LR* gegenüber Basismodell**		-34.952,4	(Verschlechterung ggü. Basismodell)		
Adjusted-ρ²		0,3141			
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\beta_{Fahrzeit}$	-0,222	0,002	-105,697	hoch-signifikant
2	β_{l_Stau}	-0,044	0,003	-15,115	hoch-signifikant
3	$\beta_{l_Stockend}$	-0,027	0,004	-6,463	hoch-signifikant
4	β_{WWW}	0,617	0,071	8,654	hoch-signifikant

Tabelle 33: Modell (10): Ergebnis der Schätzung des Basismodells mit Substitution der Konstanten durch die typische Fahrzeit (BAB, FPD).

Eine Betrachtung der Verhältnisse der *β*-Parameter in Tabelle 33 gibt Aufschluss über die Wertigkeit der Verkehrsinformationen aus Sicht der Verkehrsteilnehmer:

- β_{l_Stau} / $\beta_{Fahrzeit}$ = 0,20 → 1 km Stau ≙ Fahrzeitverlängerung von 0,20 min.
- $\beta_{l_Stockend}$ / $\beta_{Fahrzeit}$ = 0,12 → 1 km stockend ≙ Fahrzeitverlängerung von 0,12 min
- β_{l_Stau} / β_{WWW} = -2,79 → Aktive WWW ≙ Fahrzeitverkürzung von 2,79 min

Die Werte der Verkehrsmeldungen liegen deutlich unter den in Schlaich und Friedrich (2008) gemessenen, durchschnittlichen Fahrzeitverlängerungen in Höhe von 3,0 min bei gemeldetem Stau bzw. 1,6 min bei der Meldung stockend. Eine Schlussfolgerung daraus wäre, dass entweder nur ein geringer Anteil der Verkehrsteilnehmer die Verkehrsmeldungen über Radio oder ihr Navigationsgerät erhalten oder dass die Verkehrsteilnehmer die Auswirkungen der Meldungen unterschätzen.

Dass diese Schlussfolgerungen nicht anwendbar sind, zeigt eine Analyse der Schätzgüte. Dieses Modell hat auch unter Berücksichtigung der geringeren Anzahl Parameter eine deutlich geringere Schätzgüte als das Basismodell (sowohl im Bezug auf das Likelihood-Ratio als auch das *adjusted-ρ²*). Auch Modelle, die zwei (einen je Diagonale) oder vier (einen je Richtung) Parameter $\beta_{Fahrzeit}$ beinhalten, erreichen schlechtere Werte beim Log-Likelihood und *adjusted-ρ²* als das Basismodell. Das bedeutet, dass die typische Fahrzeit die dynamische Routenwahl der Verkehrsteilnehmer schlechter beschreibt als Konstanten für die einzelnen Alternativen. Daraus lässt sich wiederum ableiten, dass die Konstanten entweder neben der typischen Fahrzeit noch andere nicht erfasste Einflussgrößen (z. B. Steigungsprofil, Fahrbahnqualität, Fahrtweite etc.) beinhalten oder die Verkehrsteilnehmer die typischen Fahrzeiten nicht kennen.

4.3.6.2 Ergänzung der Konstanten

Auch wenn, wie im vorherigen Kapitel gezeigt, die typischen Fahrzeiten nicht geeignet sind, die Konstanten zu ersetzen, können sie zusätzlich zu den Konstanten das Schätzmodell verbessern. Die Schätzung mit einem gemeinsamen Parameter $\beta_{Fahrzeit}$ für alle Richtungen ergibt die in Tabelle 34 gezeigten Ergebnisse.

Anzahl Beobachtungen	1.059.712				
Null-Log-Likelihood	-734.431,6				
Log-Likelihood	-486.171,3				
Likelihood-Ratio (*LR*)	496.520,8	(hoch-signifikant)			
***LR* gegenüber Basismodell**	259,1	(hoch-signifikant)			
Adjusted-ρ²	0,3380				
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	0,828	0,051	16,276	hoch-signifikant
2	$\alpha_{KA,W-S}$	0,859	0,029	29,283	hoch-signifikant
3	$\alpha_{W,KA-HN}$	2,676	0,131	20,414	hoch-signifikant
4	$\alpha_{W,HN-KA}$	3,080	0,112	27,626	hoch-signifikant
5	β_{l_Stau}	-0,061	0,003	-19,582	hoch-signifikant
6	$\beta_{l_Stockend}$	-0,051	0,005	-11,300	hoch-signifikant
7	β_{www}	0,437	0,071	6,187	hoch-signifikant
8	$\beta_{Fahrzeit}$	-0,022	0,009	-2,404	hoch-signifikant

Tabelle 34: Modell (11): Ergebnis der Schätzung des Basismodells mit Ergänzung der typischen Fahrzeit (BAB, FPD).

Der Parameter $\beta_{Fahrzeit}$ nimmt wie im vorherigen Modell wiederum einen negativen Wert an, die anderen *β*-Parameter bleiben weitestgehend unverändert. Die sich aus dem negativen $\beta_{Fahrzeit}$ ergebenen negativen Nutzenanteile werden durch etwas geringere α–Konstanten ausgeglichen.

	β-Parameter		**Veränderung**		**Fahrzeitdifferenz [min], errechnet mit $\beta_{Fahrzeit}$**
	Basismodell	**Modell (11)**	**relativ**	**absolut**	
$\alpha_{KA,S-W}$	0,936	0,828	-11,5%	-0,108	4,9
$\alpha_{KA,W-S}$	0,924	0,859	-7,0%	-0,065	2,9
$\alpha_{W,KA-HN}$	2,968	2,676	-9,8%	-0,292	13,3
$\alpha_{W,HN-KA}$	3,328	3,080	-7,5%	-0,248	11,3

Tabelle 35: Analyse der α–Konstanten im Modell (11) im Vergleich zum Basismodell.

Tabelle 35 analysiert diese Veränderung der α–Konstanten. Die relativen Veränderungen liegen zwischen 7,0 und 11,5 %. Somit bleibt ein Großteil der Konstanten weiterhin unerklärt. Die absoluten Veränderungen sind bei der Diagonalen

Heilbronn ↔ Karlsruhe höher als bei der Diagonalen Stuttgart ↔ Walldorf. Dies liegt an der höheren mittleren Fahrzeitdifferenz dieser Diagonalen, die im Bereich der in der letzten Spalte der Tabelle 35 theoretisch aus dem Quotienten der absoluten Differenz und $\beta_{Fahrzeit}$ errechneten Werte.

4.3.7 Prüfung der Aufnahmefähigkeit des Verkehrsteilnehmers

4.3.7.1 Detaillierte Analyse der Verkehrsmeldungen

Im Basismodell werden die gemeldeten Längen der Verkehrsmeldungen unabhängig von ihrem Ort aufsummiert und in der Nutzenfunktion berücksichtigt. Da die untersuchten Routen eine Länge von ca. 100 km und damit Fahrzeiten von ca. einer Stunde aufweisen, ist es denkbar, dass der Verkehrsteilnehmer unterscheidet, ob eine Verkehrsmeldung direkt zu Beginn oder am Ende der Route liegt. Im letzteren Fall ist die Verkehrsstörung mit einer höheren Wahrscheinlichkeit beim Eintreffen nicht mehr aktuell.

Anzahl Beobachtungen	1.059.712				
Null-Log-Likelihood	-734.431,6				
Log-Likelihood	-485.862,0				
Likelihood-Ratio (*LR*)	497.139,2	(hoch-signifikant)			
***LR* gegenüber Basismodell**	877,5	(hoch-signifikant)			
Adjusted-ρ²	0,3384				
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	0,951	0,021	45,390	hoch-signifikant
2	$\alpha_{KA,W-S}$	0,873	0,025	35,390	hoch-signifikant
3	$\alpha_{W,KA-HN}$	2,917	0,050	58,014	hoch-signifikant
4	$\alpha_{W,HN-KA}$	3,334	0,057	58,959	hoch-signifikant
5	$\beta_{l_Stau_1.Abschnitt}$	-0,080	0,004	-19,813	hoch-signifikant
6	$\beta_{l_Stau_2.Abschnitt}$	-0,042	0,004	-9,344	hoch-signifikant
7	$\beta_{l_Stockend_1.Abschnitt}$	-0,073	0,007	-9,738	hoch-signifikant
8	$\beta_{l_Stockend_2.Abschnitt}$	-0,042	0,006	-7,576	hoch-signifikant
9	β_{www}	0,395	0,070	5,601	hoch-signifikant

Tabelle 36: Modell (12): Ergebnis der Schätzung des Basismodells Berücksichtigung des Ortes der Verkehrsmeldung (BAB, FPD).

Tabelle 36 zeigt die Ergebnisse einer Schätzung, bei der die Verkehrsmeldungen in zwei Abschnitte unterteilt und je Abschnitt mit *β*-Parametern versehen werden. Dabei erfolgt die Unterteilung jeweils ungefähr nach der Hälfte der Route an einer Ecke der BAB-Netzmasche.

Der absolute Betrag der *β*-Parameter für die Verkehrsmeldungen auf dem ersten Abschnitt ist sowohl bei der Meldung „Stau“ als auch bei der Meldung „Stockend“ deutlich höher als der des zweiten Abschnittes. Das *adjusted-ρ²* sowie der Likelihood-Ratio-Test haben unter allen untersuchten Varianten die größten Verbesserungen gegenüber dem Basismodell, auch die Signifikanz aller Parameter ist gegeben.

Eine weitergehende Aufteilung der Verkehrsmeldungen sowie eine Untersuchung der Reaktion der Verkehrsteilnehmer auf Warnhinweise („z.B. verlorenen Ladung“) führt zu nicht verwertbaren Ergebnissen. Auch die Meldung der Vollsperrung kann nicht detailliert ausgewertet werden, da diese zu selten auftritt (vgl. Kapitel 4.1.4).

4.3.7.2 Detaillierte Analyse der Anzeigen der Wechselwegweisung

Nachdem im vorherigen Kapitel gezeigt worden ist, dass der Verkehrsteilnehmer differenziert auf die Lage der Verkehrsmeldungen reagiert, soll nun untersucht werden, ob ebenso differenziert auf die Angaben der Wechselwegweisung zu den Gründen der Umleitung auf den dWiSta-Anzeigen reagiert wird. Diese enthalten neben eine Staulänge auch noch die Angabe des Orts des Staus. Analog zu den Verkehrsmeldungen werden dazu erst die Länge des angezeigten Staus und anschließend der Ort der Staumeldungen in Schätzungen berücksichtigt.

Anzahl Beobachtungen	1.059.712				
Null-Log-Likelihood	-734.431,6				
Log-Likelihood	-486.294,0				
Likelihood-Ratio (*LR*)	496.275,3	(hoch-signifikant)			
***LR* gegenüber Basismodell**	13,5	(hoch-signifikant)			
Adjusted-ρ²	0,3379				
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	$\alpha_{KA,S-W}$	0,936	0,021	45,184	hoch-signifikant
2	$\alpha_{KA,W-S}$	0,924	0,021	43,882	hoch-signifikant
3	$\alpha_{W,KA-HN}$	2,968	0,048	61,701	hoch-signifikant
4	$\alpha_{W,HN-KA}$	3,328	0,057	58,734	hoch-signifikant
5	β_{l_Stau}	-0,062	0,003	-19,909	hoch-signifikant
6	$\beta_{l_Stockend}$	-0,052	0,005	-11,622	hoch-signifikant
7	β_{www}	0,288	0,160	1,796	nicht signifikant (α=7,3 %)
8	$\beta_{WWW,Länge}$	0,019	0,019	1,004	nicht signifikant

Tabelle 37: Modell (13): Ergebnis der Schätzung des Basismodells Berücksichtigung der auf der Wechselwegweisung angezeigten Staulänge (BAB, FPD).

Die Länge der Staumeldung auf den dWiSta-Anzeigen ergibt zwar das erwartete positive Vorzeichen (vgl. Tabelle 37), da ein Stau auf der Route über Heilbronn die Route über Karlsruhe attraktiver erscheinen lässt. Allerdings ist die Signifikanz der *β*-Parameter der Wechselwegweisung sehr gering, obwohl der Wert für β_{www} im Basismodell hoch-signifikant ist. Die sich bei der Schätzung ergebenden *β*-Parameter lassen sich mit der mittleren angezeigten Länge von ca. 7,4 Kilometer auf den Wert für das β_{www} im Basismodell zurückführen:

$$\beta_{WWW(Basismodell)} = 0{,}428 \approx 0{,}427 = \beta_{WWW(13)} + \beta_{WWW,Länge(13)} \cdot \varnothing_{\text{Stauanzeige}}$$

Eine Schätzung, die die Lage der angezeigten Staumeldung berücksichtigt, bringt für die *β*-Parameter für die zwei Abschnitte nur minimal unterschiedliche Werte und eine entsprechend geringe Verbesserung des Modells.

Aus diesem Ergebnis darf allerdings nicht geschlossen werden, dass die dWiSta-Anzeigen keine Wirkung hätten. In anderen Veröffentlichungen (z.B. FGSV, 2007) wird davon ausgegangen, dass allein durch eine Erklärung der Umleitungsempfehlung auf den dWiSta-Anzeigen die Wirkung einer Wechselwegweisung erhöht werden kann.

Dies kann in dieser Untersuchung nicht ermittelt werden, da die Wechselwegweisung in der BAB-Netzmasche die Erläuterung der Umleitungsempfehlung immer anzeigt.

4.4 Routenwahlanalysen in der NBA Stuttgart mit Mobilfunktrajektorien

4.4.1 Stichprobenbeschreibung

Durch die in Kapitel 3.11 beschriebene lokale Verfeinerung des LAC-basierten Trajektoriengenerierungverfahrens ist es möglich, auch für die Netzbeeinflussungsanlage (NBA) Stuttgart Routenwahlanalysen durchzuführen. Das Verfahren erfasst sowohl stadtein- als auch -auswärts die Fahrzeuge auf der Hauptroute B10, der Alternativroute der NBA (B295) und der B27 als dritte Möglichkeit Richtung Heilbronn (vgl. Abbildung 73).

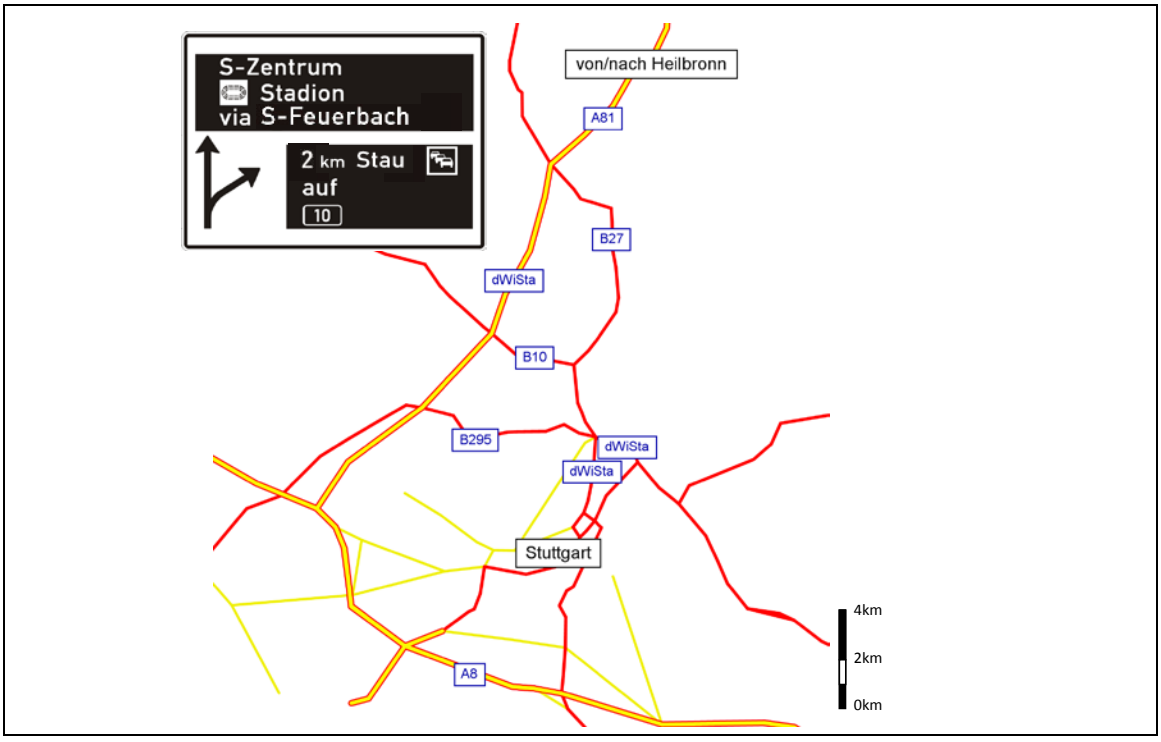


Abbildung 73: Überblick über das Untersuchungsgebiet der NBA Stuttgart.

Eine Unterscheidung der Fahrzeuge in Pkw und Lkw ist anhand der Fahrzeiten nicht möglich. Dies liegt vor allem daran, dass auf vielen Strecken im Untersuchungsgebiet die zulässige Geschwindigkeit 80 km/h nicht überschreitet und somit Lkw nicht langsamer als Pkw fahren. Abbildung 74 verdeutlicht, dass eine Trennung aufgrund der Fahrzeiten auch optisch nicht möglich ist.

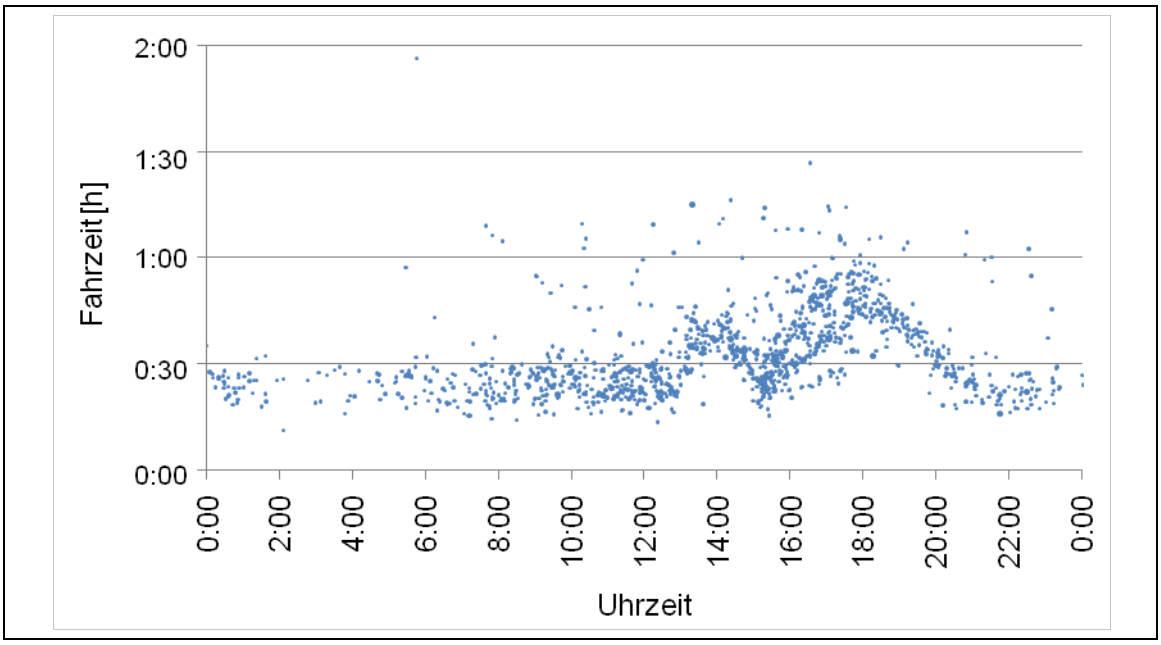


Abbildung 74: Fahrzeiten vom Stuttgarter Pragsattel über die B10 Richtung Norden.

Die folgende Tabelle 38 zeigt den Stichprobenumfang für die Routenwahlschätzungen in der Netzbeeinflussungsanlage Stuttgart. Es gehen insgesamt knapp 350.000 Mobilfunktrajektorien in die Routenwahlschätzungen ein. In beiden Richtungen dominiert die Hauptroute über die B10 mit deutlich über 90 % des Durchgangsverkehrs. Die Alternativrouten über die B295 bzw. B27 haben Anteile zwischen 2,6 und 4,7 %. Nur an wenigen Tagen haben die Alternativrouten einen Anteil von mehr als 10 % (vgl. Abbildung 75).

Richtung	Route	Anzahl Durchgangsverkehr	Anteil (je Richtung)
S-HN (stadtauswärts)	B10	160.176	93,8%
	B295	4.518	2,6%
	B27	5.984	3,5%
Gesamt S-HN (stadtauswärts)		**170.678**	
HN-S (stadteinwärts)	B10	164.826	96,6%
	B295	8.004	4,7%
	B27	5.984	3,5%
Gesamt HN-S (stadteinwärts)		**178.814**	

Tabelle 38: Übersicht über den Durchgangsverkehr (S = Stuttgart, HN = Heilbronn).

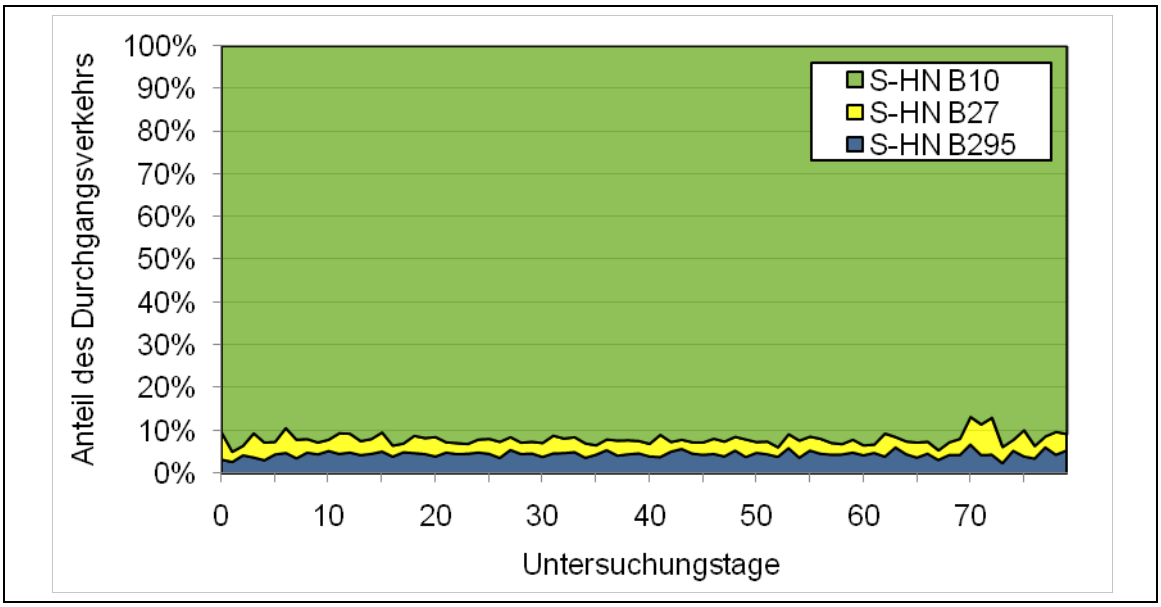


Abbildung 75: Routenwahl zwischen Stuttgart und Heilbronn.

4.4.2 Basismodell der Routenwahl in der NBA Stuttgart

Das Basismodell umfasst zwei *α*-Konstanten für die Alternativen B27 und B295 sowie *β*-Parameter für die Wechselwegweisung (nur B295) sowie die Staumeldungen auf den Bundesstraßen und Autobahnabschnitten.

Anzahl Beobachtungen		349.492			
Null-Log-Likelihood		-383.956,2			
Log-Likelihood		-105.586,9			
Likelihood-Ratio (*LR*)		556.738,5 (hoch-signifikant)			
Adjusted-ρ²		0,7250			
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	α_{B295}	-3,270	0,253	-12,942	hoch-signifikant
2	α_{B27}	-3,304	0,224	-14,785	hoch-signifikant
5	β_{l_Stau}	0,154	0,155	0,994	nicht signifikant
6	$\beta_{l_Stockend}$	-0,071	0,120	-0,587	nicht signifikant
7	$\beta_{NBA,B295}$	0,654	0,549	1,192	nicht signifikant

Tabelle 39: Modell (14): Ergebnis der Schätzung des Basismodells (NBA, FPD).

Die Ergebnisse in der Tabelle 39 zeigen, dass nur für die α-Konstanten signifikante Ergebnisse erreicht werden. Die β-Parameter dagegen haben alle eine sehr hohe Standardabweichung, die teilweise sogar höher als die Werte der Parameter sind.

Darüberhinaus hat der Parameter β_{l_Stau} nicht das erwartete Vorzeichen. Ein positiver Parameter bedeutet hier, dass die Verkehrsteilnehmer eine Staumeldung positiv in der Nutzenfunktion berücksichtigen. Dies lässt sich nicht erklären, insbesondere da die Meldung „stockend" negativ bewertet wird. Dies führt zu der Hypothese, dass Verkehrsmeldungen für den Fall der NBA mit vielen regelmäßigen Fahrern keinen Einfluss auf die Routenwahl haben. Dies kann auch darauf zurückzuführen sein, dass der Verkehrsteilnehmer bei der B295 mit ihrem teilweise innerstädtischen Charakter nicht davon ausgehen kann, dass Verkehrsstörungen auf allen Abschnitten dieser Route in den Verkehrsmeldungen enthalten sind.

Der Parameter $\beta_{NBA;B295}$ für die Netzbeeinflussungsanlage ist positiv, was bei Anwendung der Nutzenfunktion bedeutet, dass die Verkehrsteilnehmer den Empfehlungen folgen. Allerdings kann dieses Ergebnis aufgrund der hohen Standardabweichung nicht als statistisch sicher bezeichnet werden.

Die Gründe für die im Gegensatz zu den bisherigen Schätzungen des Basismodells nicht signifikanten β-Parameter lassen sich auf die folgenden zwei Gründe, die sich aus der Erhebungsmethode ergeben, zurückführen:

- Die Netzbeeinflussungsanlage ist nur selten aktiv. Insbesondere in Stadteinwärtsrichtung ist sie an der für die Routenwahlanalysen verwendeten Tagen insgesamt nur ca. 6 Stunden aktiv. Dies liegt zum einen daran, dass die Netzbeeinflussungsanlage zwischen Mitte Februar und Ende August defekt war und zum anderen daran, dass es für den 30.09.2008 mit 5 Stunden aktiver Routenempfehlung keine Versorgung mit Rohdaten aus dem Mobilfunknetz gab.
- Ein Großteil des Durchgangsverkehrs fährt zu Zeiten ohne Verkehrsmeldung im Untersuchungsbereich und beeinflusst damit nur den Wert der Konstanten.
- Die Alternativen B10 und B27 überlappen sich zwischen Pragsattel und Zuffenhausen und die Alternativen B10 und B295 zwischen Zuffenhausen und Ludwigsburg. Diese Überlappungen der drei Routen bewirken, dass einige Staumeldungen auf verschiedene Routen wirken, so dass sich die drei Alternativen im Untersuchungsgebiet geringer als drei komplett getrennte Alternativen voneinander unterscheiden.

Hier zeigt sich eine Schwäche von Daten aus tatsächlich beobachteten Routenwahlentscheidung (revealed preference, RP) gegenüber Daten aus Befragungen. Befragungen in Form von „stated preference" (SP) ermöglichen es, beliebige Situationen vorzugeben und somit abgesehen von der tatsächlichen Entscheidung des Probanden Inputdaten für die Schätzung vorzugeben.

4.4.3 Weitere Analysen der Routenwahl in der NBA Stuttgart

Wie im vorherigen Kapitel gezeigt, lassen sich für den Bereich der Netzbeeinflussungsanlage Stuttgart für das Basismodell keine statistisch gesicherten Ergebnisse für β-Parameter erreichen. Trotzdem sollen im Folgenden einige Auswertungen vorgestellt werden, die die Wirkungen der Netzbeeinflussungsanlage vor allem deskriptiv beschreiben.

Richtung	Route	Anteil der Route, wenn Netzbeeinflussungsanlage ...	
		... nicht aktiv	... aktiv
S-HN (stadtauswärts)	B10	94,0%	89,5%
	B295	2,6%	5,1%
	B27	3,5%	5,4%
Stichprobenumfang		**165.936**	**4.742**
HN-S (stadteinwärts)	B10	92,2%	76,1%
	B295	4,4%	20,4%
	B27	3,3%	3,5%
Stichprobenumfang		**178.269**	**545**

Tabelle 40: Analyse der Wirkungen der NBA Stuttgart.

Tabelle 40 fasst die Routenwahl in Abhängigkeit der Anzeige der NBA zusammen. In der Stadtauswärtsrichtung zeigt sich, dass sowohl die Route über die B295 als auch über die B27 im Falle einer aktiven NBA einen erhöhten Anteil haben. Dabei steigt der Anteil bei der von der NBA empfohlenen B295 mit +96% höher als bei der B27 (+57%). Eine mögliche Schlussfolgerung ist, auch bei der Route über die B27 einen β-Parameter in der Nutzenfunktion im Falle einer aktiven NBA zu berücksichtigen. Die Ergebnisse eines solchen Modells sind allerdings ebenfalls nicht signifikant, so dass auf eine Darstellung verzichtet wird.

In der Stadteinwärtsrichtung liegen nur 545 Beobachtungen während aktiver NBA vor. Dabei zeigt sich, dass in dieser Richtung der Anteil der Fahrzeuge über die offizielle Ausweichroute B295 bei aktiver NBA sich von 4,4 % auf 20,4 % fast verfünffacht. Der Anteil über die B27 bleibt erwartungsgemäß nahezu unverändert, da die Verkehrsteilnehmer nach Erreichen der NBA bereits an der Ausfahrt zur B27 vorbeigefahren sind. Diese deutliche Veränderung führt dazu, dass ein einfaches Modell mit Konstanten und Berücksichtigung der NBA zu signifikanten Ergebnissen, auch gegenüber einem Modell nur mit Konstanten, führt (vgl. Tabelle 41).

Anzahl Beobachtungen	178.814				
Null-Log-Likelihood	-196.447,3				
Log-Likelihood	-58.529,2				
Likelihood-Ratio (*LR*)	275.836,2 (hoch-signifikant)				
Adjusted-ρ²	0,7020				
Parameter		**Schätzergebnisse**			
Nr.	**Name**	**Wert**	**StdErr**	**t-Statistik**	**Bemerkung**
1	α_{B295}	-3,036	0,275	-11,060	hoch-signifikant
2	α_{B27}	-3,316	0,306	-10,825	hoch-signifikant
7	$\beta_{NBA,B295}$	1,701	0,634	2,684	hoch-signifikant

Tabelle 41: Modell (15): Vereinfachtes Modell für die Stadteinwärtsrichtung (NBA, FPD).

Es sollte allerdings berücksichtigt werden, dass - wie bei allem Regressionsanalysen - auch bei der Maximum-Likelihood-Schätzung ein signifikantes Ergebnis nicht zwingend einen logischen Zusammenhang zwischen unabhängiger Variable und dem beobachteten Routenwahlverhalten darstellen muss. So könnten z.B. besondere Verkehrsmeldungen einzelner Radiostationen die seltenen Schaltungen der Netzbeeinflussungsanlage begleitet haben und somit einen gewissen Anteil an dem $\beta_{NBA,B295}$ haben.

4.5 Zusammenfassung und Bewertung

4.5.1 BAB-Netzmasche

Bei den Routenwahlschätzungen in der BAB-Netzmasche wurden zwei Datenquellen analysiert:

- Daten aus ANPR-Messungen (nur Stuttgart ↔ Walldorf)
- Daten aus Mobilfunktrajektorien (beide Diagonalen)

Die Routenwahlanteile der beiden Untersuchungen liegen für die Diagonale Stuttgart ↔ Walldorf in einem ähnlichen Bereich. Daraus resultieren für die Konstanten dieser Diagonalen für beide Untersuchungen und Richtungen Werte um 0,9, wobei die Schwankung bei den ANPR-Messungen in den zwei Richtungen mehr abweichen als bei den Mobilfunktrajektorien, bei denen Hin- und Rückrichtung fast identische Konstanten im Basismodell aufweisen. Da grundsätzlich für beide Richtungen ein vergleichbares Routenwahlverhalten zu erwarten ist, kann daraus abgeleitet werden, dass die Messdauer mit ANPR-Messungen eventuell nicht ausreichend groß war, um tagesabhängige Schwankungen abzufedern.

Der zwischen den beiden Untersuchungen ebenfalls vergleichbare Parameter β_{l_Stau}, der den Einfluss von der Verkehrsmeldung „Stau“ bewertet, ist in beiden Erhebungen mit -0,060 bzw. -0,062 sehr ähnlich. Bei der Meldung „stockend“ liegen die Parameter weiter auseinander. Dies lässt sich wiederum auf den kurzen Messzeitraum bei der ANPR-Messung zurückführen, in dem die Meldung „stockend“ nur sehr selten vorgekommen ist.

Da während der ANPR-Messungen keine Schaltungen der Wechselwegweisung stattgefunden haben, können die Wirkungen dieser Schaltungen nur bei den Mobilfunktrajektorien analysiert werden. Dort zeigt sich, dass der zugehörige Parameter β_{WWW} einen Wert von 0,428 annimmt und damit der Wirkung einer Staumeldung von ca. 7km entspricht.

Die folgenden Abbildungen zeigen grafisch die Routenwahlanteile, die sich für die beiden Diagonalen ergeben. Dabei ist jeweils nur eine Richtung dargestellt, die Gegenrichtung hat bei der Diagonalen Stuttgart ↔ Walldorf einen nahezu identischen Verlauf. Bei der Diagonalen Karlsruhe ↔ Heilbronn liegen die Kurven der Gegenrichtung etwas höher, da die Konstante der Richtung Heilbronn → Karlsruhe etwas größer ist.

Abbildung 76 enthält außerdem noch den zusätzlichen Anteil der über Karlsruhe fahrenden Fahrzeuge sowie den Befolgungsgrad der Wechselwegweisung.

$$ZusätzlicherAnteil_{KA,WWW} = Anteil_{KA,WWW_aktiv} - Anteil_{KA,WWW_inaktiv}$$

$$Befolgungsgrad_{WWW} = \frac{ZusätzlicherAnteil_{KA,WWW}}{1 - Anteil_{KA,WWW_inaktiv}}$$

Dabei ist zu beachten, dass sich diese Kenngrößen auf den Durchgangsverkehr beziehen, der bezogen auf den Verkehr bei Stuttgart ca. 12 % beträgt (vgl. Kapitel 4.2.1). Die allgemeine Definition des Befolgungsgrades auf Basis der Verkehrsflüsse q_i lautet wie folgt:

$$Befolgungsgrad_{WWW} = \frac{q_{D,Alternativroute,WWW_aktiv} - q_{D,Alternativroute,WWW_inaktiv}}{q_{D,Hauptroute,WWW_inaktiv}}$$

mit

$Befolgungsgrad_{WWW}$	Befolgungsgrad einer Wechselwegweisung [-]
$q_{D,Alternativroute,WWW_aktiv}$	Durchgangsverkehr auf der Alternativroute bei aktiver Wechselwegweisung [Fz/h]
$q_{D,Alternativroute,WWW_inaktiv}$	Durchgangsverkehr auf der Alternativroute bei inaktiver Wechselwegweisung [Fz/h]
$q_{D,Hauptroute,WWW_inaktiv}$	Durchgangsverkehr auf der Hauptroute bei aktiver Wechselwegweisung [Fz/h]

Für das in Abbildung 76 eingezeichnete Beispiel bei ΔStau = 5 km ergeben sich Anteile über Karlsruhe in Höhe von 77,7 % ohne Wechselwegweisung bzw. 84,2 % bei aktiver Wechselwegweisung. Dies entspricht einem Befolgungsgrad von 29,3 % des Durchgangsverkehrs.

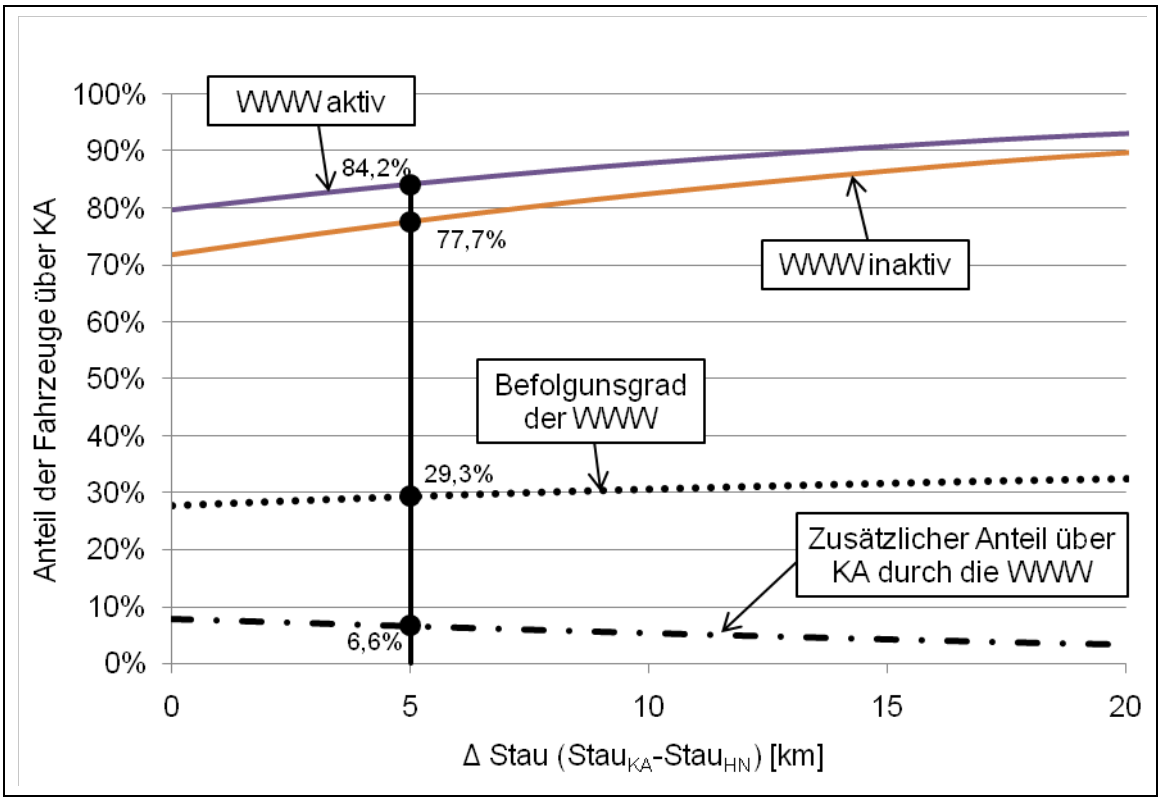


Abbildung 76: Grafische Darstellung des Basismodells für die Richtung Stuttgart → Walldorf für den Fall „mehr Stau über Karlsruhe“.

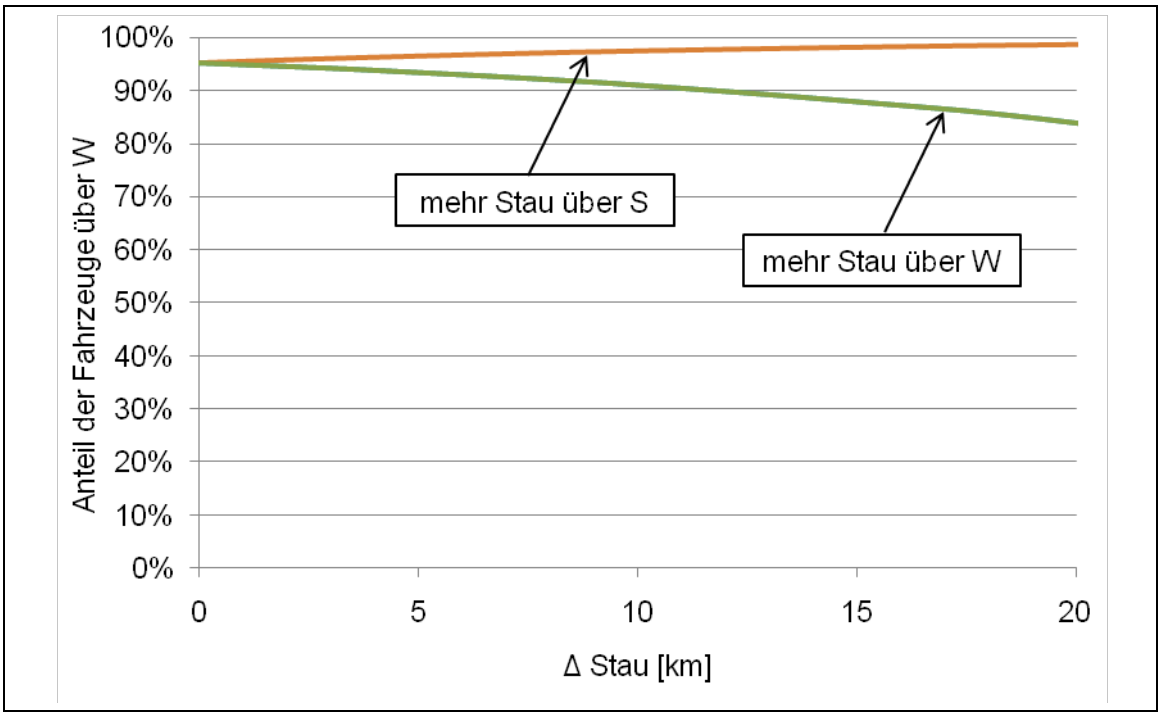


Abbildung 77: Grafische Darstellung des Basismodells für die Richtung Karlsruhe → Heilbronn.

Die folgende Tabelle 42 verdeutlicht, welchen Anteil die verschiedenen Terme zur Nutzenfunktion beitragen. Auffallend sind die hohen Anteile der dynamischen Einflussgrößen in der Richtung Walldorf → Stuttgart. Ein relativ hoher Anteil Fahrzeuge, die bei geschalteter Wechselwegweisung diese Richtung befahren, sowie die höchsten durchschnittlichen Meldungen über Stau und stockender Verkehr führen dazu, dass der Anteil der Konstanten in dieser Richtung nur bei 77 % liegt, während dieser Wert bei den anderen Richtung bei ca. 87 – 97 % liegt.

Richtung	Anzahl Beobach-tungen	Durchschnittswerte			Nutzenfunktionsbeiträge (Absolutwerte)			
		Anteil WWW	Stau [km]	Stockend [km]	Konstante	WWW	Stau	Stockend
S → W	330.891	1,3%	1,2	1,2	86,9%	0,5%	6,6%	5,9%
W → S	384.560	5,2%	3,2	1,1	77,0%	1,9%	16,3%	4,8%
KA → HN	168.793	--	3,0	1,1	92,3%	--	5,9%	1,8%
HN → KA	175.468	--	1,1	0,6	97,2%	--	1,9%	0,9%
Summe	**1.059.712**	**--**	**--**	**--**	**--**	**--**	**--**	**--**
Mittel (gewichtet)	**--**	**2,3%**	**2,2**	**1,1**	**85,9%**	**0,8%**	**9,2%**	**4,0%**

Tabelle 42: Anteile der verschiedenen Terme zur Nutzenfunktion.

Erst die längere Erhebungsdauer mit Mobilfunktrajektorien, deren Trennbarkeit in Pkw und Lkw sowie umfangreiche Erhebungen der typischen Fahrzeiten ermöglichen vertiefte Analysen mit statistisch gesicherten Ergebnissen. Die wesentlichen Ergebnisse lassen sich wie folgt zusammenfassen:

- Pkw und Lkw zeigen ein vergleichbares Routenwahlverhalten. Dabei tendieren Lkw etwas stärker zur kürzeren Route, während Pkw etwas stärker auf Verkehrsmeldungen und die Wechselwegweisung reagieren.

- Die Konstanten haben einen wesentlichen Einfluss auf die Routenwahl und lassen sich nur sehr eingeschränkt durch die typischen, zu erwartenden Fahrzeiten auf den Alternativen ersetzen.

- Die Verkehrsteilnehmer bewerten Verkehrsmeldungen, die sie früher auf ihrer Route erreichen, höher als Meldungen auf einem späteren Abschnitt ihrer Route. Bei einer Unterteilung in zwei Abschnitte der ca. 100 km langen Alternativen werden die Verkehrsmeldungen des ersten Abschnittes fast doppelt so hoch bewertet wie die des zweiten Abschnitts.

- Die Verkehrsteilnehmer, die der Routenempfehlung der Wechselwegweisung folgen, tun dies unabhängig von der Anzeige der Staulänge bzw. des Stauortes.

4.5.2 Netzbeeinflussungsanlage Stuttgart

Bei der Routenwahl im Bereich der Netzbeeinflussungsanlage Stuttgart sind die Ergebnisse auf Basis von FPD-Trajektorien weniger eindeutig. Während zwischen der Routenwahl und den Verkehrsmeldungen kein Zusammenhang festgestellt werden kann, ergibt sich bei den Schätzungen ein erklärbarer Zusammenhang zwischen Routenwahl und der Empfehlung der NBA. In Stadteinwärtsrichtung erhöht die Routenempfehlung der NBA den Anteil der Route über die B295. In Stadtauswärtsrichtung wirkt eine Empfehlung der NBA in kleinerem Umfang auf beide Alternativrouten B27 und B295.

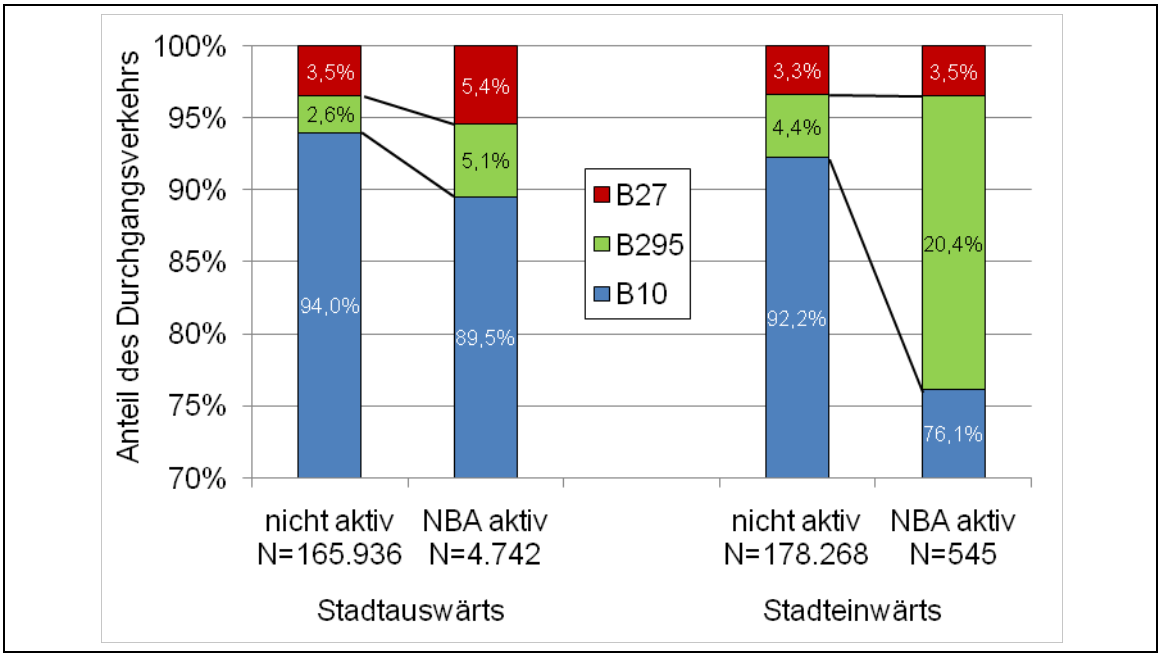


Abbildung 78: Routenwahlverhalten im Bereich der NBA Stuttgart in Abhängigkeit der NBA (aus Tabelle 40).

Abbildung 78 zeigt das Routenwahlverhalten in Abhängigkeit der Schaltzustände der NBA. Aus den Werten ergeben sich bei Vernachlässigung der B27 Befolgungsgrade von 2,7 % (stadtauswärts) bzw. 17,3 % (stadteinwärts).

Für statistisch gesicherte Analysen der Wirksamkeit der NBA Stuttgart ist ein längerer Untersuchungszeitraum mit regelmäßigen und häufigeren Schaltungen der NBA notwendig.

4.5.3 Beurteilung der Wirksamkeit von dynamischen Routenempfehlungen

Im Folgenden soll die Wirksamkeit der untersuchten Wechselwegweisung in der BAB-Netzmasche und der NBA Stuttgart beurteilt werden. Für die NBA Stuttgart werden dabei die für die zwei Richtungen getrennt ermittelten Befolgungsgrade verwendet, auch wenn diese nicht so statistisch gesichert sind wie die der Wechselwegweisung in der BAB-Netzmasche.

In der Literatur gibt es unterschiedliche Definitionen für den Befolgungsgrad (vgl. Everts, 1978; Kayser & Krause, 1986), was vor allem darauf zurück zu führen ist, dass mit der jeweiligen Erhebungsmethode nur bestimmte Kenngrößen ermittelt werden konnten. Die unterschiedlichen Kenngrößen erschweren die Vergleichbarkeit der Ergebnisse. Im Rahmen dieser Untersuchung wird der Befolgungsgrad wie in Kapitel 4.5.1 definiert, der den durch die aktive Wechselwegweisung umgeleiteten Verkehr ins

Verhältnis zum Durchgangsverkehr auf der Hauptroute bei inaktiver Wechselwegweisung setzt. Dabei ergeben sich die folgenden Befolgungsgrade:

- BAB-Netzmasche: ca. 30 %
- NBA Stuttgart (stadteinwärts): 17,3 %
- NBA Stuttgart (stadtauswärts): 2,7 %

In aktuellen Hinweisen der Forschungsgesellschaft für Straßen- und Verkehrswesen (FGSV, 2007, Kurzfassung in Trapp & Feldges, 2009) werden „Erfahrungswerte von Befolgungsraten“ angegeben (vgl. Tabelle 43), wobei durch zusätzliche Angaben auf dWiSta-Tafeln die jeweils höheren Werte der angegebenen Bandbreite erreicht werden sollen. Nach Trapp (2009) beziehen sich die Angaben auf die in dieser Arbeit verwendete Definition des Befolgungsgrades

Charakterisitk des Netzes	Maschengröße	Umwegfaktor	Befolgungsraten
kleine Masche, kurzer Umweg	< 50 km	< 1,5	30 – 40 %
große Masche, kurzer Umweg	> 50 km	< 1,5	20 – 30 %
kleine Masche, großer Umweg	< 50 km	> 1,5	10 – 20 %
große Masche, großer Umweg	> 50 km	> 1,5	0 – 10 %

Tabelle 43: Befolgungsraten bei NBA (FGSV, 2007, Tabelle 6).

Die Wechselwegweisung in der BAB-Netzmasche fällt in die Kategorie „große Masche, kurzer Umweg“, somit ist laut den FGSV-Hinweisen bei Anwendung von dWiSta-Tafeln ein Befolgungsgrad von 30 % zu erwarten. Dieser passt in den Ergebnissen der Erhebung mit FPD-Trajektorien, wobei bei einem Umwegfaktor von ca. 0,9 (die Alternativroute ist etwas kürzer) eventuell etwas höhere Werte zu erwarten wären.

Die NBA Stuttgart fällt mit einem Umwegfaktor von 1,7 in die Kategorie „kleine Masche, großer Umweg“, so dass nach Tabelle 43 bei Anwendung von dWiSta-Tafeln ein Befolgungsgrad von 20 % zu erwarten ist. Dieser Wert wird für die Stadteinwärtsrichtung mit 17,3 % in etwa erreicht, wobei dieses Ergebnis auf nur 545 Datensätzen während aktiver Routenempfehlung basiert. In der Stadtauswärtsrichtung mit fast 10fach höherer Stichprobe während aktiver Routenempfehlung (vgl. Tabelle 40) liegt der beobachtete Befolgungsgrad mit 2,7 % dagegen deutlich unter diesen Werten.

Allerdings sind Befolgungsgrade nicht die maßgebende Größe zur Beurteilung der Wirksamkeit einer Wechselwegweisung. Vielmehr ist es entscheidend, um wie viel Kfz die Hauptroute entlastet werden kann. Dafür sind neben dem Befolgungsgrad noch zwei weitere Kenngrößen entscheidend:

- *Anteil Durchgangsverkehr am Gesamtverkehr* (*$Anteil_{Durchgangsverkehr}$*)*:*
  Von dem Verkehr, der eine Wechselwegweisung passiert, ist häufig nur ein geringer Teil Durchgangsverkehr durch die untersuchte Netzmasche. Die anderen Fahrzeuge haben andere Ziele und sind daher nicht durch die Wechselwegweisung beeinflussbar.
- *Anteil der Hauptroute am Durchgangsverkehr bei deaktivierter Wechselwegweisung* (*$Anteil_{Hauptroute}$*):
  Durchgangsverkehr kann nur dann durch eine Empfehlung der Wechselwegweisung auf eine Alternativroute umgeleitet werden, wenn im Fall einer inaktiven Wechselwegweisung die Hauptroute der Wechselwegweisung befährt.

Für die Wechselwegweisung in der BAB-Netzmasche ergibt sich ein Anteil umgelenkter Fahrzeuge bezogen auf den zufließenden Verkehr (*$Anteil_{umgelenkt,Zufluss}$*) von ca. 1,1 %:

$$Anteil_{umgelenkt,Zufluss} = Anteil_{Durchgangsverkehr} \cdot Anteil_{Hauptroute} \cdot Befolgungsgrad_{WWW}$$
$$= 12\ \% \cdot 30\ \% \cdot 30\ \% = 1{,}08\ \%$$

Dieser Wert stellt allerdings eher eine Obergrenze dar, da er für den Fall ohne Staumeldungsdifferenz gilt. Je größer die Staumeldungen über Heilbronn, von der bei einer aktiven Wechselwegweisung auszugehen sind, desto geringer wird der Anteil des Durchgangsverkehrs auf der Hauptroute (*$Anteil_{Hauptroute}$*) und damit auch der Wert von *$Anteil_{umgelenkt,Zufluss}$*. Somit sind die in der folgenden Tabelle 44 angegebene Werte für die Anzahl umgelenkter Fahrzeuge für verschiedene Belastungen in der Zufahrt als Obergrenze zu verstehen.

Gesamtverkehr Zufahrtsquerschnitt [Kfz/h]	Durchgangsverkehr [Kfz/h]	Durchgangsverkehr auf Hauptroute ohne Wechselwegweisung [Kfz/h]	Anzahl umgelenkter Fahrzeuge [Kfz/h]
1.000	120	36	11
2.000	240	72	22
4.000	480	144	43
6.000	720	216	65

Tabelle 44: Durch Schaltung der Wechselwegweisung umlenkbarer Verkehr (Diagonale Stuttgart ↔ Walldorf, ohne Staumeldungen).

In der anderen Diagonale, die nicht mit einer Wechselweisung ausgestattet ist, ist das Potential unter der Annahme eines identischen Befolgungsgrades ähnlich hoch. Zwar wird dort die Hauptroute vom Durchgangsverkehr deutlich mehr genutzt (ca. 95 %), dagegen ist aber der aus FPD-Trajektorien ermittelte Durchgangsverkehrsanteil nur ca. 4 %.

Gesamtverkehr Zufahrtsquerschnitt [Kfz/h]	Durchgangsverkehr [Kfz/h]	Durchgangsverkehr auf Hauptroute ohne Wechselwegweisung [Kfz/h]	Anzahl umgelenkter Fahrzeuge [Kfz/h]
1.000	40	38	11
2.000	80	76	23
4.000	160	152	46
6.000	240	228	68

Tabelle 45: Durch Schaltung der Wechselwegweisung umlenkbarer Verkehr (Diagonale Karlsruhe ↔ Heilbronn, ohne Staumeldungen).

4.5.4 Beurteilung von FPD-Trajektorien zur Analyse der Routenwahl

Die FPD-Trajektorien auf den LAC-basierten Trajektoriengenerierungsverfahren eignen sich sehr gut für die Analyse der Routenwahl, wobei für den Bereich der NBA Stuttgart manuelle Erweiterungen notwendig sind. Die Routenwahlanteile, die sich aus den FPD-Trajektorien ergeben, liegen im Bereich von vergleichbaren Messungen mit automatischen Kennzeichenerfassungssystemen.

Das tatsächlich realisierte Routenwahlverhalten bietet für die Analyse von ausgewählten Untersuchungsgebieten realistischere Ergebnisse als solche, die aus Befragungen und Modellversuchen in diesem oder in vergleichbaren Gebieten erreicht werden können. Somit können z.B. die Wirkungen der Wechselwegweisung sehr genau analysiert werden. Dies setzt allerdings voraus, dass die dynamischen Einflussgrößen wie z.B. Verkehrsmeldungen und Wechselwegweisungen hinreichend oft geschaltet sind. Dies war im Untersuchungszeitraum bei der NBA Stuttgart nicht der Fall, so dass hier keine gesicherten Aussagen über die Wirksamkeit der NBA gemacht werden können.

Die angewendeten Methoden sind sowohl im Bereich der Trajektoriengenerierung als auch bei den Routenwahlanalysen auf andere und größere Untersuchungsgebiete übertragbar. So würde z.B. eine deutschlandweite Untersuchung die Datenbasis insbesondere für die Analyse von Wechselwegweisungen deutlich erhöhen und somit differenzierte Analysen ermöglichen, die u.a. auch einen Zusammenhang zwischen Umwegfaktor zwischen Alternativ- und Hauptroute und dem Befolgungsgrad aufzeigen könnten.

Deutschlandweite FPD-Trajektorien aus dem LAC-basierten Trajektoriengenerierungsverfahren würden es darüberhinaus ermöglichen, die Routenwahlparameter von klassischen Umlegungsmodellen zu überprüfen und ggf. anzupassen. Diese basieren im Wesentlichen auf der Fahrzeit im belasteten Netz (siehe dazu auch Kapitel 6.2) und ein an gemessenen Verkehrsstärken geeichtes Aufteilungsmodell wie das Logit-Modell.

5 Generierung von Quelle-Ziel-Matrizen (AP6)

In diesem Kapitel werden Methodik und Ergebnisse zur Erzeugung von typischen Quelle-Ziel-Matrizen aus Mobilfunktrajektorien beschrieben. Die Methodik gliedert sich dabei zunächst in die folgenden beiden parallel laufenden Prozesse:

- Erzeugen von Clustern typischer Tage (z.B. Mo-Do, So/Feiertag, usw.) aus den Ähnlichkeiten der Ganglinien an Richtungsmessquerschnitten (RMQ).
- Tagesweises Hochrechnen der vorab erzeugten Mobilfunktrajektorien an den Zähldaten der jeweiligen RMQs.
- Zusammenfassen der hochgerechneten Mobilfunktrajektorien eines Clusters zu einem typischen Tag.

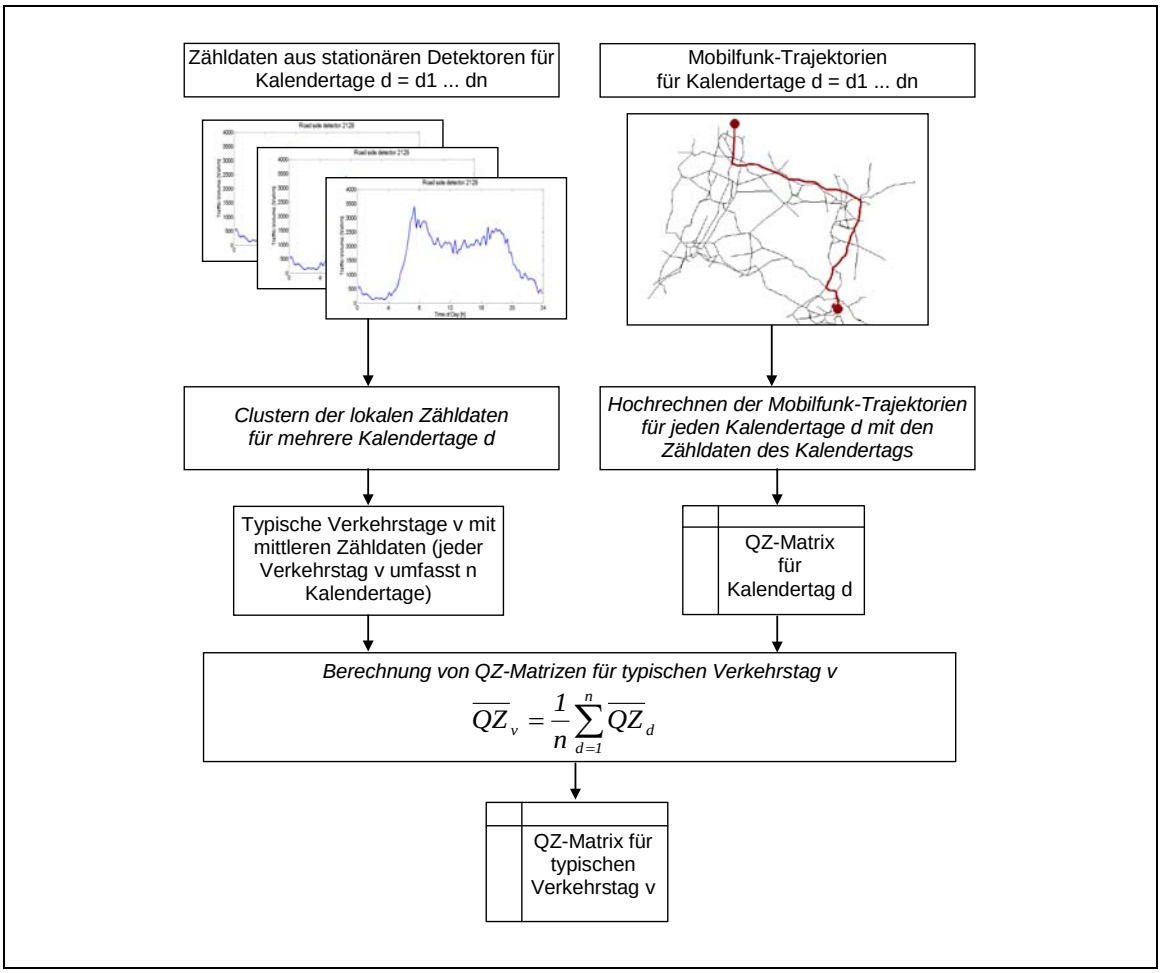


Abbildung 79: Schema für die Generierung von typischen QZ-Matrizen.

Die verwendeten Clusterverfahren sowie die Hochrechnung werden in den Kapiteln 5.1 und 5.2 genauer erläutert.

5.1 Durchführung von Clusterverfahren

Das Ziel einer Clusterung ist es, mehrere Objekte (hier Netzganglinien) in Cluster aufzuteilen, so dass in jedem Cluster ähnliche Objekte zusammengefasst werden. Das Ergebnis der Clusterung ist eine Gruppierung in verkehrstagtypische Tage mit ähnlichen Netzganglinien, sowie verkehrstagtypische Ganglinien, die aus den gemittelten Verkehrsstärken und Geschwindigkeiten berechnet werden. Die hier verwendete Clusterung besteht aus 3 grundlegenden Bestandteilen (Vortisch 2005).

- Vorklassifizierung: Aufteilung der Objekte entsprechend ihrem TagTyp in Gruppen
- Abstandsfunktion: Bestimmung der Ähnlichkeit zwischen zwei Objekten
- Clusterung: Aufteilung aller Objekte je Gruppe entsprechend ihrer Abstandswerte zueinander.

In diesem Kapitel werden die drei Bestandteile der Clusterung separat erläutert. Eingangsdaten

5.1.1 Vorklassifizierung

Unter einer Vorklassifizierung versteht man die Zuordnung aller Kalendertage eines Jahres zu Tagtypen, von denen man annehmen kann, dass sie sich im Bezug auf die Verkehrsnachfrage ähnlich verhalten. Die Zuordnung eines Kalendertages erfolgt im Wesentlichen aus Eigenschaften, die dem Kalender zu entnehmen sind. Eigenschaften eines Kalendertags sind:

- Wochentag (Montag, Dienstag, …)
- Feiertag (ja/nein)
- Ferientag (ja/nein)
- Brückentag (ja/nein)

Zusätzlich können Eigenschaften verwendet werden, die sich nicht aus einem Kalender ergeben:

- Großveranstaltung (Messe/Sport),
- Wetter,
- Höhere Gewalt (Bahnstreik, längere Streckensperrungen),
- Baustellen oder Verkehrsmeldungen.

In den Abbildungen 80 und 81 sind beispielhaft zwei Vorklassifizierungen dargestellt.

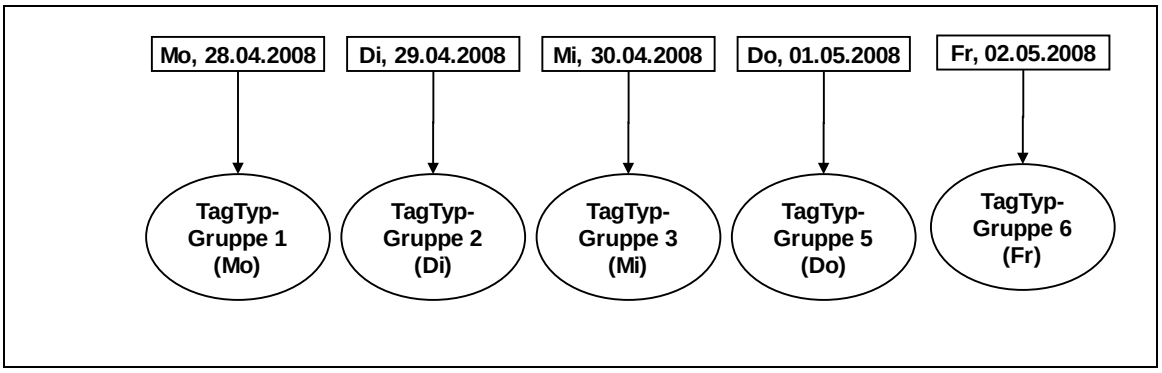


Abbildung 80: Ergebnis einer beispielhaften Vorlassifizierung von 5 Tagen nach dem Attribut „Wochentag“.

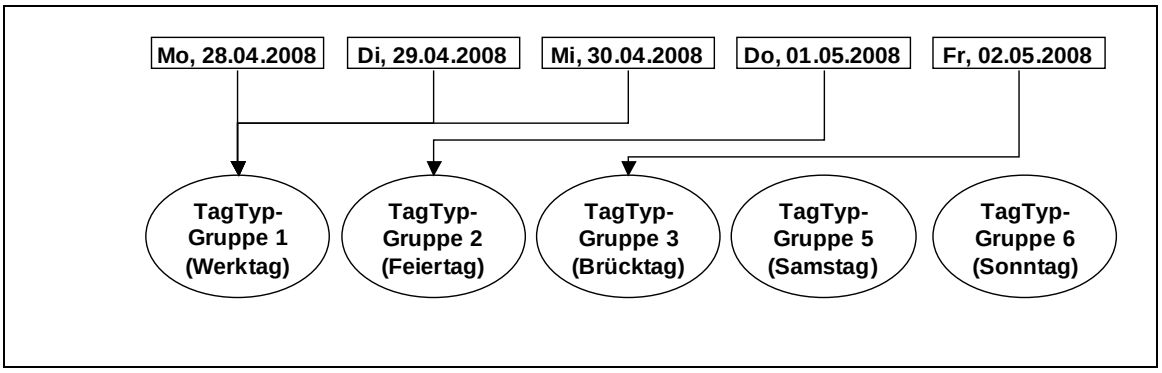


Abbildung 81: Ergebnis einer beispielhaften Vorlassifizierung von 5 Tagen nach den Attributen „Werktag“, „Feiertag“, „Brückentag“, „Samstag“ und „Sonntag“.

Bei einer ungünstigen Vorklassifizierung mit sehr vielen Attributen werden TagTyp-Gruppen erzeugt, denen im Untersuchungszeitraum wenige oder gar keine Kalendertage zugeordnet sind. Eine Clusterung dieser Gruppen würde nicht mehr zuverlässig ähnliche Tage identifizieren, wenn sich ähnliche Tage bereits durch die sehr feine Vorklassifizierung in unterschiedlichen TagTyp-Gruppen befinden.

5.1.2 Abstandsfunktion

Die Wahl der Abstandsfunktion hat auf das Ergebnis der Clusterung einen entscheidenden Einfluss, da der Funktionswert der Abstandsfunktion die Größe ist, nach der die Clusterung durchgeführt wird. In diesem Kapitel werden die Abstandsfunktionen an beispielhaft ausgesuchten RMQ-Ganglinien betrachtet, um die Unterschiede näher zu erläutern.

In den Abbildungen 82 und 83 sind jeweils 3 Ganglinien dargestellt, mit denen der Einfluss folgender Abstandsberechnungen beispielhaft dargestellt wird.

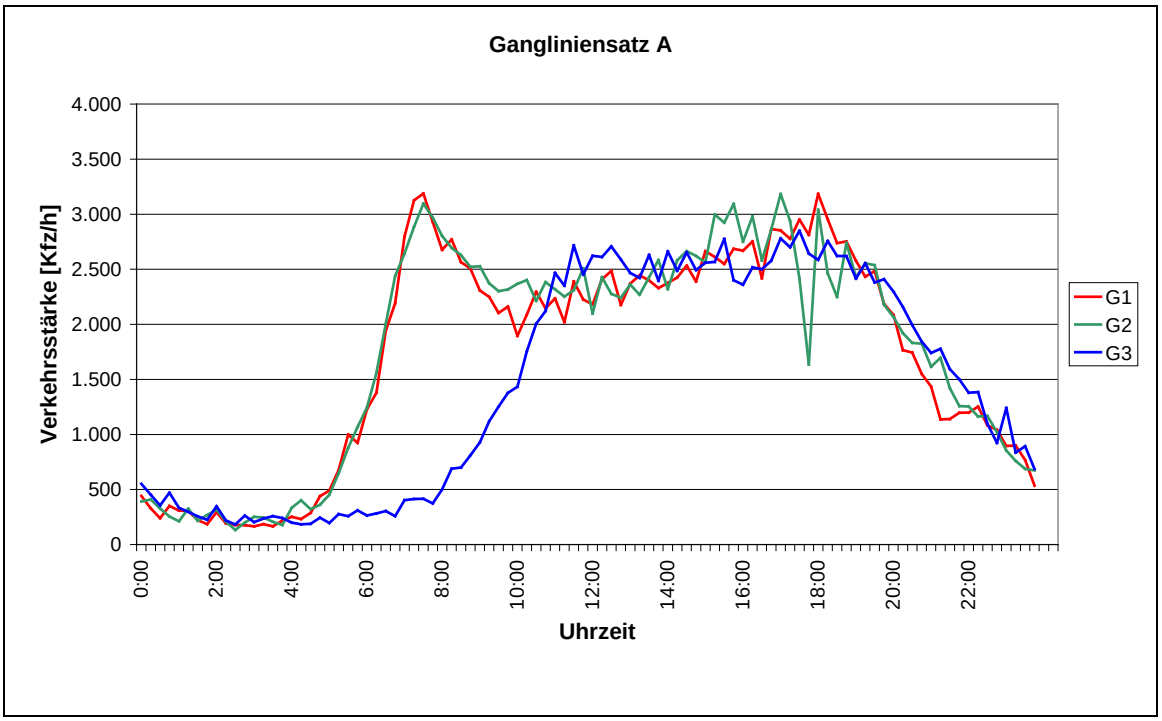


Abbildung 82: Gangliniensatz A aus 3 RMQ-Ganglinien

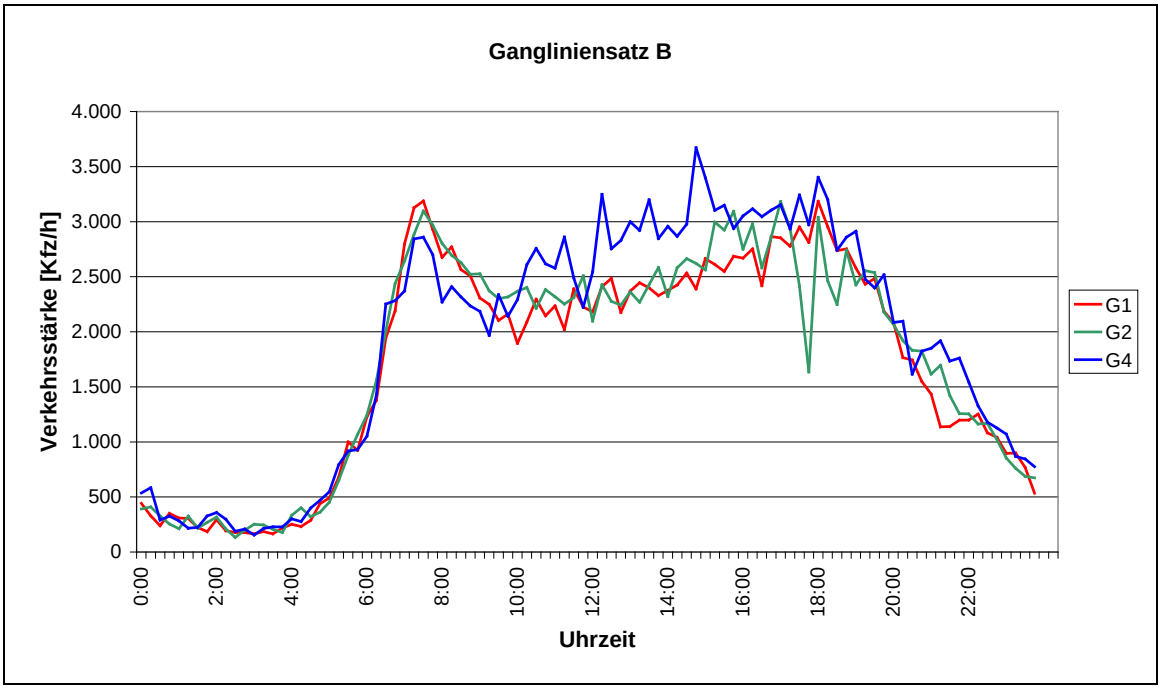


Abbildung 83: Gangliniensatz B aus 3 RMQ-Ganglinien

5.1.2.1 Summe der Abstandsquadrate (SAQ)

Die SAQ-Funktion berechnet für jedes Intervall zweier RMQ-Ganglinien das Quadrat der Differenz und summiert diese zu einem Gesamtwert.

$$D_{SAQ}(G(j,d),G(k,d)) = \sum_{Intervall\ T=1}^{96} \left(G(j,d).Q_{Kfz}(T) - G(k,d).Q_{Kfz}(T)\right)^2$$

mit

$G(j,d)$: Ganglinie am RMQ j am Tag d.

$G(j,d).Q_{Kfz}(T)$: Verkehrsstärke der Ganglinie im Zeitintervall T.

Dabei wird eine große Abweichung der Werte, wie in Abbildung 84 zu sehen ist, mit einem besonders hohen Abstand bewertet, geringe Differenzen fallen kaum ins Gewicht.

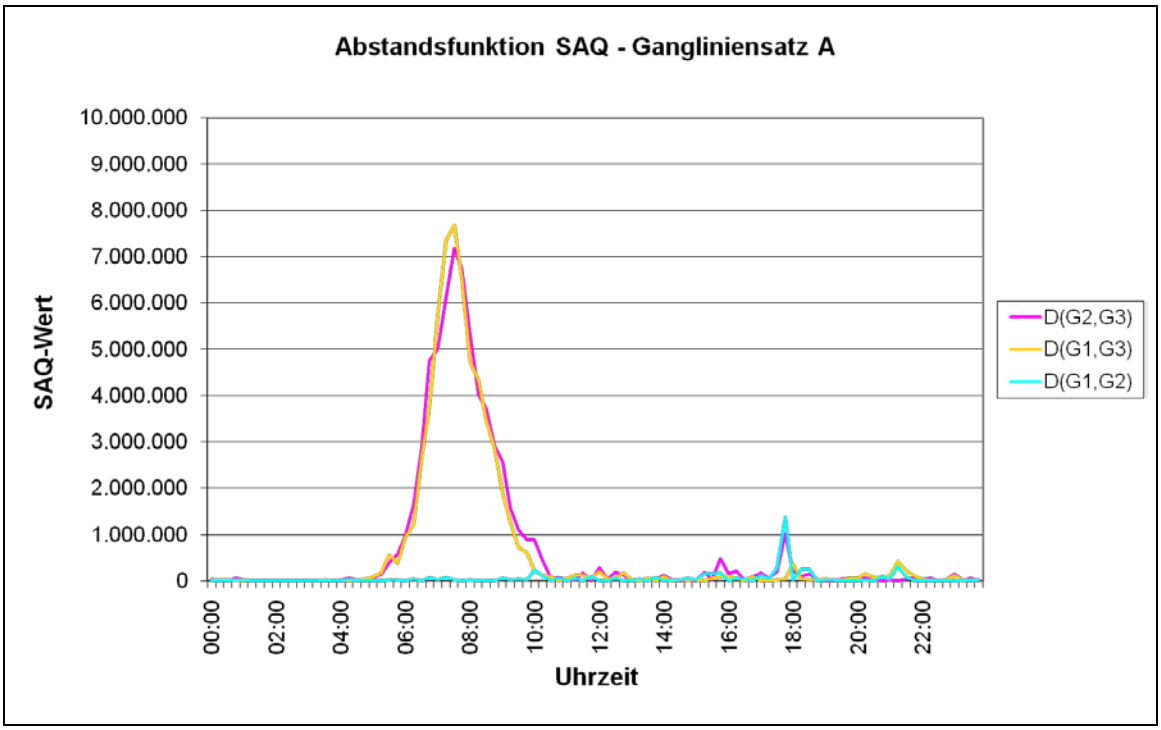


Abbildung 84: Gangliniensatz A – Werte der Teilergebnisse für die Summe der Differenzquadrate.

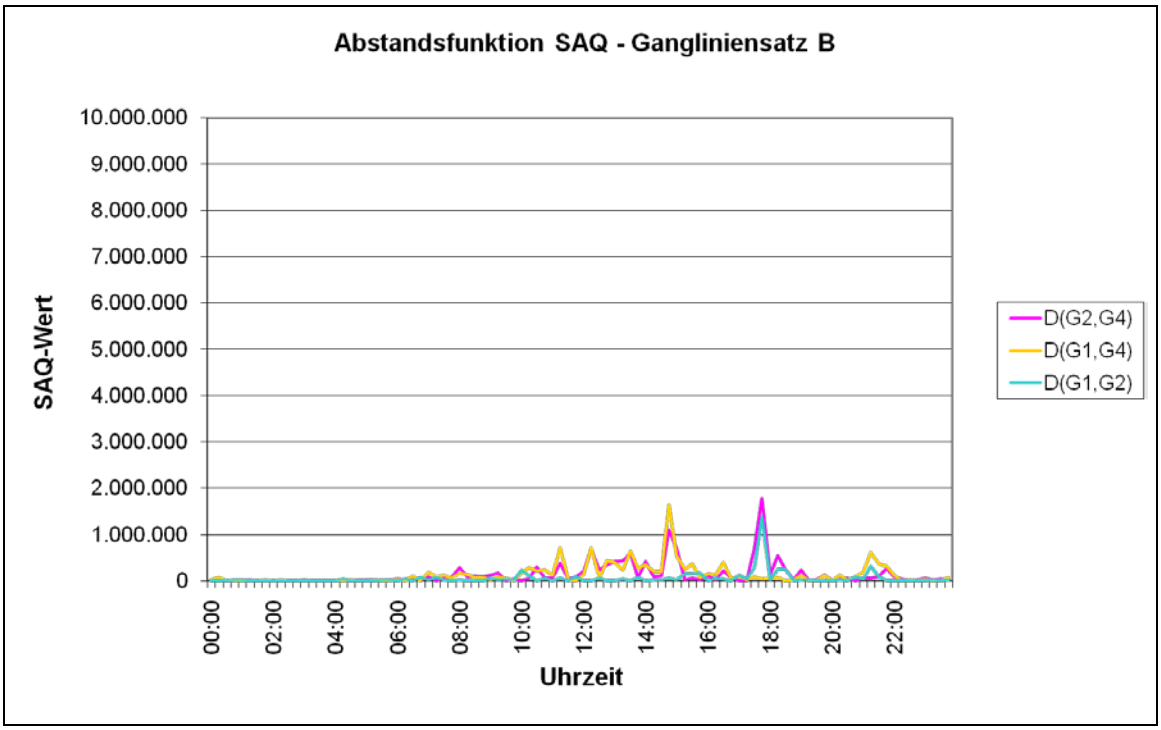


Abbildung 85: Gangliniensatz B – Werte der Teilergebnisse für die Summe der Differenzquadrate.

Aus Tabelle 46 wird ersichtlich, dass in Gangliniensatz A G3 jeweils einen größeren Abstandswert zu G1 und G2, als diese untereinander. Beim Gangliniensatz B kann man erkennen, dass die RMQ-Ganglinie G4 von den übrigen 2 einen größeren Abstand hat.

Gangliniensatz A	Gemittelter SAQ-Wert	Gangliniensatz B	Gemittelter OLS-Wert
D_{SAQ} (G1,G2)	49.636	D_{SAQ} (G1,G2)	49.636
D_{SAQ}(G2,G3)	675.085	D_{SAQ}(G2,G4)	132.005
D_{SAQ} (G1,G3)	633.545	D_{SAQ} (G1,G4)	131.764

Tabelle 46: D_{SAQ}-Werte der Gangliniensätze A und B.

Der besonders große Abstand zwischen G3 und den übrigen RMQ-Ganglinien entsteht durch die bei G3 nicht vorhandene Morgenspitze zwischen 6 und 9 Uhr. Das Quadrat dieser Abweichung hat eine viel größere Änderung des Abstandswertes verursacht als eine geringe und länger andauernde Differenz der Werte, wie die geringeren Morgenwerte und höheren Mittagswerte von G4 bezüglich G1 und G2. Die Maximale Abweichung ist daher der Hauptbestandteil zur Bestimmung des Abstandswertes. Weil die SAQ-Funktion nur die absoluten Differenzen betrachtet, hat sie außerdem die Eigenschaft, dass RMQ-Ganglinien mit hohen Verkehrsströmen bei einer Abweichung von 10% einen höheren Abstandswert bekommen, als RMQ-Ganglinien mit niedrigen Verkehrsströmen bei einer Abweichung von 10%.

5.1.2.2 Prozentuelle Abstandsfunktion

Für prozentuelle Abstandsfunktionen wird jeweils der Abstand von 2 Werten als prozentuelle Abweichung bezüglich deren Mittelwert berechnet.

$$D_{Percent}\left(G(j,d),G(k,d)\right)=\frac{\sum_{Intervalle\ T}\left(\frac{200\cdot\left|G(j,d)\cdot Q_{Kfz}(T)-G(j,d)\cdot Q_{Kfz}(T)\right|}{\left(G(j,d)\cdot Q_{Kfz}(T)+G(j,d)\cdot Q_{Kfz}(T)\right)}\right)}{n_{Intervalle}}$$

mit

$G(j,d)$: Ganglinie am RMQ j am Tag d.

$G(j,d).Q_{Kfz}(T)$: Verkehrsstärke der Ganglinie im Zeitintervall T.

Diese Funktion hat gegenüber der SAQ-Funktion den Vorteil, dass diese auf die Werte zwischen 0 und 100 skaliert sind und damit sowohl der beste Wert als der auch schlechteste Wert definiert ist. Dies ermöglicht z.B. eine Vorgabe bei der Clusterung, dass in einem Cluster nur Tagesganglinien zusammen sein dürfen, wenn deren Abstandswert kleiner als ein bestimmter Schwellenwert ist (siehe Kapitel 5.1.3.2). In den Abbildungen 86 und 87 sind die Abstandswerte der Prozent-Funktion für die Ganglinensätze A und B grafisch dargestellt.

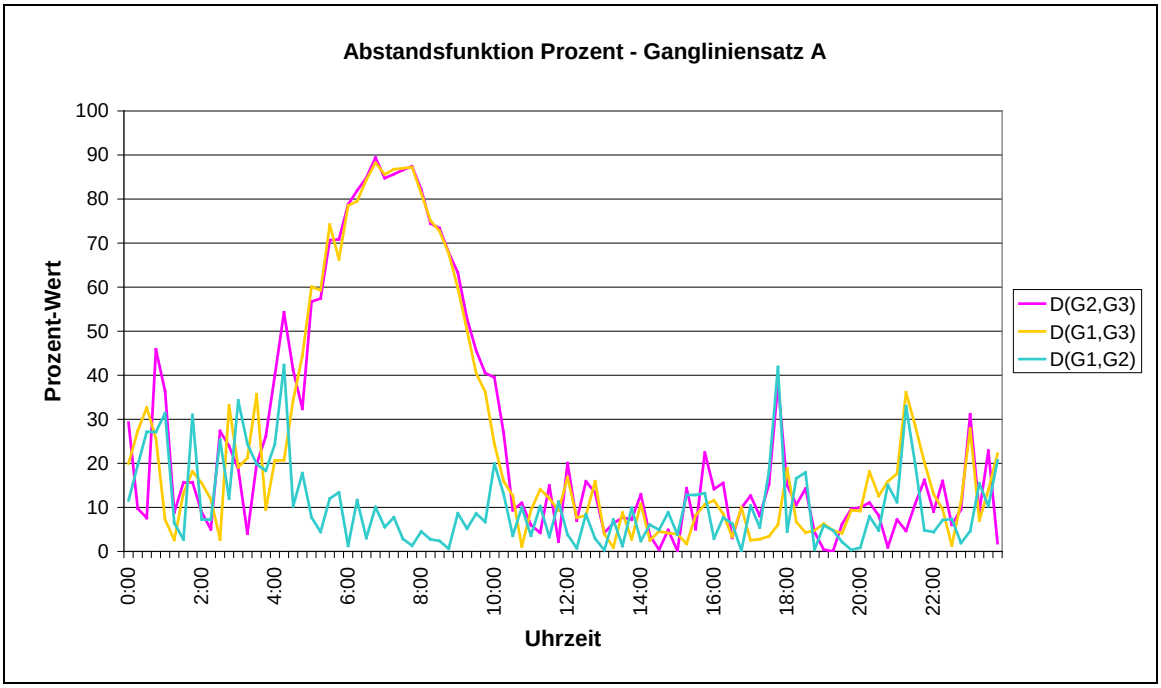


Abbildung 86: Gangliniensatz A – Werte der Teilergebnisse für den Prozent-Wert.

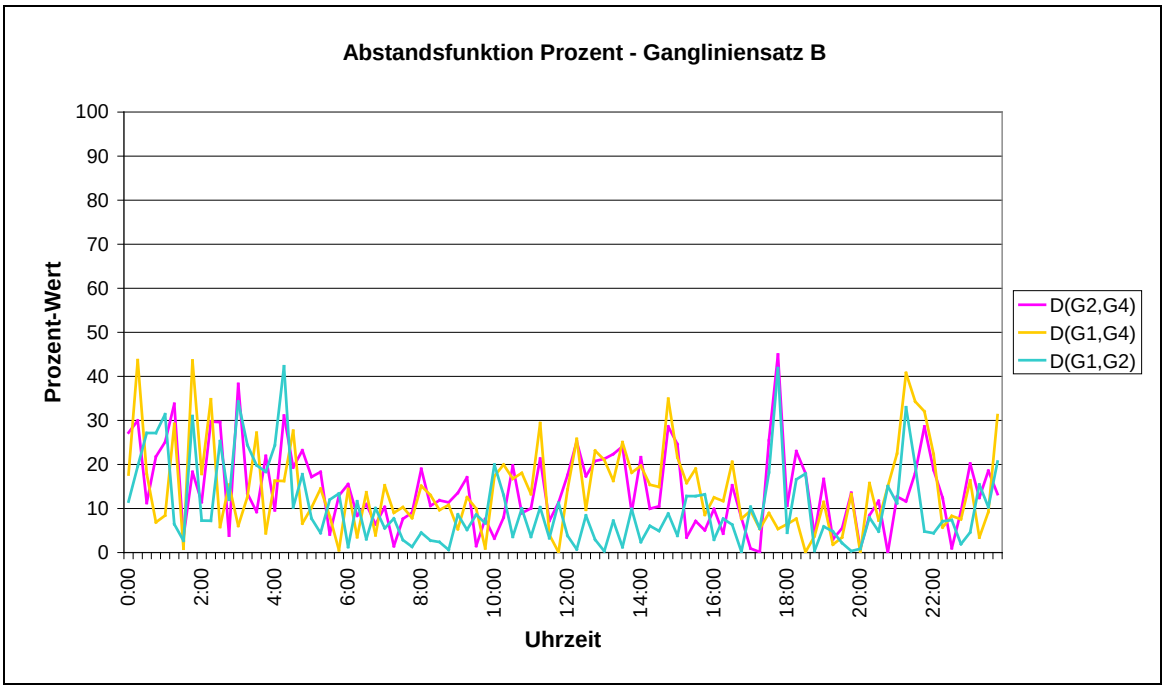


Abbildung 87: Gangliniensatz B – Werte der Teilergebnisse für den Prozent-Wert.

Auch hier sind wie bei der SAQ-Funktion in beiden Gangliniensätzen die Abstände zwischen den RMQ-Ganglinien G1 und G2 geringer, als deren Abstände zu G3 bzw. G4. Ebenso ist, wie bei der SAQ-Funktion, der Abstand von G1 geringer als der Abstand von G2 zu den RMQ-Ganglinien G3 bzw. G4.

Gangliniensatz A	Gemittelter %-Wert	Gangliniensatz B	Gemittelter %-Wert
$D_{\%}$ (G1,G2)	10,52	$D_{\%}$ (G1,G2)	10,52
$D_{\%}$ (G2,G3)	26,22	$D_{\%}$ (G2,G4)	14,44
$D_{\%}$ (G1,G3)	25,19	$D_{\%}$ (G1,G4)	14,29

Tabelle 47: $D_{\%}$-Werte der Gangliniensätze A und B.

Nachteilig an dieser Abstandsfunktion ist, dass analog zum Beispiel in der SAQ-Funktion hier eine Abweichung von 10% sowohl bei sehr geringen Verkehrsstärken, als auch bei sehr hohen Verkehrsstärken mit dem Funktionswert „10“ bewertet wird. Dabei sollte eine Abweichung von 300 Kfz/h bei einer Verkehrsstärke von 3.000 Kfz/h schlechter bewertet werden, als eine Abweichung von 5 Kfz/h bei einer Verkehrsstärke von 50 Kfz/h.

5.1.2.3 GEH-Funktion

Die Berechnung des GEH-Wertes stammt von Geoffrey E. Havers aus dem Jahr 1970, der die Funktion für den Vergleich von Verkehrsstärken entwickelt hat (siehe auch UK Highway Agency, 2008).

$$D_{GEH}\left(G(j,d),G(k,d)\right)=\frac{\sum_{Intervalle\ T}\sqrt{\frac{2\left(G(j,d).Q_{Kfz}(T)-G(j,d).Q_{Kfz}(T)\right)^2}{G(j,d).Q_{Kfz}(T)+G(j,d).Q_{Kfz}(T)}}}{n_{Intervalle}}$$

mit

$G(j,d)$: Ganglinie am RMQ j am Tag d.

$G(j,d).Q_{Kfz}(T)$: Verkehrsstärke der Ganglinie im Zeitintervall T.

Die Funktion ist selbst-skalierend und verringert daher den Effekt, der besonders hohen Funktionswerte bei Messwertdifferenzen mit hohen Verkehrsstärken, wie er bei der SAQ-Funktion aufgetreten ist. In der Berechnung der Abweichung wird jedoch die absolute Abweichung mit berücksichtigt, so dass im Vergleich zur Prozent-Funktion eine prozentual identische Abweichung bei hohen Verkehrsstärken einen höheren Funktionswert der GEH-Funktion bekommt, als bei niedrigen Verkehrsstärken.

Wie in den Abbildungen 88 und 89 dargestellt, ist das Ergebnis der GEH-Funktion ähnlich der Prozent-Funktion in Kapitel 5.1.2.2.

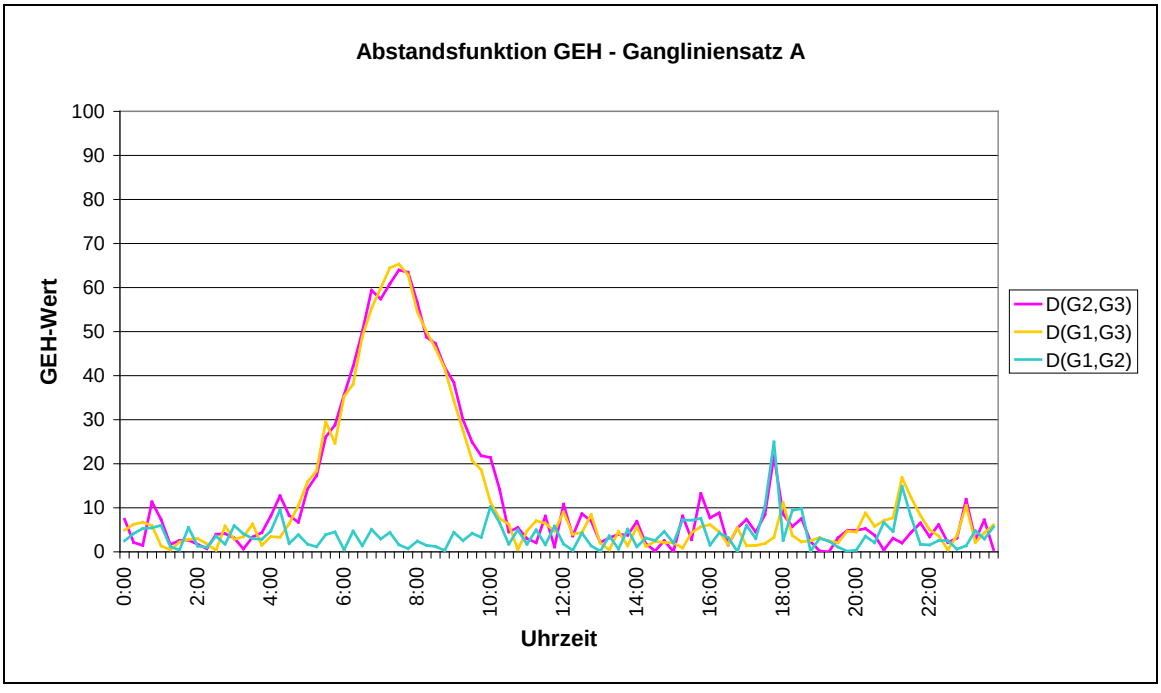


Abbildung 88: Gangliniensatz A – Werte der Teilergebnisse für den GEH-Wert.

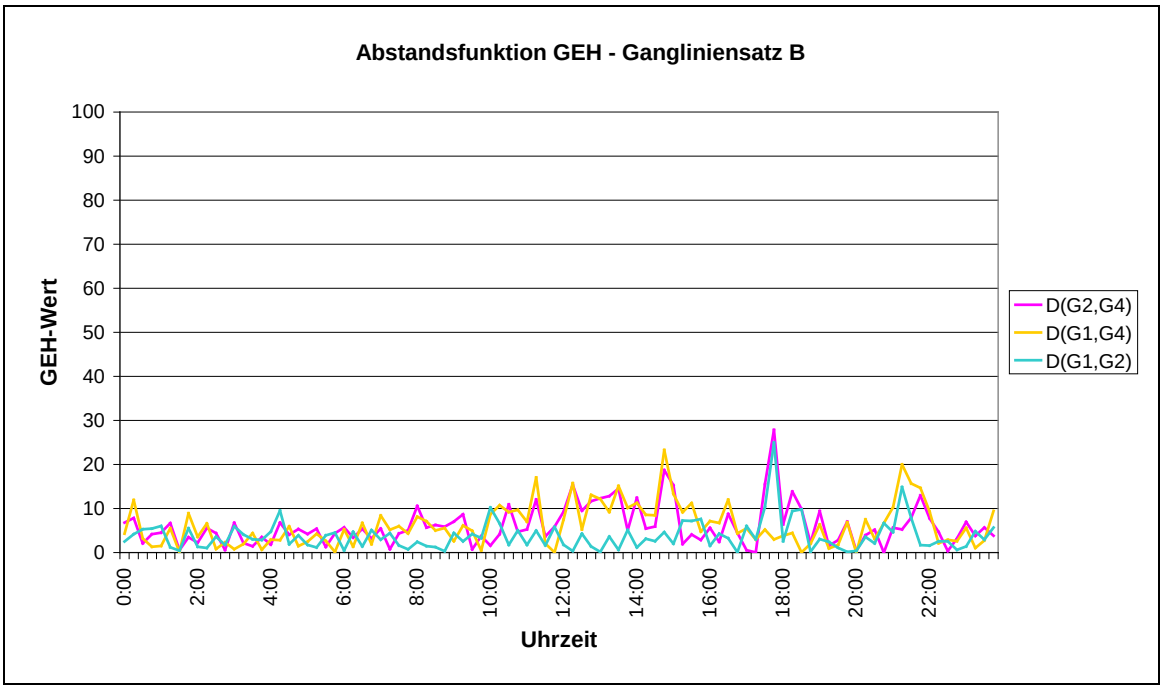


Abbildung 89: Gangliniensatz B – Werte der Teilergebnisse für den GEH-Wert.

Die Werte in Tabelle 56 zeigen die Ähnlichkeit der RMQ-Ganglinien G1, G2 und G3 des Gangliniensatzes A sowie der RMQ-Ganglinien G1, G2 und G4 des Gangliniensatzes B, sowie den erhöhten Abstand der übrigen Tagespaare.

Gangliniensatz A	Gemittelter GEH-Wert	Gangliniensatz B	Gemittelter GEH-Wert
D_{GEH} (G1,G2)	3,79	D_{GEH} (G1,G2)	3,79
D_{GEH} (G2,G3)	12,79	D_{GEH} (G2,G4)	6,03
D_{GEH} (G1,G3)	12,13	D_{GEH} (G2,G4)	6,13

Tabelle 48: D_{GEH}-Werte der Gangliniensätze A und B.

Durch die Skalierung der GEH-Funktion ist es möglich anhand des Funktionswertes eine Aussage über die Ähnlichkeit zu treffen. Die Ähnlichkeit zweier RMQ-Ganglinien ist dann vorhanden, wenn der gemittelte GEH-Wert unter 5 liegt. Bei einem mittleren GEH-Wert zweier Ganglinien von über 10, wie z.B. zwischen den Ganglinien G2 und G3 kann von zwei unähnlichen RMQ-Ganglinien ausgegangen werden.

5.1.2.4 Bewertung

Für die Clusterung von Ganglinien zeigt sich, dass alle untersuchten Funktionen für eine Klassifizierung geeignet sind. Jedesmal ist der Abstand zwischen den Ganglinien G1 und G2 geringer als zwischen G1 bzw. G2 und G3 bzw. G4. Die hohen Werte der SAQ-Funktion bei größeren Messwertabweichungen sorgen jedoch dafür, dass durch diese extremen Werte der gesamte Abstandswert zweier Ganglinien benachteiligt ist,

sobald es in einem einzigen Intervall zu Abweichungen z.B. durch Messfehler kommt. Für die Clusterung von RMQ-Ganglinien und Netzganglinien wird die GEH-Funktion verwendet, da sie eine gute Skalierung hat und auch für Abbruchkriterien eingesetzt werden kann.

5.1.3 Clusterung von Netzganglinien

Unter Clusterung von Netzganglinien versteht man deren Aufteilung in Cluster (Gruppen) mit ähnlichen Eigenschaften und die Erzeugung einer verkehrstagtypischen Netzganglinie aus allen Ganglinien eines Clusters. Jedes Cluster wird durch eine Referenz auf alle seine Netzganglinien und die verkehrstagtypische Ganglinie bestimmt.

Unabhängig von der Vorklassifizierung und der Wahl der Abstandsfunktion gibt es für eine Clusterung von Zähldaten im Untersuchungsgebiet zwei Herangehensweisen. Die separate Clusterung von Ganglinien an jedem einzelnen RMQ und die gemeinsame Clusterung von Tagen.

Clusterung von RMQ-Ganglinien

Bei der Clusterung von RMQ-Ganglinien wird jeder RMQ für sich geclustert, so dass aus jedem RMQ unabhängig von den anderen RMQs mehrere Cluster entstehen. Dadurch kann ein bestimmter Tag für verschiedene RMQ auch in verschiedenen Clustern auftreten, d.h. dass für einen Tag kein typisches Cluster gewählt werden kann, welches bei allen RMQ für den Tag repräsentativ wäre.

Clusterung von Netzganglinien

Für die Berechnung des Abstandswertes zwischen zwei Netzganglinien werden die Abstandswerte der RMQ-Ganglinien berechnet, die an beiden Tagen vollständig mit Messdaten versorgt sind. RMQ-Ganglinien von RMQs, die nur an einem der zwei Tage versorgt sind, werden nicht berücksichtigt. Der Abstandswert der zwei Netzganglinien ist der Mittelwert der Abstandswerte der entsprechenden RMQ-Ganglinien.

$$D_{Typ}\left(G\left(\vec{l}_1, d_1\right), G\left(\vec{l}_2, d_2\right)\right) = \frac{1}{n_{l_1=l_2}} * \sum_{l_1=l_2} D_{Typ}\left(G\left(l, d_1\right), G\left(l, d_2\right)\right)$$

mit

$G(l, d_1)$: RMQ-Ganglinie am RMQ l am Tag d_1.

$G\left(\vec{l}_1, d_1\right)$: Netzganglinie an den RMQs $\vec{l}$ am Tag d_1

$\vec{l}$: Eine Menge von RMQs

Als Clusterganglinie wird die Netzganglinie der gemittelten RMQ-Ganglinien bezeichnet. Die gemittelte RMQ-Ganglinie entsteht dabei aus den RMQ-Ganglinien eines RMQ an allen Tagen, die dem Cluster zugeordnet sind. Wurde ein RMQ an keinem einzigen Tag des Tagesclusters mit einer Ganglinie versorgt, wird auch keine gemittelte RMQ-Ganglinie für diesen RMQ bestimmt.

Im weiteren Verlauf dieses Kapitels werden die unterschiedlichen Clusterverfahren anhand von RMQ-Ganglinien gezeigt. Bei einer Clusterung von Netzganglinien wird analog verfahren, wobei die Berechnung der Abstandswerte zweier RMQ-Ganglinien und der gemittelten RMQ-Ganglinien durch die entsprechenden Verfahren für Netzganglinien ersetzt werden.

5.1.3.1 Clusterung nach k-Means

Durchführung

Die k-Means Methode zur Clusterung wurde 1967 von MacQueen entwickelt und ist ein einfacher, iterativer Prozess, bei dem nach einer Initialbelegung von k Clustern durch k zufällig ausgewählte RMQ-Ganglinien reihum jede weitere RMQ-Ganglinie betrachtet und einem Cluster zugeordnet wird, zu dem sie den geringsten Abstandswert hat. Der Ablauf dieses Verfahrens sieht wie folgt aus:

1. Bestimme k RMQ-Ganglinien die k Clustern zugeordnet werden und für die Methode als Initialcluster dienen.
2. Bestimme für jede RMQ-Ganglinie und jedes Cluster den Abstandswert.
3. Bestimme für jede RMQ-Ganglinie das Cluster mit dem geringsten Abstandswert.
4. Setze jede RMQ-Ganglinie zu seinem neuen Cluster (mit dem geringsten Abstandswert).
5. Berechne für jedes Cluster die neue mittlere RMQ-Ganglinie.
6. Wiederhole ab Schritt 2 solange keine Abbruchbedingung erreicht ist.

Abbruchbedingung

Die Iteration ist beendet, sobald nach einem Durchlauf keine RMQ-Ganglinie neu zugeordnet wurde. Da bei diesem Verfahren die Gefahr besteht, in eine alternierende Endlosschleife zu kommen, ist es notwendig, einen Abbruch nach einer endlichen Zahl von Iterationen zu erzwingen. Hierbei kann eine Endlosschleife erkannt werden, wenn nach jeder Iteration der Zustand der Zuordnungen aller Ganglinien zu ihren Clustern gespeichert und mit den bisher erreichten Zuständen verglichen wird. Endet eine Iteration in einem Zustand mit bereits bekannten Zuordnungen, bewegt sich der Algorithmus in einer Endlos-Schleife und wird abgebrochen.

Bewertung

Bei der k-Means-Methode wird jede RMQ-Ganglinie mit der gemittelten RMQ-Ganglinie jedes Clusters verglichen, so dass nach jeder neuen Zuweisung einer RMQ-Ganglinie zu einem Cluster die Abstandswerte aller RMQ-Ganglinien zum alten und neuen Cluster neu berechnet werden müssen.

Gleichzeitig ist das Ergebnis dieser Cluster-Methode abhängig von der Auswahl der Initialcluster, so dass bei identischen Eingangsdaten und unterschiedlichen Initial-clustern auch unterschiedliche Cluster entstehen können.

Bei sehr ähnlichen RMQ-Ganglinien und einer ungeeigneten Auswahl von Initialclustern kann es vorkommen, dass ein Cluster vollständig entfällt, sobald alle dort vorhandenen RMQ-Ganglinien innerhalb von einer Iteration anderen Clustern zugeordnet werden. In den darauffolgenden Iterationen kann wegen einer fehlenden Abstandsberechnung zu diesem Cluster keine RMQ-Ganglinie mehr zugeordnet werden. Dies trifft besonders dann zu, wenn z.B. bei 3 Initialclustern die ersten zwei aus fast identischen RMQ-Ganglinien bestehen. Dann ist es möglich, dass in der nächsten Iteration alle RMQ-Ganglinien aus einem Cluster einem anderen, fast identischen Cluster zugeordnet werden. Durch eine geeignete Wahl der Initialcluster würde sich eine ungünstige Clusterung vermeiden lassen.

5.1.3.2 Single-Linkage (Nearest-Neighbour)

Durchführung

Die Single-Linkage-Methode zur Clusterung ist ein Prozess, bei dem zuerst jede RMQ-Ganglinie als eigenständiges Cluster betrachtet wird. Durch das nachfolgende Verfahren wird die Anzahl der Cluster auf die gewünschte Menge reduziert.

1. Ordne jede RMQ-Ganglinie einem eigenen Cluster zu.
2. Bestimme für jedes Paar von Clusterganglinien C_i,C_j einen Abstandswert und lege diesen in einer NxN-Matrix ab, wobei N die Anzahl aller Cluster ist.
3. Finde die Cluster C_i,C_j die voneinander den kleinsten Abstand besitzen und vereine diese zu einem gemeinsamen Cluster C_{ij}. Lösche die Cluster Ci und C_j.
4. Setze die Abstandswerte zu den übrigen Clustern C_k auf den kleineren der zwei Werte $min\left(D_{Typ}\left(G(i,d),G(k,d)\right),D_{Typ}\left(G(j,d),G(k,d)\right)\right)$
5. Gehe zu Schritt 3 solange die Abbruchbedingung nicht erfüllt ist.

In Abbildung 90 ist eine durchgeführte Clusterung beispielhafter RMQ-Ganglinien dargestellt. Sie umfasst 2 Cluster (blau, rot). Die ähnlichen RMQ-Ganglinien des blauen Clusters beinhalten die Werktage Montag bis Donnerstag und die RMQ-Ganglinien des roten Clusters beinhalten den Freitag.

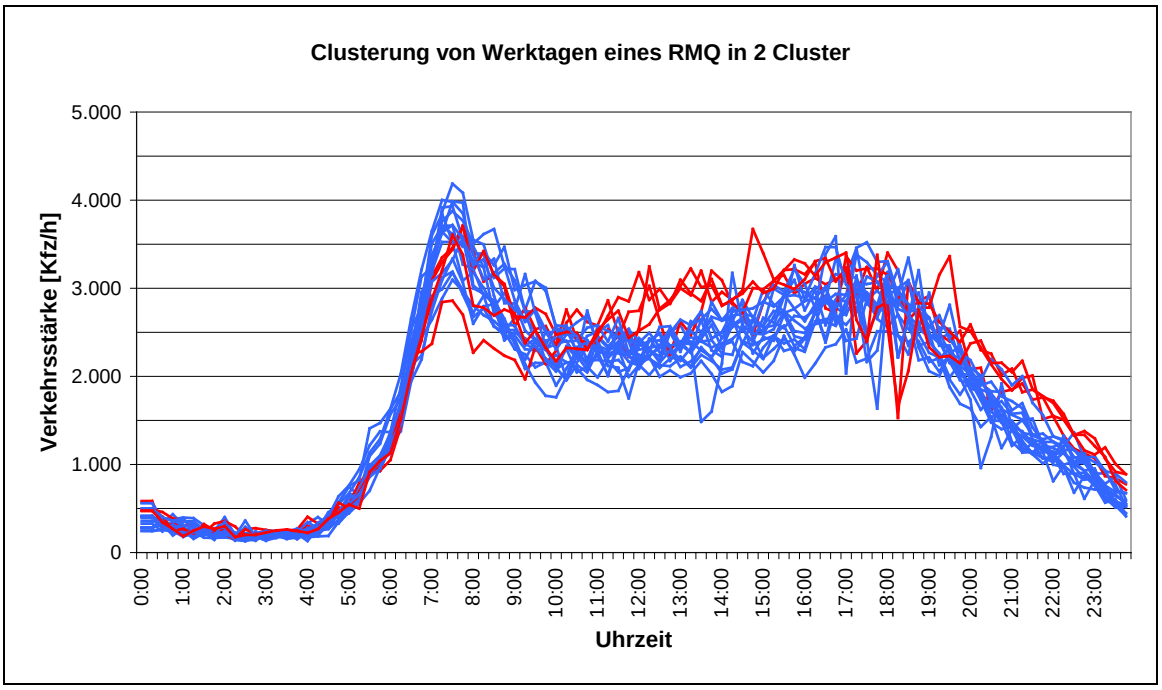


Abbildung 90: Clusterung von Ganglinien von Werktagen in zwei Cluster.

Wird eine Clusterung derselben RMQ-Ganglinien in 3 Cluster durchgeführt, ändert sich das Ergebnis, wie in Abbildung 91 gezeigt.

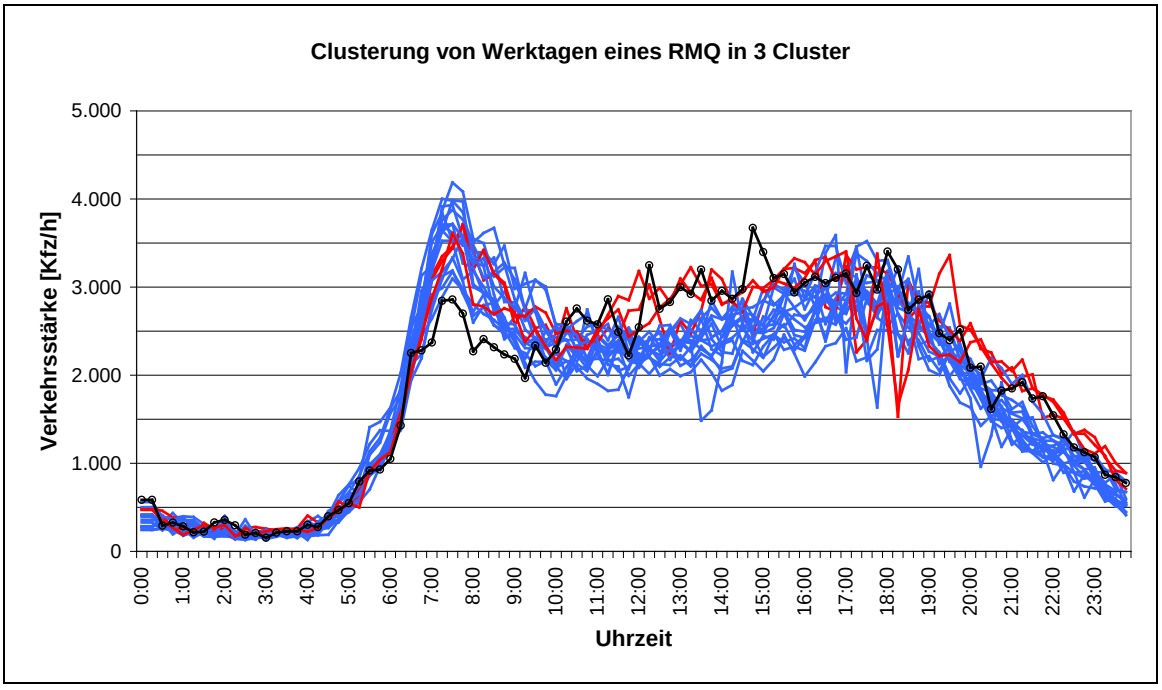


Abbildung 91: Clusterung von Ganglinien von Werktagen in drei Cluster.

Die RMQ-Ganglinien der Werktage Montag bis Donnerstag bleiben unverändert in einem Cluster (blau), dafür wird die RMQ-Ganglinie eines Freitages einem eigenen Cluster (schwarz) zugeordnet. Wie in Abbildung 91 zu sehen ist, handelt es sich um

eine RMQ-Ganglinie mit einer auffällig geringeren Morgenspitze im Vergleich zu allen RMQ-Ganglinien der anderen 2 Cluster.

Abbruchbedingung

Bei jedem Durchlauf der Schritte 3 bis 5 wird die Anzahl der Cluster durch Verschmelzung um 1 reduziert. Wird eine definierte Anzahl von k Clustern gewünscht, so müssen die Schritte 3 bis 5 genau

$$n_{Iterationen} = n_{Ganglinien} - n_{Cluster}$$

mit

$n_{Ganglinien}$: Anzahl der zu clusternden Ganglinien.

$n_{Cluster}$: Anzahl an gewünschten Clustern.

mal durchgeführt werden.

Alternativ zum Erzeugen einer vorgegebenen Anzahl von Clustern kann als Abbruchbedingung zur Verschmelzung zweier Cluster ein maximaler Abstandswert gegeben sein. Dadurch werden abstandsnächste Cluster nur so lange miteinander verschmolzen, bis die abstandsnächsten Cluster zu weit auseinander liegen. Durch dieses Vorgehen wird bei sehr ähnlichen RMQ-Ganglinien in einer Gruppe nur ein Cluster und bei großen Abstandswerten in einer anderen Gruppe mehrere Cluster erzeugt. Eine Eichung der Abbruchbedingung wird durchgeführt, sobald die Durchführung der Clusterung auf eine Clustermethode festgelegt wird.

Bewertung

Der Vorteil dieser Methode besteht darin, dass die Cluster unabhängig von der Sortierung der eingehenden RMQ-Ganglinien immer dieselbe Aufteilung von Clustern liefern. Zudem ist sichergestellt, dass 2 RMQ-Ganglinien, die voneinander den geringsten Abstandswert besitzen, auch tatsächlich in dasselbe Cluster zugeordnet werden.

Bei diesem Verfahren können bei ungünstigen RMQ-Ganglinien allerdings unerwünschte Ergebnisse entstehen, wie im Folgenden durch Beispielganglinien gezeigt wird. Für die Ähnlichkeit von zwei Clustern wird der beste Ähnlichkeitswert aller RMQ-Ganglinienpaar aus den Clustern als Referenz gewählt. Bei einer Clusterung von 6 RMQ-Ganglinien, wie sie in Abbildung 92 zu sehen sind, werden die Abstände aller RMQ-Ganglinien zueinander berechnet. Diese sind in Tabelle 49 gezeigt. Nun werden die abstandnächsten Cluster C_1 und C_2 im ersten Schritt verschmolzen. Danach folgen entsprechend den Abständen C_{12} mit C_3, C_{123} mit C_4 und C_{1234} mit C_5. Dadurch sind

nach 4 Iterationen die Ganglinien C_5 und C_1 in einem Cluster, obwohl die Ganglinie C_5 viel ähnlicher zur Ganglinie in C_6 ist, als zu den Ganglinien in C_1 bis C_3.

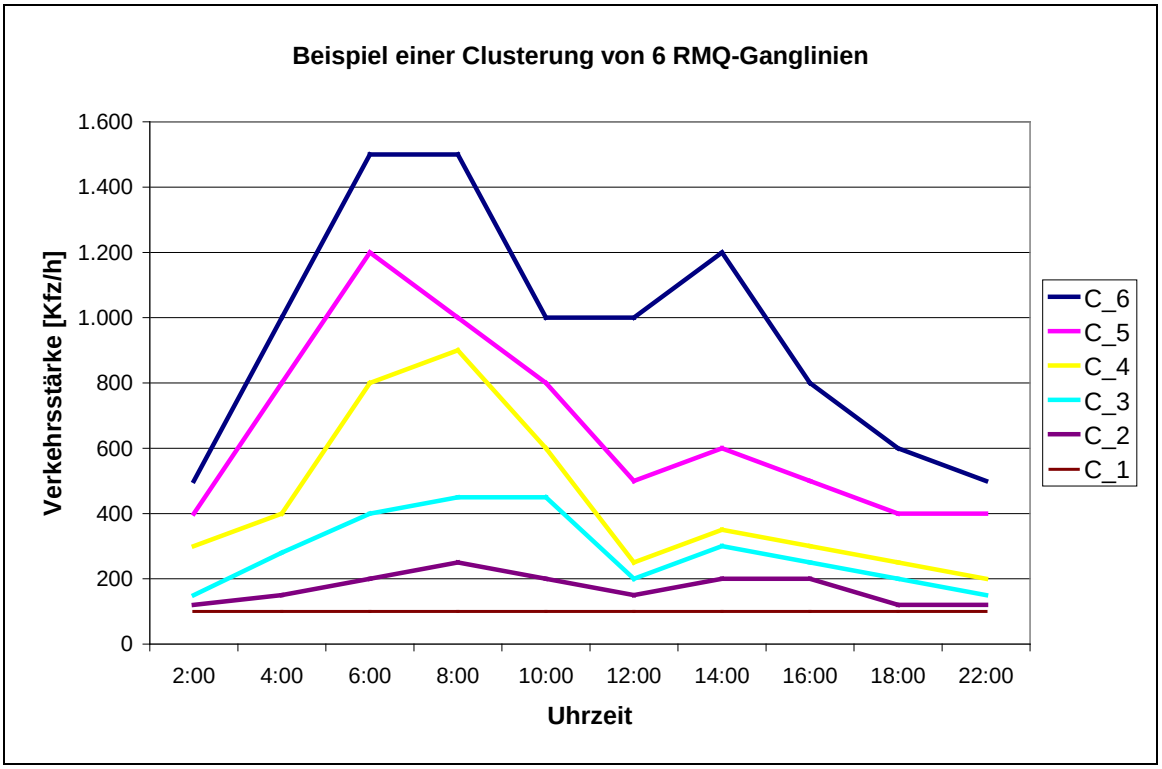


Abbildung 92: 6 Beispielganglinien zur Veranschaulichung der Clusterung.

	C_1	C_2	C_3	C_4	C_5	C_6
C_1	0	**5,9**	27,7	19,0	12,5	36,4
C_2	5,9	0	**7,0**	14,0	23,2	32,3
C_3	27,7	7,0	0	**7,3**	17,0	26,7
C_4	19,0	14,0	7,3	0	**9,9**	20,0
C_5	12,5	23,2	17,0	9,9	0	**10,4**
C_6	36,4	32,3	26,7	20,0	10,4	0

Tabelle 49: Abstandswerte aus GEH-Funktion aller Beispielganglinien.

5.1.3.3 Average Linkage

Durchführung

Die Clusterung nach der Average-Linkage Methode funktioniert ähnlich der Single-Linkage-Methode aus Kapitel 5.1.3.2. Als Kriterium der Verschmelzung von zwei Clustern dient hier jedoch nicht der kleinste Abstand von Tagesganglinien aus je einem Cluster, sondern der Abstandswert der 2 gemittelten Clusterganglinien.

Zusätzlich zum Vorgehen der Single-Linkage-Methode wird hier in Schritt 4 nach der Verschmelzung zweier Cluster als neuer Abstand aller Tagesganglinien zum neuen Cluster nicht jeweils der kleinere Abstand als neuer Abstandswert bestimmt, sondern für jedes Cluster ein neuer Abstandswert zum neuen Cluster berechnet.

Eine Clusterung der Beispielganglinien in Abbildung 92 führt hier zu folgendem Ergebnis. Die initialen Abstandswerte in Tabelle 49 führen zu einer Verschmelzung von C_1 mit $C_{2.}$ Die neuen Abstandswerte sind in Tabelle 50 markiert.

	C_{12}	**C_3**	**C_4**	**C_5**	**C_6**
C_{12}	0	**9,6**	**16,4**	**25,4**	**34,3**
C_3	**9,6**	0	7,3	17,0	26,7
C_4	**16,4**	7,3	0	9,9	20,0
C_5	**25,4**	17,0	9,9	0	10,4
C_6	**34,3**	26,7	20,0	10,4	0

Tabelle 50: Abstandswerte der Average-Linkage-Methode nach 1 itteration.

Im zweiten Schritt wird nun abweichend von der Single-Linkage-Methode wegen dem geringeren Abstandswert C_3 mit C_4 verschmolzen. Für das verschmolzene Cluster werden erneut neue Abstandswerte berechnet (Tabelle 51)

	C_{12}	**C_{34}**	**C_5**	**C_6**
C_{12}	0	**13,3**	25,4	34,3
C_{34}	**13,3**	**0**	**13,3**	**23,2**
C_5	25,4	**13,3**	0	10,4
C_6	34,3	**23,2**	10,4	0

Tabelle 51: Abstandswerte der Average-Linkage-Methode nach 1 Iteration.

Der dritte Schritt verschmilzt C_5 mit C_6, so dass aus den neuen Abstandswerten in Tabelle 52 die Cluster C_{12} und C_{34} verschmolzen werden.

	C_{12}	**C_{34}**	**C_{56}**
C_{12}	0	13,3	**30,2**
C_{34}	13,3	0	**18,6**
C_{56}	**30,2**	**18,6**	**0**

Tabelle 52: Abstandswerte vor der letzten Iteration.

Danach gibt es zwei Cluster, wobei die Ganglinie G_5 und G_6 aus den ursprünglichen Clustern C_5 und C_6 zusammen in einem Cluster sind.

Bewertung

Diese Methode erzeugt auch aus ungünstigen Ganglinien passende Cluster. Allerdings müssen dadurch in jeder Iteration für jedes Cluster neue Abstandswerte berechnet werden, was zu einem erheblichen Anstieg der Rechenzeit führt. Da die Rechenzeit für die Clusterung vorerst kein relevanter Faktor ist, wird die Average-Linkage-Methode zur Clusterung von Ganglinien verwendet.

5.1.4 Ergebnisse

Clusterung von Netzganglinien

Es werden alle Ganglinien des Jahres 2007 (25.02. – 31.12.), die für Autobahnquerschnitte im Untersuchungsgebiet vorliegen, als Eingangsdaten für die Clusterung benutzt.

Vorklassifizierung

Für die Vorklassifizierung der Ganglinien eines ganzen Jahres werden die Kalendertage anhand der folgenden Regeln in drei Klassen aufgeteilt:

- Klasse 1: Werktag Ferien (56 Tage)
- Klasse 2: Werktag Schule (148 Tage)
- Klasse 3: Samstag, Sonntag, Feiertag (51 Tage)

Bestimmung der Parameter zur Clusterung

Die Clusterung wird anhand der „Average-Linkage“-Methode im Kapitel 5.1.3.3 mit der Abstandsfunktion „GEH“ aus Kapitel 5.1.2.3 durchgeführt.

Für die Bestimmung eines Abbruchkriteriums werden zuerst zwei Clusterungen mit den Abbruchkriterien GEH=5 und GEH=10 durchgeführt. Eine Verteilung der Ganglinien auf die unterschiedlichen Cluster beider Durchgänge ist in den Abbildungen 93 und 88 zu sehen. Dabei entspricht jede Säule in der Abbildung einem Cluster. Die Höhe der Säule entspricht der Anzahl an Ganglinien in dem Cluster und die 7 Farben entsprechen den einzelnen Wochentagen. Die Cluster sind nach Vorklassifizierung (1 Links, 2 Mitte, 3 Rechts) vorsortiert.

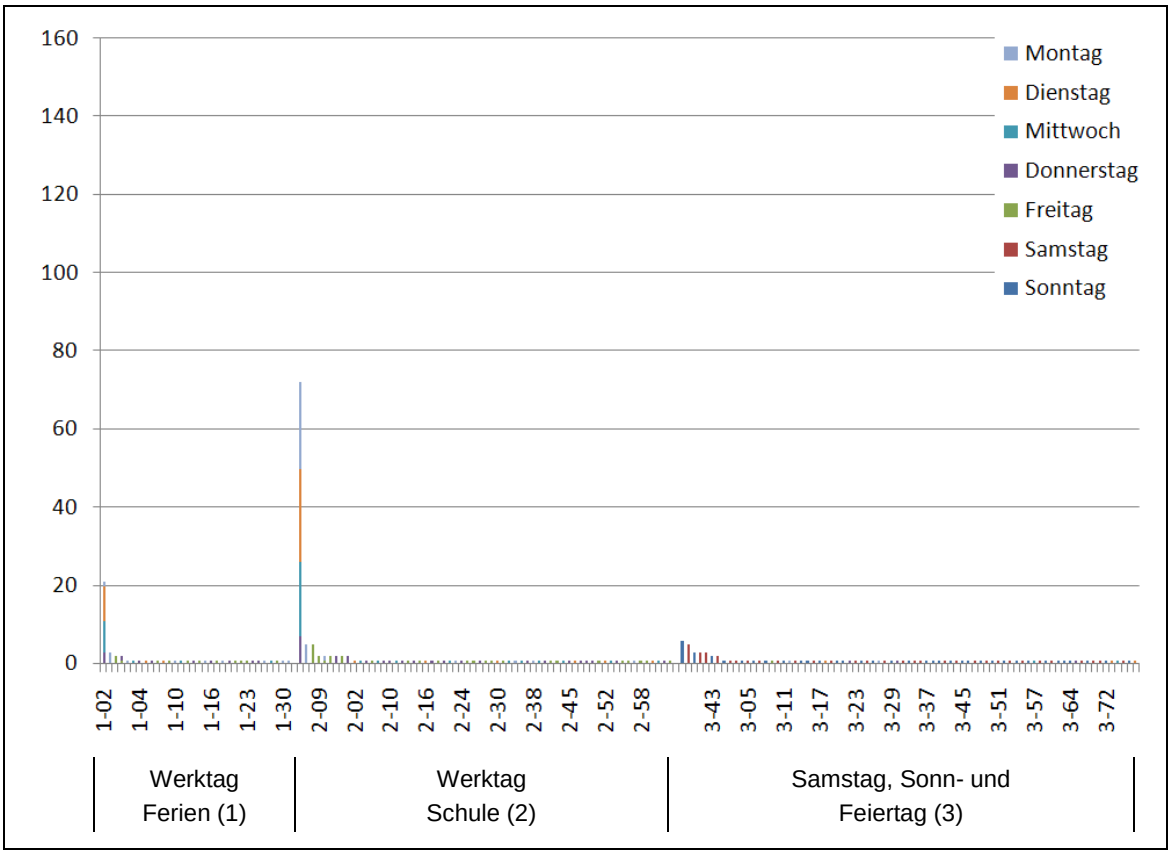


Abbildung 93: Clusterverteilung mit GEH-Abstandsfunktion, Clustermethode Average Linkage und Abbruchbedingung GEH=5.

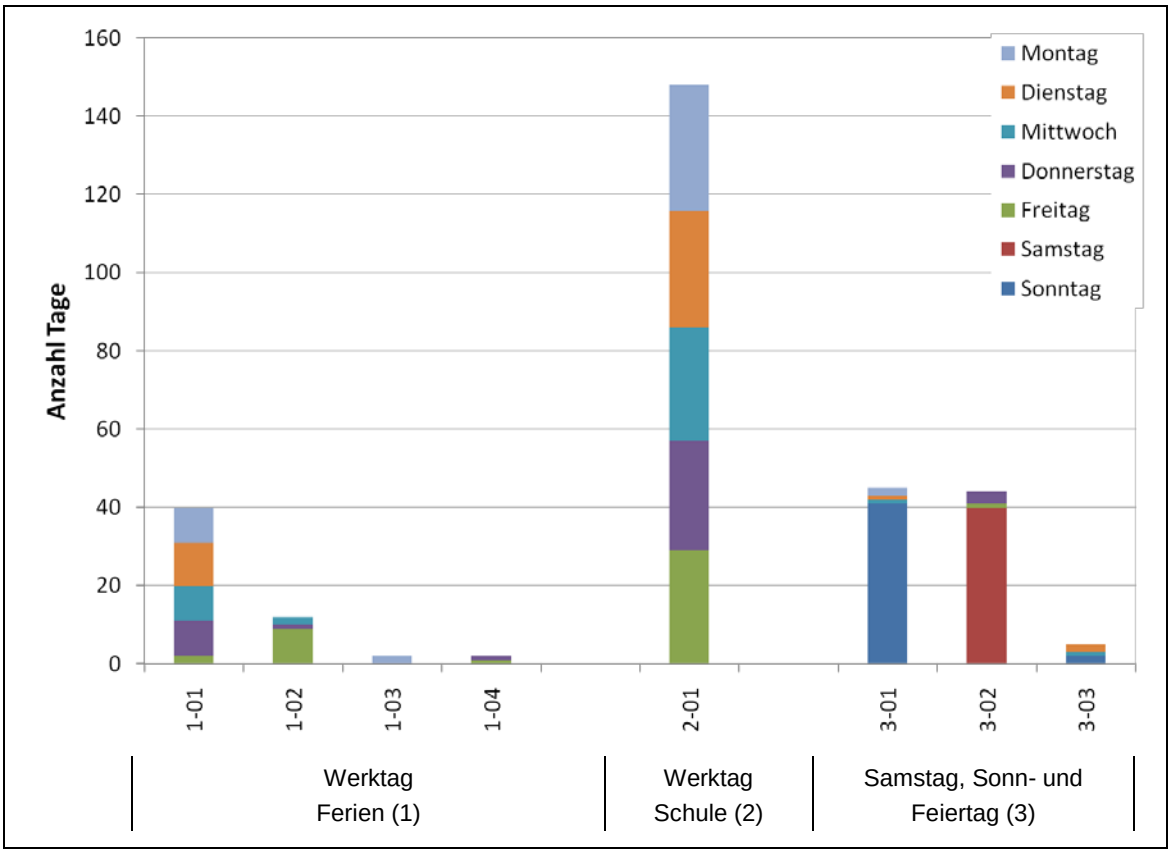


Abbildung 94: Clusterverteilung mit GEH-Abstandsfunktion, Clustermethode Average Linkage und Abbruchbedingung GEH=10.

Beim Vergleich der zwei Clusterverfahren ist zu erkennen, dass bei einer Clusterung mit dem Abbruchkriterium GEH=5 ein großes Cluster entsteht mit den Werktagen Montag bis Donnerstag und der Freitag als eigenständiges Cluster erkannt wird. Desweiteren gibt es insgesamt 153 der 172 Cluster, die aus genau einer Ganglinie bestehen. Diese Einzelcluster unterscheiden sich voneinander durch einen GEH-Wert von mehr als 5 und werden beim „Average Linkage“-Verfahren nicht zusammengefügt.

Bei dem Clusterverfahren mit einem Abbruchkriterium von GEH=10 fallen in der Vorklassifizierung der Werktage ohne Schulferien ausnahmslos alle Ganglinien zum Cluster 2-01. Dennoch gibt es bei den Werktagen in den Schulferien zwei Cluster, die aus jeweils zwei Ganglinien bestehen. Dies sind Ausreißer, an denen ein besonderes Verkehrsverhalten vorliegt. Das Cluster 1-03 besteht aus dem 24.12. und 31.12. und das Cluster 1-04 aus dem 27.12. und 28.12. Beide Cluster sind zu Recht vom Hauptcluster 1-01 bzw. 1-02 getrennt worden, da an diesen Tagen eine veränderte Verkehrsnachfrage vorliegt.

Das Abbruchkriterium GEH=5 ist zu stark, da es zu früh abbricht und sehr viele Einzelcluster übrig lässt, wohingegen das Kriterium GEH=10 zu schwach ist, da es zu spät abbricht. Ein erfolgreiches Verfahren hat als Abbruchbedingung einen Wert zwischen GEH=5 und GEH=10. Die Clusterverteilungen der Abbruchkriterien GEH=6 bis GEH=9 mit der Schrittweite 1 ist in den Abbildungen 89 bis 92 dargestellt.

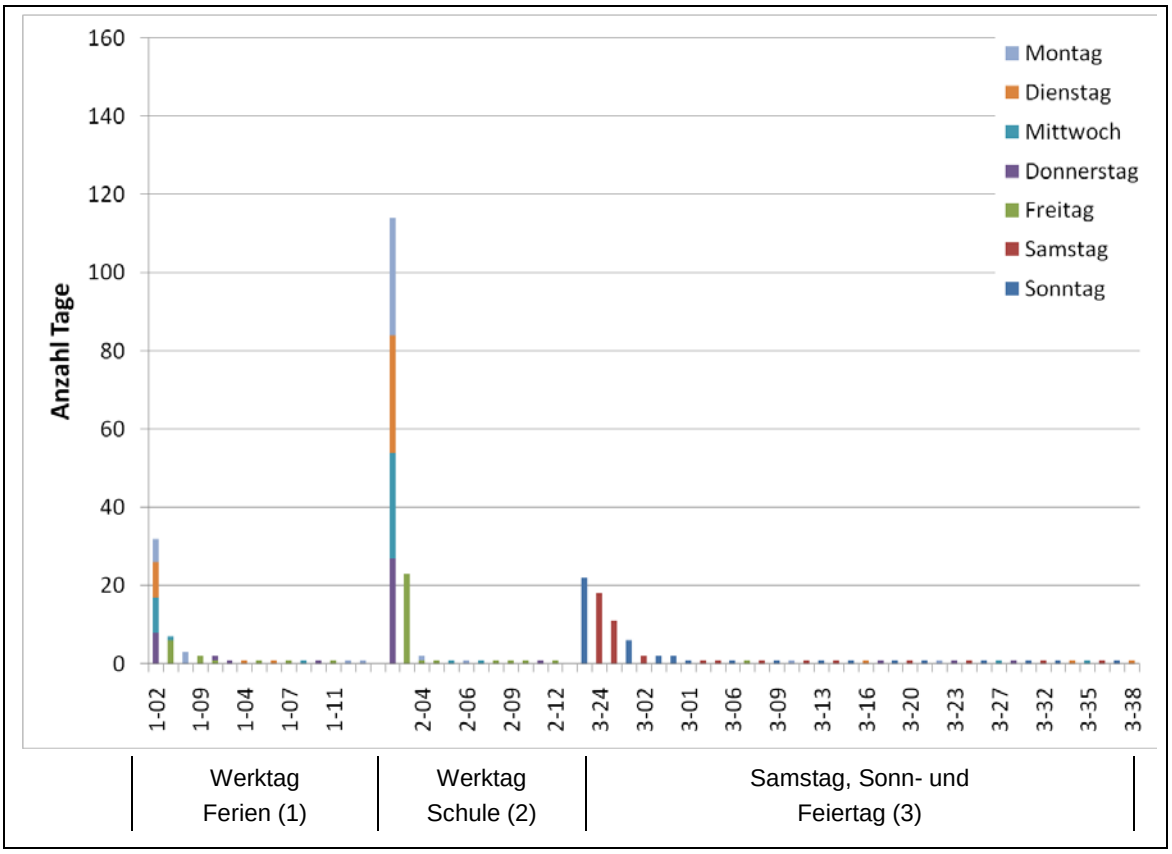


Abbildung 95: Clusterverteilung mit GEH-Abstandsfunktion, Clustermethode Average Linkage und Abbruchbedingung GEH=6.

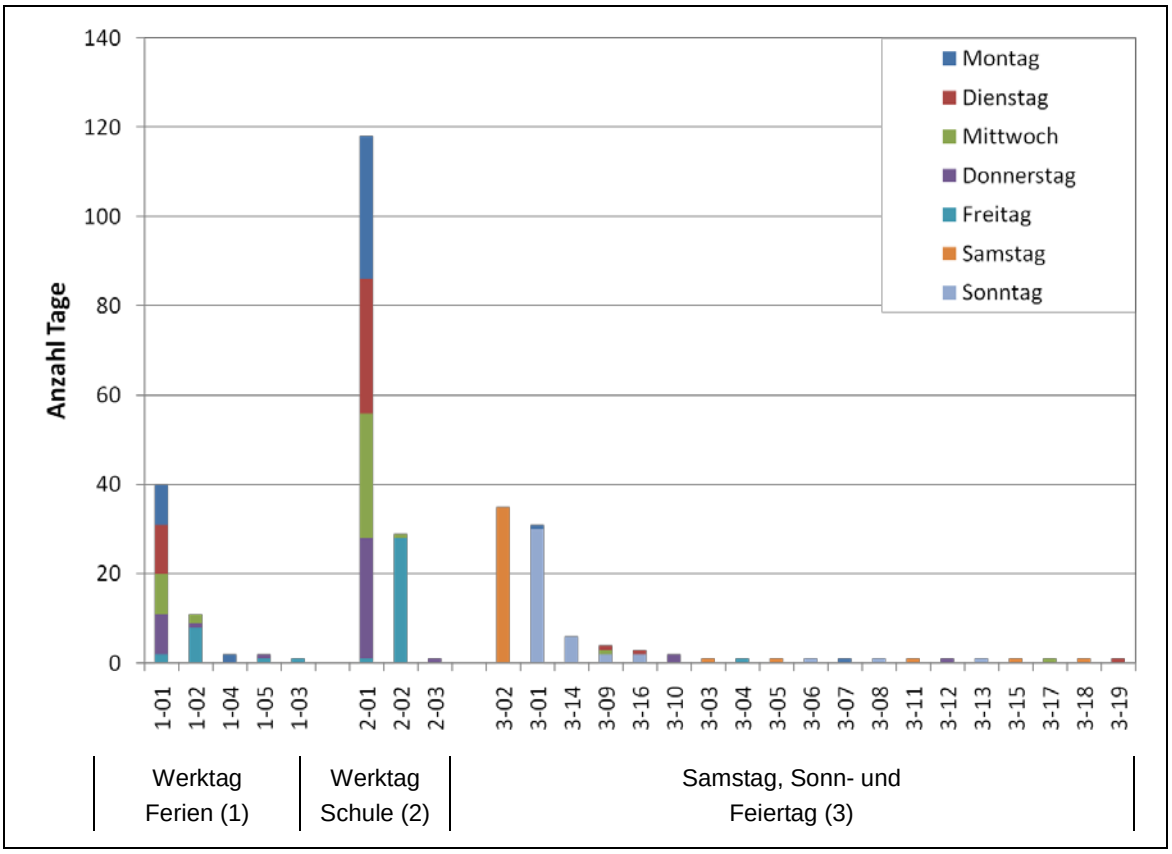


Abbildung 96: Clusterverteilung mit GEH-Abstandsfunktion, Clustermethode Average Linkage und Abbruchbedingung GEH=7.

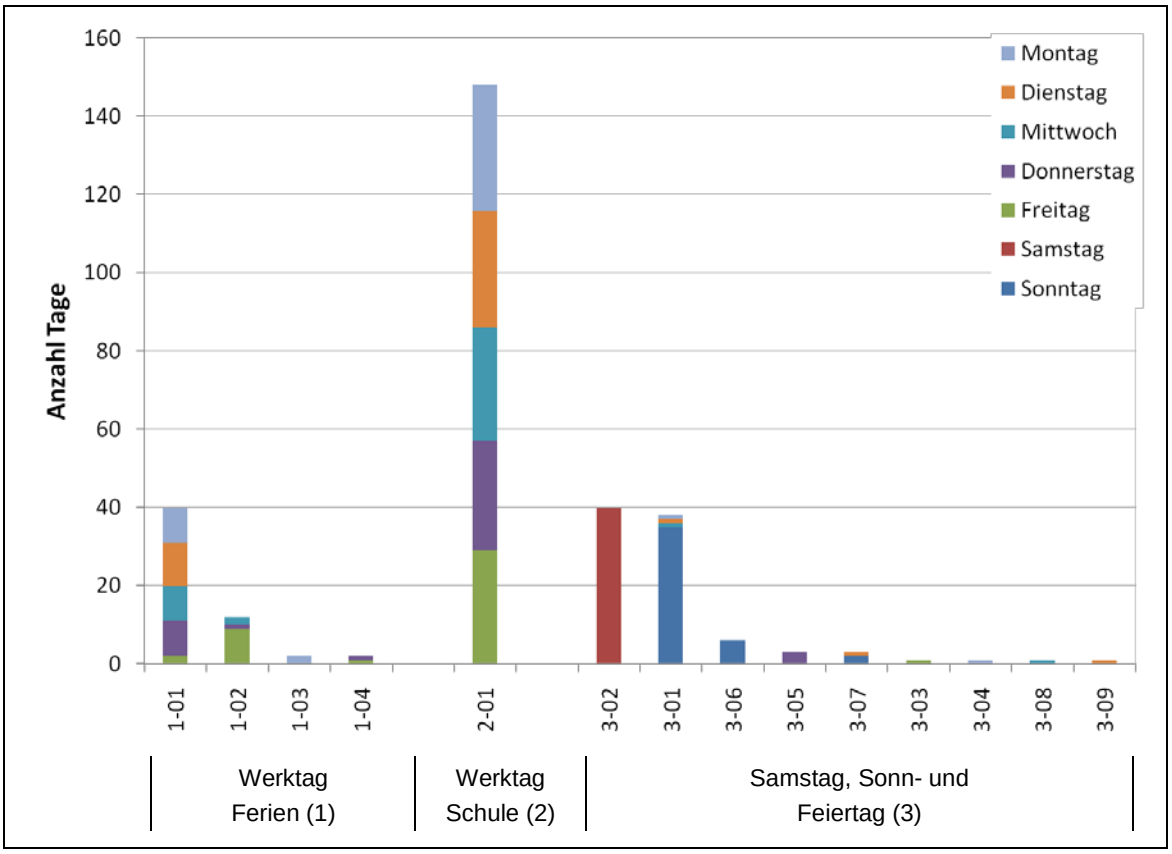


Abbildung 97: Clusterverteilung mit GEH-Abstandsfunktion, Clustermethode Average Linkage und Abbruchbedingung GEH=8.

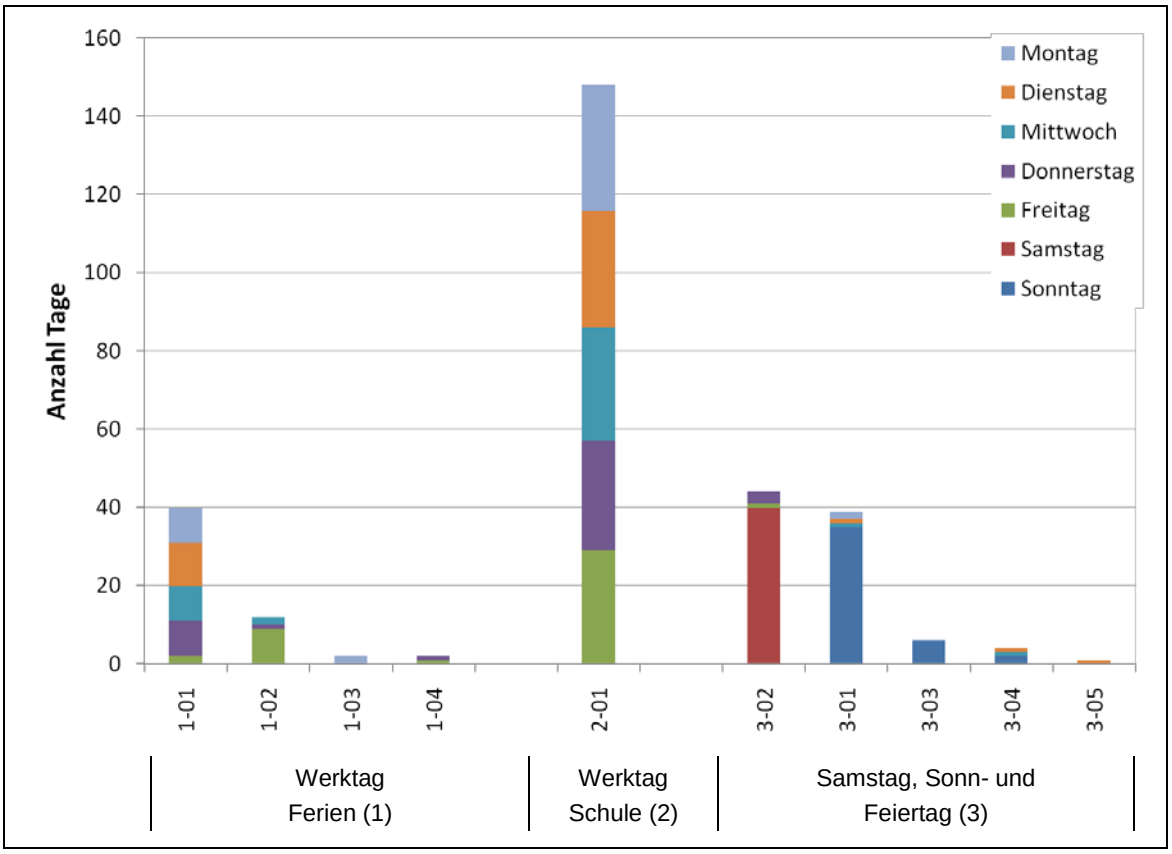


Abbildung 98: Clusterverteilung mit GEH-Abstandsfunktion, Clustermethode Average Linkage und Abbruchbedingung GEH=9.

Der Verlauf der Clusterverteilung zeigt, dass ab einem Abbruchkriterium mit GEH=8 eine Unterteilung zwischen den Wochentagen Montag bis Donnerstag und dem Freitag nicht mehr vorhanden ist. Jedoch sind beim Verfahren mit dem Abbruchkriterium GEH=7 einige Einzelcluster vorhanden. Für einige Einzelcluster ist nach dem Kalender nicht eindeutig zu erkennen, warum bei ihnen eine veränderte Verkehrsnachfrage vorliegt. Ein Unterschied im Verkehrsverhalten kann durch verschiedene externe Einflüsse begründet sein, z.B. Unwetter oder Bahnstreik. Wenn diese Einflüsse nicht eindeutig identifiziert werden können, ist es ebenfalls nicht möglich ein daraus resultierendes Cluster für eine Prognose zu benutzen. Um eine genauere Clusterung durchführen zu können, ist eine manuelle Analyse der einzelnen Cluster notwendig.

Manuelle Analyse

Für eine genauere Unterteilung der Ganglinien in Cluster wird eine manuelle Analyse der Cluster durchgeführt. Für die Analyse werden die Daten des Clusterverfahrens mit der „Average Linkage“-Methode und einem GEH-Wert 7 als Abbruchkriterium als Eingangsdaten verwendet. Dabei werden die einzelnen Cluster und die darin vorhandenen Tage analysiert. In Tabelle 53 sind alle durch das Clusterverfahren entstandenen Cluster kalendarisch beschrieben.

Cluster	Tage	Beschreibung
1-01	40	Ferien, Werktag, Mo-Do | Ferien; Brückentag, Fr
1-02	11	Ferien, Werktag, Freitag | Ferien, Werktag, Mi-Do vor Feiertag
1-03	1	Freitag in den Pfingstferien – kein kalendarischer Unterschied zu 1-02
1-04	2	Heilig Abend und Silvester
1-05	2	Werktage zwischen Heilig Abend und Silvester
2-01	118	Schultag, Werktag, Mo-Do | Schultag, Brückentag, Fr
2-02	29	Schultag; Werktag, Freitag | Schultag, Werktag, Mi-Do vor Feiertag
2-03	1	Donnerstag vor Weihnachten. (Freitag wie jeder Freitag und Verschmelzung mit 2-01 knapp verfehlt) – keine Unterscheidung zu 2-01
3-01	31	Sonntag | Feiertag, Montag – (jeweils Mrz.-Okt.)
3-02	35	Samstag
3-03	1	Beginn der Osterferien – keine kalendarische Unterscheidung zu 3-02
3-04	1	Karfreitag
3-05	1	Samstag nach Karfreitag
3-06	1	Ostersonntag
3-07	1	Ostermontag
3-08	1	Sonntag – keine kalendarische Unterscheidung zu 3-01
3-09	4	Feiertag Di-Mi
3-10	2	Feiertag Do-Fr
3-11	1	Samstag – keine kalendarische Unterscheidung zu 3-02
3-12	1	Feiertag Do-Fr – keine kalendarische Unterscheidung zu 3-10
3-13	1	Samstag – keine kalendarische Unterscheidung zu 3-02
3-14	6	Sonntag und Feiertag; Montag (jeweils Nov.-Dez.)
3-15	1	Samstag vor Weihnachten
3-16	3	1. Weihnachtsfeiertag | Sonntage zwischen 22.12. – 31.12.
3-17	1	2. Weihnachtsfeiertag
3-18	1	Samstag, 29.12.2007 – keine kalendarische Unterscheidung zu 3-15
3-19	1	Neujahr

Tabelle 53: Kalendarisch analysierte Cluster des „Average Linkage“-Verfahrens mit Abbruchkriterium GEH=7.

Alle kalendarischen Zuordnungen werden anschließend in Tabelle 54 zu Regeln zusammengefasst. Darin sind 9 Hauptregeln (C01 bis C09) enthalten, die alle Kalendertage beinhalten. Die weiteren Regeln identifizieren besondere Cluster für die Ostertage (C10 bis C13) und Weihnachtstage (C14 bis C19). In Tabelle 53 vorhandene Cluster, die keine kalendarische Regel besitzen, werden zu den Hauptclustern hinzugefügt. Hierbei handelt es sich um Netzganglinien, deren Unterscheidung zum Hauptcluster nicht eindeutig erklärt werden kann.

Die Cluster 3-01 und 3-14 decken alle Sonntage im Bereich von März bis Dezember ab. Die Monate Januar und Februar sind aufgrund fehlender Netzganglinien im Jahr

2007 nicht vorhanden. Eine Aufteilung der Sonntage in die zwei Cluster ist eine Folge des jahreszeitlich bedingt unterschiedlichen Verkehrsverhaltens. (Wetter, Tageslänge, …). Die Sonntage der fehlenden Monate Januar und Februar werden entsprechend dieser Begründung zum „Wintercluster“ 3-14 gezählt. Die entsprechende manuelle Zuordnung von Kalendertag zu Cluster wird aus den Regeln in Tabelle 54 erstellt.

Cluster	Beschreibung
C01	Werktag – Schule Montag bis Donnerstag
C02	Werktag – Schule Freitag und Mi/Do vor Feiertag
C03	Werktag – Ferien Montag bis Donnerstag
C04	Werktag – Ferien Freitag oder Mi/Do vor Feiertag
C05	Samstag
C06	Sonntag und Montag Feiertag (März bis Oktober)
C07	Sonntag und Montag Feiertag (November bis Februar)
C08	Feiertag Dienstag und Mittwoch
C09	Feiertag Donnerstag und Freitag
C10	Ostern – Karfreitag
C11	Ostern – Samstag nach Karfreitag
C12	Ostern – Ostersonntag
C13	Ostern – Ostermontag
C14	Samstag nach 21.12.
C15	Sonntag nach 21.12.
C16	Heilig Abend | Silvester
C17	Werktage zwischen Heilig Abend und Silvester
C18	1. Weihnachtsfeiertag | 2. Weihnachtsfeiertag
C19	Neujahr

Tabelle 54: Kalendarisch analysierte Cluster des „Average Linkage“-Verfahrens mit Abbruchkriterium GEH=7.

Die Zuordnung der Regeln erfolgt überschreibend, so dass für das Datum einer Ganglinie die unterste gültige Regel die darüber gültigen Regeln überschreibt. So wird z.B. ein Werktag am 27.12. dem Cluster C03 zugeordnet, jedoch wird diese Zuordnung vom Cluster C24 überschrieben, so dass die Ganglinie im Cluster C24 gespeichert wird. Ein Verfahren, welches auf den definierten Regeln basiert, erzeugt die in Abbildung 99 dargestellte Clusterverteilung.

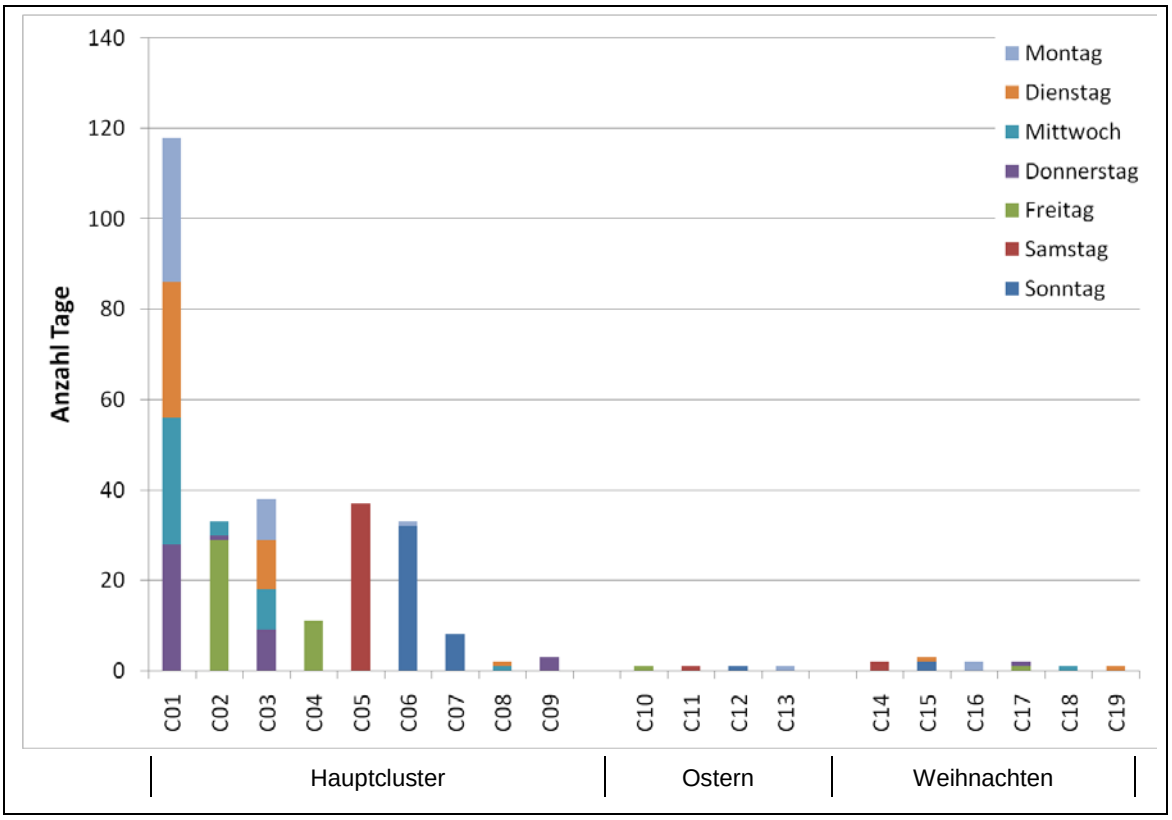


Abbildung 99: Clusterverteilung der manuellen Clusterung.

Zum Vergleich der manuellen Clusterung mit dem bisherigen Clusterverfahren wurden in Tabelle 55 die gewichteten durchschnittlichen Abstände zwischen den Ganglinien und ihren zugeordneten Clustern dargestellt. Hierzu wurden zusätzlich vier Clusterverfahren mit feiner unterteilten Abbruchbedingungen durchgeführt (GEH=7,2 bis GEH=7,8 mit Schrittweite 0,2)

Clusterverfahren	**Anzahl Cluster**	**Gemittelte Abstand aller Ganglinien zu ihrem Cluster**
GEH-5	172	1,91
GEH-6	65	3,95
GEH-7	27	4,87
GEH-7,2	*25*	*4,91*
GEH-7,4	*23*	*4,96*
GEH-7,6	*17*	*5,45*
GEH-7,8	*15*	*5,55*
GEH-8	14	5,59
GEH-9	10	5,71
GEH-10	8	5,85
GEH-Manuell	19	5,09

Tabelle 55: Vergleich der Clusteranzahl und der gemittelten Abstände in den Clusterverfahren.

Ein Vergleich der manuellen Clusterung bezüglich des Clusterverfahrens mit der „Average Linkage“-Methode und der Abstandsfunktion GEH zeigt in Abbildung 100, dass die manuelle Clusterung ein besseres Ergebnis liefert, als das „Average Linkage“-Verfahren, obwohl es von genau diesem abgeleitet wurde. Eine Clusterung mit „Average Linkage“ und einem gemittelten Abstand von 5,09 würde anhand des Diagramms aus 22 Clustern bestehen. Einige dieser Cluster sind Einzelcluster, die nicht als eigenständige Cluster gezählt werden sollten. Da diese durch die manuellen Clusterung nicht mehr entstehen, wird die Anzahl der Cluster reduziert, der gemittelte Abstand durch die anteilsmäßig wenigen schlechter bewerteten Ganglinien aber nur geringfügig verschlechtert.

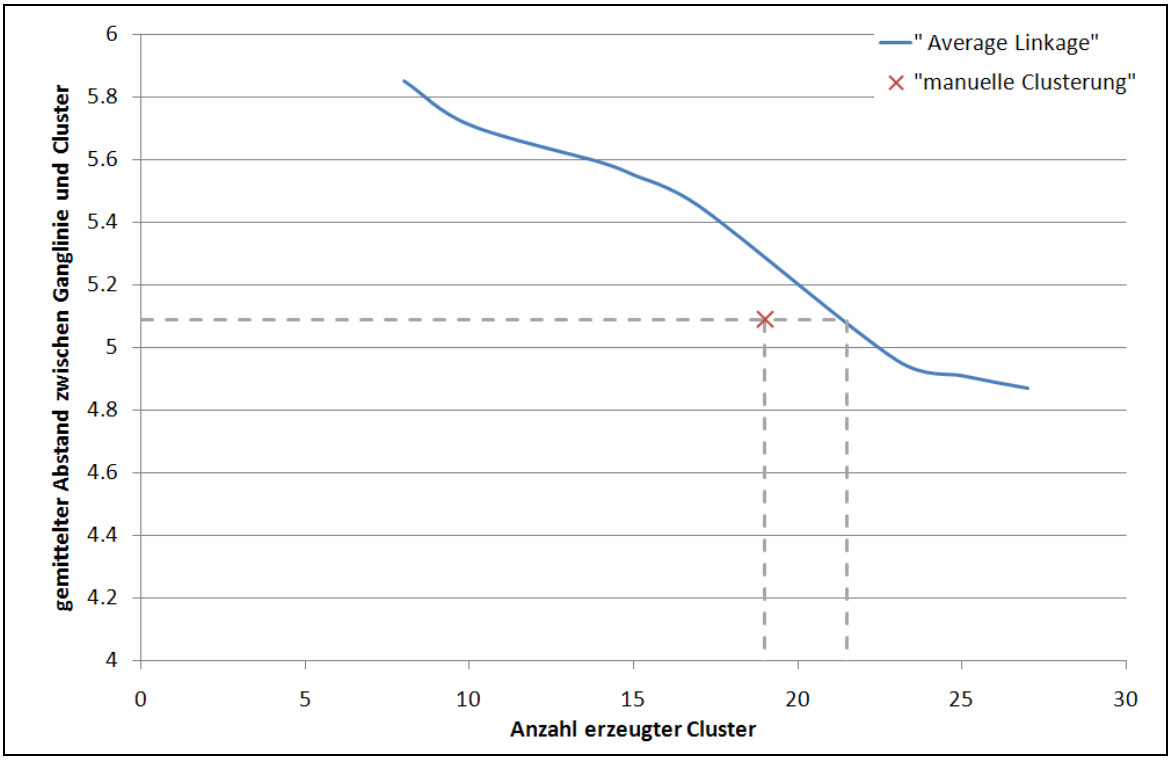


Abbildung 100: Vergleich der manuellen Clusterung mit „Average Linkage“.

Prognose

Durch die in Tabelle 54 beschriebenen Regeln lassen sich beliebige zukünftige Tage zu einem gegebenen Cluster zuordnen. Die in diesem Cluster vorhandenen Ganglinien oder QZ-Matrizen werden als Grundlage für die Prognose angewendet.

Evaluierung der Prognose für das erste Halbjahr 2008

Die in der manuellen Clusterung erzeugten Cluster basieren auf den Verkehrsstärken auf Bundesautobahnen an den Tagen des Jahres 2007. Für eine Prognose werden die

Regeln auf das Jahr 2008 angewendet und die Tage des Jahres 2008 in dem Cluster des Jahres 2007 zugeordnet.

Jedes Cluster beinhaltet vergangene Tage aus dem Jahr 2007 und zukünftige Tage des Jahres 2008 für die die Verkehrsstärken noch nicht bekannt sind. Somit lässt sich für jeden Tag eine Prognose berechnen, indem für einen Tag in der Zukunft (im Jahr 2008) die gemittelte Ganglinie des Clusters ausgewählt wird, in dem sich der Tag befindet.

Für das erste Halbjahr des Jahres 2008 werden die tatsächlichen Verkehrsstärken mit den prognostizierten Verkehrsstärken verglichen. Für die Evaluierung wird eine Abweichung der Prognose von GEH≤10 einzelner 15-Minuten-Intervalle als akzeptabel angesehen. Die Abweichung zwischen prognostizierten und gemessenen Verkehrsstärken wird als Prognosedifferenz bezeichnet. In Abbildung 101 ist die Summenhäufigkeit der Prognosedifferenzen dargestellt. Hierbei ist zu sehen, dass die Prognose durch verkehrstagtypische Ganglinien im Mittel ein besseres Ergebnis liefert, als eine Prognose basierend auf kalendertagtypische Ganglinien.

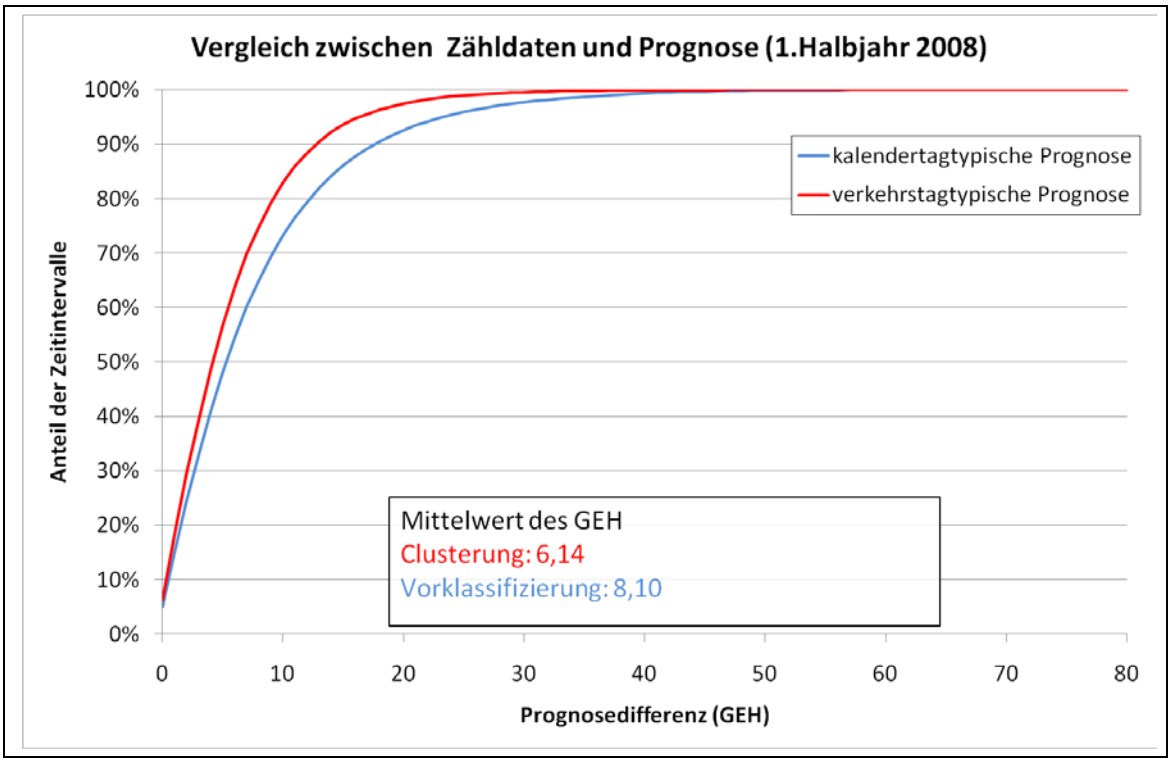


Abbildung 101: Summenhäufigkeit der Prognosedifferenzen.

An der Summenhäufigkeit der Prognosedifferenzen durch Clusterdaten ist zu sehen, dass 17% aller prognostizierten Messwerte einen Fehler von mehr als GEH=10 haben.

Prognosedifferenzanalyse

In den Abbildungen 102 und 103 sind die Anteile der nicht akzeptabel prognostizierten Messintervalle (Prognosedifferenz GEH>10) bezogen auf Verkehrsstärke, Datum und Uhrzeit dargestellt.

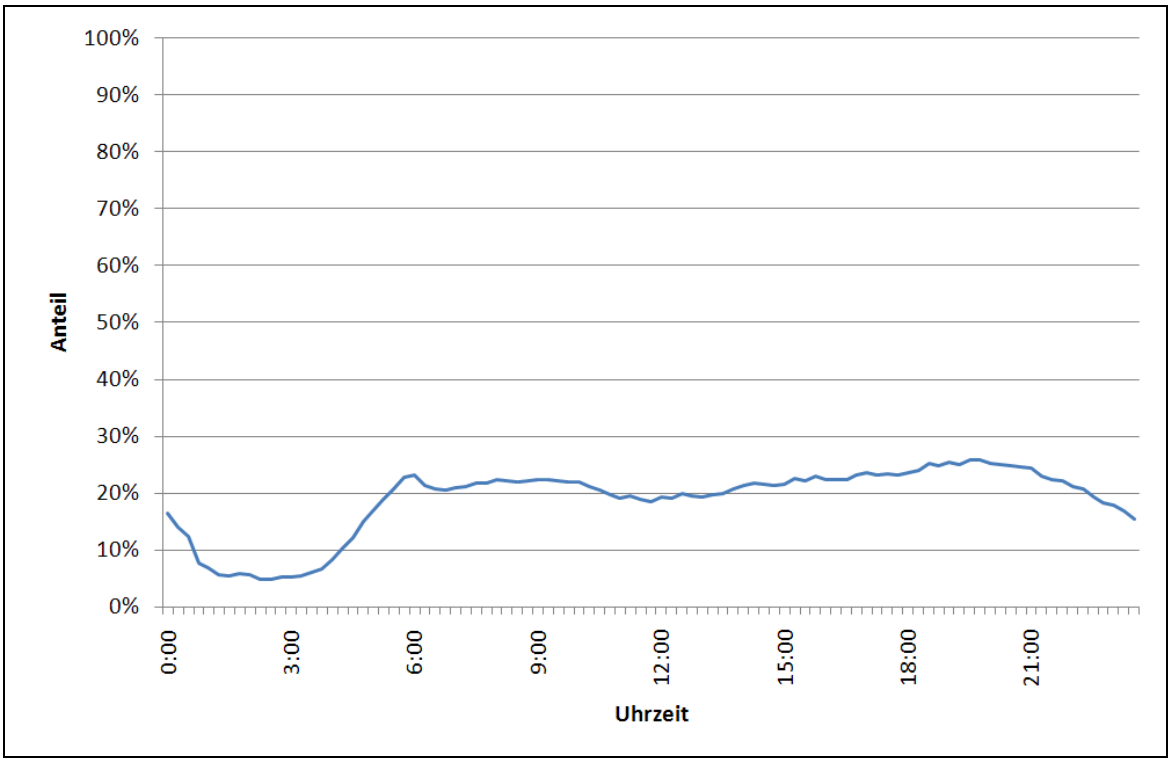


Abbildung 102: Häufigkeit der Prognosedifferenzen mit GEH>10 bezogen auf die Uhrzeit.

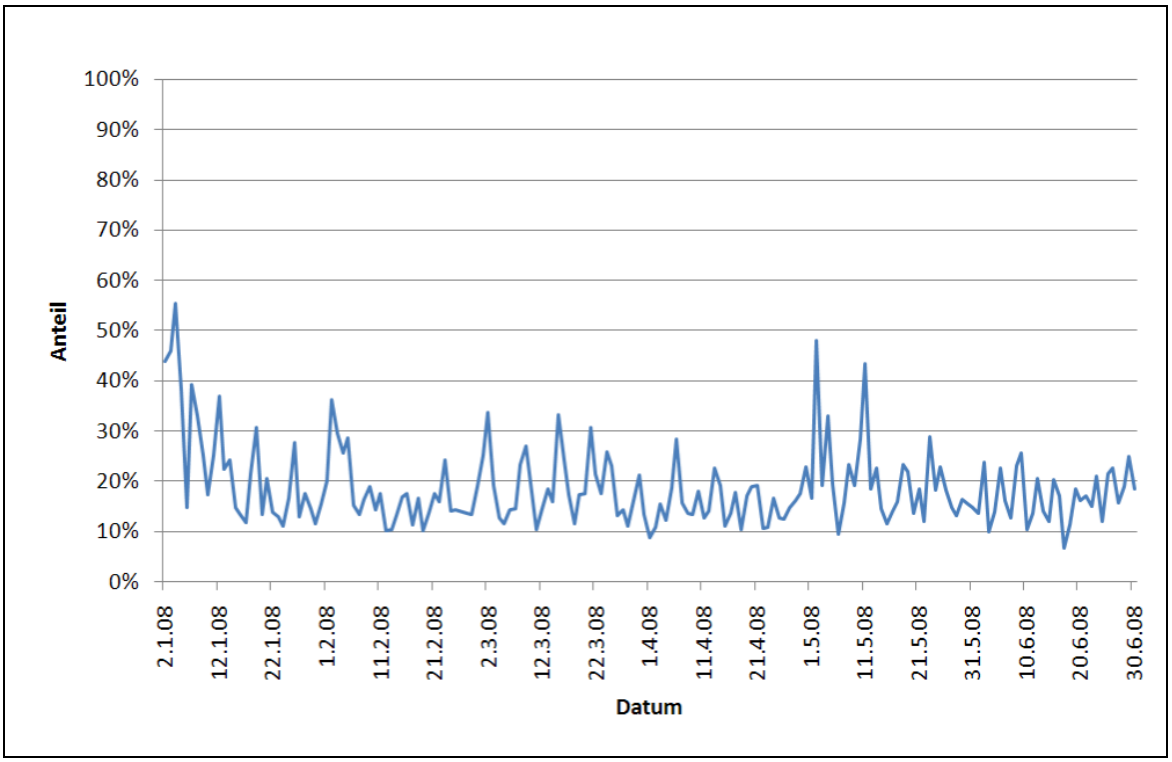


Abbildung 103: Häufigkeit der Prognosedifferenzen mit GEH>10 bezogen auf das Datum.

Die Häufigkeit der Prognosedifferenzen über GEH=10 in Abbildung 102 ist in der Nacht sehr gering. Tagsüber weichen die prognostizierten Verkehrsstärken an 20% der Messstellen so weit ab, dass der GEH größer als 10 ist. Erhöhte Häufigkeitswerte gegen 6 und 21 Uhr lassen sich auf die steil aufsteigenden und abfallenden Flanken der Spitzenstunden zurückführen. Eine zeitliche Verschiebung der Verkehrsströme an einem Tag erzeugt gerade an den Flanken eine größere Differenz zwischen der prognostizierten und tatsächlich gemessenen Verkehrsstärke.

Im Datumsverlauf in Abbildung 103 ist eine leichte Erhöhung der Prognosedifferenzen über GEH=10 in der ersten Woche des Jahres 2008, besonders am Freitag, den 04.01.2008 zu beobachten. Eine Erklärung ist hierfür, dass diese Woche bezogen auf die Nachfrage keine normalen Ferientage beinhaltet. Da für das Jahr 2007 jedoch keine Messdaten der Monate Januar und Februar vorlagen, konnte bei der Clusterung der Ferientage keine Unterscheidung speziell auf dieses jährlich einmalige Cluster erkannt und durchgeführt werden.

An den Feiertagen 1. Mai und Christi-Himmelfahrt sind die Prognosedifferenzen ebenfalls höher.

Prognosebewertung

Für die Bewertung der Clusterung und der damit verbundenen Prognose eines Kalendertages wird ein Bewertungskriterium eingeführt. Dabei wird ein Tag dann als gut prognostiziert, sobald der Anteil der akzeptablen Prognosedifferenzen (GEH≤10) den Wert 70% überschreitet. In Abbildung 104 sind beispielhaft zwei Ganglinien mit derselben Prognose dargestellt.

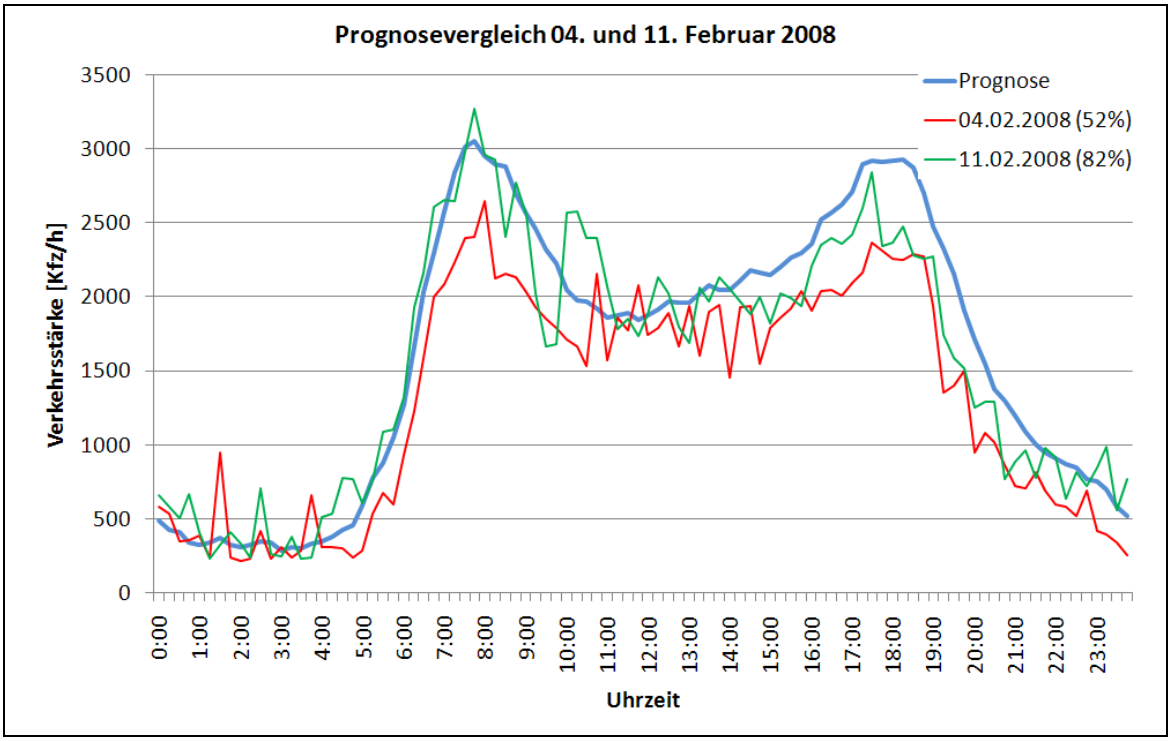


Abbildung 104: Vergleich zweier Tage mit einem Anteil an Messstellen mit akzeptablen Prognosedifferenzen über (11.02.2008) und unter (04.02.2008) 70% am Beispiel von genau einem RMQ.

Bei der roten Ganglinie vom 04.02.2008 haben nur 52% aller Messwerte eine akzeptable Prognosedifferenz von den Messwerten der prognostizierten Ganglinie. Näher betrachtet ist die tatsächliche Verkehrsstärke in den Spitzenstunden um etwa 500 Kfz/h geringer, als die vorausgesagte. Bei der grünen Ganglinie vom 11.02.2008 gibt es auch einige Abweichungen von der Prognose, jedoch ist bis auf einige wenige die gesamte Ganglinie bis etwa 14 Uhr gleich der prognostizierten. Auch danach gibt es nur eine geringe Abweichung. Diese hat über den gesamten Tag insgesamt 82% aller Messwerte innerhalb der Grenze einer akzeptablen Prognosedifferenz.

Anhand der Summenverteilung in Abbildung 105 ist ersichtlich, dass 91% aller Tage im ersten Halbjahr 2008 diesem Kriterium entsprechen.

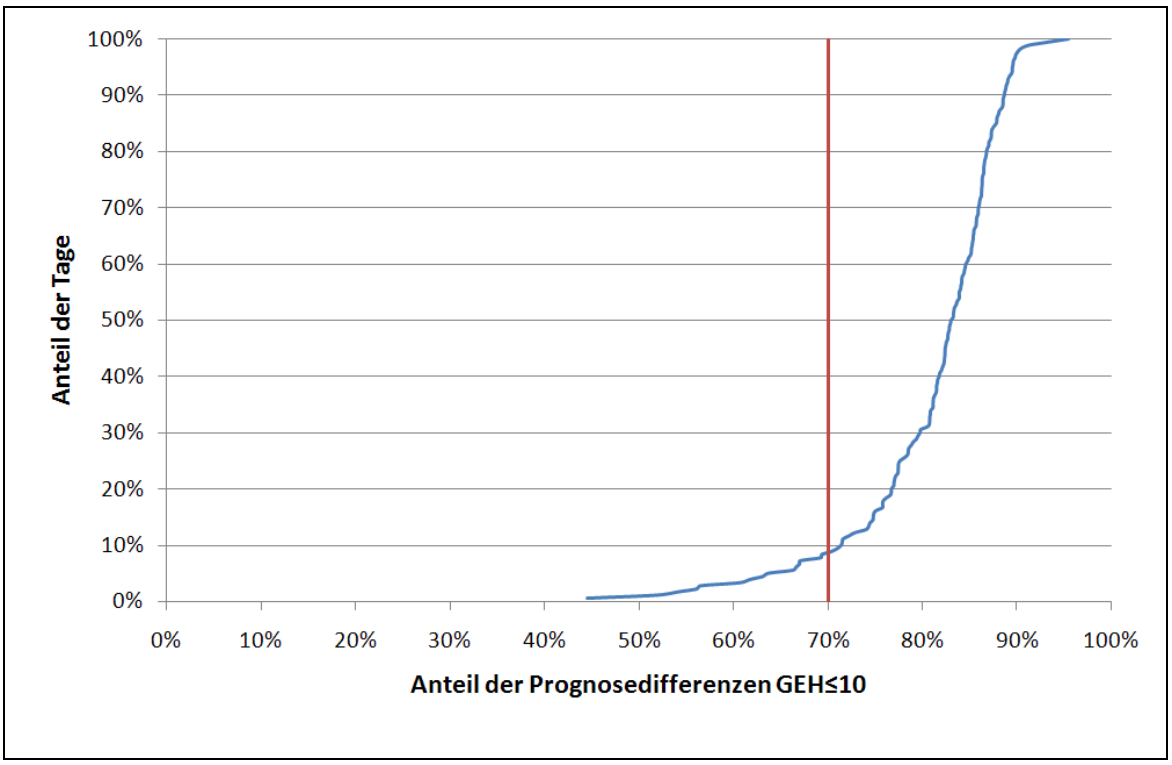


Abbildung 105: Summenverteilung der Tage mit Prognosedifferenz GEH≤10.

5.2 Hochrechnung von Quelle-Ziel-Matrizen

Es wird zwischen zwei Grundtypen von Quelle-Ziel-Matrizen unterschieden. Zum einen klassische QZ-Matrizen die eine Nachfrage in Kfz zu einem bestimmten Zeitpunkt bzw. für einen bestimmten Zeitraum definieren. Diese Matrizen werden im Folgenden als QZ-Matrizen (Kfz) bezeichnet. Die Einheit der Matrizen ist [Kfz/t], wobei t die Dauer des Gültigkeitszeitraums ist. Zum anderen gibt es Matrizen die sich aus den Fahrtbeziehungen der FPD-Trajektorien ableiten (bzw. aus den VPN-Trajektorien deren Start- und Endpunkte Verkehrszellen im VPN entsprechen). Diese Matrizen werden als QZ-Matrizen (FPD) bezeichnet und entstehen durch das Aufsummieren aller Fahrten die innerhalb eines bestimmten Zeitintervalls beginnen. Die Einheit ist [Mobilfunkortsveränderungen/t].

Die QZ-Matrizen (FPD) stehen direkt aus den eingelesenen VPN-Trajektorien zur Verfügung. Hier ist neben dem Aufsummieren aller Fahrtbeziehungen mit gleicher Quelle und gleichem Ziel innerhalb desselben Startzeitintervalls keine weitere Aufbereitung oder Korrektur der Daten notwendig. QZ-Matrizen (Kfz) stehen nicht direkt zur Verfügung und müssen anhand von Zähldaten aus den VPN-Trajektorien oder den QZ-Matrizen (FPD) korrigiert (Kapitel 5.2.1) oder hochgerechnet (Kapitel 5.2.2) werden.

5.2.1 QZ-Matrizen (Kfz) aus Korrektur von QZ-Matrizen (FPD)

Matrixkorrekturverfahren, wie z.B. VStromFuzzy setzen eine umgelegte QZ-Matrix und Zähldaten an bestimmten Strecken im Netz voraus (VISUM Benutzerhandbuch, 2006, Kap. 5.5). Die Verfahren verwenden die Abweichung zwischen Zähldaten und Umlegungsergebnis um die Nachfrage der Fahrtbeziehungen, die eine gezählte Strecke passieren, über einen Korrekturfaktor vorteilhaft anzupassen. Nachfragebeziehungen, die über keine Strecke mit Zähldaten verlaufen, können dabei mit einem mittleren Korrekturfaktor verrechnet werden.

Die notwendige vorangehende Umlegung stellt für diese Verfahren eine zusätzliche mögliche Fehlerquelle der Eingangs-Daten dar. Für die Korrektur von QZ-Matrizen (FPD) ist es möglich diese Umlegung zu umgehen, da die VPN-Trajektorien, aus denen die QZ-Matrizen (FPD) erzeugt werden, bereits einer durchgeführten Umlegung entsprechen. D.h. QZ-Matrix (FPD) und die zugehörige Anfangsumlegung stehen beide direkt zur Verfügung. Allerdings werden die Belastungen nach einer Umlegung der korrigierten Matrix im Allgemeinen nicht die geforderten Zähldaten an den Strecken im Netz erfüllen. Die korrigierte, geänderte Nachfrage führt zu anderen Belastungen im Netz, zu anderen Fahrzeiten auf den Strecken und damit zu anderen Routenwahlentscheidungen. Folglich müssen die Matrixkorrekturverfahren als iterative Prozesse verstanden werden, in denen abwechselnd die QZ-Matrix korrigiert und umgelegt wird,

so lange bis die Abweichungen zwischen gezählten und berechneten Verkehrsstärken innerhalb gewünschter Grenzen liegen oder erkannt wird, dass diese nicht erreicht werden können.

Das Beispielnetz in Abbildung 106 soll die Abhängigkeit von Matrixkorrektur und Umlegung verdeutlichen. Das Netz besteht aus der Zelle 100 und der Zelle 200 zwischen denen 2 unterschiedlich lange Routen existieren (Route 1 oben, Route 2 unten). Die Nachfrage (grüner Balken) von 1.000 Kfz verteilt sich, nach einer Umlegung, entsprechend den blauen Balken mit 44,8% auf Route 1 und 55,2% auf Route 2. An der Strecke oben rechts liegt ein Zählwert von 1000 Kfz vor. Aus diesem ergibt sich für die obere Route, wie folgt ein Korrekturfaktor:

$$K_{1,1} = \frac{Z(1)}{Q(1)} = \frac{1000}{448} \approx 2{,}232$$

mit

$K_{1,1}$ Korrekturfaktor für die Belastung der Route 1 von Zelle 100 nach Zelle 200

$Z(1)$ Zählwert an der Route 1 von Zelle 100 nach Zelle 200

$Q(1)$ Belastung der Route 1 von Zelle 100 nach Zelle 200

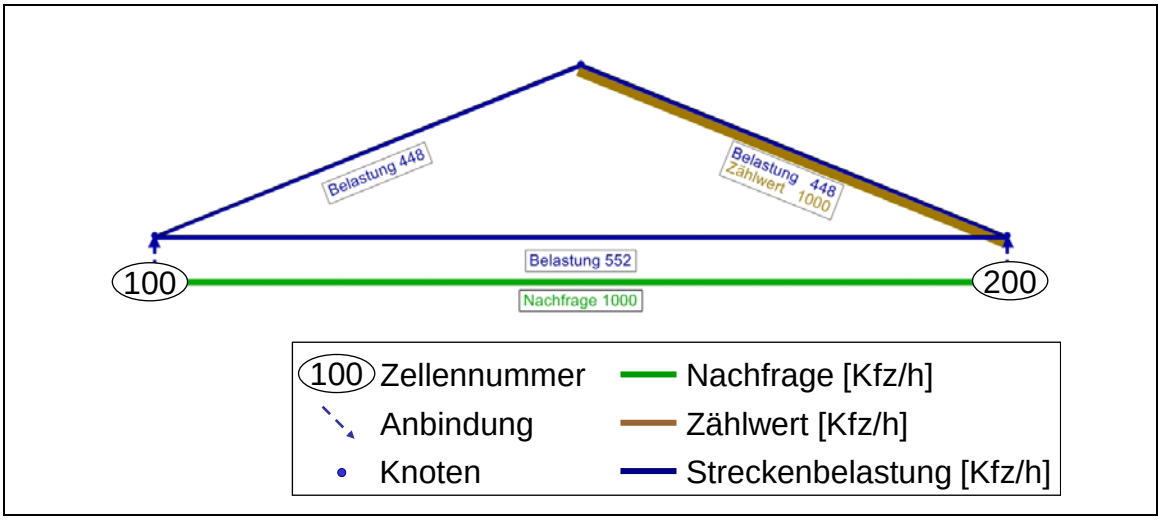


Abbildung 106: Beispielnetz vor der Korrektur durch VStromFuzzy.

Da für die andere Route keine Zählwerte vorliegen, ergibt sich der Gesamtkorrekturfaktor K_1 direkt aus $K_{1,1}$. Die neue Nachfrage beträgt folglich $1.000\text{Kfz} \cdot 2{,}232 = 2.232\text{Kfz}$. Das Ergebnis für die Gleichgewichtsumlegung der korrigierten Matrix ist in Abbildung 107 dargestellt und zeigt, dass der Zählwert nicht exakt getroffen wird, da jetzt auf die Route 1 48% der Nachfrage entfallen.

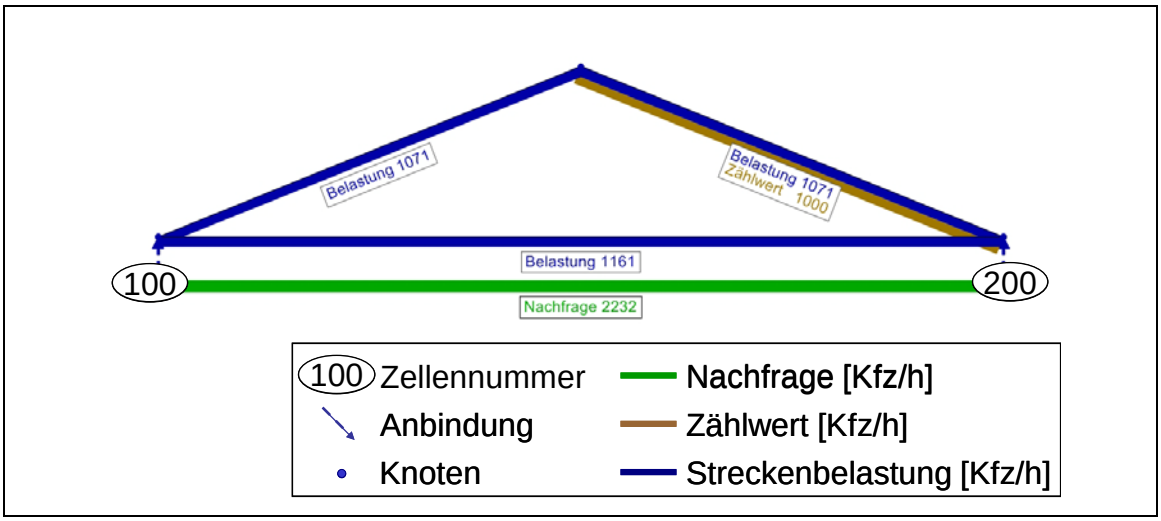


Abbildung 107: Beispielnetz nach der Korrektur durch VStromFuzzy und Umlegung der Ergebnismatrix.

In komplexeren Netzen fällt die Anzahl notwendiger Iterationen höher aus und die gefundene Lösung kann vom verwendeten Umlegungsverfahren abhängen. Außerdem ist es nicht möglich, Weg-Zeit-Informationen in die Matrixkorrektur mit einfließen zu lassen. Diese beiden Punkte sind der Grund für die Entwicklung eines eigenen Hochrechnungsverfahrens im Rahmen des Projekts Do-iT.

5.2.2 QZ-Matrizen (Kfz) aus Hochrechnung von Trajektorien

Zur optimalen Nutzung aller Weg-Zeit-Informationen, die durch die Trajektorien zur Verfügung stehen, wird ein dynamisches Hochrechnungsverfahren entwickelt, das anhand Abbildung 108 erläutert wird. Zudem wird das Verfahren in den folgenden Kapiteln anhand verschiedener Testszenarien in einem überschaubaren Beispielnetz analysiert.

Die dynamische Hochrechnung bestimmt für jede Trajektorie i (am Beispiel der roten hervorgehobenen Trajektorie) einen Hochrechnungsfaktor. Dieser entspricht dem mittleren Korrekturfaktor K_i aller Strecken l mit einem Detektor, die die jeweilige Trajektorie passiert. Ein einzelner Korrekturfaktor $K_{i,j}$ ergibt sich aus der relativen Abweichung zwischen dem Zählwert $Z_{Kfz}(l,t)$ und der Summe der Belastungen aller Trajektorien $Q_{FPD}(l,t)$ die innerhalb desselben Zeitintervalls t diesen Detektor passieren wie folgt:

$$K_{i,j} = \frac{Z_{Kfz}(l,t)}{Q_{FPD}(l,t)}$$

mit

$K_{i,j}$ Korrekturfaktor j für die Trajektorie i, wobei $j \in [1,2,...,n]$ und n der Anzahl passierter, aktiver (d.h. mit Zählwert) Detektoren entspricht

$Z_{Kfz}(l,t)$ Zählwert an der Strecke l im Zeitintervall t

$Q_{FPD}(l,t)$ Summe aller Trajektorienbelastungen an der Strecke l im Zeitintervall t

Der mittlere Korrekturfaktor K_i einer Trajektorie i ergibt sich dann zu:

$$K_i = \frac{1}{n} \cdot \sum_{j=1}^{n} K_{i,j}$$

Nachdem für alle Trajektorien ein Hochrechnungsfaktor bestimmt ist, werden die Belastungen aller Trajektorien mit diesem Faktor multipliziert, wobei die Belastungen von Trajektorien die über keinen Detektor verlaufen mit der mittleren Änderung aller hochgerechneten Trajektorienbelastungen multipliziert werden (können). Infolgedessen ändern sich die Werte für $Q_{FPD}(l,t)$ wodurch in einem iterativen Prozess neue Hochrechnungsfaktoren bestimmt werden können.

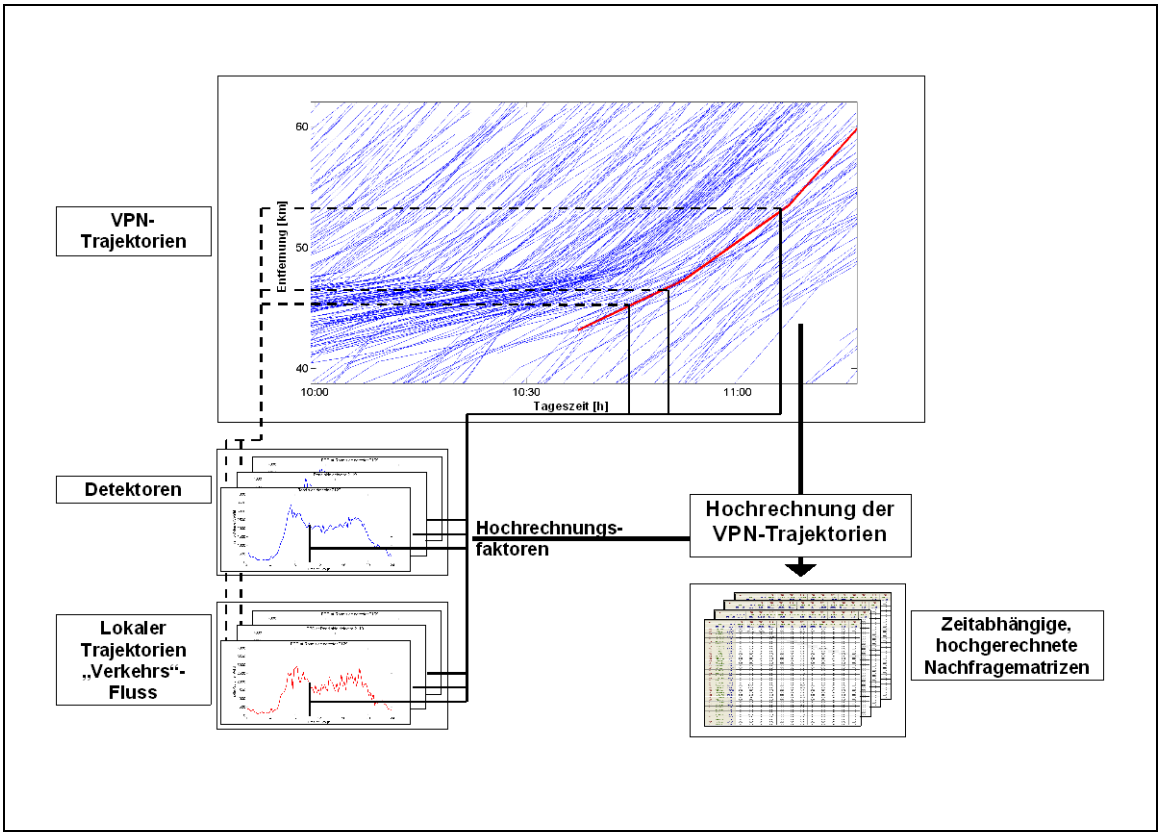


Abbildung 108: Schematische Darstellung der dynamischen Hochrechnung.

Die Anzahl durchzuführender Iterationen hängt dabei davon, was für eine maximale Abweichung zwischen allen $Q_{FPD}(l,t)$ und $Z_{Kfz}(l,t)$ erreicht werden soll und ob diese erreicht werden kann. Liegen widersprüchliche Zähldaten vor oder fehlen für die Hochrechnung notwendige Quelle-Ziel-Beziehungen, so ist es unter Umständen nicht möglich eine gewünschte maximale Abweichung zu erreichen. Der Iterationsprozess des Hochrechnungsverfahrens wird deshalb über die folgenden Werte gesteuert:

- Maximale Anzahl durchzuführender Iterationen
- Erreichen gewünschter maximaler oder mittlerer Abweichung anhand des GEH (siehe auch Kapitel 5.1.2.3)
- Änderungen bestimmter Kennzahlen (Summe aller Fahrten, minimaler/maximaler Hochrechnungsfaktor, relative Abweichung, usw.) sind im Vergleich zur letzten Iteration größer als bestimmter Schwellenwerte.

Nachdem der Iterationsprozess abgeschlossen ist, lassen sich aus den hochgerechneten VPN-Trajektorien, wie in Kapitel 5.1 beschrieben, für beliebige Zeitintervalle Nachfragematrizen aufsummieren. Diese entsprechen aufgrund der Hochrechnung an den Messwerten der Trajektorien einer QZ-Matrix (Kfz).

Die dynamische Hochrechnung verwendet eine diskrete Zeitachse, deren Zeitintervalle frei wählbar sind. Die ebenfalls diskrete Ortsauflösung ergibt sich aus den Strecken des Verkehrsnetzes (VISUM-Modell). Eine Trajektorie befindet sich innerhalb eines bestimmten Zeitintervalls auf einer Strecke, wenn die Trajektorie den Anfangsknoten der Strecke in diesem Zeitintervall passiert. Zähldaten werden entsprechend der Spezifikationen für Ganglinien eingelesen. Im Rahmen des Projekts Do-iT wird die dynamische Hochrechnung einen Tag von 0:00 bis 24:00 Uhr, mit einer zeitlichen Auflösung von 15 Minuten hochrechnen (entsprechend der zeitlichen Auflösung der Gangliniendaten).

Die dynamische Hochrechnung von Trajektorien hat gegenüber einem Matrixkorrekturverfahren den Vorteil, dass keine Umlegungen berechnet werden müssen, d.h. die Routenwahl muss nicht modelliert werden, sondern kann aus den beobachteten VPN-Trajektorien übernommen werden. Die Hochrechnung korrigiert dann lediglich die Belastungen der VPN-Trajektorien entsprechend den Zählwerten. Die Belastungen auf den Strecken des Netzes ergeben sich dadurch direkt aus den VPN-Trajektorien und nicht aus der Umlegung einer QZ-Matrix. Die QZ-Matrix selbst wird aus den Trajektorien aufsummiert.

Es ist zu beachten, dass die VPN-Trajektorien mehr Informationen als ein Umlegungsergebnis und die zugehörige QZ-Matrix enthalten. Die VPN-Trajektorien legen nicht nur die Belastungen auf den Strecken fest, sondern auch Quelle, Ziel, Abfahrts-, Ankunfts- und Zwischenzeitpunkte. Ein weiterer Vorteil der dynamischen Hochrechnung ist, dass durch das Verwenden einer Zeitachse berücksichtigt wird, wie sich die Belastungen im

Netz zeitlich ausbreiten, wodurch die Weg-Zeit-Informationen, die über die Mobilfunk-trajektorien zur Verfügung stehen optimal genutzt werden.

5.2.2.1 Analyse der dynamischen Hochrechnung

Abbildung 109 zeigt das Beispielnetz für die Analyse der dynamischen Hochrechnung.

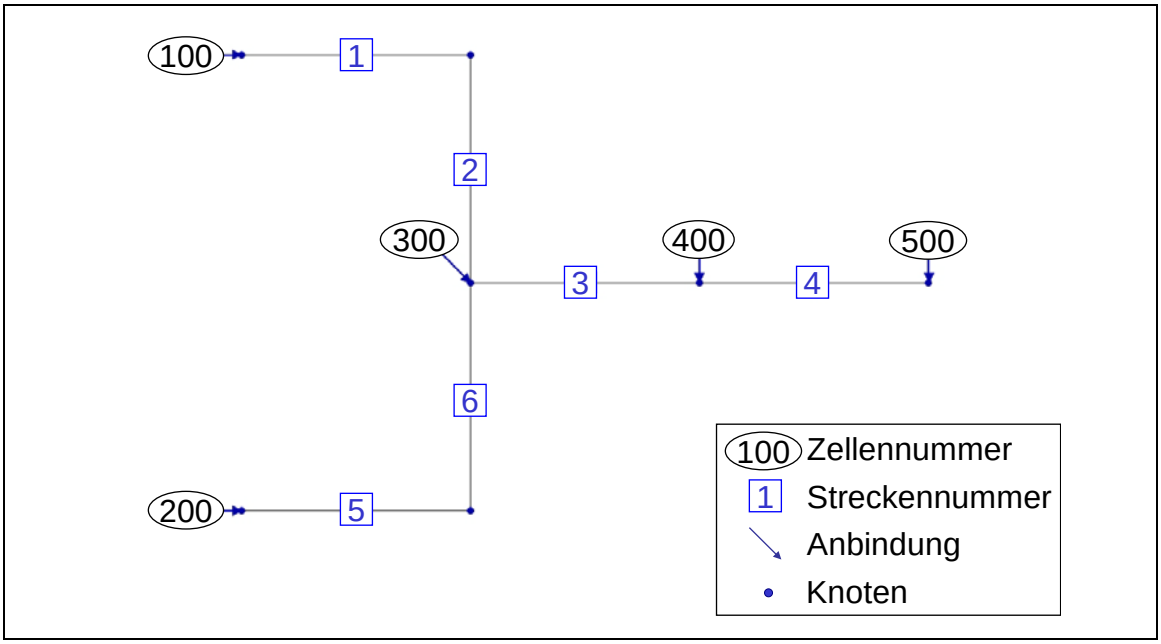


Abbildung 109: Testnetz der Hochrechnungsanalyse.

Alle Strecken sind identisch und haben eine Länge von fünf Kilometern. Die zeitliche Auflösung der dynamischen Hochrechnung ist für alle Tests dieses Kapitel auf 1 Minute eingestellt.

Zur Erzeugung von Eingangsdaten wird das in VISUM erstellte Testnetz nach VISSIM exportiert. Die dort implementierten Auswertungstools erlauben es die Trajektorien aller simulierten Fahrzeuge in eine Access-Datenbank zu exportieren. Diese Daten emulieren für die dynamische Hochrechnung die Mobilfunktrajektorien. Belastungen und Zählwerte werden dabei von Szenario zu Szenario variiert.

5.2.2.1.1 Szenario 1

Im ersten Szenario liegt eine Nachfrage von 1.000 Kfz/h von Zelle 100 nach Zelle 500 vor. Der Simulations- und der Hochrechnungszeitraum betragen jeweils eine Stunde. An den Strecken 1 und 3 liegen fiktive Zählwerte vor, deren Verlauf Abbildung 110 zu entnehmen ist.

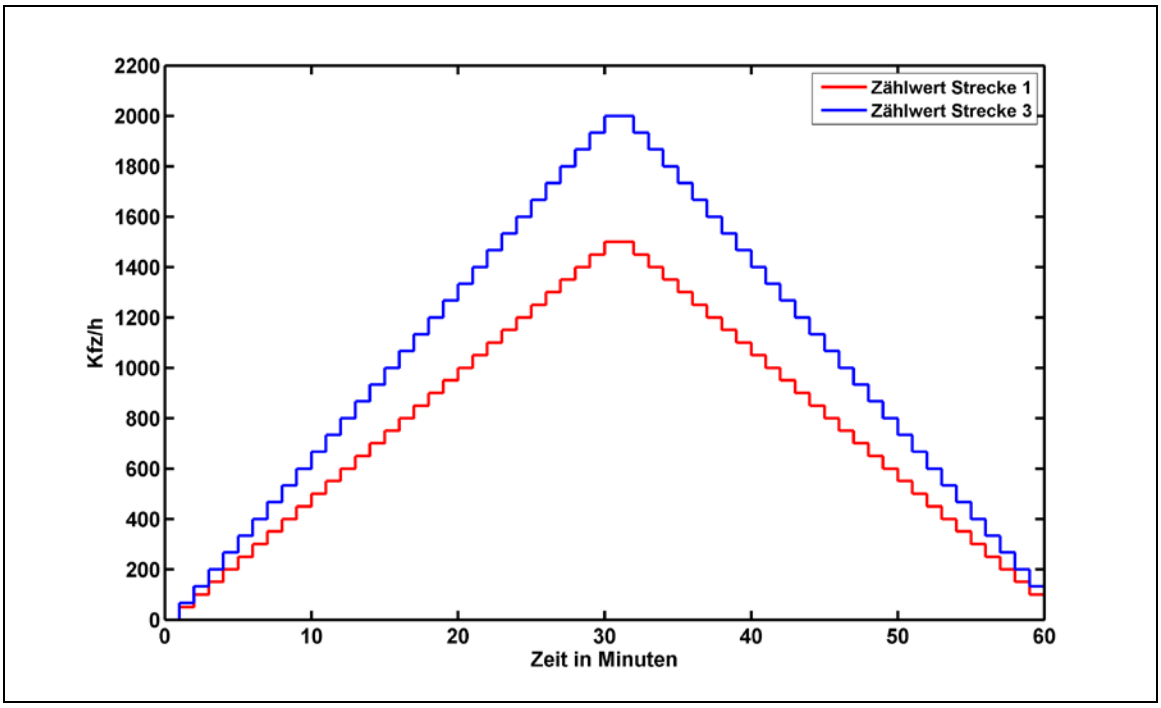


Abbildung 110: Szenario 1, Zählwerte.

Die Belastungen auf den entsprechenden Strecken, die sich durch Aufsummieren der einzelnen Trajektorien aus der Simulation in VISSIM ergeben, ist in Abbildung 111 dargestellt. Von zufälligen Schwankungen und einer Einlaufphase, die sich durch die Null-Belastung der ersten Minuten auf Strecke 3 zeigt, abgesehen, bewegen sich diese um den Nachfragewert von 1.000 Kfz/h.

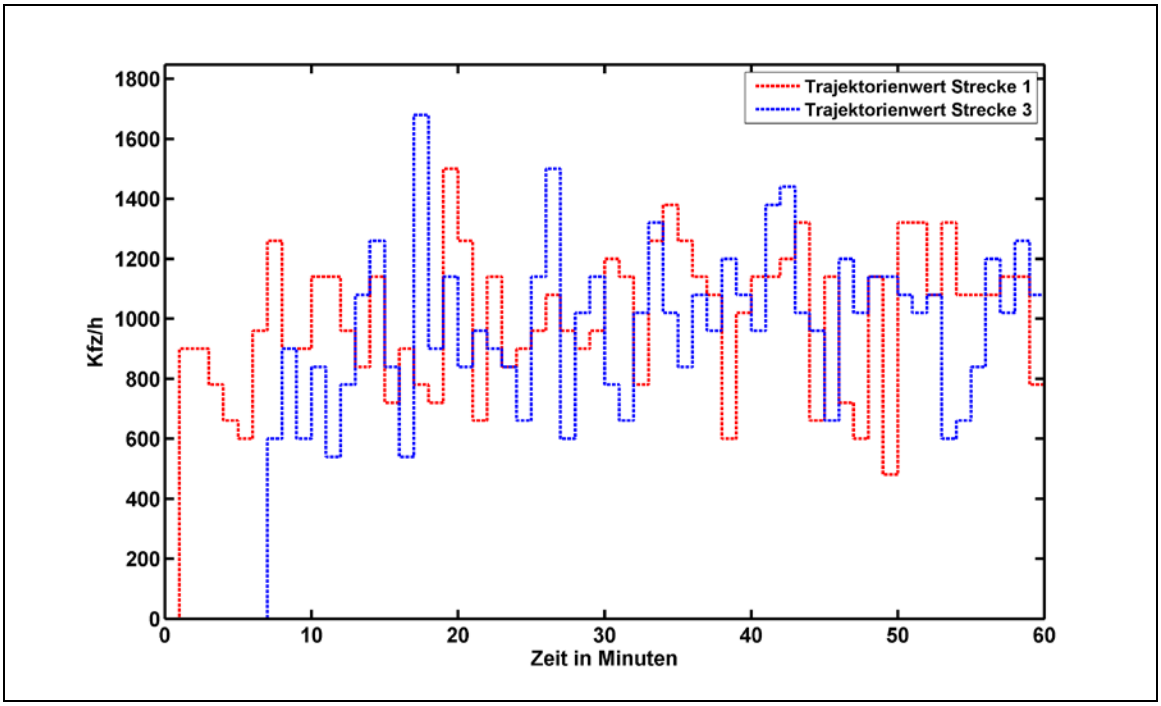


Abbildung 111: Szenario 1, Trajektorienstreckenbelastung, Anfangszustand.

Abbildung 112 und Abbildung 113 zeigen die Zählwerte und die Streckenbelastungen der hochgerechneten Trajektorien nach 10 bzw. 100 Iterationen. Weitere Iterationen führen zu keiner sichtbaren Änderung der Streckenbelastungen.

Es fällt auf, dass sich die Streckenbelastungen an die geforderten Zählwerte angenähert haben, diese aber nicht ganz erreichen. Grund hierfür ist, dass sich Forderungen der Zähldaten unter den gegebenen Rand- und Anfangsbedingungen nicht erfüllen lassen. Zum einen berücksichtigen die fiktiven Zählwerte nicht, dass das Netz zur Minute 0 leer ist und selbst die schnellsten Trajektorien die Strecke 3 erst nach ca. 7 Minuten erreichen. Zum anderen widersprechen sich die Zählwerte. Die Zählwerte auf Strecke 3 sind immer um ca. 33% höher als die der Strecke 1, eine Forderung die von einer einzigen Quelle-Ziel-Relation nicht erfüllt werden kann. Hinzu kommt, dass selbst bei identischen Zählwerten nicht berücksichtigt wird, dass die flussabwärts liegende Strecke eine zeitlich nach hinten verschobene Belastung aufweisen muss (was an den um ca. 6-7 Minuten verschobenen Streckenbelastungs-kurven deutlich wird).

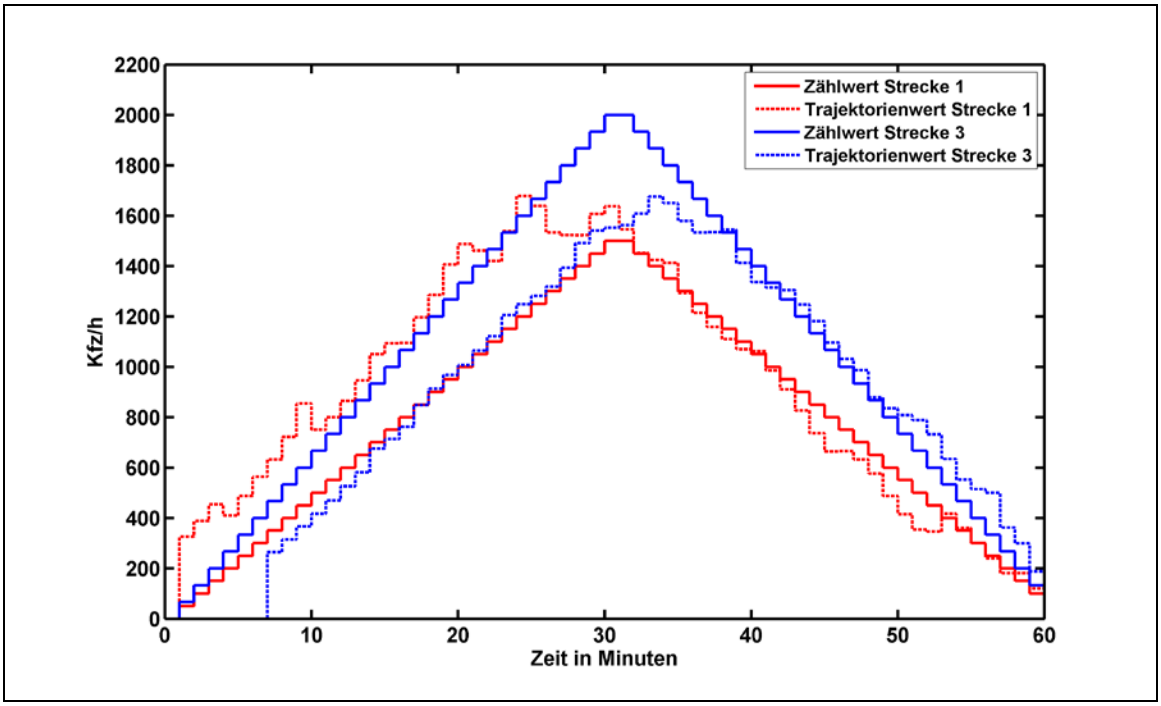


Abbildung 112: Szenario 1, Zählwerte und Trajektorienstreckenbelastung, Iteration 10.

Folgende Punkte können festgestellt werden:

- Die Zählwerte auf Strecke 1 werden ab Minute 53 sehr gut reproduziert. Der Grund hierfür ist, dass keine Trajektorie, die ab Minute 53 startet, die Strecke 3 innerhalb der 60 Minuten erreicht.
- Die Lösung der dynamischen Hochrechnung pendelt sich, unter Berücksichtigung der Fortbewegung der Fahrzeuge im Netz, auf die mittlere Belastung der wider-

sprüchlichen Zählwerte ein. Dies zeigt sich unter Anderem an der Summe aller Fahrten. 903 Kfz/h entsprechen der mittleren Zählwertsumme beider Strecken und 909 Kfz ist die Summe aller Trajektorienbelastungen die innerhalb der 60 Simulationsminuten beginnen.

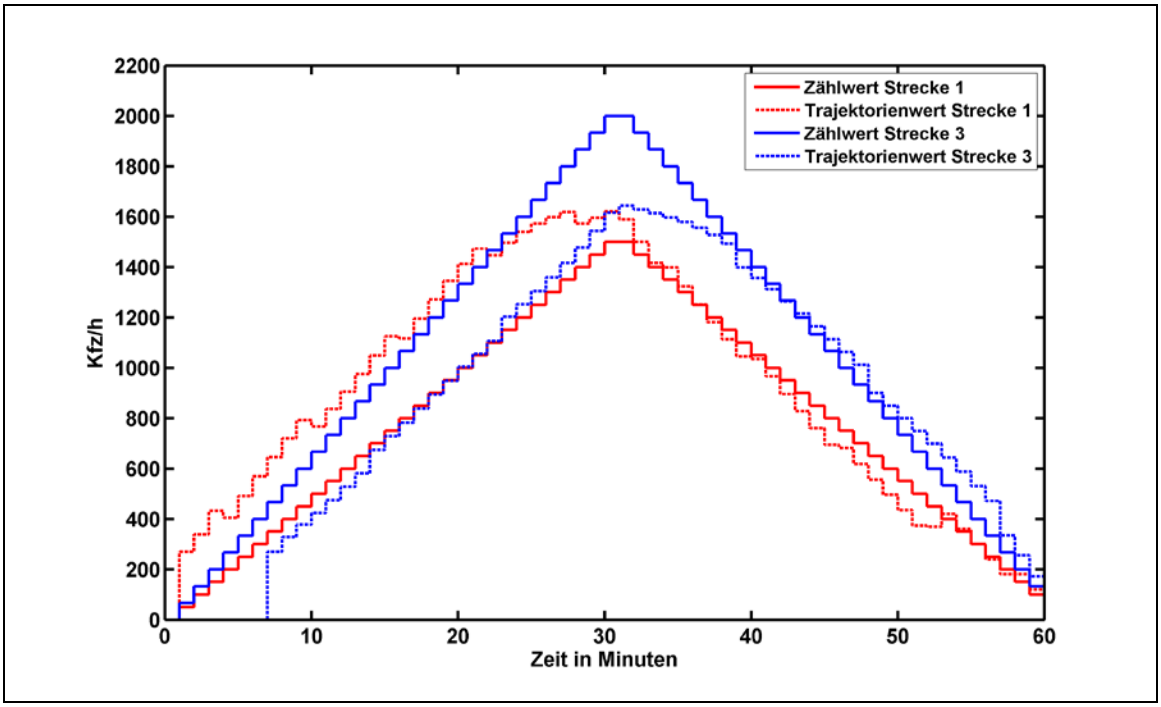


Abbildung 113: Szenario 1, Zählwerte und Trajektorienstreckenbelastung, Iteration 100.

Um das Konvergenzverhalten genauer zu betrachten, zeigt Abbildung 114 den Verlauf der Belastungen aller Trajektorien über die ersten 10.000 Iterationen. Diese sind erst ab ca. der 2.000. Iteration konstant. Die Änderungen sind allerdings in der Form, dass sie sich schon viel früher gegenseitig ausgleichen. Wie zuvor erwähnt, ändert sich die Belastung auf den Strecken schon ab ca. der 100. Iteration nicht mehr merklich. Dies bedeutet, dass die minimale oder maximale Änderung der Trajektorien von Iteration zu Iteration nicht als Abbruchkriterium für die Steuerung des Iterationsprozesses geeignet ist und stattdessen globale, akkumulierte Größen, wie z.B. die Änderung der Streckenbelastung verwendet werden sollten.

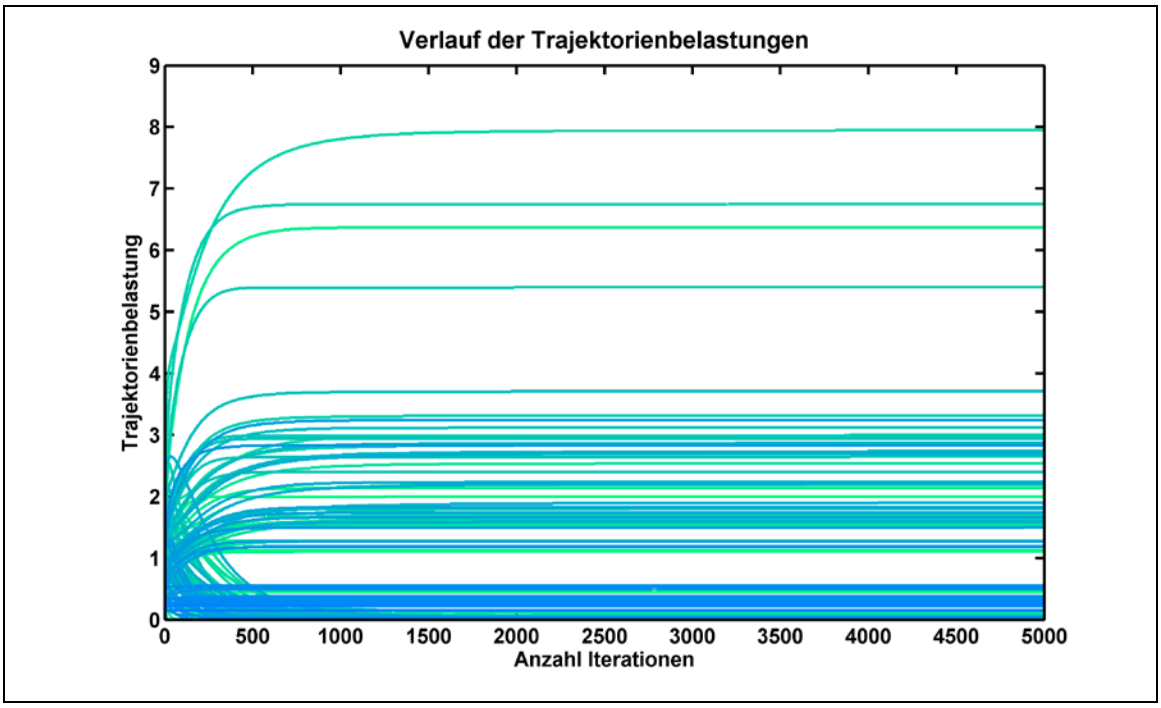


Abbildung 114: Szenario 1, Konvergenzverhalten.

5.2.2.1.2 Szenario 2

Im zweiten Szenario wird die Nachfrage entsprechend Tabelle 56 geändert.

Quellzelle	Zielzelle	Nachfrage [Kfz/h]
100	500	500
200	500	500

Tabelle 56: Szenario 2, Anfangsnachfrage.

Zählwerte liegen auf den Strecken 1, 3 und 5 vor und sind, wie auch die Belastungen der Simulation, Abbildung 115 zu entnehmen. In diesem Szenario widersprechen sich Zählwerte und Anfangsnachfrage nicht. Einzig die nicht berücksichtigte zeitliche Ausbreitung des Verkehrs im Netz verhindert eine exakte Lösung.

Abbildung 116 zeigt die durch die Trajektorien entstehenden Belastungen nach 100 Iterationen. Die Resultate decken sich dabei mit den zuvor gewonnenen Erkenntnissen. Zählwerte von Detektoren die näher am Beginn der Trajektorien liegen, werden am Ende des Hochrechnungszeitraums besser getroffen und die nicht vorhandene zeitliche Verschiebung der Zählwerte verhindert eine exakte Lösung.

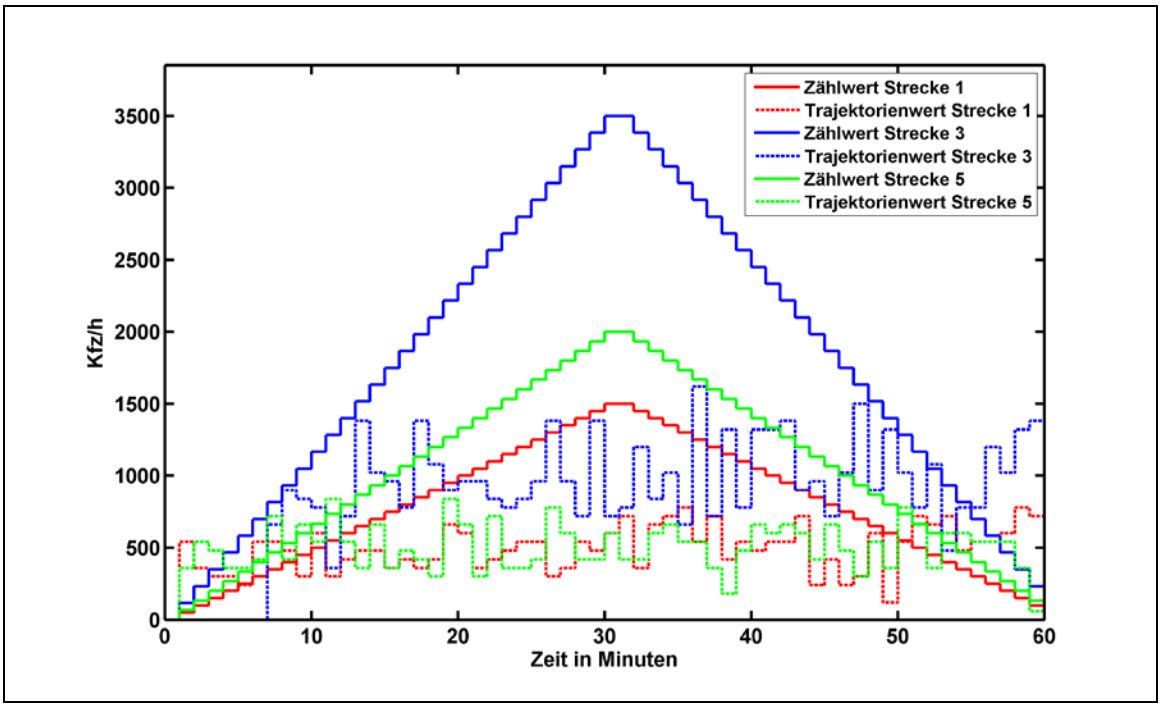


Abbildung 115: Szenario 2, Zählwerte und Trajektorienstreckenbelastung, Anfangszustand.

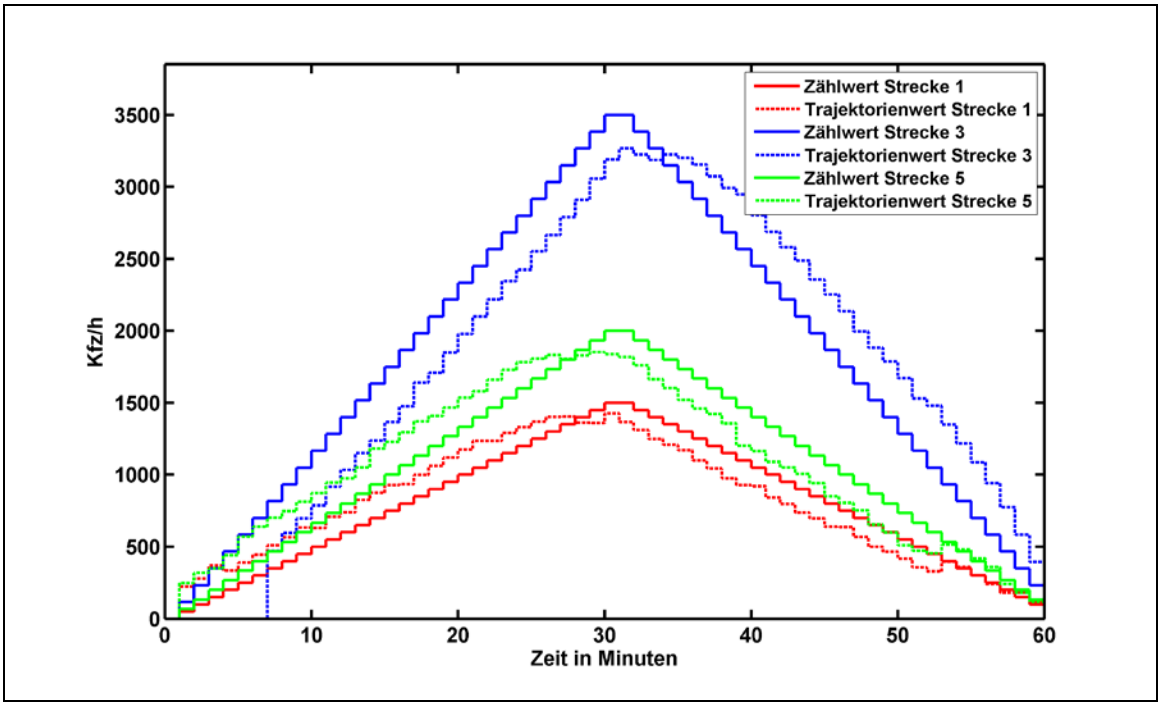


Abbildung 116: Szenario 2, Zählwerte und Trajektorienstreckenbelastung, Iteration 100.

Im Folgenden wird untersucht, wie sich unterschiedliche Rechengenauigkeiten und das Zusammenfassen ähnlicher Trajektorien auf die Qualität der Lösung und die Anzahl notwendiger Iterationen auswirken. Das Zusammenfassen ähnlicher Trajektorien bedeutet, dass die Belastungen von Trajektorien die sich in jedem Zeitintervall auf

derselben Strecke befinden, addiert werden. Anschließend werden alle Trajektorien bis auf eine gelöscht. Dies hat zur Folge, dass redundante Berechnungen der dynamischen Hochrechnung verhindert werden, da ähnliche Trajektorien immer denselben Hochrechungsfaktor erhalten. Die knapp 1.000 Trajektorien des zweiten Szenarios werden dadurch auf 370 unabhängige Trajektorien reduziert.

Die Resultate sind im Bereich der numerischen Genauigkeit identisch. D.h. die Belastungen einzelner Trajektorien ändern sich früher nicht mehr, wenn sie mit einem Wert nahezu identisch 1 multipliziert werden, als dies bei der Summe der Trajektorien der Fall ist. Die Rechenzeit ist aufgrund der geringeren Anzahl einzelner Trajektorien geringer, wobei auf eine exakte Angabe verzichtet wird, da diese nicht nur von der Anzahl Trajektorien abhängt, sondern viel mehr von der durchschnittlichen Länge der Trajektorien und den Verknüpfungen zwischen Trajektorien und Zählstellen.

Das Reduzieren der Rechengenauigkeit, d.h. Speichern der Trajektorienbelastungen mit einer geringeren Anzahl Dezimalstellen (bisher Datentyp Double) soll bewirken, dass sich die Belastungen der Trajektorien schneller auf einen stabilen Wert einstellen. Abbildung 117 bis Abbildung 119 zeigen den Verlauf der Belastungen aller Trajektorien über die Anzahl Iterationen für die Rechengenauigkeiten Double, eine Dezimalstelle und Integer. Hier ist zu beachten, dass die Skalierung der Achsen variiert.

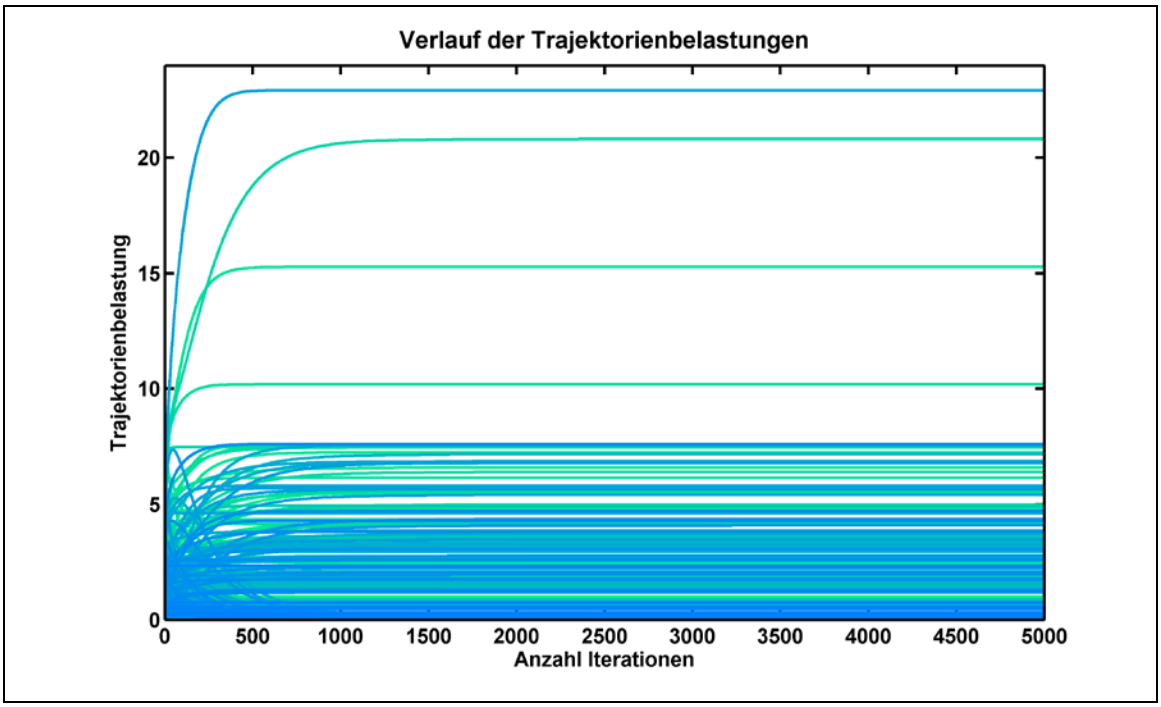


Abbildung 117: Szenario 2, Konvergenzverhalten bei hoher Rechengenauigkeit (Trajektorienbelastungswert als Datentyp Double).

Aus den Abbildungen wird ersichtlich, dass die Anzahl Iterationen, die dafür notwendig ist, dass die Belastungen konstante Werte annehmen, mit abnehmender Genauigkeit sinkt. Ganz allgemein reduzieren sich die Abweichungen zwischen Zählwerten und

Trajektorienbelastungen mit jeder Iteration. Dies hat zur Folge, dass sich die Hochrechnungsfaktoren dem Wert 1 annähern. Für kleinere Rechengenauigkeiten bedeutet dies, dass eine Änderung schneller abgeschnitten wird und die Trajektorienbelastungen früher einen konstanten Wert annehmen.

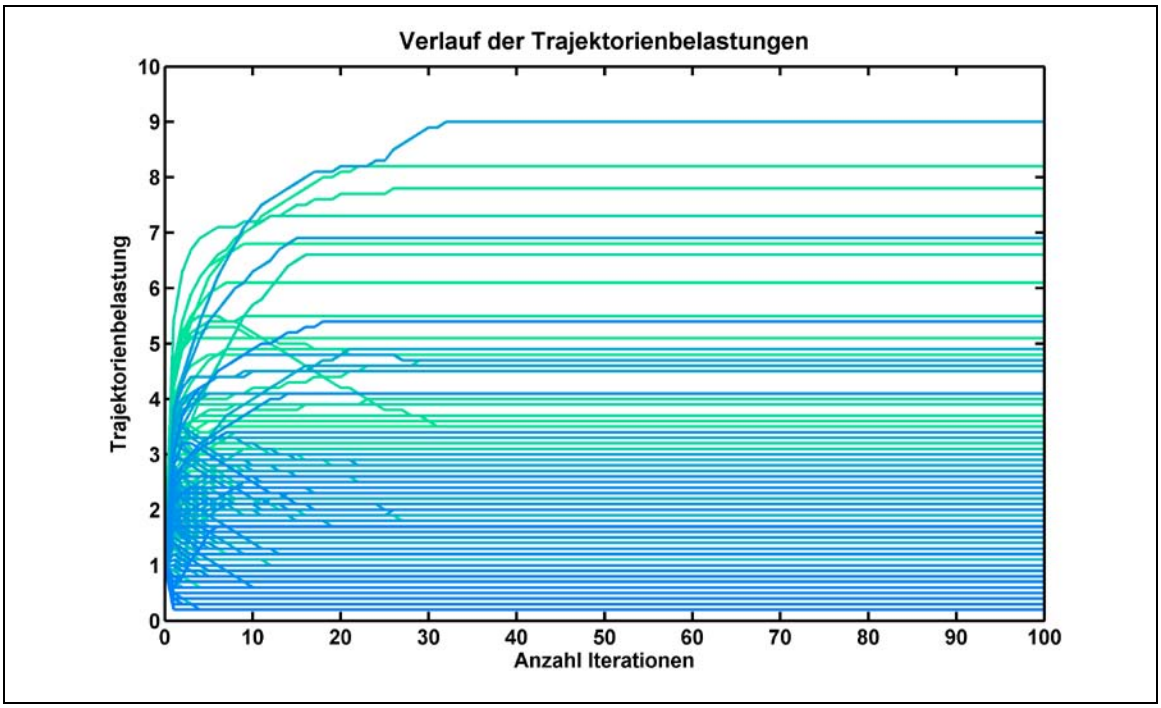


Abbildung 118: Szenario 2, Konvergenzverhalten bei mittlerer Rechengenauigkeit (Trajektorienbelastungswert wird auf eine Dezimalstelle genau gespeichert).

Des Weiteren fällt auf, dass die maximalen Trajektorienbelastungen für größere Rechengenauigkeiten höhere Werte erreichen. Der Grund hierfür ist derselbe wie beim schnelleren Erreichen konstanter Werte. Die geforderten Zählwerte dieses Szenarios lassen sich aufgrund der nicht vorhandenen zeitlichen Verschiebung der Zählwerte nicht erfüllen. D.h. die Hochrechnungsfaktoren werden den Wert 1 nie ganz erreichen. Manche Trajektorienbelastungen werden mit einem Hochrechnungsfaktor kleiner als 1 multipliziert und werden mit jeder Iteration kleiner. Andere Belastungen werden im selben Maß immer wieder mit einem Faktor größer 1 multipliziert und steigen stetig an. Die Änderungen werden dabei immer kleiner, wie zuvor festgestellt ist die Summe aller Belastungen schon sehr viel früher nahezu konstant. Irgendwann erreicht allerdings auch die Rechengenauigkeit des Zahlentyps Double ihre Grenzen, die Änderungen werden verschluckt und die Werte bleiben konstant.

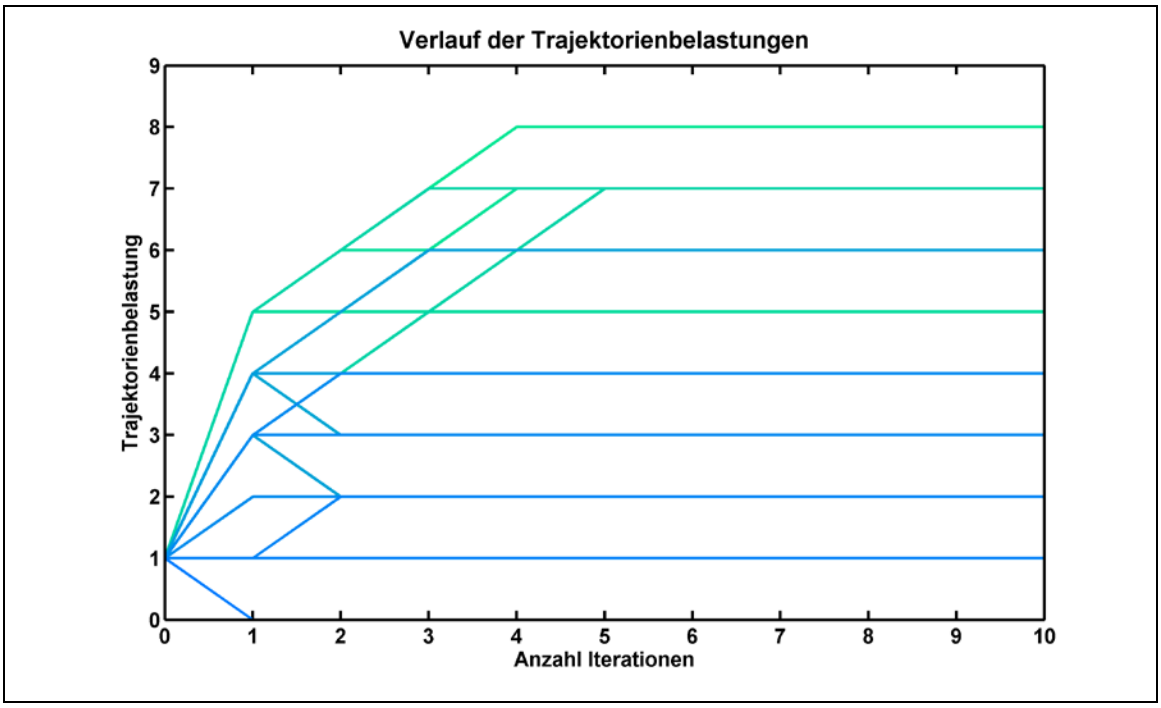


Abbildung 119: Szenario 2, Konvergenzverhalten bei niedriger Rechengenauigkeit (Trajektorienbelastungswert als Datentyp Integer).

Die Verläufe der aus den Trajektorien resultierenden Streckenbelastungen sind in Abbildung 120 und Abbildung 121 dargestellt. Es zeigt sich, dass die Qualität der Lösungen im Vergleich zur vollen Genauigkeit (Abbildung 116) niedriger ausfällt und die Abweichungen zu den Zählwerten größer bleiben (siehe dazu auch Tabelle 57). Infolgedessen werden Ansätze, die die Rechengenauigkeit reduzieren, um eine schnellere Konvergenz zu erzwingen, wieder verworfen.

	Double	Zehntel	Integer
Summe Fehlerquadrat [Kfz^2/min^2]	2.768,046	3.225,388	5.188,545

Tabelle 57: Summe der Fehlerquadrate für verschiedene Rechengenauigkeiten.

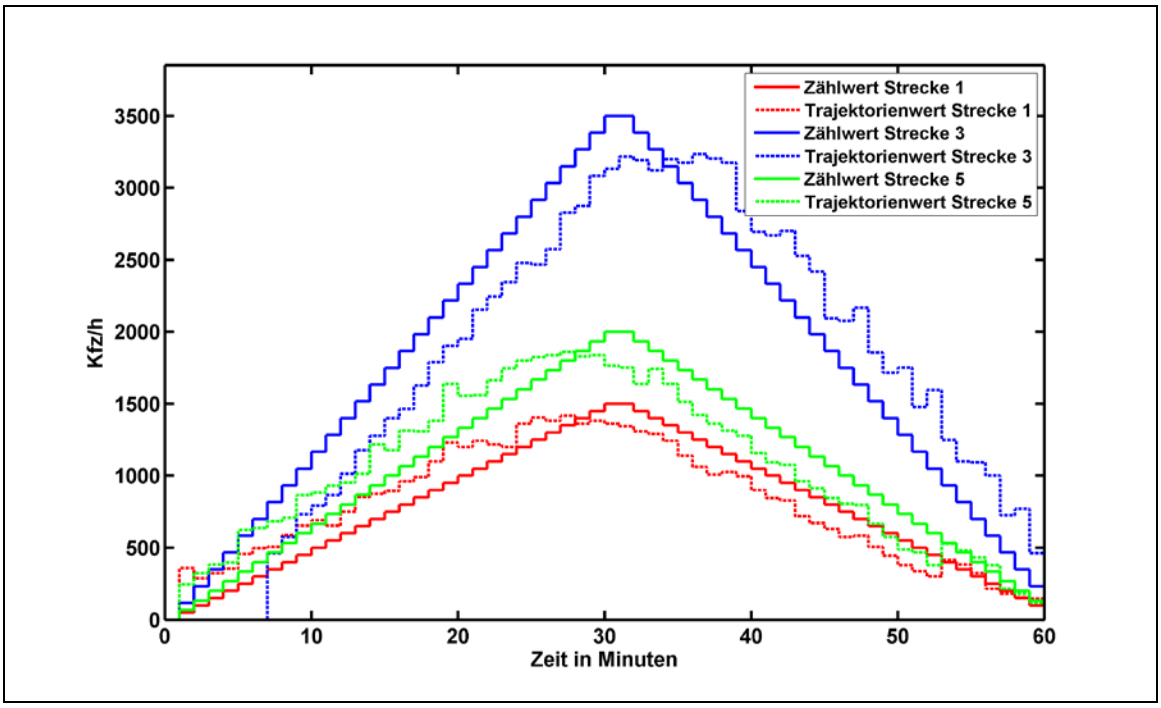


Abbildung 120: Szenario 2, Zählwerte und Trajektorienstreckenbelastung, Zehntel, stabile Lösung.

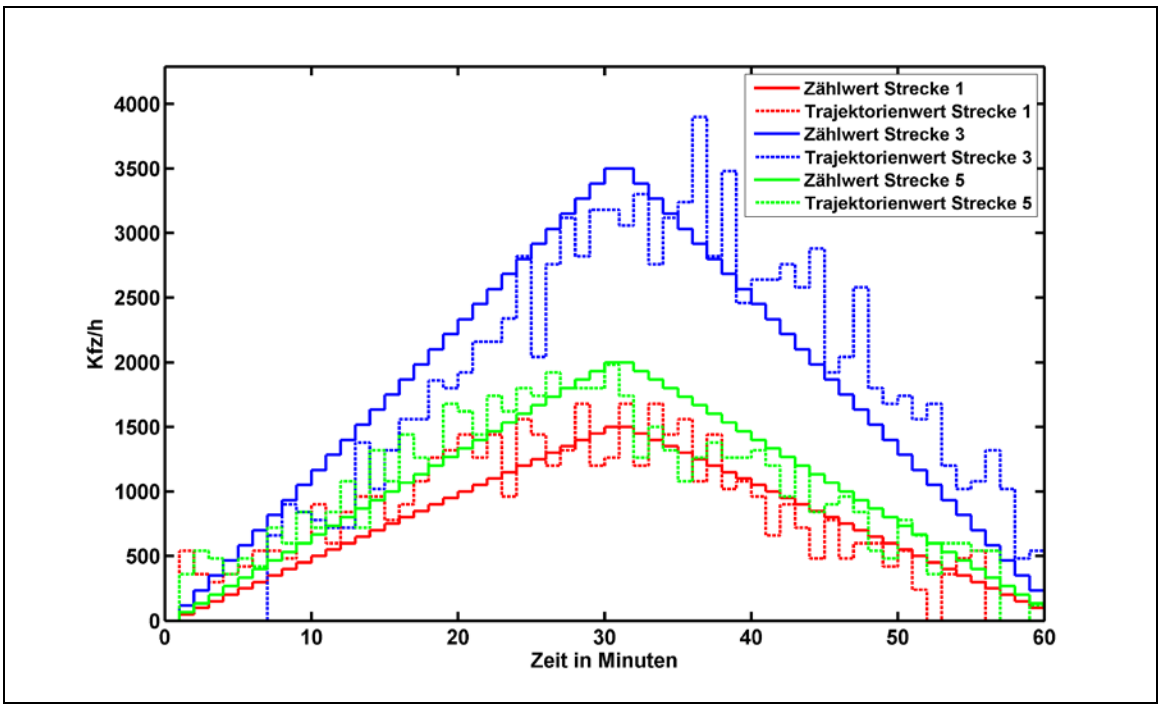


Abbildung 121: Szenario 2, Zählwerte und Trajektorienstreckenbelastung, Integer, stabile Lösung.

5.2.2.1.3 Szenario 3

Für dieses Szenario wird die Zahl der belegten Quelle-Ziel-Relationen erhöht. Die Nachfrage wird in 15-Minuten-Intervallen schwanken (siehe Tabelle 58). Um Zähldaten zu erzeugen, die die räumliche Ausbreitung im Netz korrekt abbilden, werden die Streckenbelastungen der mikroskopischen Simulation verwendet und mit einem Faktor manipuliert. Die Strecken mit Zählwerten sind für dieses Szenario die Strecken 2, 4 und 6.

Insgesamt werden 3 verschiedene Teilszenarien untersucht:

- Änderung der Zähldaten, Hochrechnung aller Trajektorien
- Änderung der Zähldaten, nur Hochrechnung der Trajektorien mit eigener Information
- Änderung der Trajektorien, Hochrechnung aller Trajektorien

In den ersten beiden Fällen wird, wie in den vorigen Kapiteln, versucht die Belastungen der Trajektorien an die Vorgaben bestimmter Zähldaten anzupassen. In diesem Fall bedeutet das, die Streckenzählwerte der mikroskopischen Simulation zu verändern. Im Vergleich zu den vorangegangen Szenarien existieren nun zum ersten Mal Quelle-Ziel-Beziehungen und damit Trajektorien, die über keine Zählstelle verlaufen und folglich nicht hochgerechnet werden können (die Trajektorien von Zelle 300 nach Zelle 400 verlaufen nur auf der Strecke 3, auf der in diesem Szenario keine fiktiven Zählwerte angenommen werden). Abhängig von den vorliegenden Gegebenheiten kann es sinnvoll sein, diese Trajektorien mit dem mittleren Hochrechnungsfaktor der anderen Trajektorien mit hochzurechnen. Für das Projekt Do-iT und die Verwendung von Mobilfunktrajektorien als Eingang ist davon auszugehen, dass alle Quelle-Ziel-Beziehungen in ähnlichem Maße über- oder unterschätzt werden und es sinnvoll ist alle Trajektorien hochzurechnen. Die vorgestellten Tests dienen lediglich zur Verdeutlichung dieser Option.

Der dritte Testfall ändert als Eingangsgröße nicht die Zähldaten sondern die Anzahl erfasster Trajektorien. D.h. es werden zufällig Trajektorien gelöscht, die der Hochrechnung nicht zur Verfügung stehen und das Ziel ist es, die ursprünglichen Streckenbelastungen wieder zu reproduzieren.

Von Zelle	Nach Zelle				
0:00-0:15 Uhr	**100**	**200**	**300**	**400**	**500**
100	0	0	75	75	38
200	0	0	56	56	56
300	0	0	0	38	38
400	0	0	0	0	75
500	0	0	0	0	0
0:15-0:30 Uhr	**100**	**200**	**300**	**400**	**500**
100	0	0	76	76	38
200	0	0	57	57	57
300	0	0	0	38	38
400	0	0	0	0	76
500	0	0	0	0	0
0:30-0:45 Uhr	**100**	**200**	**300**	**400**	**500**
100	0	0	180	180	300
200	0	0	150	150	300
300	0	0	0	120	120
400	0	0	0	0	240
500	0	0	0	0	0
0:45-1:00 Uhr	**100**	**200**	**300**	**400**	**500**
100	0	0	19	19	9
200	0	0	14	14	14
300	0	0	0	9	9
400	0	0	0	0	18
500	0	0	0	0	0

Tabelle 58: Szenario 3, Nachfragedaten [Kfz/15min] (gerundet).

Analyse der Varianten mit und ohne Hochrechnen von Trajektorien ohne Informationen

Abbildung 122 zeigt den Verlauf der Zählwerte und der Streckenbelastungen aus den Trajektorien vor der Hochrechnung. Die Zählwerte entsprechen dabei dem halben Wert der Messungen aus der mikroskopischen Simulation. Hier ist zu beachten, dass in VISSIM die Funktion der Streckenauswertung verwendet wird. Diese berücksichtigt alle Fahrzeuge die sich innerhalb eines definierten Streckenabschnitts befinden. Im Gegensatz dazu entsteht die Kurve der Trajektorienbelastungen aus den Zeitpunkten an denen eine Trajektorie den Anfangsknoten einer Strecke passiert. Dies hat zur Folge, dass die Zählwertkurven im Vergleich zu den Trajektorienstreckenbelastungs-kurven geglättet sind.

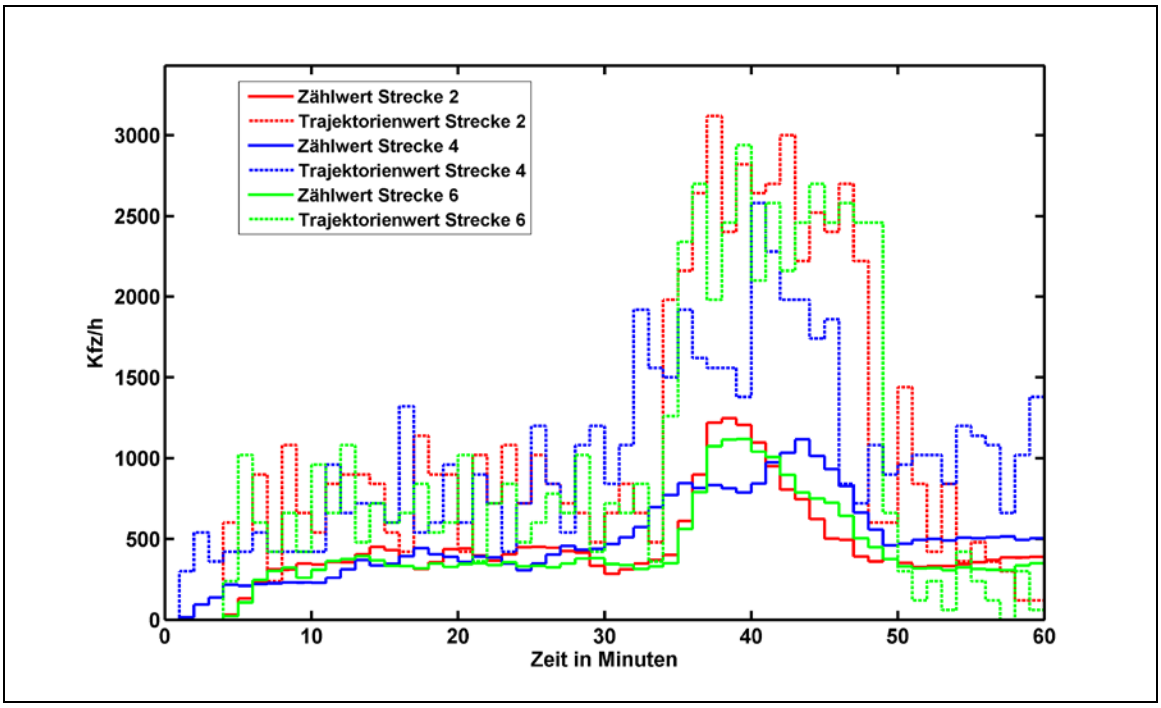


Abbildung 122: Szenario 3_12, Zählwerte und Trajektorienstreckenbelastung, Anfangszustand.

In Abbildung 123 (andere Y-Achsen-Skalierung) ist das Ergebnis der Hochrechnung nach 10 Iterationen dargestellt. Im Vergleich zu den vorigen Szenarien stellt sich eine gute Lösung sehr viel schneller ein, da eine exakte Lösung, aufgrund einer realistischen zeitlichen Verschiebung der Zählwerte, existiert. Es zeigt sich, dass die Anzahl notwendiger Iteration sehr viel mehr von Rand- und Anfangsbedingungen als von Rechengenauigkeiten oder der Anzahl Trajektorien (für dieses Szenario 2.870) abhängt.

Die Belastung von 0 Kfz/h in der 57. Minute auf Strecke 6 erklärt sich dadurch, dass in diesem Zeitintervall keine Trajektorie den Startknoten dieser Strecke passiert. Auf diese Problematik wird im folgenden Beispiel detaillierter eingegangen.

Der Verlauf der Belastungen auf den Strecken mit Zählwerten ist dabei unabhängig davon, ob Trajektorien ohne eigene Informationen mit hochgerechnet werden. Im vorliegenden Szenario ist die einzige Quelle-Ziel-Relation, die über keine Strecke mit Zählwerten verläuft, die von Zelle 300 nach Zelle 400. D.h. zur genaueren Betrachtung dieser Option ist die Strecke 3 von Bedeutung (Abbildung 124).

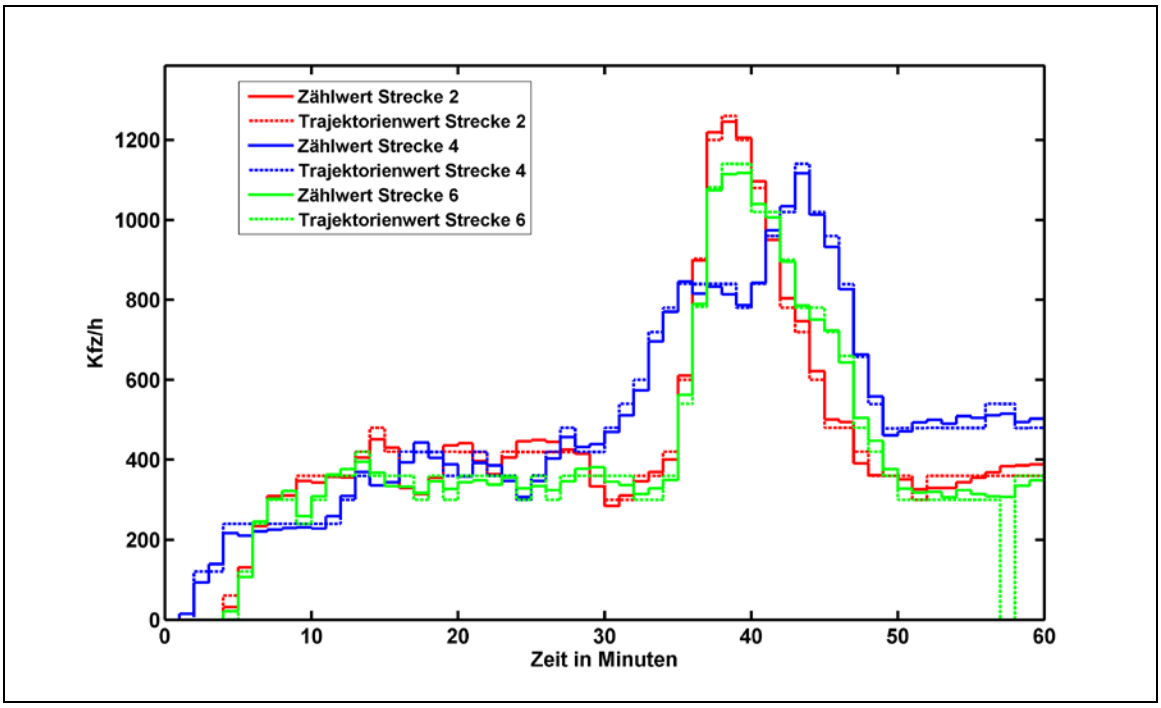


Abbildung 123: Szenario 3_12, Zählwerte und Trajektorienstreckenbelastung, Iteration 10.

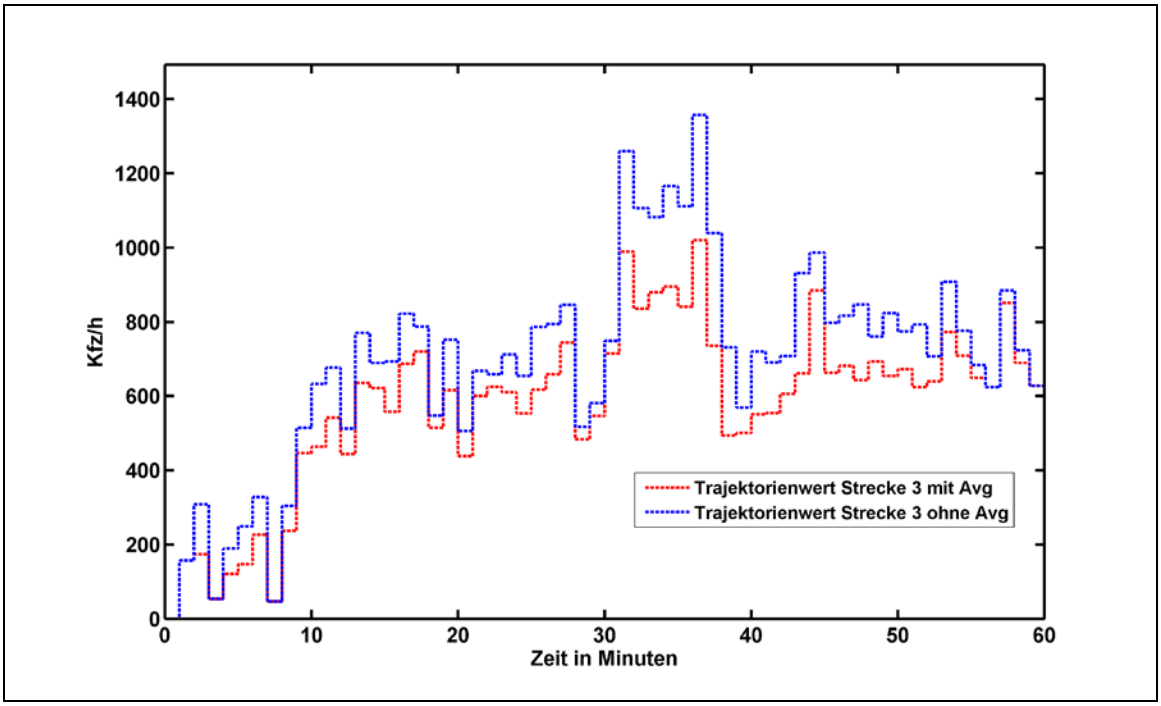


Abbildung 124: Szenario 3_12, Trajektorienstreckenbelastung mit und ohne Hochrechnung aller Trajektorien, Iteration 10.

Hier wird deutlich, dass die Belastung auf der Strecke Nummer 3 höher ist, wenn die Trajektorien ohne eigene Informationen nicht hochgerechnet werden (die Zählwerte sind niedriger als die Anfangsbelastungen). Die Abweichung entspricht dabei dem

Anteil an der Belastung auf Strecke 3 den die Trajektorien ohne Informationen im Anfangszustand haben (ca. 11%), da genau diese Trajektorien unverändert bleiben.

Für das Projekt Do-iT in dem im gesamten Netz verteilte Zählstellen und ungefähr gleichmäßig über- oder unterschätzte Mobilfunktrajektorien vorliegen ist eine Hochrechnung aller Trajektorien mit Sicherheit die sinnvollere Variante.

Analyse der Hochrechnung mit zufällig ausgewählten Trajektorien

Mit diesem Beispiel soll untersucht werden, wie sich die Hochrechnung verhält, wenn zufällig Trajektorien aus der Menge zur Verfügung stehender Trajektorien gelöscht werden. Diese Variante hat große Ähnlichkeiten zum späteren Anwendungsfall, da die Mobilfunktrajektorien nicht alle durchgeführten Fahrten im VPN widerspiegeln werden.

Aus den 2.870 Trajektorien der mikroskopischen Simulation werden zufällig 1.675 (58%) als Eingang für die Hochrechnung ausgewählt. Die Zähldaten entsprechen den Werten der mikroskopischen Simulation und sind doppelt so hoch wie in den vorangegangenen Tests. Die jeweiligen Verläufe sind Abbildung 125 zu entnehmen.

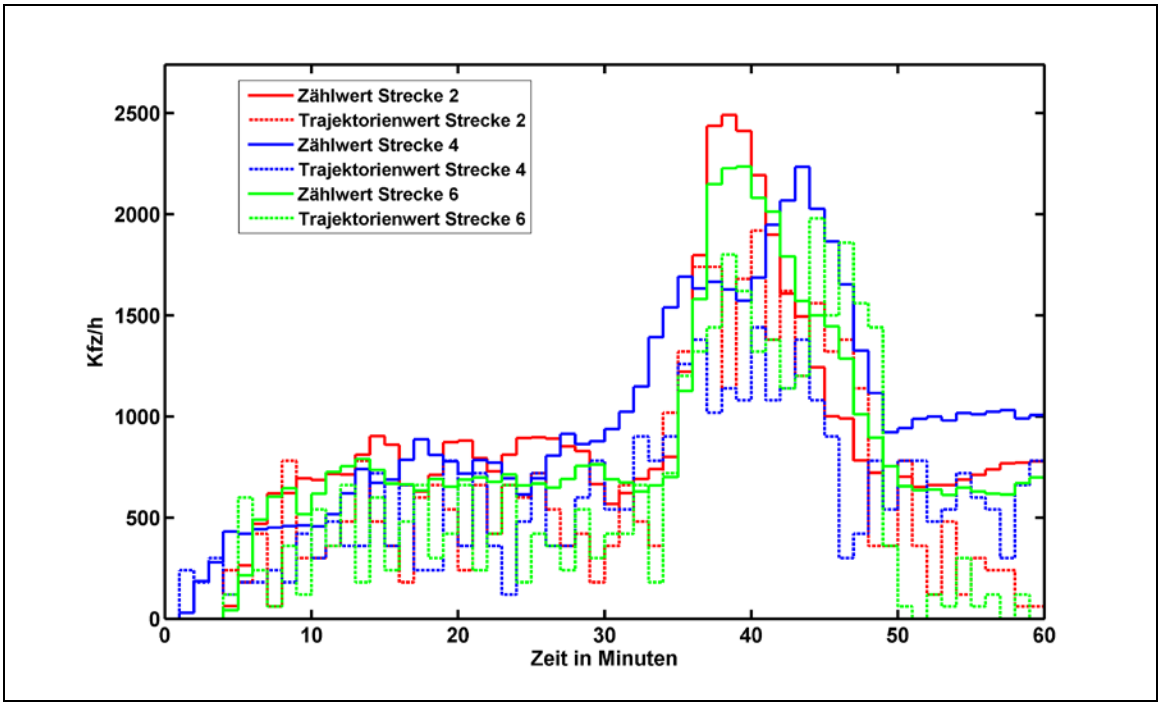


Abbildung 125: Szenario 3_3, Zählwerte und Trajektorienstreckenbelastung, Anfangszustand.

Abbildung 126 zeigt die Streckenbelastungen der Trajektorien nach der 10. Iteration der Hochrechnung. Ein Zustand, der die Zähldaten sehr gut reproduziert, ist wieder sehr schnell gefunden. Auffällig sind die 0-Belastungen auf Strecke 6 in den Minuten 51, 57 und 59. Diese sind, wie Abbildung 125 zeigt, Teil des Anfangszustands und

entstehen durch den unglücklichen Zufall, dass, verstärkt durch die zufällige Entfernung von Trajektorien, in diesen Minuten keine Trajektorie den Anfangsknoten der Strecke 6 passiert. Infolgedessen gibt es keine Trajektorie, die von der Hochrechnung verwendet werden kann um sich den Zählwerten anzunähern. D.h. abhängig von den zur Verfügung stehenden Eingangsdaten ist es nicht immer möglich eine exakte Lösung zu finden, was wiederum für die Steuerung des Iterationsprozess von Bedeutung ist.

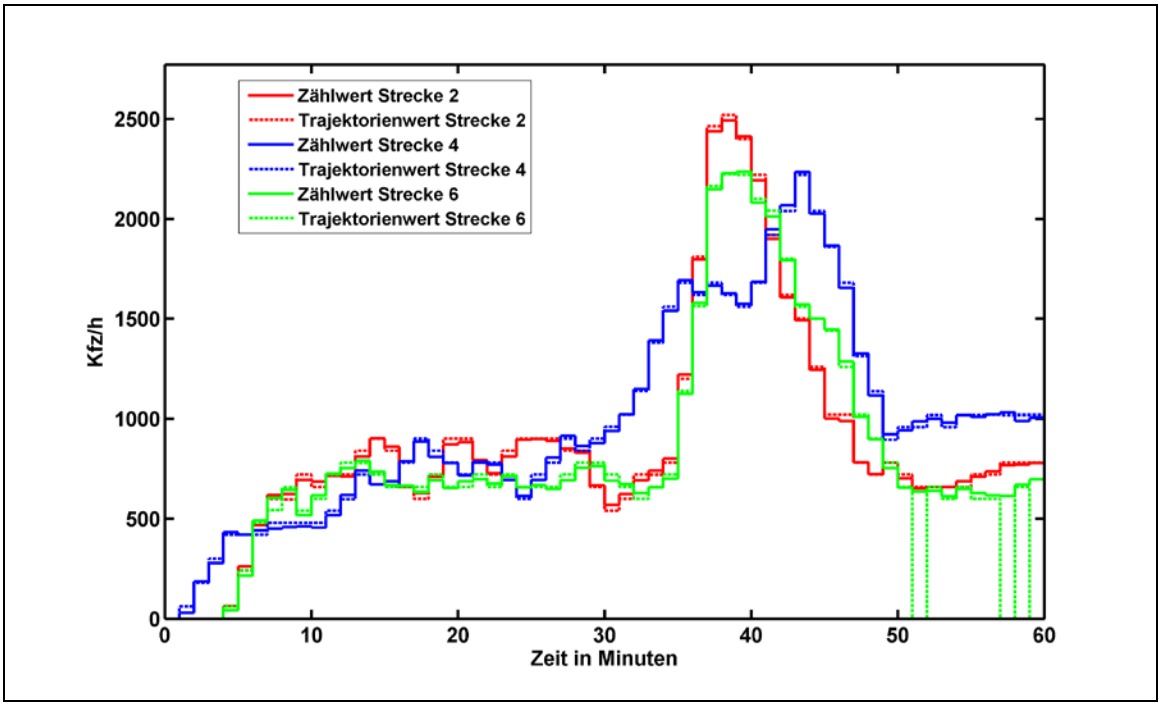


Abbildung 126: Szenario 3_3, Zählwerte und Trajektorienstreckenbelastung, Iteration 10.

5.2.2.2 Randwertproblem

Für mehrere Hochrechnungen über einen Zeitraum von mehreren Tagen ergibt sich ein Randwertproblem. Die dynamische Hochrechnung hat gegenüber statischen Verfahren den Vorteil, dass die räumlich-zeitliche Ausbreitung des Verkehrs im Netz berücksichtigt wird. Allerdings ist es nur möglich, dies innerhalb eines bestimmten (Hochrechnungs-)Zeitraums zu berücksichtigen. Am Beispiel eines Tages B von 0:00 bis 24:00 Uhr werden Trajektorien existieren, die am Vortag A beginnen und erst z.B. um 1:00 Uhr an Tag B ihr Ziel erreichen. Genauso wird es Trajektorien geben die um 23:45 Uhr an Tag B beginnen und erst irgendwann am folgenden Tag C ihr Ziel erreichen. Dabei ist es möglich, dass die Trajektorien an unterschiedlichen Tagen, d.h. in unterschiedlichen Hochrechnungszeiträumen, eine Zählstelle passieren und sich für die unterschiedlichen Hochrechnungszeiträume verschiedene Hochrechnungsfaktoren ergeben. Prinzipiell gibt es 3 Möglichkeiten dieses Problem zu behandeln:

- Den Hochrechnungszeitraum maximieren. D.h. berücksichtigen aller Trajektorien und aller Zählstellendaten. Dieser Ansatz stößt allerdings an praktische Grenzen, denn weder Speicher noch Rechenzeit stehen unbegrenzt zur Verfügung.
- Eine weitere Iterationsschleife auf den Hochrechnungsprozess aufsetzen. In dieser werden unterschiedliche Hochrechnungszeiträume abwechselnd berechnet, wobei das Ergebnis jeder Hochrechnung den Eingang der nächsten beeinflusst. Dieser Ansatz wird enorme Rechenzeit in Anspruch nehmen und muss für jeden Tag an dem neue Daten hinzukommen, komplett wiederholt werden.
- Trajektorien einem Tag fest zuweisen, z.B. über den Abfahrtszeitpunkt. In diesem Fall wird in Kauf genommen, dass nicht alle Trajektorien mit den bestmöglichen Daten hochgerechnet werden. Wenn eine Trajektorie erst am nachfolgenden Tag eine Zählstelle passiert, so wird diese nur mit dem durchschnittlichen Faktor aller anderen hochgerechnet. Diese Variante ist die Praktikabelste und wird im Rahmen des Projekts Do-iT zur Anwendung kommen.

5.2.2.3 Steuerung des Iterationsprozesses

Zur Steuerung bzw. Terminierung eines Iterationsprozesses ist es notwendig Abbruchbedingungen zu definieren. Für den hier betrachteten Prozess soll das Erreichen der folgenden Zustände als Abbruchkriterium genügen:

1. Erreichen des gewünschten Ziels durch Erfüllen eines bestimmten Gütekriteriums.
2. Erreichen eines stabilen Zustandes ohne Änderungen der die gewünschten Zielvorstellungen nicht erfüllt. D.h. ein Fortsetzen der Berechnungen führt zu keiner weiteren Verbesserung.
3. Erreichen eines nicht definierten Zustandes der durch die vorigen Punkte nicht abgedeckt wird.

Im Folgenden werden Kenngrößen, die zur Bestimmung dieser Zustände geeignet sein könnten, analysiert. Diese Untersuchungen werden beispielhaft an den Daten des 17. September 2008 (Freitag) erläutert. Die daraus abgeleiteten Kenngrößen werden anschließend auf alle berechneten Tage angewendet und die resultierenden Ergebnisse in Kapitel 5.2.3 dargestellt.

Definition eines Gütekriteriums

Die Sollwerte der Hochrechnung sind die Zählwerte der RMQs. Demzufolge ist es sinnvoll, ein Gütekriterium aus der Abweichung zwischen diesen Zählwerten und den hochgerechneten Werten der Trajektorienbelastungen abzuleiten. Wie auch bereits bei der Clusterung der Zähldaten wird hierfür der GEH (siehe auch Kapitel 5.1.2.3)

berechnet. Im Vergleich zur Clusterung ist zu beachten, dass der GEH für die Hochrechnung aus der Verkehrsstärke je 15 Minuten berechnet wird.

Abbildung 127 zeigt den Verlauf des maximalen (blau) und des durchschnittlichen (rot) GEHs für die 4.704 Zählwerte (49 RMQs mit 96 15-Min-Intervallen) des 17. September 2008 über 200 Iterationen der Hochrechnung. In schwarz ist der Verlauf der Summe aller Fahrten dargestellt - eine Kenngröße, deren Verlauf gewisse Rückschlüsse auf das Verhalten des Prozesses zulässt.

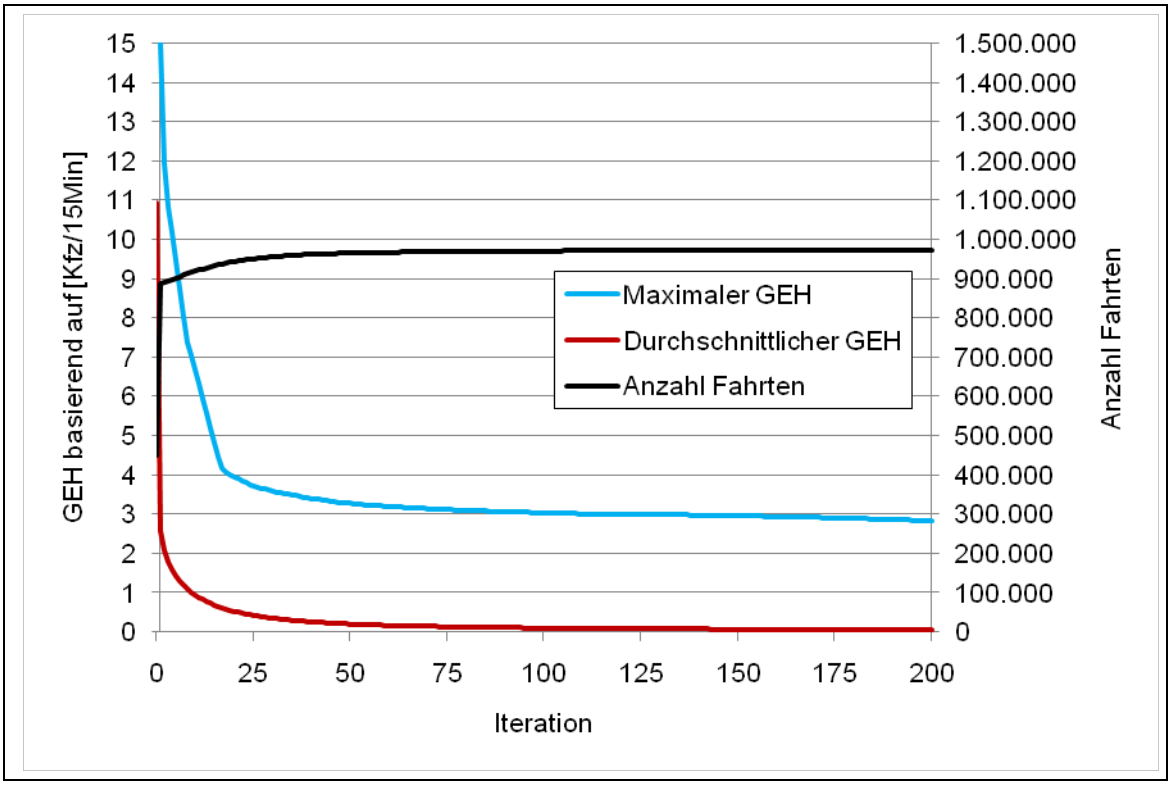


Abbildung 127: Verlauf von durchschnittlichem und maximalem GEH und der Anzahl Fahrten für die Hochrechnung des 17.09.2008.

Die Anzahl Fahrten steigt mit jeder Iteration weiter an, wobei die Steigung immer kleiner wird und davon auszugehen ist, dass eine obere Grenze existiert. Auf den Zusammenhang zwischen berechneten Iterationen, Anzahl Fahrten und der Fahrtweitenverteilung wird im Laufe der Analyse noch detaillierter eingegangen.

Der maximale GEH sinkt, besonders zu Beginn, langsamer als der durchschnittliche GEH und bleibt nach der 200. Iteration um ca. Faktor 46 über dem durchschnittlichen GEH von 0,062. Im Anfangszustand ist der maximale GEH nur um Faktor 3,1 größer als der Mittlere. Das heißt, der maximale Fehler hat sich weitaus weniger verbessert als der mittlere Fehler. Eine Analyse zeigt, dass die maximalen Abweichungen meist durch widersprüchliche, wahrscheinlich fehlerhafte, Zählwerte entstehen, die nicht erfüllt werden können bzw. sollten. Ein Beispiel hierfür ist das Verschwinden unrealistisch vieler Fahrzeuge an einer Autobahnausfahrt oder Raststätte. Diese

Umstände führen zu der Erkenntnis, dass es nicht zwingend sinnvoll ist, die fehlerbehafteten Zählwerte mit allen zur Verfügung stehenden Mitteln zu reproduzieren. Hier ist anzumerken, dass Zählwerte, die in die Hochrechnung einfließen, diverse Filter und Plausibilitätsprüfungen bestehen müssen. Trotzdem darf nicht davon ausgegangen werden, dass die Zählwerte hundertprozentig mit der Realität übereinstimmen.

Der durchschnittliche GEH ist weniger von solchen Ausreißern beeinflusst und wird, um eine sinnvolle Grenze für die Abbruchbedingung zu bestimmen, im Folgenden weiter untersucht. Abbildung 128 zeigt die Verteilung der GEH-Werte für den Anfangszustand (Iteration 0) in blau entlang der Achsen rechts und oben sowie die 200. Iteration in schwarz entlang der Achsen links und unten.

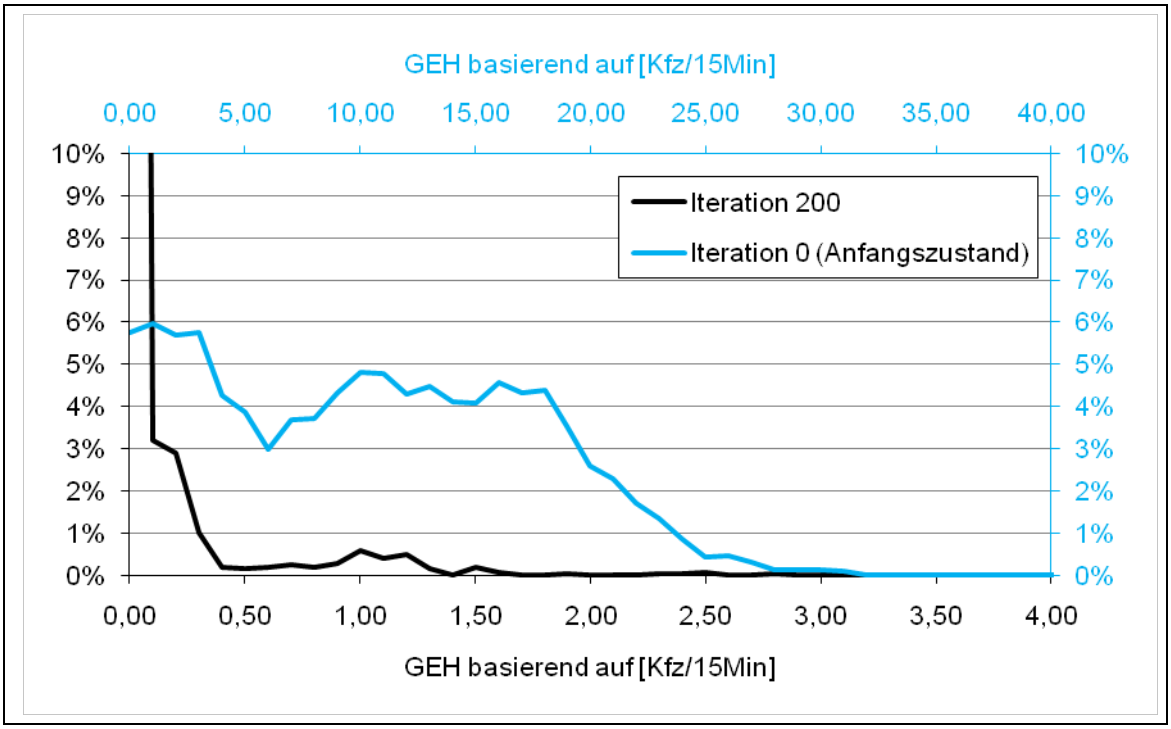


Abbildung 128: Verteilung der GEH-Werte in Klassen zu 0,1 bzw. 1,0 für die Hochrechnung des 17.09.2008.

Da die Werte aufgrund der Hochrechnung nach der 200. Iteration sehr viel kleiner sind als zu Beginn, ist es notwendig die beiden Verteilungskurven und die Klassen unterschiedlich zu skalieren. Interessant ist hier allerdings nicht die Größenordnung, sondern die unterschiedliche Form und damit die Verteilung der GEH-Werte dieser beiden Kurven. Während im Anfangszustand die GEH-Werte zwischen 0 und 20 eher gleichverteilt sind, weisen nach der 200. Iteration nur ca. 3% aller Werte einen GEH größer 0,5 auf - die Kurve ist viel steiler und fast alle Werte befinden sich am linken Rand der Achse. Das heißt, im Anfangszustand gibt es in etwa gleich viele große wie kleine Abweichungen zwischen Zählwerten und Trajektorienbelastungen, während nach der Hochrechnung fast nur noch kleine Abweichungen vorhanden sind. Dies bedeutet wiederum, dass sich der Großteil der Zählwerte durch den Hochrechnungs-

prozess in einer ähnlichen Qualität reproduzieren lässt. Lediglich der kleine Peak zwischen 1,0 und 1,2 deutet auf eine Gruppe von Zählwerten, die sich anders verhält und deswegen detaillierter analysiert wird.

Abbildung 129 zeigt die mittleren GEH-Werte des gesamten Tages für alle RMQs, wobei die rot umrandeten Werte als klare Ausreißer identifiziert werden können. Diese beiden RMQs (800 551 und 800 561) bilden zusammen mit dem RMQ 800 571 eine Abfolge von unrealistischen Zählwerten. Der Verbindung zu dem dritten RMQ lässt sich bei einer geringeren Anzahl Iterationen oder über den Zusammenhang mit der Netzstruktur (siehe Abbildung 130) identifizieren. Dabei ist anzumerken, dass dieses Zusammenspiel nicht nur für den hier betrachteten Tag auftritt.

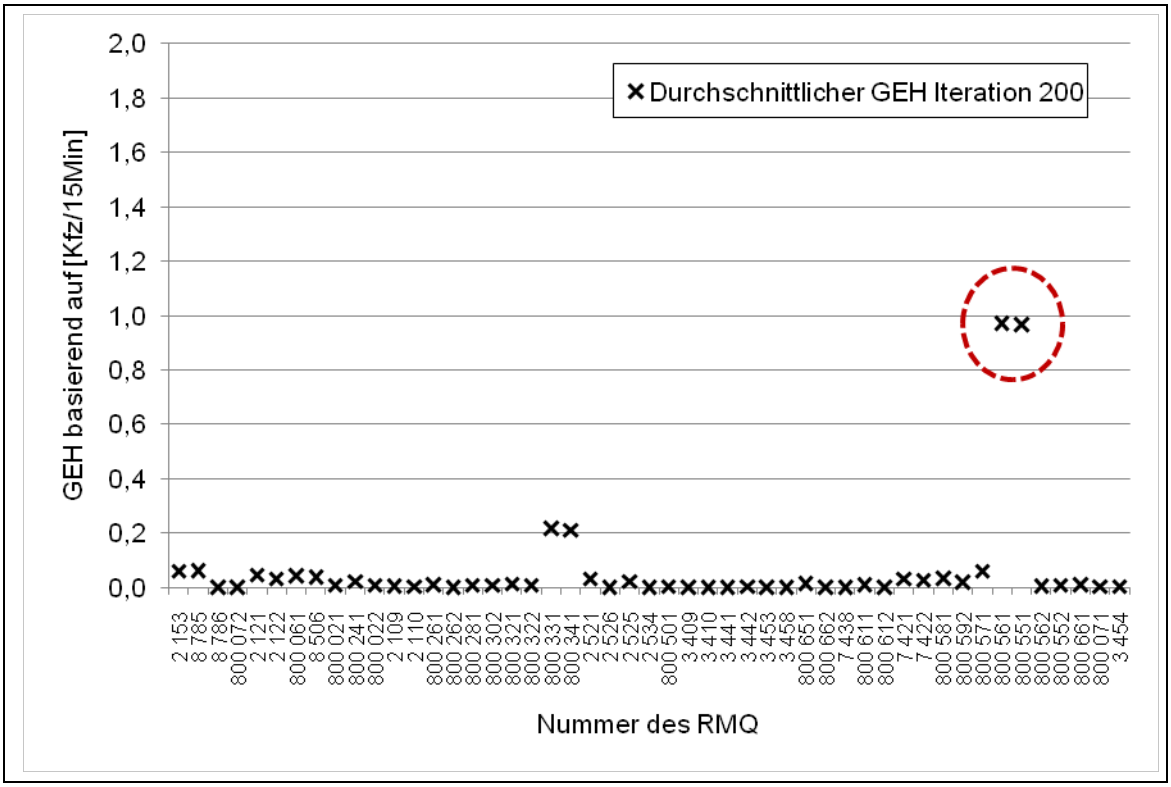


Abbildung 129: Mittlere GEH-Werte je RMQ für die Hochrechnung des 17.09.2008.

Abbildung 130 zeigt den Abschnitt der A81 dieser drei RMQs, der sich über die Anschlussstellen Mundelsheim, die Raststätte Wunnenstein und die Anschlussstellen Ilsfeld und Heilbronn/Untergruppenbach erstreckt. Die Tagesverkehrsstärke der durch die Raststätte Wunnenstein getrennten RMQs 800 571 mit 47.000 Kfz und 800 561 mit 50.700 Kfz müsste aus Bilanzierungsgründen nahezu identisch sein. Eine kleine Differenz ist theoretisch denkbar, wobei sich Fahrzeuge, die die Autobahn an dieser Raststätte verlassen oder befahren in derselben Größenordnung bewegen dürften. Mit Sicherheit ist allerdings auszuschließen, dass über diese Raststätte ungefähr 3.700 Fahrzeuge am Tag auf die A81 fahren. Folglich ist davon auszugehen, dass diese Zählwerte nicht korrekt sind.

Wie zuvor bereits angemerkt, werden alle in die Hochrechnung einfließenden Zählwerte gefiltert und Plausibilitätsprüfungen unterzogen. Die drei betrachteten RMQs weisen alle grundsätzlich mögliche Daten auf, so dass nicht festgestellt werden kann, welcher fehlerhaft ist. Realistischer weise ist davon auszugehen, dass kein RMQ absolut korrekt zählt, wobei Abweichungen von einigen Prozent ausreichend genau sind. Für die Hochrechnung ist dies nur dann problematisch, wenn auf diese Weise ein logisches Problem entsteht (Fahrzeuge verlassen die Autobahn nicht an einer Raststätte), das Zielwerte formuliert die nicht erreicht werden können.

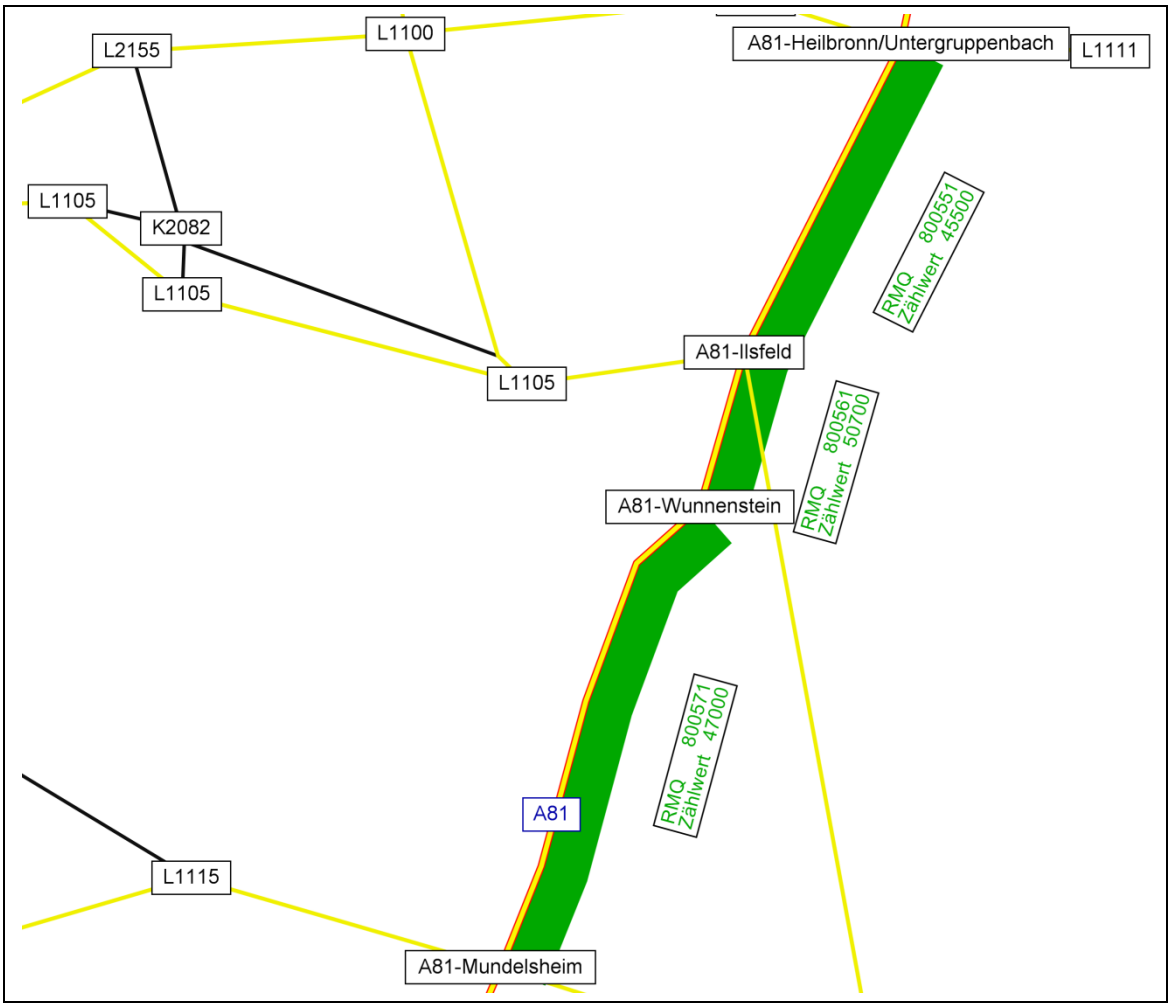


Abbildung 130: Ausgewählte RMQs und deren Tagesverkehrsstärke am 17.09.2008.

Abbildung 131 zeigt die Verkehrsstärke und den GEH zwischen den beiden widersprüchlichen RMQs über den Verlauf des Tages (der mittlere GEH beträgt 3,3). Die Darstellung macht deutlich, dass die Abweichungen der RMQs untereinander und demzufolge auch im Vergleich zur Realität weitaus größer sind als zu den Ergebnissen der Hochrechnung. Dementsprechend ist es nicht unbedingt sinnvoll, die Abweichungen zu den Zählwerten auf null zu reduzieren. Eine übermäßig große Anzahl Iterationen führt vermutlich dazu, dass den Trajektorien Belastungen aufgeprägt werden, die nicht in Zusammenhang mit der Realität stehen.

Abbildung 132 zeigt den Verlauf der Anzahl GEH-Werte in Klassen zu 0,1 nach der 200. Iteration. Dabei entspricht die schwarze Kurve dem Verlauf der relativen Darstellung aus Abbildung 128 und die blaue Kurve dem Verlauf ohne die GEHs der widersprüchlichen RMQs 800 571, 800 561 und 800 551. Die Darstellung macht deutlich, dass diese drei RMQs für nahezu alle GEH-Werte über 0,7 verantwortlich sind.

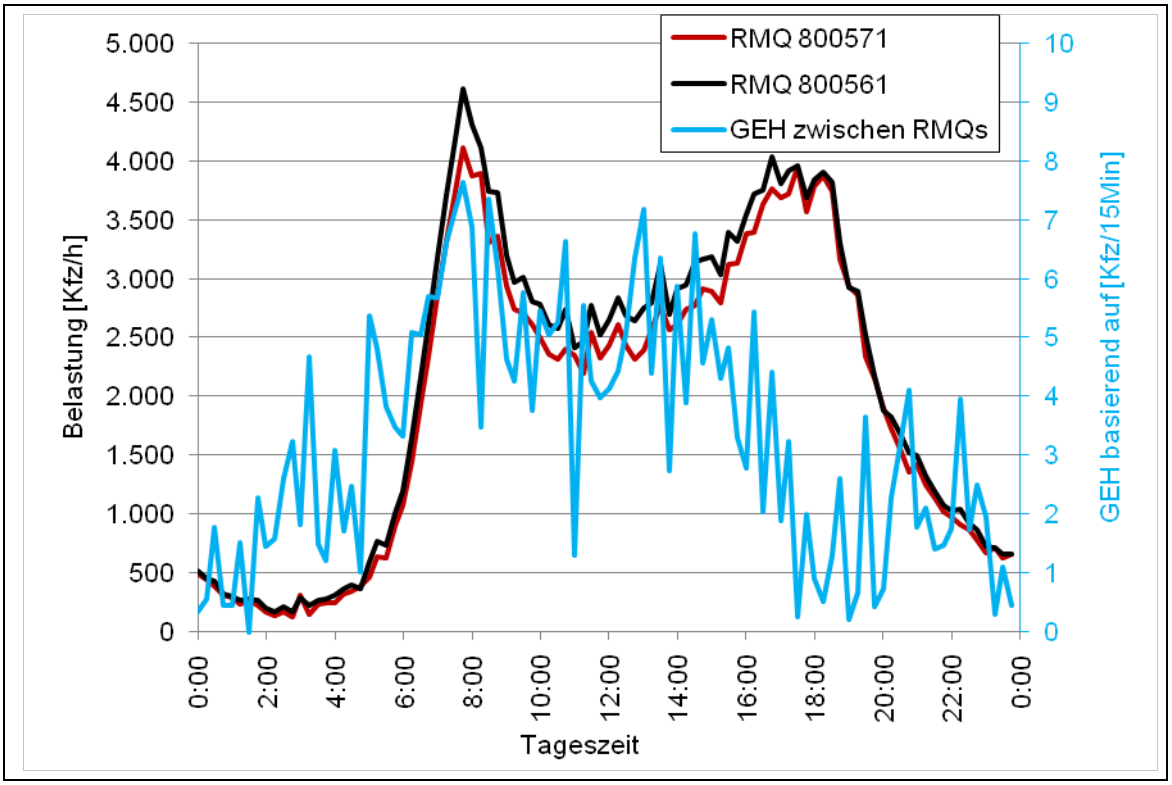


Abbildung 131: Verkehrsstärke und GEH zwischen ausgewählte RMQs am 17.09.2008.

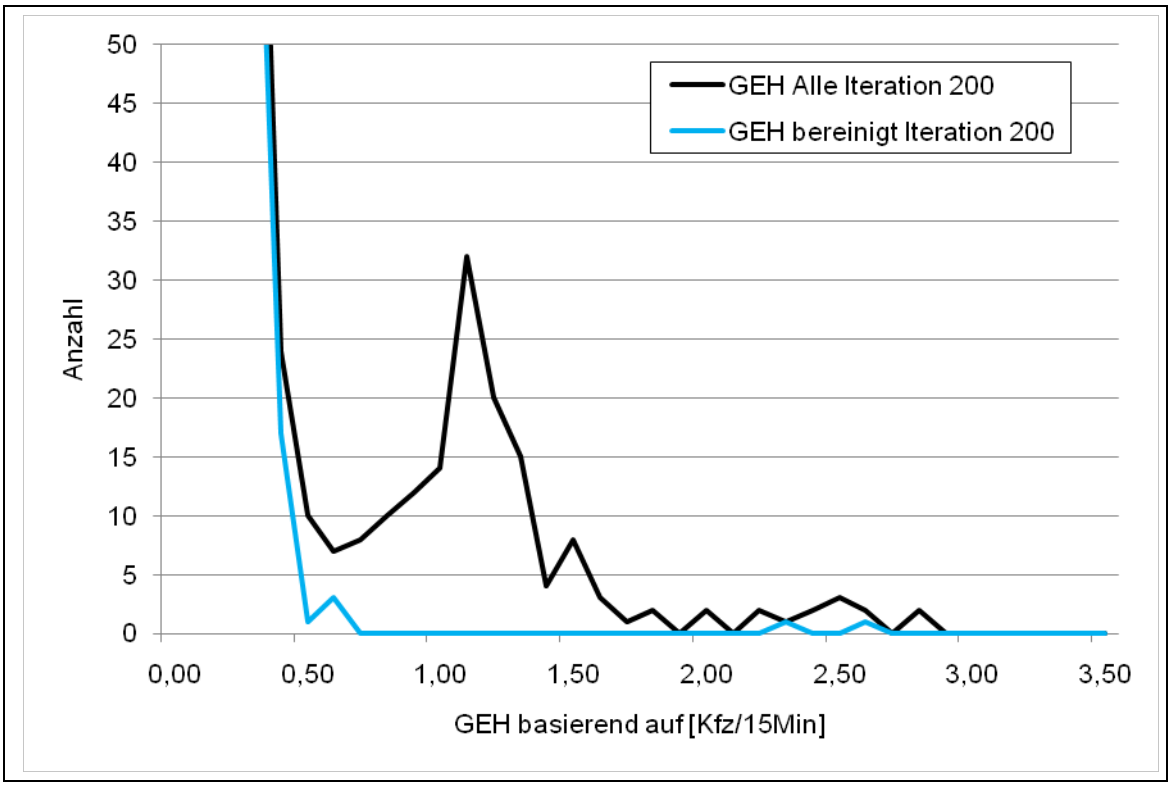


Abbildung 132: Verteilung der GEH-Werte in Klassen zu 0,1 für die Hochrechnung des 17.09.2008.

Abbildung 134 verdeutlicht anhand des Verlaufs der gemessenen und aus Mobilfunktrajektorien erzeugten Verkehrsstärke, dass die erzielte Übereinstimmung dennoch um ein Vielfaches besser ist, als das was herkömmliche Matrixkorrekturverfahren in der Lage sind zu erreichen.

In Abbildung 135 sind die Verteilungen der GEH-Werte der 42. und der 200. Iteration gegenübergestellt. Diese zeigen, dass die Peaks für höhere GEHs bei der 42. Iteration weniger ausgeprägt sind, was als ein Indiz dafür interpretiert wird, dass der Einfluss widersprüchlicher Zählwerte nicht dominiert und die Wahl der Abbruchbedingung eines mittleren GEHs von 0,25 bestätigt.

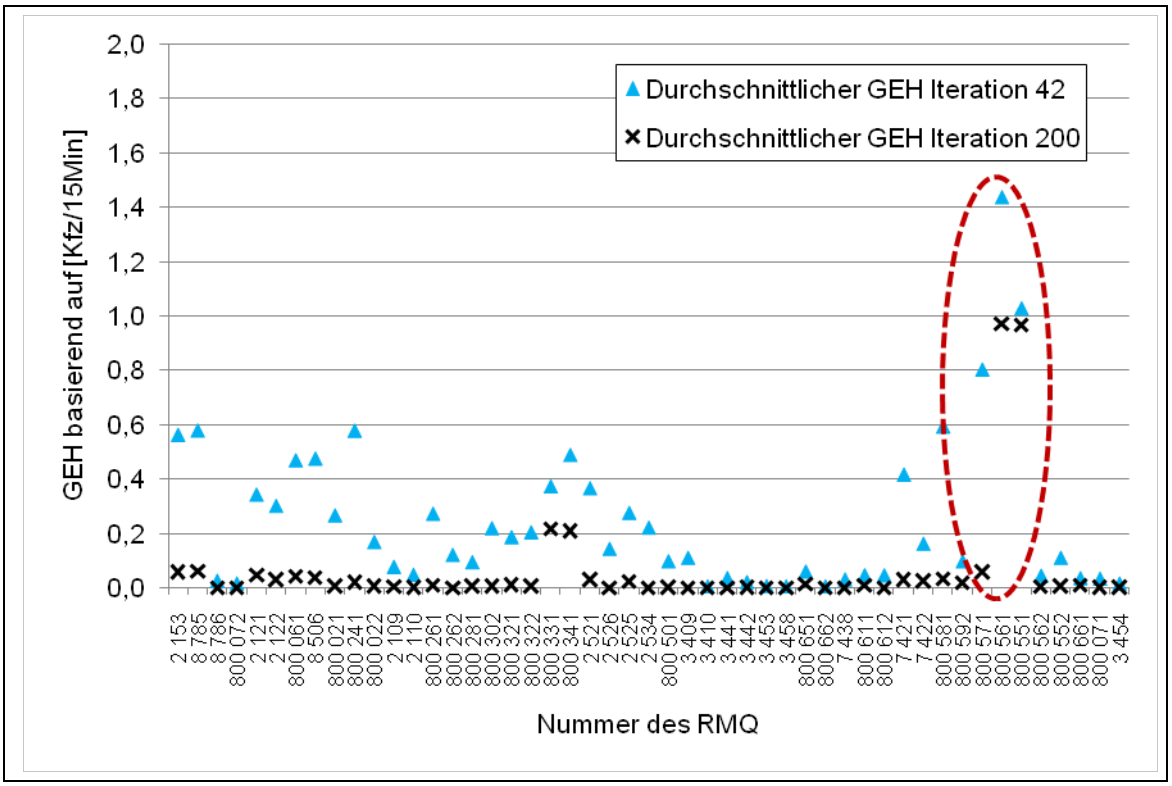


Abbildung 133: Mittlere GEH-Werte je RMQ für die Hochrechnung des 17.09.2008.

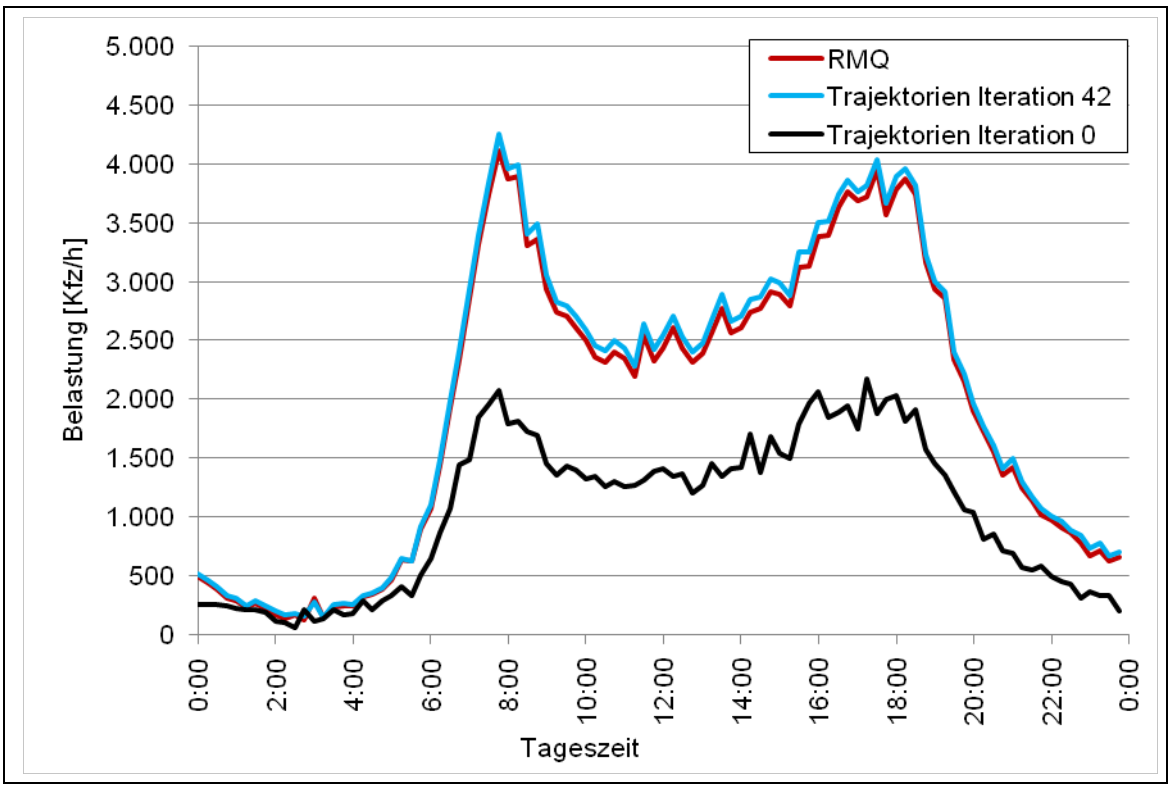


Abbildung 134: Verkehrsstärken am RMQ 800 571 am 17.09.2008.

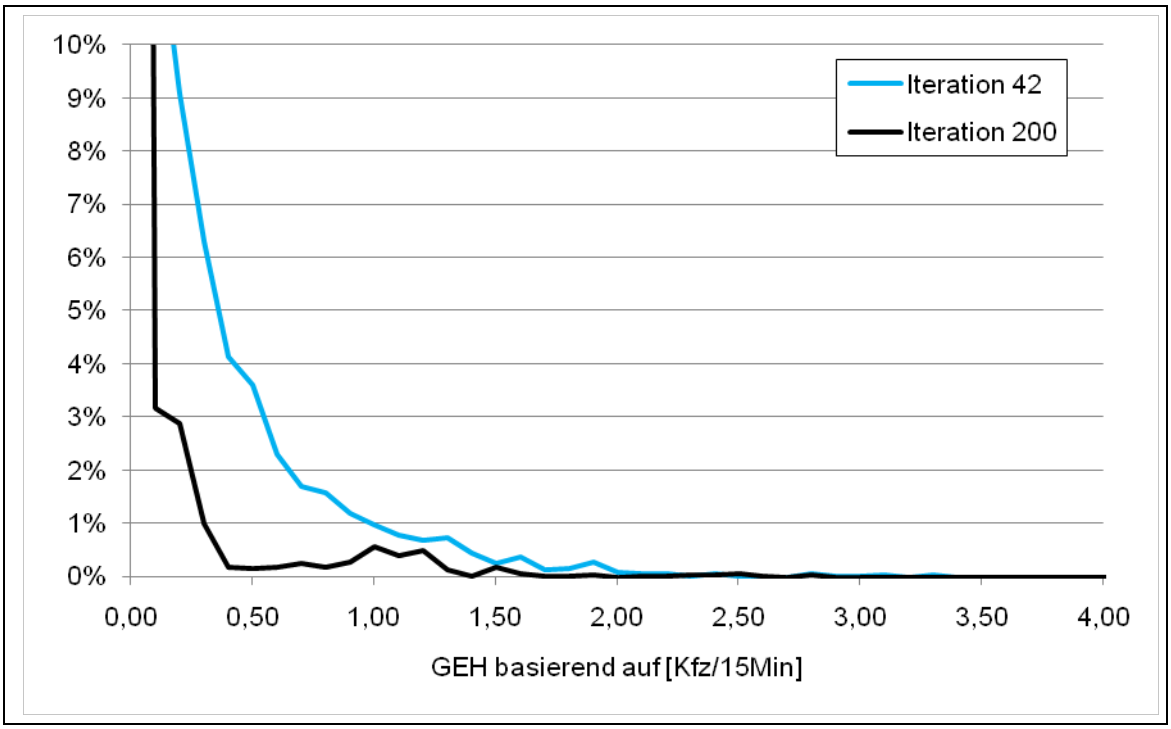


Abbildung 135: Verteilung der GEH-Werte in Klassen zu 0,1 für die Hochrechnung des 17.09.2008.

Abbruchbedingung für einen stabilen Zustand

Um die Berechnung weiterer kaum zielführender Iterationen zu verhindern, ist es notwendig die Änderung der relevanten Größen zu beobachten. Abbildung 136 zeigt den Verlauf der Änderungen der bereits zuvor betrachteten Größen. Aus Gründen der Übersichtlichkeit ist der Verlauf der ersten Iterationen nicht für alle Kenngrößen dargestellt. Insbesondere für die erste Iteration ergeben sich aufgrund der verhältnismäßig hohen Abweichung zwischen Anfangszustand und Zielgröße sehr viel größere Änderungen als in allen nachfolgenden Schritten. Die gestrichelte Linie markiert die 42. Iteration, die nach dem zuvor festgelegten Abbruchkriterium eines mittleren GEHs von 0,25 das Ende des Hochrechnungsprozesses darstellt.

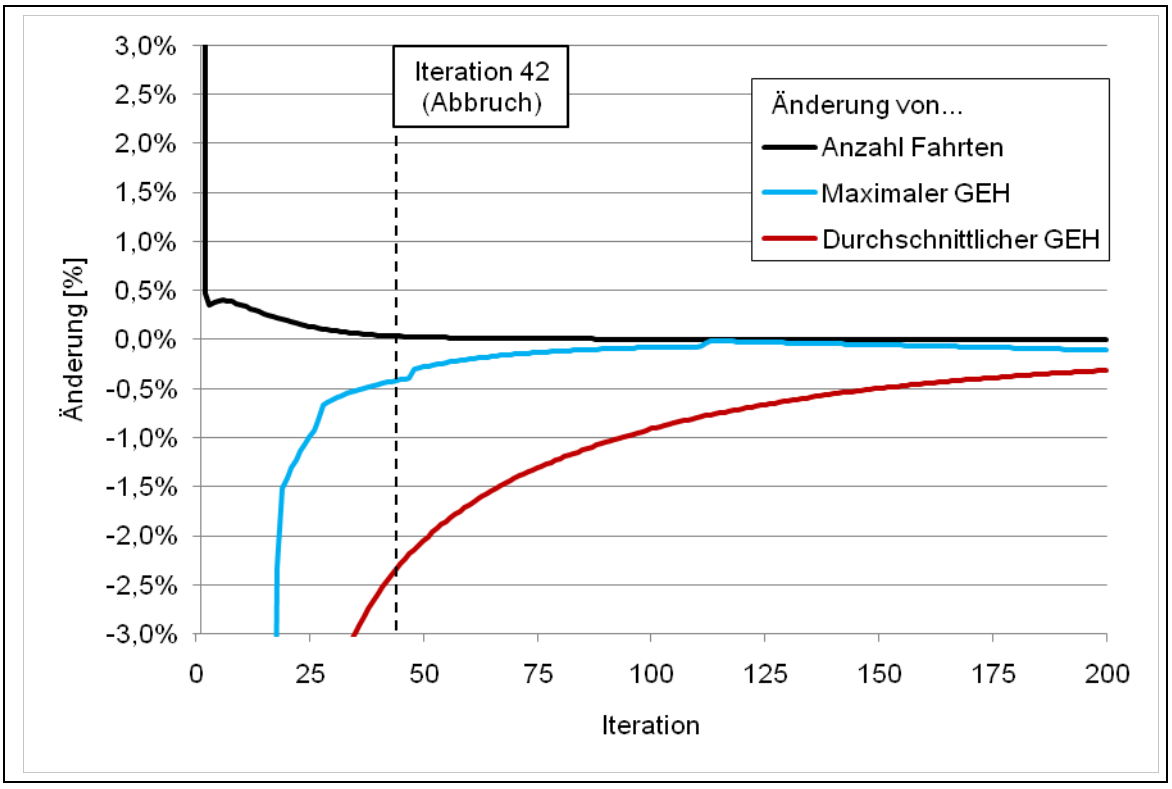


Abbildung 136: Verlauf der Änderungen von durchschnittlichem und maximalem GEH und der Anzahl Fahrten für die Hochrechnung des 17.09.2008.

Von den betrachteten Größen nähert sich die Änderung der Anzahl aller Fahrten am schnellsten dem Wert 0 an. Die im Vergleich dazu höheren Änderungen der Gütekriterien weisen darauf hin, dass es bis zu einem gewissen Punkt sinnvoll ist, Fahrten zu verlagern. Das heißt, eine Verbesserung der Fehlerkenngrößen wird, abgesehen vom ersten Iterationsschritt, größtenteils durch eine Verlagerung der Fahrten erreicht und nicht durch eine Erhöhung der Fahrten.

Die Änderungen des maximalen GEHs weisen Sprünge auf, die immer dann entstehen, wenn sich das Zeitintervall und der RMQ, an dem der maximale GEH vorliegt, ändern. Diese Eigenschaft ist ein weiterer Grund, warum der maximale GEH nicht als Abbruchkriterium geeignet ist (siehe auch im vorigen Abschnitt „Definition eines Gütekriteriums“).

Entsprechend dem zuvor definierten Abbruchkriterium für das Erfüllen der Sollwerte durch einen durchschnittlichen GEH von 0,25 wird ein weiteres Abbruchkriterium festgelegt, dass den Hochrechnungsprozess beendet, wenn die letzten 5 Iterationen alle eine Änderung des durchschnittlichen GEHs aufweisen, deren Betrag kleiner 2% ist. Für den 17. September 2008 würde diese Abbruchbedingung nach der 56. Iteration greifen und ist damit so gewählt, dass das Erreichen des gewünschten Gütekriteriums mit hoher Wahrscheinlichkeit vorher eintritt, was sich in den nachfolgenden Betrachtungen bestätigt (siehe auch Kapitel 5.2.3).

Abbruchbedingung für einen nicht definierten Zustand

Um zu verhindern, dass der Iterationsprozess nicht endlos läuft, wird zusätzlich die maximal erlaubte Anzahl Iterationen auf 200 begrenzt. Dieser Wert stellt keine unerreichbaren Grenzen an die Rechenzeit und ist außerdem groß genug, dass er für die untersuchten Tage niemals die Hochrechnung terminiert (siehe auch Kapitel 5.2.3). Die Definition einer solchen oberen Grenze ist aus praktischen Gründen trotzdem sinnvoll, wobei weitere Analysen notwendig sind, wenn der Iterationsprozess tatsächlich deswegen beendet wird.

5.2.2.4 Binnenverkehrsproblematik

Wie in Kapitel 2.3 bereits angeführt, ist das hier verwendete Verfahren zur Generierung von Mobilfunktrajektorien nicht in der Lage, Fahrten unterhalb einer bestimmten Fahrtweite zu generieren. Dies führt zu einer gewissen Schiefe in der Stichprobe die aus allen tatsächlichen Fahrten gezogen wird und als Eingang für die Hochrechnung dient. Die verwendeten Mobilfunktrajektorien repräsentieren so eine andere Fahrtzusammensetzung, als die Zählwerte der verwendeten RMQs, deren Erkennungsrate unabhängig von der Fahrtweite der gezählten Fahrzeuge ist.

Diese Problematik wird unter anderem dadurch entschärft, dass für die Hochrechnung nur RMQs auf Bundesautobahnen verwendet werden, die ebenfalls eine hohe Fahrtweitenverteilung aufweisen. Dies hat zur Folge, dass von den Ergebnissen der Hochrechnung nur auf Autobahnen eine hohe Qualität gefordert werden kann.

Außerdem lässt sich feststellen, dass der Hochrechnungsprozess die Fahrtweitenverteilung zu Gunsten der kurzen Fahrten verlagert. Abbildung 137 zeigt die Fahrtweitenverteilungen der Mobilfunktrajektorien vor der Hochrechnung in rot (Iteration 0), nach der 42. (endgültigen) Iteration in blau und für die 200. Iteration in einer schwarz gestrichelten Linie.

Die Darstellung zeigt eine Verkürzung der Fahrtweitenverteilung für die 42. und die 200. Iteration. Eine detaillierte Analyse zeigt, dass die Verlagerung mit ansteigender Anzahl Iterationen immer langsamer wird - bereits in dieser Darstellung ist kein Unterschied zwischen der 42. und der 200. Iteration, wohl aber zum Anfangszustand auszumachen. Diese Beobachtung führt zu dem Schluss, dass der Hochrechnungsprozess die kurzen Trajektorien bevorzugt, weil diese in der Stichprobe unterrepräsentiert sind und somit der Schiefe der Stichprobe entgegen wirkt.

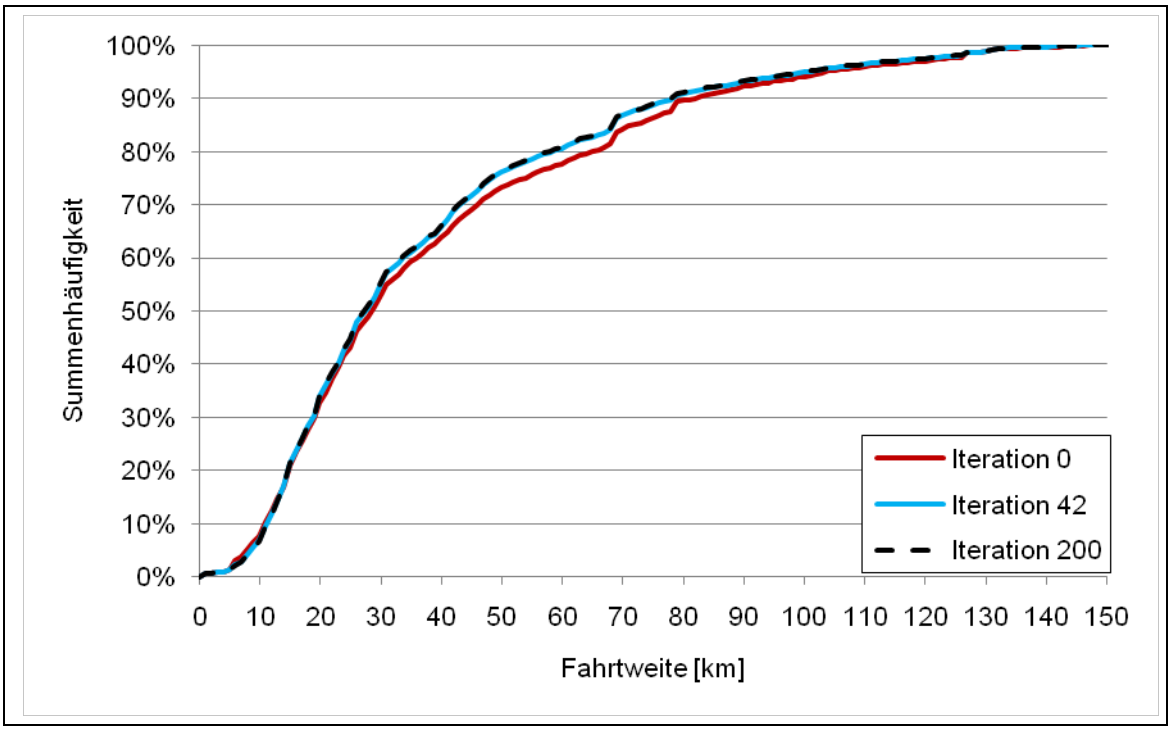


Abbildung 137: Fahrtweitenverteilungen für die Hochrechnung des 17.09.2008.

5.2.3 Langzeitergebnisse

Das vorgestellte dynamische Hochrechnungsverfahren wird im VPN für insgesamt 68 der 78 vorhandenen Tage durchgeführt. Die Einschränkung auf 68 Tage ergibt sich dadurch, dass nur Tage hochgerechnet werden, bei denen mindestens 1.350 A-Daten-Rohdateien (theoretisches Maximum: 1.440) vorhanden sind. Diese Bedingung ermöglicht eine hohe Abdeckung des Netzes durch Trajektorien (siehe nachfolgende Kapitel). Mit den in Kapitel 5.2.2.3 bestimmten Abbruchbedingungen wird die Hochrechnung durchschnittlich nach der 42. Iteration bzw. nach 198 Minuten (auf einem Athlon 64 X2 6000+) abgeschlossen. Dies führt zu einer Gesamtrechenzeit von ca. 9,4 Tagen, die zwar hoch ist, sich aber noch in einem vertretbaren Rahmen befindet. Die Zeitschrittweite des Verfahrens beträgt 15 Minuten, folglich werden für die Hochrechnung eines Tages 96 Zeitintervalle verwendet.

Im Folgenden werden zunächst die Eingangsdaten und somit die Start- und Randbedingungen für die Hochrechnung analysiert. Anschließend wird das Verhalten des Verfahrens geprüft und am Ende werden die erzeugten Ergebnisse dargestellt, erläutert und bewertet.

5.2.3.1 Analyse der Eingangsdaten

Richtungsmessquerschnitte

Aufgrund der Fahrtweitenverteilung der in Kapitel 3 erzeugten Mobilfunktrajektorien werden zur Hochrechnung nur Daten von Detektoren verwendet, die sich auf Bundesautobahnen befinden. Deren Verfügbarkeit schwankt über die hochgerechneten Tage. Dies liegt an einer umfangreichen Vorfilterung, die zum einen die Qualität der Messwerte sicherstellen soll, und zum anderen einen RMQ nur dann für die Hochrechnung auswählt, wenn für den kompletten Tag aufbereitete Daten vorhanden sind. Ferner kann die Hochrechnung für jede VPN-Strecke nur genau einen RMQ berücksichtigen. Da sich auf den VPN-Strecken teilweise mehrere RMQs befinden, reduziert sich die Anzahl verwendeter RMQs aufgrund der Wahl des jeweils zuverlässigsten RMQs weiter. Abbildung 138 zeigt auf der linken Achse (schwarze Kreuze) die Anzahl der verwendeten RMQ, die zwischen 38 und 53 liegt. Diese Werte sind aber nur bedingt für die Abdeckung für die Hochrechnung aussagekräftig, da dafür auch die Lage der verfügbaren RMQ wichtig ist. Deshalb werden in einer weiteren Analyse die vier betrachteten Autobahnen A5, A6, A8 und A81 in jeweils zwei Abschnitte pro Richtung unterteilt und geprüft, wie viele dieser insgesamt 16 Abschnitte durch mindestens einen RMQ abgedeckt sind. Das Ergebnis ist ebenfalls in Abbildung 138 abgedruckt (rechte Achse, blaue Dreiecke). Hier liegen die Werte zwischen 81 und 100 %. Im ungünstigsten Fall sind damit drei Abschnitte ohne verfügbaren RMQ, eine Hochrechnung der Trajektorien, die nur über diese Abschnitte fahren, erfolgt daher mit dem durchschnittlichen Hochrechnungsfaktor.

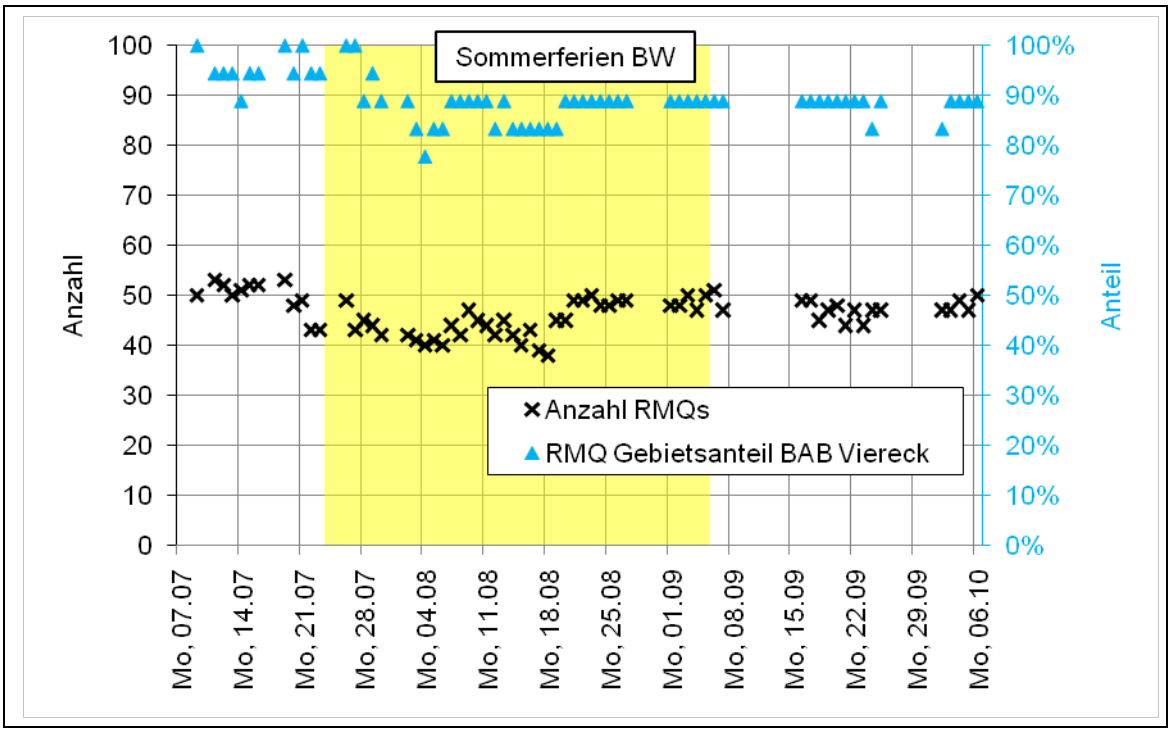


Abbildung 138: Anzahl RMQs und Abdeckung der BAB-Netzmasche für jeden Tag.

Eine andere Form der Visualisierung ist Abbildung 139 zu entnehmen. Diese zeigt das VPN, wobei die Strecken, die an mindestens 50% aller Hochrechnungstage mit Messwerten versorgt sind, mit grünen richtungsbezogenen Balken hervorgehoben werden. Die Abbildung zeigt, dass alle 4 Kanten der BAB-Netzmasche mit Messdaten versorgt sind. Lediglich die A8 von Stuttgart nach Pforzheim ist häufig ohne aktiven Richtungsmessquerschnitt.

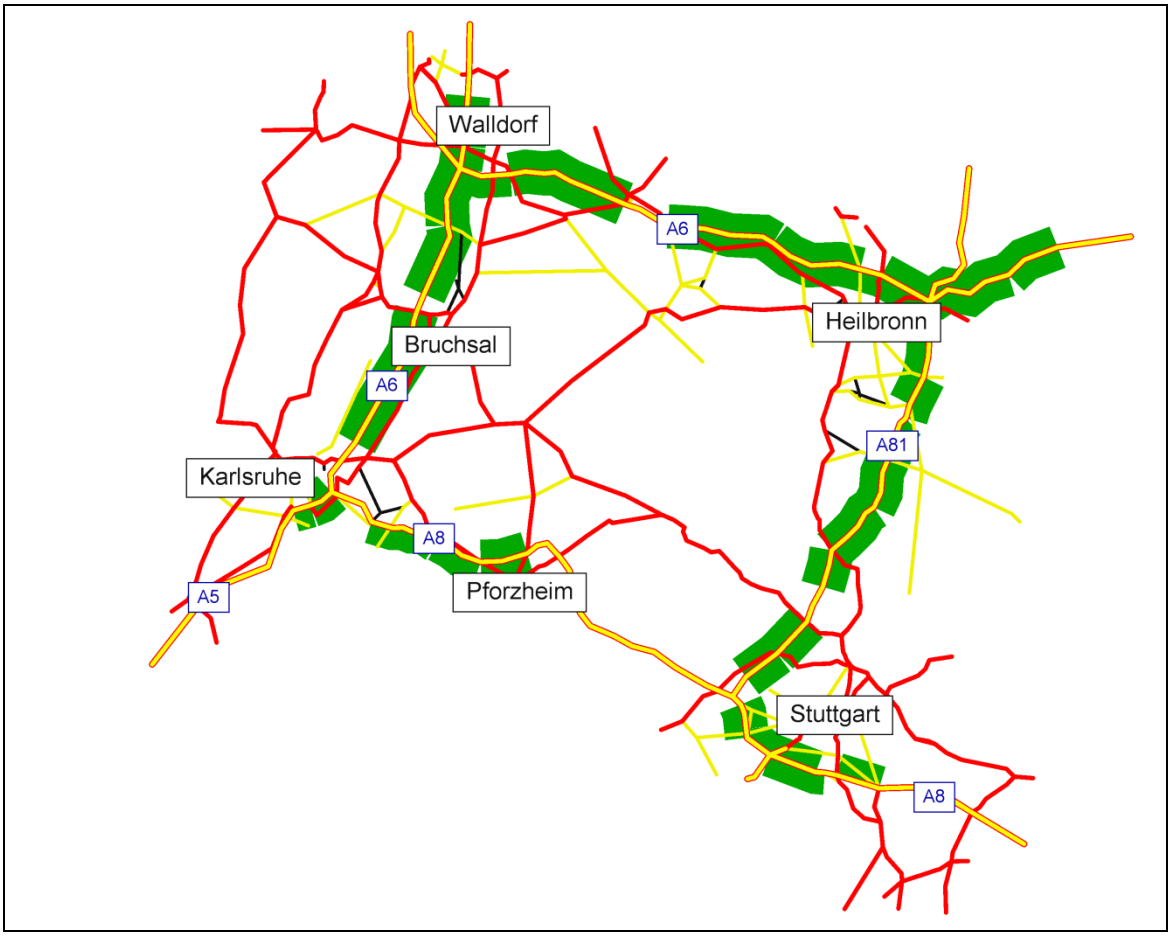


Abbildung 139: Strecken mit RMQs, die für mindestens 50% aller Hochrechnungen verwendet werden.

Mobilfunktrajektorien

Zu Beschreibung der Vollständigkeit der Trajektorien aus dem LAC-basierten Trajektoriengenerierungsverfahren wird geprüft, in wie vielen 15min-Intervallen es an für diesen Tag verfügbaren RMQs keine Trajektorienbelastung gibt. Bei einer durchschnittlichen Anzahl von ca. 47 verwendeten RMQ ergeben sich ca. 4.500 RMQ-Intervalle, die geprüft werden.

Abbildung 140 zeigt als Ergebnis die Anzahl RMQ-Intervalle ohne Trajektorienbelastung (schwarze Kreuze). Es zeigt sich, dass über 90 % der Tage (62 von 68 Tagen) eine vollständige Abdeckung mit Trajektorien haben. Die verbleibenden 6 Tage

haben nur vereinzelte RMQ-Intervalle ohne Belastung während der Nacht. Diese hohe Abdeckung entsteht dadurch, dass nur Tage hochgerechnet werden, bei denen mindestens 1.350 A-Daten-Rohdateien vorhanden sind. Die hohe Abdeckung zeigt aber trotzdem auch die Qualität des LAC-basierten Trajektoriengenerierungsverfahrens, das zu jeder Tageszeit Trajektorien auf allen Strecken erzeugt.

Die blauen Dreiecke geben wiederum den Anteil der Trajektorien an, die über mindestens einen RMQ verlaufen. Das heißt, diese Trajektorien werden mit einer individuellen Korrektur versehen, während der Rest von ca. 28 % mit den mittleren Hochrechnungsfaktoren korrigiert wird. Der große Anteil an Mobilfunktrajektorien, die über die Bundesautobahn fahren, ist im Einklang mit den Analysen zur Fahrtweitenverteilung (siehe Kapitel 3.9.1). Der Hauptgrund dafür, dass eine Trajektorie nicht über einen RMQ fährt, ist dass zwischen Pforzheim und Stuttgart nur selten verwertbare Zähldaten vorhanden sind (vergleiche Abbildung 139).

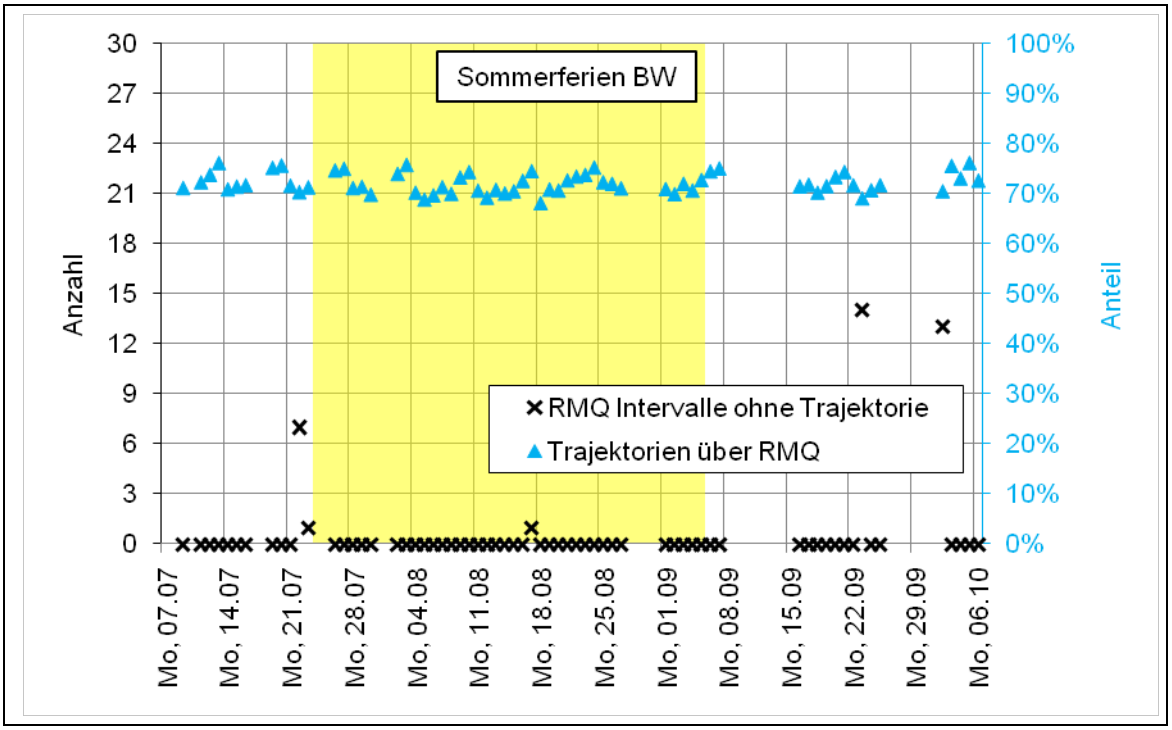


Abbildung 140: Gegenseitige Abdeckung von RMQs und Trajektorien.

5.2.3.2 Analyse des Hochrechnungsverfahrens

Abbildung 141 gibt einen Gesamtüberblick über die Ergebnisse der einzelnen Hochrechnungen. Der Verlauf der Summe aller hochgerechneter Fahrten (schwarze Kreuze) weist ein typisches 7-Tage-Muster mit hohen Werten von Montag bis Freitag und niedrigeren Werten am Wochenende auf und während der Sommerferien in Baden-Württemberg geht die Gesamtsumme aller Fahrten leicht zurück.

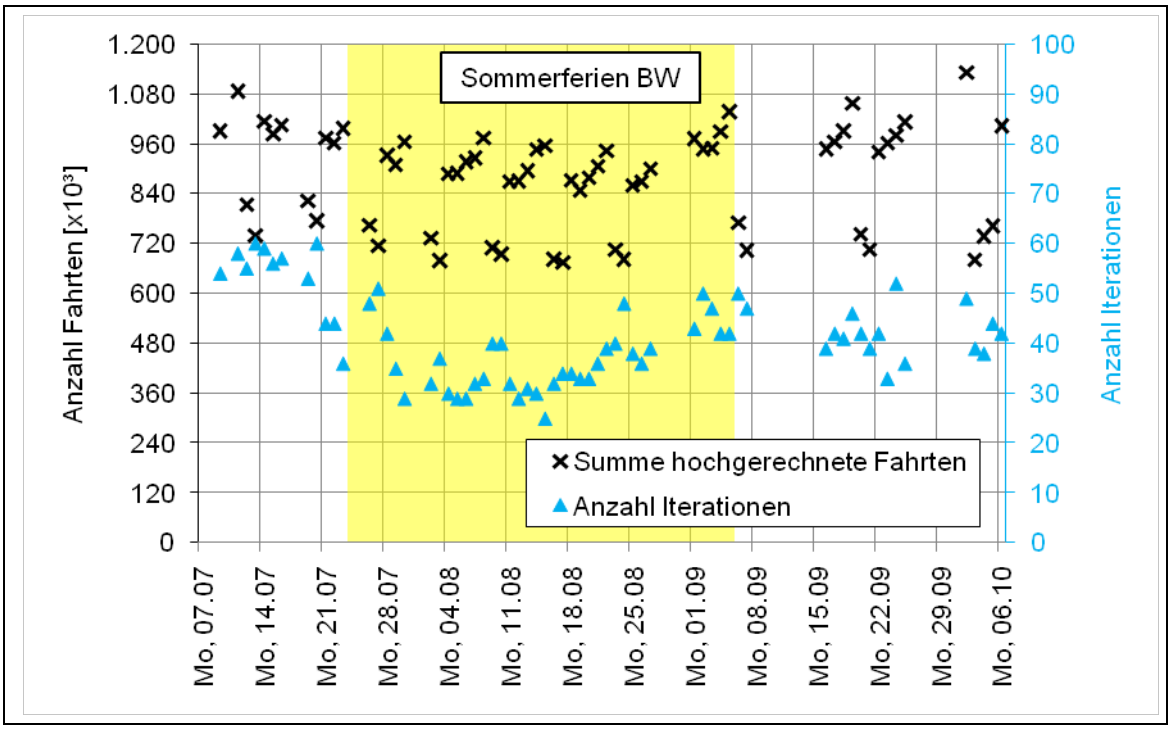


Abbildung 141: Summe hochgerechneter Fahrten und Anzahl Iterationen für jeden Tag.

Die Anzahl notwendiger Iterationen (blaue Dreiecke) bewegt sich zwischen 25 und 60, wobei für Schwankungen nach oben ein gewisser Zusammenhang mit einem höheren Abdeckungsgrad und mehr aktiven RMQs (Abbildung 138) besteht - in diesem Fall müssen mehr Sollwerte erfüllt werden. Schwankungen nach unten hängen mit einer Anzahl RMQ-Intervalle ohne Trajektorien größer Null und weniger aktiven RMQs zusammen (Abbildung 140). Hier sind entweder die Verbesserungen schneller zu klein oder es müssen weniger Sollwerte erfüllt werden.

Abbildung 142 zeigt den durchschnittlichen GEH (schwarze Kreuze) und dessen Änderung (blaue Dreiecke) für jeden hochgerechneten Tag. Die Skalierung der Y-Achsen erlaubt, den Niedrigeren der beiden Werte direkt als das verwendete Abbruchkriterium für den Hochrechnungsprozess zu identifizieren (siehe Kapitel 5.2.2.3).

Da das erste Kriterium häufiger zur Beendigung der Hochrechnungen führt, liegt der durchschnittliche GEH im Normalfall bei ca. 0,25. Aber auch, wenn das zweite Abbruchkriterium maßgebend ist, liegt der durchschnittliche GEH bis auf eine Ausnahme in diesem Bereich. Alle erreichten Werte liegen in einem so geringen Bereich, der in der Regel bei anderen verkehrsplanerischen Anwendungen nicht erreicht wird und vermutlich auch über der Genauigkeit der Zählstellen liegt.

Zusammenfassend kann festgestellt werden, dass das Verfahren auch für eine große Anzahl Tage ein stabiles Verhalten aufweist.

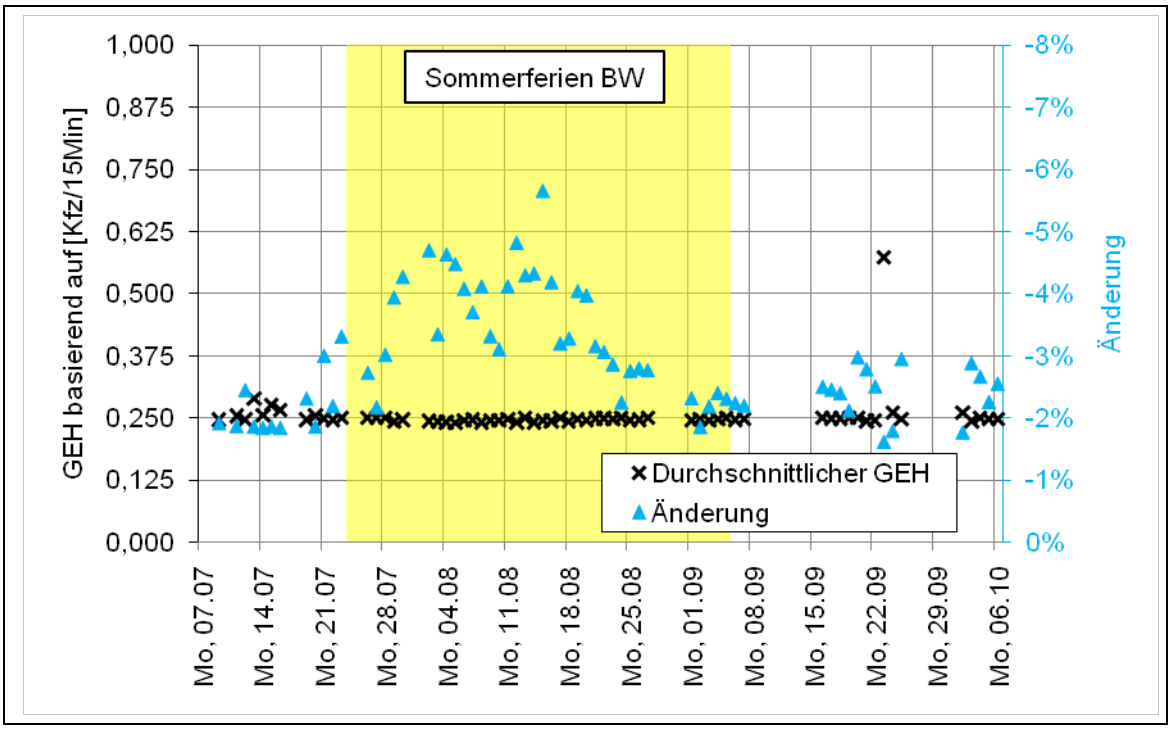


Abbildung 142: Durchschnittlicher GEH und Änderung für jeden Tag.

5.2.3.3 Analyse und Bewertung der Ergebnisse

Die einzelnen hochgerechneten Tage werden entsprechend den Ergebnissen der Clusterung (Kapitel 5.1) zu typischen Tagen zusammengefasst. Das Berechnen einer typischen QZ-Matrix wird durch eine einfache Mittelwertbildung der Belastungen aller Quelle-Ziel-Beziehungen bzw. Trajektorien realisiert. Da die Datenbasis von 68 Tagen im Vergleich zu den RMQ-Daten relativ klein ist, ist nur die Bildung der in Tabelle 59 aufgeführten Cluster möglich, wobei Cluster 9 nur der Vollständigkeit halber aufgeführt ist und lediglich dem 3.Oktober 2008 enthält.

Cluster	Beschreibung	Anzahl Tage	Anzahl Fahrten
1	Schule von Montag bis Donnerstag	16	983.880
2	Schule Freitag oder Mittwoch/Donnerstag vor Feiertag	3	1.091.880
3	Ferien von Montag bis Donnerstag	22	908.890
4	Ferien Freitag oder Mittwoch/Donnerstag vor Feiertag	4	977.582
5	Samstag	11	749.950
6	Sonntag oder Feiertag an einem Montag	11	713.932
9	Feiertag an einem Donnerstag oder Freitag (im Untersuchungszeitraum nur der 3.Oktober 2008)	1	680.592

Tabelle 59: Cluster der Hochrechnungstage.

Im Folgenden wird das Cluster 1 eingehender betrachtet, da dieses zum Einen eine große Basis an Tagen enthält und zum Anderen der Tagesgruppe angehört, die über den Verlauf eines Jahres hinweg am häufigsten vorkommt.

Abbildung 143 zeigt die resultierenden Belastungen für diesen typischen Montag, Dienstag, Mittwoch oder Donnerstag außerhalb der baden-württembergischen Schulferien und Abbildung 144 veranschaulicht die dazugehörige Fahrtweiten-verteilung, die einen typischen Verlauf zeigt.

Insgesamt kann festgehalten werden, dass die erzeugten typischen Tage plausible Ergebnisse aufweisen und auf einer ausreichend großen Datenbasis stehen, um als sinnvolle Eingangsgröße für die Tagesprognose (Kapitel 6) oder weiterer Verfahren zu dienen.

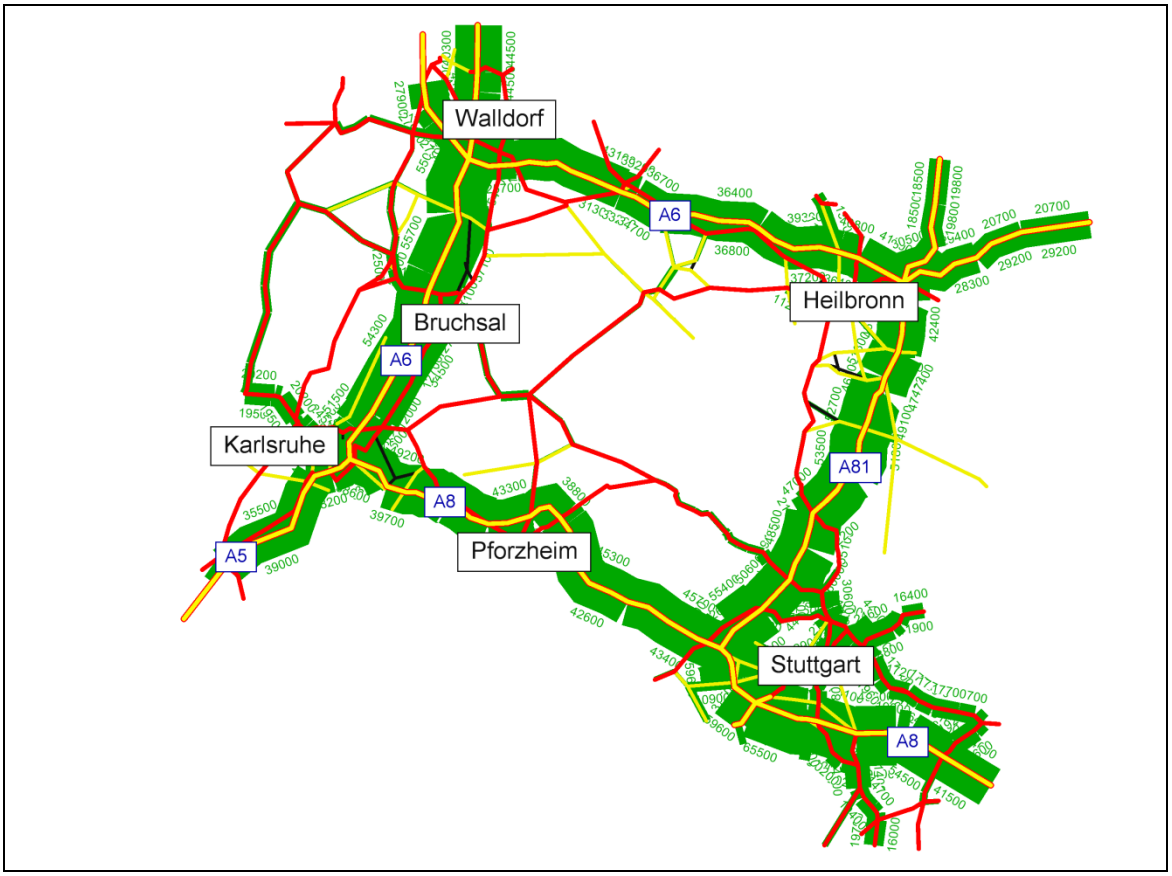


Abbildung 143: Belastungen für den typischen Tag „Schule von Montag bis Donnerstag“.

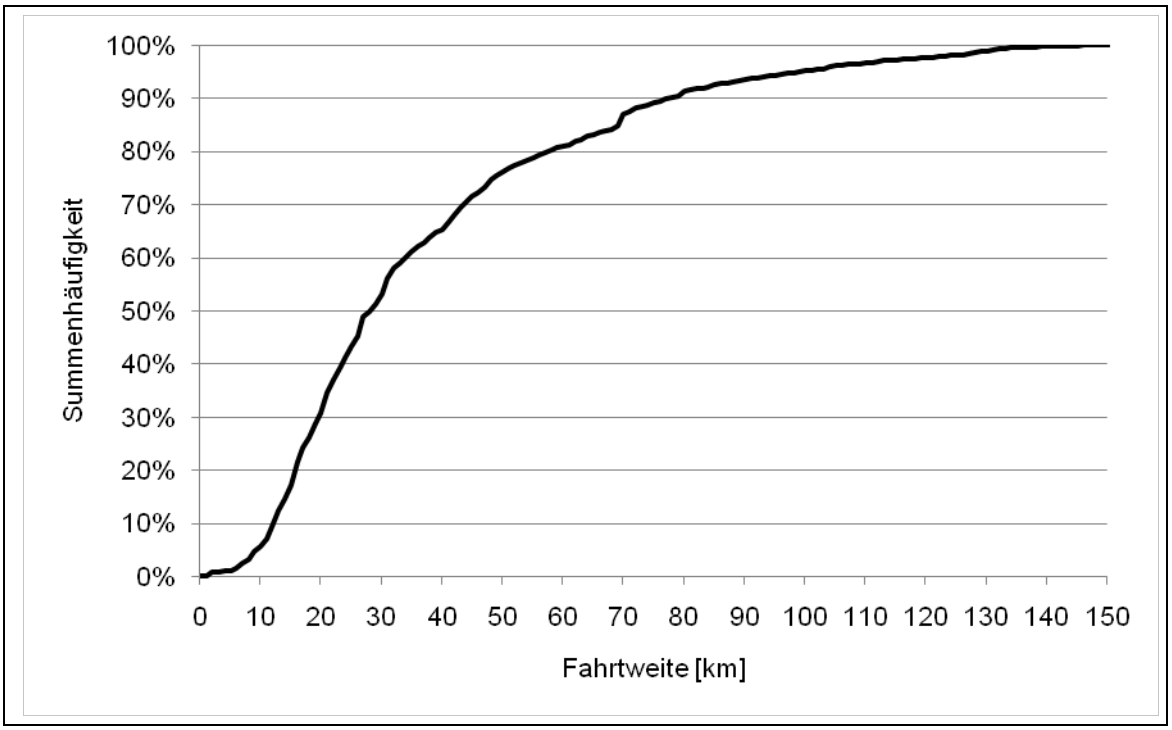


Abbildung 144: Fahrtweitenverteilung für den typischen Tag „Schule Montag bis Donnerstag“.

6 Tagesprognose (AP7-B)

Ziel einer Tagesprognose ist es, den Verkehrszustand im Untersuchungsnetz für einen in der Zukunft liegenden Kalendertag stundenfein zu prognostizieren. Dabei werden die typische zeitabhängige Verkehrsnachfrage für den Kalendertag und bekannte Kapazitätsreduzierungen durch Baustellen berücksichtigt:

- Verkehrsnachfrage: Als Nachfragedaten werden die Quelle-Ziel-Matrizen verwendet, die mit dem im Kapitel 5.2 beschriebenen Verfahren ermittelt wurden. Für jeden Kalendertag liegt eine spezifische Quelle-Ziel-Matrix vor, die aus 24 stundenfeinen Teilmatrizen besteht. Quellen und Ziele sind TMC-Locations. In einer zweiten Version dienen die nach dem in Kapitel 5.1 vorgestellten Verfahren geclusterten Quelle-Ziel-Matrizen als Nachfragedaten. Auch hier handelt es sich um 24 stundenfeine Teilmatrizen. Allerdings ist dies keine für einen ausgewählten Prognosetag spezifisch ermittelte Matrix, sondern entspricht vielmehr einem Erwartungswert der Verkehrsbelastung, mit dem an ähnlichen Tagen wie dem Prognosetag zu rechnen ist.
- Verkehrsangebot: Als Grundlage für die Tagesprognose wird das im vereinfachten Projektnetz (VPN, siehe Kapitel 2.1) abgebildete Verkehrsnetz verwendet. In diesem Netz ist jede Strecke mit der zulässigen Geschwindigkeit und einer Stundenkapazität attributiert. Am Kalendertag vorhandene Baustellen werden über eine Reduktion der Kapazität und der Geschwindigkeit berücksichtigt.

Im Folgenden werden zuerst vier Ansätze für eine Tagesprognose vorgestellt. Dann wird dargestellt, wie die Geschwindigkeiten und die Kapazitäten im Projektnetz gesetzt werden und wie der Level of Service bestimmt wird. Abschließend werden die Ergebnisse der Tagesprognose gezeigt und bewertet.

6.1 Ansätze für die Tagesprognose

6.1.1 Verkehrszustände aus historischen Ganglinien

Die Zähldaten der Richtungsmessquerschnitte im Projektnetz liefern Ganglinien der Verkehrsstärke. Clustert man diese Ganglinien mit den in Kapitel 5.1 beschriebenen Verfahren erhält man typische Ganglinien. Sie enthalten Informationen über täglich wiederkehrende Ereignisse, die die Verkehrsnachfrage beeinflussen. Fusioniert man die Zähldaten mit den FPD über die in Kapitel 5.2 erläuterte Hochrechnung, dann liegen die Verkehrsstärken nicht nur für die Strecken mit Zählstellen sondern für das gesamte Projektnetz vor. Allerdings nimmt die Zuverlässigkeit der so ermittelten Verkehrsstärken mit der Entfernung von den Zählstellen ab.

Diese Verkehrsstärken können nun mit den spezifischen Kapazitäten eines Kalendertages verglichen werden. Der daraus berechnete Auslastungsgrad der einzelnen

Strecken kann als Ergebnis der Tagesprognose dem Verkehrsteilnehmer als Level of Service (LOS) angeboten werden.

Bei der Tagesprognose anhand historischer Ganglinien wird die Annahme getroffen, dass die aktuelle Baustellensituation am Kalendertag keinen Einfluss auf die Verkehrsnachfrage, insbesondere auf die Abfahrtszeitwahl und Routenwahl hat. Eine eventuelle räumliche oder zeitliche Verlagerung wird nicht berücksichtigt.

6.1.2 Verkehrszustände aus einer statischen Umlegung

Verkehrsumlegungen bilden das Routenwahlverhalten der Verkehrsteilnehmer nach. Sie verteilen die Verkehrsnachfrage auf die im Netz verfügbaren Routen und ermitteln so Verkehrsstärken für jede Strecke. Um die Routenwahl nachzubilden, wird häufig von einem sogenannten Nutzergleichgewicht (Deterministic User Equilibrium, DUE, vgl. Cascetta, 2001) ausgegangen. Es ist dann gegeben, wenn sich die Verkehrsteilnehmer so verteilen, dass die objektiv messbare Fahrtdauer auf allen alternativen Routen gleich ist und jeder Wechsel auf eine andere Route die persönliche Fahrzeit erhöhen würde. Der Ansatz dieses deterministischen Nutzergleichgewichts kann auf ein stochastisches Nutzergleichgewicht (Stochastic User Equilibrium, SUE) erweitert werden, bei dem die Verkehrsteilnehmer ihre subjektiv empfundene Fahrtdauer optimieren.

Führt man mit den vorliegenden Quelle-Ziel-Matrizen eine Verkehrsumlegung durch, erhält man für jede Strecke Verkehrsstärken. Da die Routenwahl neu berechnet wird, ist es nun möglich, Auswirkungen von Baustellen, wie beispielsweise Geschwindigkeitsbegrenzungen und Verringerung der Kapazität, in die Tagesprognose einzubeziehen. Anders als im ersten Ansatz erfolgt die Routenwahl im Modell nun nach den Widerständen, in denen sich die aus den Baustellen resultierenden Behinderungen niederschlagen. Damit wird davon ausgegangen, dass der Verkehrsteilnehmer über die Behinderung durch die Baustelle auf der entsprechenden Strecke informiert ist und in seine Routenwahl mit einbezieht. Da es sich um ein statisches Verfahren handelt, geht der durch Überlastungen verursachte Rückstau im Netz nur durch eine Erhöhung der Fahrzeit an der überlasteten Strecke in die Prognose ein. Die räumliche Ausbreitung eines Rückstaus wird nicht berücksichtigt. Genauso wird die Entlastung der Strecken hinter der Engstelle und damit eventuelle Fahrzeitverkürzungen für Verkehrsteilnehmer, die nur Straßen nach der Engstelle befahren, nicht modelliert.

Um stundenfeine Verkehrsstärken zu erhalten, werden die 24 Quelle-Ziel-Matrizen eines Kalendertages in 24 unabhängigen Schritten auf das Verkehrsnetz umgelegt. Überlastungen einer Stunde werden nicht in den nächsten Zeitschritt übertragen. Wie im ersten Ansatz wird aus dem Auslastungsgrad der einzelnen Strecken der LOS abgeleitet.

6.1.3 Verkehrszustände aus einer statischen Umlegung mit Rückstau

Dieses Verfahren unterscheidet sich von den vorherigen durch die Modellierung der Rückstauausbreitung an Engstellen. Zunächst wird eine statische Umlegung, wie sie oben beschrieben wurde, durchgeführt. Danach wird entlang jeder Route ein Anteil der Nachfrage von einer Strecke zur nächsten solange weiter gereicht, bis die Kapazität auf keiner der Strecken mehr überschritten wird. Dabei werden die folgenden Regeln eingehalten:

- Über jede Strecke fließt höchstens so viel Verkehr, wie die Kapazität angibt. Es zählt die Verkehrsstärke, die die Strecke verlässt (Engpass am Streckenende).
- Auf jeder Strecke steht höchstens so viel Verkehr, wie die Standkapazität der Strecke zulässt. Die Standkapazität ergibt sich aus der Staudichte (Fzg/km).
- Existiert ein Stau auf einer Strecke, so kann kein Verkehr daran vorbeifließen, auch wenn die betroffene Route nicht über den Engpass führt, der den Stau verursacht.

Im Anschluss daran kann aus dem Rückstau über ein Warteschlangenmodell die Zeit berechnet werden, die benötigt wird, bis sich der Stau abgebaut hat. Der LOS einer Strecke kann nun entweder aus der Verlustzeit, aus der mittleren Wartezeit oder aus der Verkehrsstärke des Rückstaus abgeleitet werden.

Um stundenfeine Verkehrsstärken zu erhalten, werden wie im zweiten Ansatz die 24 Quelle-Ziel-Matrizen eines Kalendertages in 24 unabhängigen Schritten auf das Verkehrsnetz umgelegt. Überlastungen einer Stunde werden nicht in den nächsten Zeitschritt übertragen.

6.1.4 Verkehrszustände aus einer dynamischen Umlegung

Dynamische Umlegungsverfahren unterscheiden sich von den in den Ansätzen 2 und 3 verwendeten statischen Verfahren durch die explizite Modellierung der Zeitachse. Sie umfassen neben den Regeln für die Routenwahl (→ Routenwahlmodell) zusätzliche Regeln, die festlegen, wie sich die Verkehrsnachfrage, d.h. die Verkehrsteilnehmer im Netz fortbewegen (→ Verkehrsflussmodell). Wie bei der statischen Umlegung besteht die Aufgabe der dynamischen Umlegung in der Berechnung der Verkehrsstärken im Netz aus der Verkehrsnachfrage, aber für den Fall einer zeitlich veränderlichen Nachfrage F(t) und einem zeitlich veränderlichen Verkehrsangebot A(t). Als Ergebnis erhält man somit zeitabhängige Verkehrsstärken und Fahrzeiten. Für die dynamische Umlegung im Individualverkehr gibt es eine Vielzahl von Lösungsansätzen. Für die Tagesprognose wird ein Ansatz vorgeschlagen, der nur auf einem einfachen, kapazitätsabhängigen Verkehrsflussmodell basiert (Friedrich et al., 2000). Die Zeitachse wird dabei in 96 Intervalle a 15-Minuten diskretisiert. Der zeitliche Fahrtverlauf durch das Netz wird entlang der Route abgetragen, wobei für jedes Netzelement die Fahrzeit des Zeitintervalls relevant ist, in dem das Netzelement befahren wird.

Im Gegensatz zu den ersten drei Ansätzen, werden die zeitabhängige Verkehrsstärke und die Fahrzeit in einem geschlossenen Verfahren bestimmt, bei dem Rückstau aus Überlastungen in nachfolgende Zeitintervalle übernommen wird. Der LOS einer Strecke kann nun entweder aus der Verlustzeit oder aus dem Auslastungsgrad abgeleitet werden.

6.2 Umsetzung der Tagesprognose

6.2.1 Verwendete Software

Die für die Tagesprognose erforderlichen Umlegungen werden mit dem Softwareprogramm VISUM der PTV AG (PTV, 2006) durchgeführt. Hier sind die in Kapitel 6.1 dargestellten Verfahren implementiert. Auch die LOS-Darstellungen werden mit VISUM erzeugt. Dazu wird für jede Stunde eines Kalendertages eine jpg-Datei erstellt. Die Steuerung der tagesspezifischen Streckenattribute, die Berechnung der Umlegung und die Erstellung der LOS-Darstellungen erfolgen mit einem Visual Basic Skript.

6.2.2 Widerstände bei der Routenwahl

Routenwahlmodelle basieren auf in der Regel monokriteriellen Entscheidungsmodellen, d.h. die verschiedenen Einflussfaktoren werden zu einem Wert zusammengefasst. Dieser Wert kann als Widerstand bezeichnet werden. Oft wird hierfür auch der Begriff generalisierte Kosten oder Nutzen (Widerstand = negativer Nutzen) verwendet.

Der Widerstand einer Route setzt sich aus den Widerständen der einzelnen Netzelemente (Strecken, Knotenpunkte) zusammen. Bei einer statischen Betrachtung ohne Berücksichtigung der Zeitachse setzt sich der Widerstand w_r einer Route r aus der Summe der Widerstände aller S Netzelemente (Strecken, Knotenpunkte) zusammen, über die die Route führt:

$$w_r = \sum_{s=1}^{S} w_s$$

mit

w_r Widerstand einer Route r

S Anzahl Netzelemente der Route r

w_s Widerstand für Netzelement s

Der Widerstand w_s einer Strecke ist eine Funktion die maßgeblich von folgenden Variablen beeinflusst wird:

- Fahrzeit im belasteten Netz *t* [s],
- Streckenlänge *l* [m],
- streckentypabhängige Abminderungsfaktoren.

Streckentypabhängige Abminderungsfaktoren ermöglichen es, spezielle Streckentypen zu bevorzugen. Das ist insbesondere bei hochrangigen Straßen (Autobahnen) sinnvoll, da diese Straßen durch eine gute Beschilderung den nicht ortskundigen Fahrern oft als einzige Alternative bekannt sind. Damit hat eine Widerstandsfunktion für eine Strecke im Projektnetz folgende Form:

$$w_s(Q) = \beta_{StrTyp} \cdot t_s(Q) + \beta_{Lge} \cdot l_s$$

mit

$w_s(Q)$	Widerstand der Strecke s bei Belastung Q [s]
$t_s(Q)$	Fahrzeit der Strecke s bei Belastung Q [s]
l_s	Länge [m]
β_{StrTyp}	streckentypspezifischer Faktor [-]
β_{Lge}	Längenfaktor [s/m]

Die Parameter der Widerstandsfunktion β_{StrTyp} und β_{Lge} werden anhand der beobachteten Routenwahl so kalibriert, dass die Routenwahl in der Umlegung soweit wie möglich mit der beobachteten Routenwahl übereinstimmt. Mit den beobachteten Routen ergeben sich folgende Werte:

$$\beta_{BAB} = 0{,}95$$
$$\beta_{BAB_Baustelle} = 0{,}8$$
$$\beta_{Sonstige\,Straßen} = 1{,}0$$
$$\beta_{Lge} = 0{,}02$$

Wie bei den Autobahnen wird auch der Streckenfaktor für die Baustellenabschnitte auf einen Wert kleiner 1 gesenkt. Damit wird die Toleranz der Verkehrsteilnehmer gegenüber Baustellen auf Autobahnen im Modell erhöht. Die Interpretation ist, dass eine Fahrzeitverlängerung aufgrund einer Autobahnbaustelle die Verkehrsteilnehmer nicht im gleichen Maße zu einer Änderung ihrer Routenwahl veranlasst, wie eine Fahrzeitverlängerung im untergeordneten Netz beispielsweise durch viele Ortsdurchfahrten.

6.2.3 Ermittlung der belastungsabhängigen Fahrzeit bei freier Strecke

Zur Berechnung der belastungsabhängigen Fahrzeit, die einen Teil des Netzwiderstandes darstellt, werden sogenannte CR-Funktionen (Capacity Restraint) der Form

$$t_{ist} = t_0 \cdot (1 + a \cdot \left(\frac{Q}{C \cdot c} \right)^{b}) \quad für\, Q \le C$$

$$t_{ist} = t_0 \cdot (1 + a \cdot \left(\frac{Q}{C \cdot c} \right)^{b'}) \quad für\, Q > C$$

mit

C	Kapazität des Netzelementes
Q	Verkehrsstärke des Netzelements
t_{ist}	Fahrzeit im belasteten Netz
t_0	Fahrzeit bei Q=0
a,b,b',c	Parameter

verwendet. Mittels dieser ergibt sich in Abhängigkeit der Auslastung die Fahrzeit für eine Strecke. Die in der Tagesprognose verwendeten Parameter a und b der CR-Funktionen werden mit den Messdaten der Verkehrsstärke und der Geschwindigkeit aus den Richtungsquerschnitten geschätzt. Als Fahrzeit t_{soll} wird die Fahrzeit im freien Verkehrsfluss angenommen. Diese Fahrzeit errechnet sich aus der Länge und der freien Geschwindigkeit eines Streckenabschnitts in 15-Minuten-Zeitintervallen über mehrere Monate. Aus den Fahrzeiten dieser Intervalle wird die Zeit als t_{soll} festgelegt, die von 85% der Fahrzeiten aller Intervalle nicht überschritten wird. Je nach Anzahl der Fahrstreifen ergeben sich unterschiedliche Parametersets der CR-Funktion für zwei- und dreistreifige Autobahnen.

CR-Parameter	zweistreifige Autobahnen	dreistreifige Autobahnen
a	0,17	0,2
b	1,0	2,5
b'	3,99	5,1
c	1,0	1,0

Tabelle 60: CR-Parameter für zwei- und dreistreifige Autobahnen.

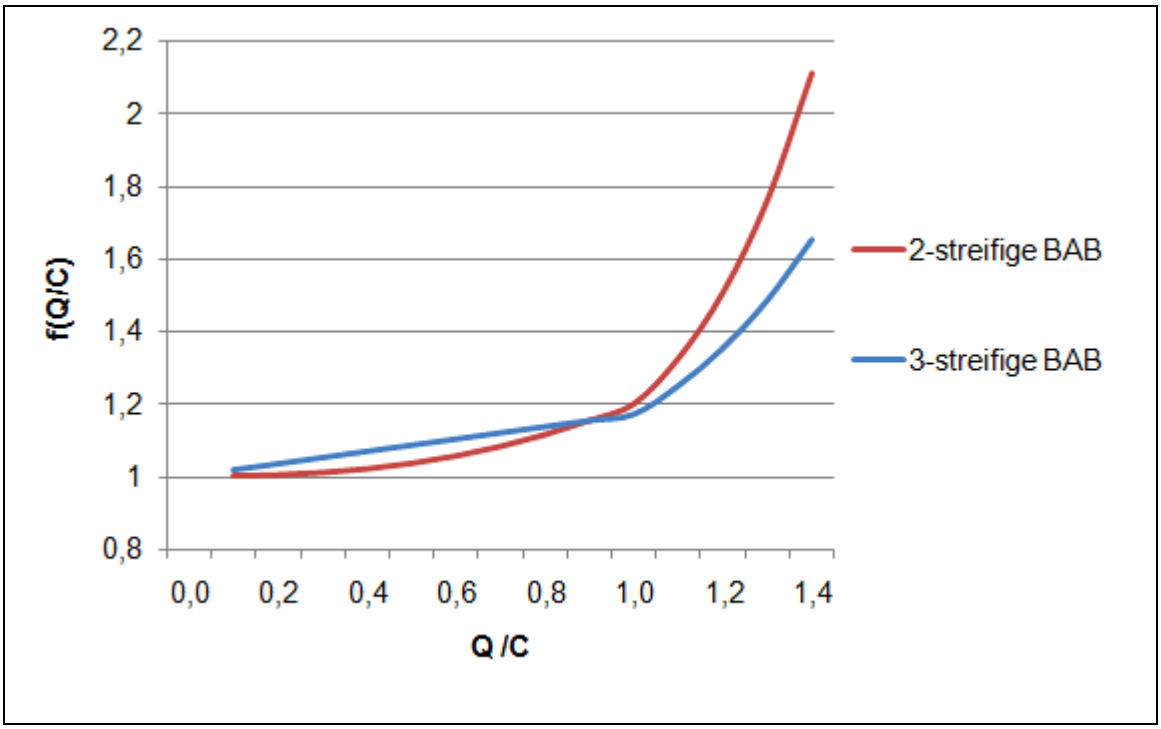


Abbildung 145: Verlauf der verwendeten CR-Kurven.

6.2.4 Ermittlung der belastungsabhängigen Fahrzeit bei Baustellen

Zur Berechnung der belastungsabhängigen Fahrzeit bei Baustellen werden nicht die oben beschriebenen CR-Funktionen verwendet, sondern eine für Baustellen angepasste CR-Funktion. Die Besonderheit dieses Funktionstyps ist, dass im ausgelasteten Zustand ein zusätzlicher Zeitzuschlag je Fahrzeug angegeben werden kann.

$$t_{ist} = t_0 \cdot (1 + a \cdot \left(\frac{Q}{C \cdot c} \right)^b) \quad für\ Q \le C$$

$$t_{ist} = t_0 \cdot (1 + a \cdot \left(\frac{Q}{C \cdot c} \right)^{b'}) + (Q - C) \cdot d \quad für\ Q > C$$

Die CR-Parameter werden für die Baustellenstrecken angepasst. Zunächst werden die Fahrzeiten für den Baustellenabschnitt und den restlichen Teil der Strecke für verschiedene Belastungen getrennt berechnet. Im nächsten Schritt werden für Autobahnen die CR-Parameter so festgelegt, dass die Gesamtfahrzeit auf beiden Streckenteilen abgebildet wird. Je nach Verhältnis von freier Strecke zu Baustellenstrecke, zugelassener Geschwindigkeit im Baustellenbereich und Anzahl der Spuren je Richtungsfahrbahn variiert der Parameter a. Der Parameter b bleibt konstant. Abbildung 146 zeigt beispielhaft die verwendeten Parameter a auf Baustellenstrecken dreistreifiger Autobahnen, die auf zwei Fahrstreifen reduziert wurden.

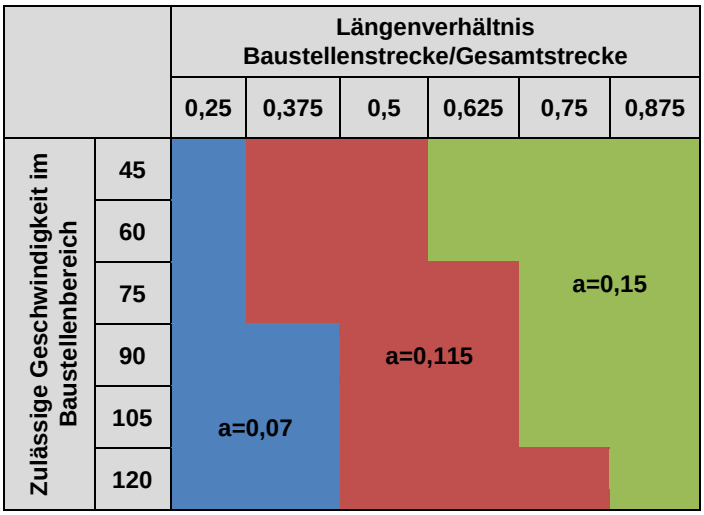


Abbildung 146: CR-Parameter a nach Längenverhältnis und Geschwindigkeit auf dreistreifigen Autobahnen mit zwei Fahrstreifen im Baustellenbereich.

Beispielhaft wird im Folgenden die Wirkung der CR-Funktion auf die Berechnung der Fahrzeit auf Baustellenabschnitten gezeigt. Dazu werden die CR-Funktionen eines fiktiven dreistreifigen Autobahnabschnittes mit und ohne Baustelle gegenübergestellt. Als Kapazität wird in beiden Fällen die Baustellenkapazität dreistreifiger Autobahnen angenommen, um die Vergleichbarkeit zu gewährleisten. Die Eckdaten sind in folgender Tabelle 61 aufgeführt:

	Baustellenstrecke			**Vergleichs-strecke ohne Baustelle**
	Freie Strecke	Baustelle	Gesamt	
Länge [km]	4	1	5	5
Zulässige Geschwindigkeit v [km/h]	130	60	-	130
Fahrzeit bei Q=0; t0 [s]	111	60	171	138
Kapazität C [Fzg/h]	5.400	3.060	-	5.400
Längenverhältnis Baustellen-strecke/Gesamtstrecke			0,2	-
CR Parameter			a=0,07 b=3,8 b'=3,8 c=1 d=1,5	a=0,2 b=2,5 b'=5,1 c=1

Tabelle 61: Kenndaten einer Beispielstrecke mit und ohne Baustelle.

Abbildung 147 zeigt den Verlauf der entsprechenden CR-Kurven. Aufgrund der Geschwindigkeitsbegrenzung im Baustellenbereich ist die Fahrzeit auf der Baustellen-strecke länger. Solange die Verkehrsstärke unterhalb der Kapazitätsgrenze liegt, entwickelt sich die Fahrzeit auf beiden Strecken gleich. Sobald die Kapazität erreicht wird, steigt die Fahrzeit sowohl auf der Baustellen- als auch auf der Vergleichsstrecke

stark an. Aufgrund der unterschiedlichen Fahrstreifenanzahl wird diese Kapazitätsgrenze auf der Baustellenstrecke bei weniger Fahrzeugen überschritten als auf der Vergleichsstrecke. Darüber hinaus ist der Anstieg der Fahrzeit auf der Baustellenstrecke höher.

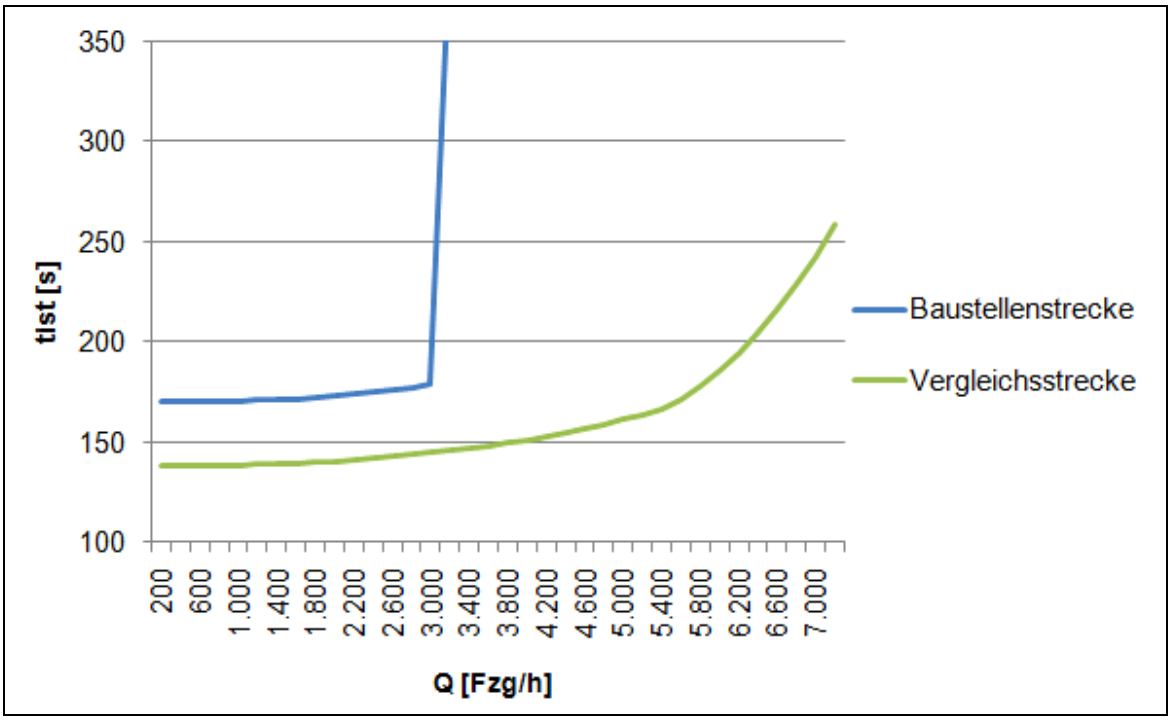


Abbildung 147: Verlauf der CR-Kurven auf einem sechsstreifigen Autobahnabschnitt (5km) ohne Baustelle und mit Baustelle (Reduktion um je einen Richtungsfahrstreifen).

6.2.5 Verwaltung von Baustellen

In die Tagesprognose werden die Auswirkungen von Baustellen auf den Verkehrsablauf mit einbezogen. Informationen über Länge, Standort und Dauer einer Baustelle, sowie die damit verbundene Verkehrsführung und eventuelle Geschwindigkeitsbegrenzungen werden von der Landesstelle für Straßentechnik des Regierungspräsidiums Tübingen zur Verfügung gestellt. Tabelle 62 zeigt die Baustellenverwaltung für eine Beispielstrecke mit Hin- und Rückrichtung.

	Strecke	
Streckennummer	3.030	3.030
Von Knotennr.	11.890	11.891
Nach Knotennr.	11.891	11.890
Länge [km]	9,886	9,725
Kapazität Ursprünglich [Fzg/h]	3.600	3.600
Länge Baustelle [km]	9,0	9,0
v Baustelle [km/h]	80	80
Streckentyp	62	62
V_{soll} korreliert [km/h]	82,82	81,49
Kapazität korreliert [Fzg/h]	3.060	3.060
von Datum	02.04.2009	02.04.2009
bis Datum	22.06.2009	22.06.2009

Tabelle 62: Verwaltung von Baustellen.

Die Darstellung der Baustellen im Netzmodell wird durch die Veränderung der Streckeneigenschaften realisiert. Gibt es am gewählten Prognosetag eine Baustelle, wird die entsprechende Strecke aufgrund ihrer Streckennummer und den Anfangs- und Endknotennummern in VPN identifiziert, um ihr neue, korrigierte Baustellenattribute zuzuordnen. Als Kapazität wird dabei die aufgrund von Spurreduktion veränderte Baustellenkapazität, in der Tabelle als „Kapazität korreliert“ bezeichnet, angenommen. Diese liegt nach einer Untersuchung der Universität Hannover für Fahrstreifenreduktionen von drei auf zwei Fahrstreifen auf der freien Strecke von Autobahnen im Mittel um 15% unter der einer durchgängig zweistreifigen Richtungsfahrbahn (Friedrich et al., 2003).

Die Streckengeschwindigkeit (v_{soll} korreliert) entspricht einer über die gesamte im Netzmodell enthaltene Strecke gewichteten Geschwindigkeit. Damit wird der Tatsache Rechnung getragen, dass die Länge der Baustelle in der Realität nicht der Länge einer Strecke im Netzmodell entspricht.

6.2.6 Ableitung des LOS

Die Qualitätsstufen des Verkehrsablaufs (LOS) werden je nach Verfahren aus der Auslastung oder aus dem Verhältnis der Fahrzeit im belasteten Netz t_{ist} zur Fahrzeit im unbelasteten Netz t_{soll} abgeleitet. Tabelle 63 zeigt die Klassifizierung des LOS.

Qualitätsstufe	Auslastung	Verhältnis t_{ist}/t_{soll}
frei	< 80%	< 1,5
stockend	80% bis 99%	1,5 bis 2,0
gestaut	> 99 %	> 2,0

Tabelle 63: Ableitung des LOS.

6.2.7 Netzkalibrierung

Zur Netzkalibrierung wurden die Umlegungsparameter der verschiedenen Umlegungsverfahren variiert. Die im folgenden Kapitel präsentierten Ergebnisse stammen aus Umlegungen mit der Parameterkombination, die in dem jeweiligen Verfahren die besten Ergebnisse liefert.

6.3 Ergebnisse

In diesem Kapitel wird beispielhaft dargestellt, wie die Ergebnisse einer Tagesprognose aussehen. Als Input werden in einer ersten Version die Nachfragematrizen des entsprechenden Prognosetages verwendet. Anschließend dienen die durch die Clusterung ermittelten Erwartungswerte als Nachfragematrizen. Da bei der Hochrechnung der Matrizen nur Zähldaten der Autobahnen berücksichtigt werden, beschränkt sich die Tagesprognose auf diese Straßenklasse.

Als Ergebnis der Verkehrsumlegung erhält man ein Netz mit Verkehrsbelastungen. Entsprechend Kapitel 6.2.6 können daraus Auslastungsgrade bzw. der Level of Service bestimmt werden.

6.3.1 Bewertung der Ergebnisse

6.3.1.1 Bewertung der absoluten Abweichung

Die absolute Abweichung der Verkehrsstärken in der Realität und in der Tagesprognose kann mittels eines Differenznetzes abgebildet werden. Abbildung 148 zeigt dieses für den 23.06.2008, für die Zeit zwischen 7:00 Uhr und 8:00 Uhr. Auf grün gekennzeichneten Strecken werden höhere, auf roten Strecken niedrigere Verkehrsbelastungen prognostiziert, als tatsächlich aufgetreten sind.

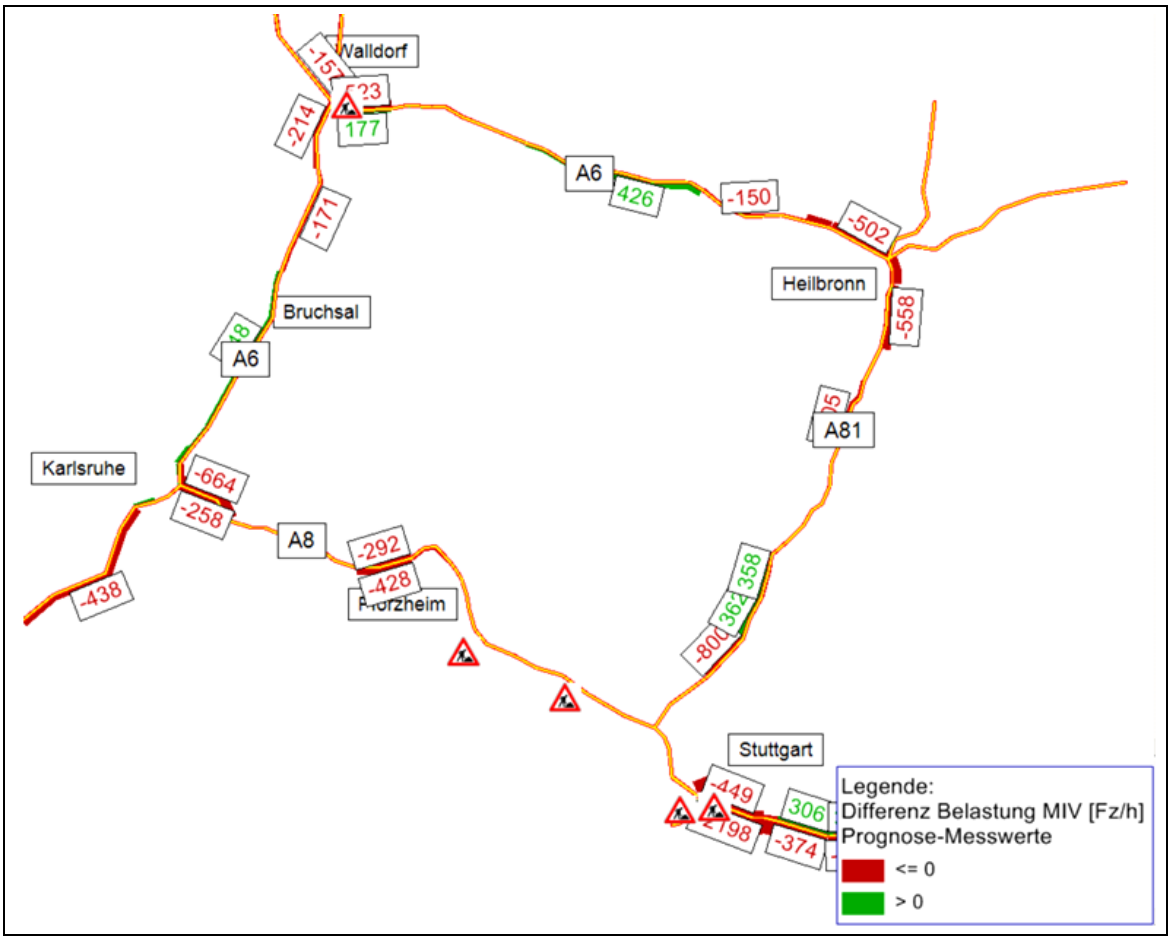


Abbildung 148: Differenzen der Verkehrsstärken zwischen Modellprognose und Messwerten am 23.06.2008 von 7:00 Uhr bis 8:00 Uhr.

Nachteilig an dieser Darstellung ist, dass nur die absoluten Abweichungen zu sehen sind. Der relative Unterschied zwischen den Vergleichswerten, der für eine Bewertung der Ergebnisse nötig ist, ist nicht ersichtlich. Aus diesem Grund lassen sich nur recht allgemeine Aussagen ableiten. Zu den höchsten Abweichungen kommt es an den Rändern des Untersuchungsgebietes. Hier sind nicht alle in der Realität vorhandenen Straßen modelliert. Daher erfolgt die Verkehrsaufteilung im Modell unter anderen Netzbedingungen. Die beobachteten Abweichungen sind damit unvermeidbar.

6.3.1.2 Bewertung der relativen und absoluten Abweichung

Um die Qualität der Verkehrsprognose bewerten und verschiedene Umlegungsverfahren miteinander vergleichen zu können, werden Kontrollkriterien eingeführt. Dabei werden Verkehrsstärken aus der Tagesprognose mit den in der Realität gemessenen Verkehrsstärken verglichen. Die Abweichung wird mittels eines GEH-Faktors bestimmt.

$$GEH = \sqrt{\frac{(M-C)^2}{(M+C)/2}}$$

mit

M	modellierte Verkehrsbelastung
C	gemessene Verkehrsbelastung

Wie in Kapitel 5.1.2.3 bereits beschrieben, bewertet dieser Wert nicht nur die relative Abweichung der Vergleichswerte sondern auch deren absolute Differenz.

Das britische Highway Agency empfiehlt die Verwendung des GEH-Faktors zur Bewertung der Qualität eines Verkehrsmodells. In einem Netz sollte der GEH-Faktor bei mindestens 85 % der Strecken unter fünf liegen.

Zur Kontrolle der Qualität der Tagesprognose wurden zwei verschiedene Verfahren entwickelt, in denen der GEH-Wert verwendet wird. Beim ersten Kontrollverfahren werden die gezählten und prognostizierten Streckenbelastungen auf acht ausgewählten, in Abbildung 149 blau hervorgehobenen Kontrollstrecken verglichen.

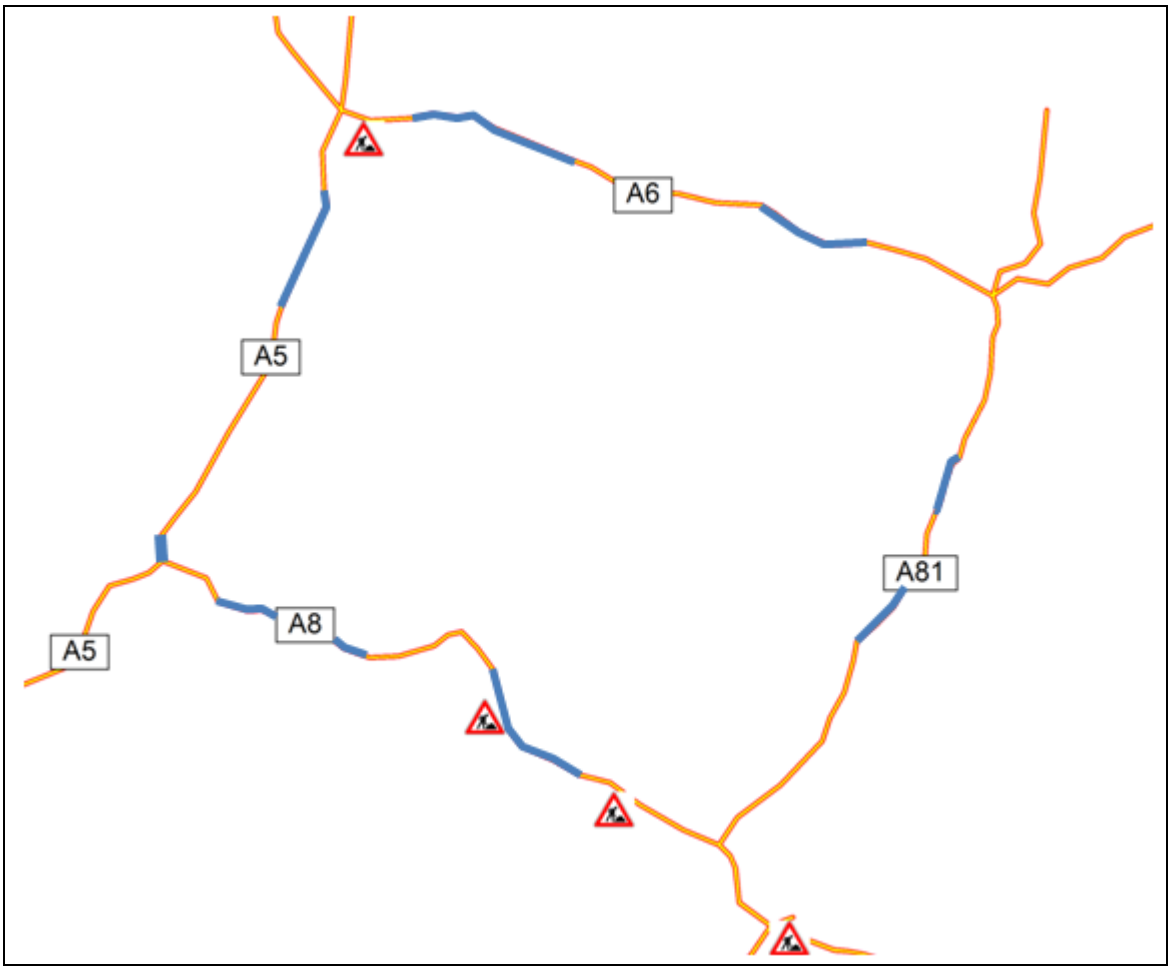


Abbildung 149: Kontrollstrecken zur Bewertung der Umlegungsqualität.

Im Ergebnis erhält man, wie in Tabelle 64 dargestellt, für jede Stunde 16 GEH-Werte, zwei je Strecke (Hin-und Rückrichtung). Daraus wird jeweils der Mittelwert berechnet, so wie der Anteil der GEH-Werte, die die Grenze von 5 überschreiten.

			Zeit		8-9	Zeit		9-10
			GEH max		8,7	GEH max		5,2
			GEH mittel		2,4	GEH mittel		1,8
			Anteil>5		13%	Anteil>5		13%
Strecken-nr.	**von KnotNr.**	**nach KnotNr.**	**Tages-prognose [Fzg/h]**	**Mess-werte [Fzg/h]**	**GEH**	**Tages-prognose [Fzg/h]**	**Mess-werte [Fzg/h]**	**GEH**
1050	11598	11599	4.217	4.152	1,0	3.453	3.468	0,2
1050	11599	11598	4.903	4.313	8,7	3.750	3.439	5,2
1055	11593	11594	4.080	4.431	5,4	3.375	3.684	5,2
1055	11594	11593	2.863	2.925	1,2	2.522	2.518	0,1
3033	11887	11888	2.648	2.755	2,1	2.301	2.395	1,9
3033	11888	11887	2.376	2.496	2,4	1.993	1.963	0,7
3037	11883	11884	1.920	1.933	0,3	1.843	1.823	0,4
3037	11884	11883	2.977	3.160	3,3	2.519	2.485	0,7
5041	12518	12519	2.504	2.353	3,1	2.057	1.932	2,8
5041	12519	12518	2.293	2.231	1,3	2.301	2.318	0,4
5045	12514	12515	3.053	2.965	1,6	2.542	2.444	2,0
5045	12515	12514	3.585	3.760	2,9	3.162	3.341	3,1
7016	12664	12665	3.532	3.652	2,0	3.090	3.258	3,0
7016	12665	12664	3.363	3.356	0,1	2.728	2.688	0,8
7018	12662	12663	3.213	3.197	0,3	2.901	2.790	2,1
7018	12663	12662	3.519	3.351	2,9	3.145	3.096	0,9

Tabelle 64: Auswertungsformular zur Berechnung der mittleren GEH-Werte der Gesamtverkehrsbelastung.

Beim zweiten Kontrollverfahren werden die Verkehrsbelastungen auf den bereits beschriebenen Kontrollstrecken verglichen, wenn nur die Verkehrsströme ausgewählter Relationen im Netz betrachtet werden. Eine Relation hat einen festen Start- und Zielpunkt. In diesem Kontrollverfahren werden ausschließlich die Fahrten betrachtet, die am Startpunkt der Relation beginnen und an deren Zielpunkt enden. Ein Beispiel für so einen Verkehrsstrom ist in Abbildung 150 dargestellt. Hierfür werden die 62 verkehrsstärksten Relationen ausgewählt. Dies liefert eine Aussage darüber, ob die in der Realität beobachtete Routenwahl der in der Tagesprognose modellierten Routenwahl entspricht. Da es auf einigen der ausgewählten Relationen keine bzw. nur sehr wenige Alternativen für eine Route gibt, wird die mittlere Abweichung der GEH-Werte für alle Relationen einer Stunde sehr niedrig und ist wenig aussagekräftig. Hier ist die in einer Stunde auftretende maximale Abweichung auf einer Relation als Kontrollgröße geeigneter. Sie liefert Aufschluss darüber, auf welcher Relation die modellierte Routenwahl von der in der Realität beobachteten Routenwahl am stärksten abweicht und wie groß diese Abweichung ist.

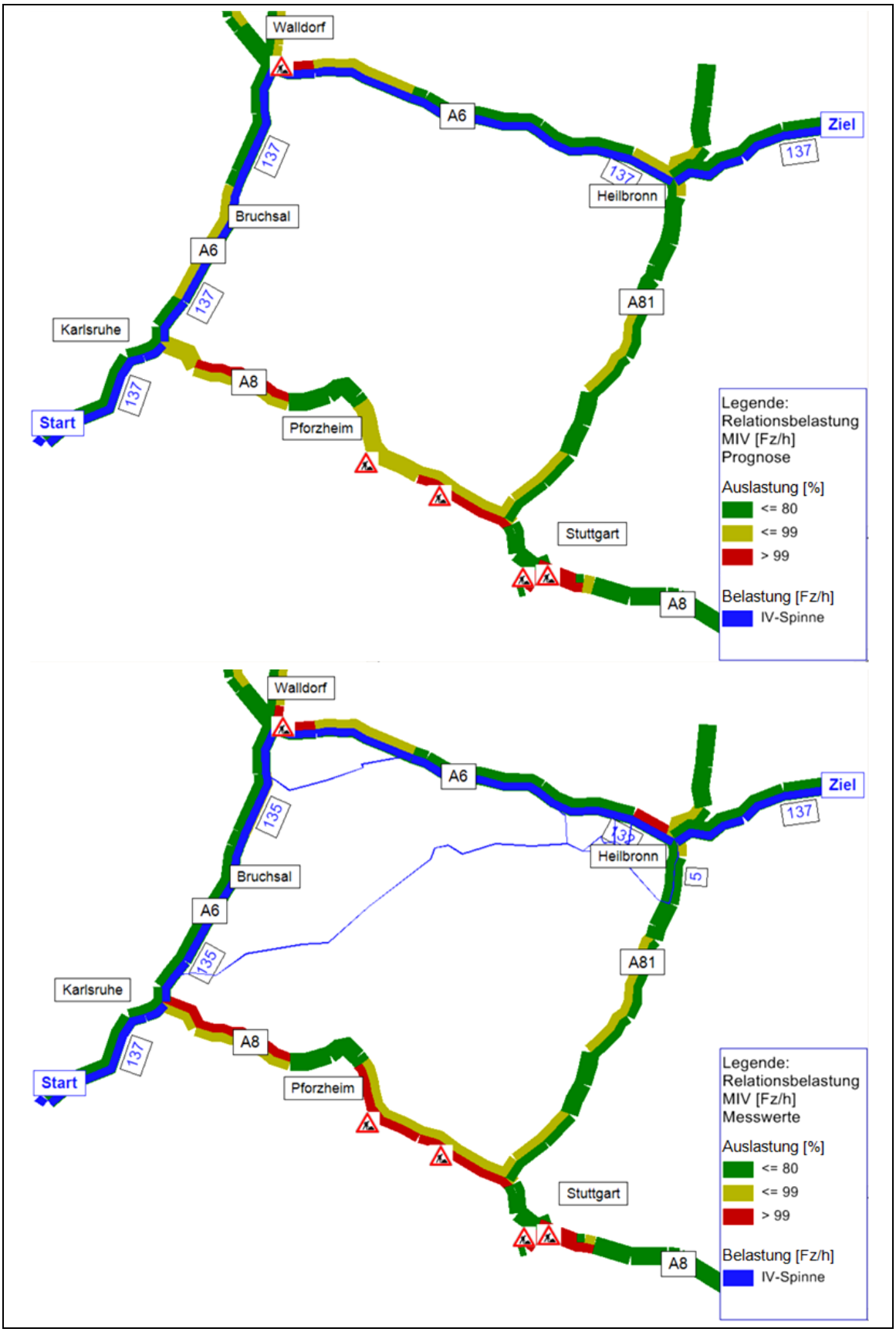


Abbildung 150: Verkehrsstrom auf einer ausgewählten Relation in der Tagesprognose (oben) und in der Realität (unten) am 23.06.08, 7:00 Uhr - 8:00 Uhr.

6.3.2 Ergebnisse der Zustandsprognose aus historischen Ganglinien

Im Rahmen der Untersuchungen wird für 68 Tage eine Tagesprognose erstellt. Eine Möglichkeit zum Erstellen einer Tagesprognose besteht darin, die Streckenbelastungen aus typischen, geclusterten Tagesganglinien abzuleiten. Um die Qualität der Tagesprognose zu bewerten, werden die daraus resultierenden Verkehrsbelastungen mit den tatsächlich an dem Tag gemessenen Verkehrsstärken verglichen. Es werden die typischen Verkehrsstärken auf den Teststrecken, die an allen dem Cluster zugehörigen Tagen gemessen wurden, den Verkehrsstärken des Prognosetages gegenüber gestellt und mit dem ersten der vorgestellten Verfahren bewertet. Abbildung 151 zeigt die Ergebnisse für drei zufällig ausgewählte Tage. Diese drei Testtage gehören laut Aufteilung in Kapitel 5.1 zum Cluster C 01, Werktag – Schule Montag bis Donnerstag (siehe Tabelle 54).

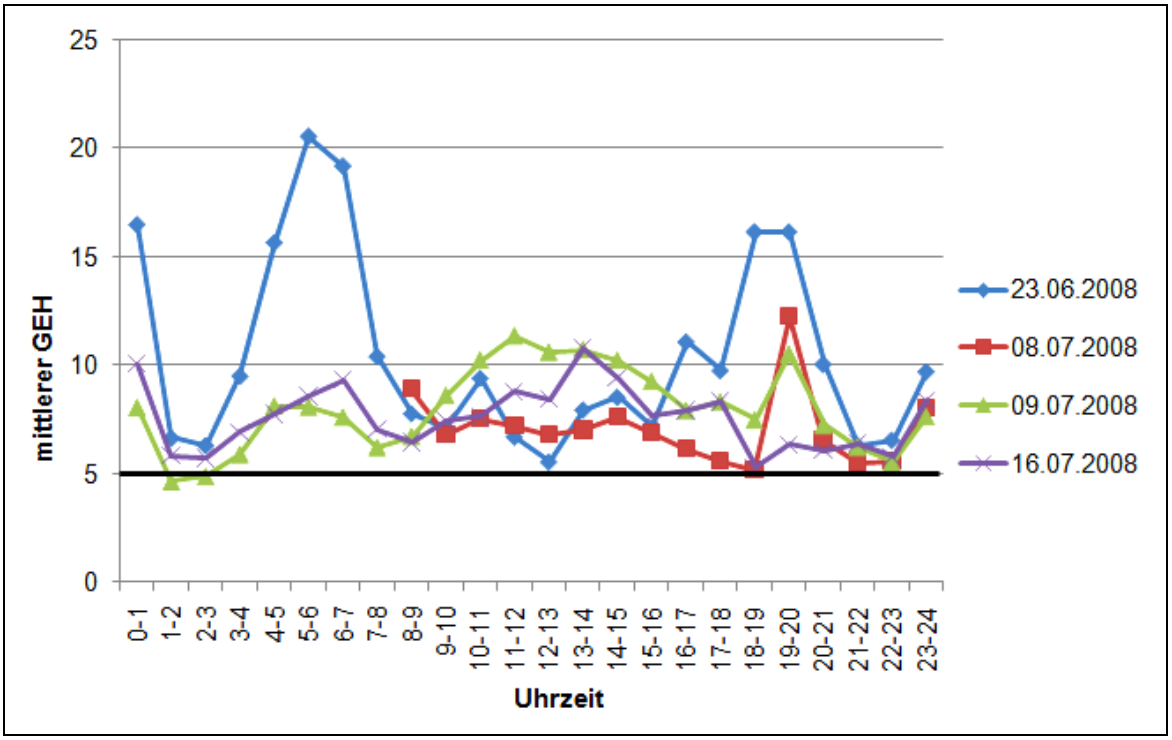


Abbildung 151: Prognosequalität für mehrere Testtage mit historischen Ganglinien.

Die als GEH-Werte ausgedrückten stündlichen Mittelwerte der Abweichungen zwischen den Verkehrsstärken liegen an jedem Prognosetag über der empfohlenen Grenze von 5. Damit ist die Prognosequalität als niedrig einzuschätzen, die tatsächlich auftretenden Verkehrsstärken können nur ungenau durch das Modell prognostiziert werden. Ein Grund dafür ist, dass durch dieses Verfahren Baustellen, die damit verbundenen Störungen des Verkehrsablaufs und in der Folge auftretende Veränderungen in der Routenwahl nicht in die Prognose einbezogen werden. Eine andere Ursache für die Ungenauigkeiten in der Prognose könnte in der Methodik der Clusterbildung liegen. Dabei werden nur die wichtigsten Einflüsse auf die Verkehrsnachfrage wie beispielsweise die Ferienzeiten in Baden-Württemberg, sowie Feiertage

und Wochenenden als Kriterium zur Zuordnung der einzelnen Tage zu den Clustern berücksichtigt. Zahlreiche andere Einflussfaktoren, wie Ferien in angrenzenden Bundesländern oder Witterungseinflüsse z.B. durch starken Schneefall können hingegen nicht in die Clusterbildung einbezogen werden. Die entstehenden Mittelwerte der Verkehrsstärken sind dadurch starken Streuungen unterworfen, so dass die Abweichungen zu einem einzelnen beliebigen Tag aus dem Cluster sehr hoch ausfallen können.

6.3.3 Ergebnisse der Zustandsprognose aus statischem Umlegungsverfahren

Als statische Verfahren werden das Gleichgewichtsverfahren, das Lernverfahren und die stochastische Umlegung getestet. Dabei dienen für jedes Cluster die aus den hochgerechneten Fahrten ermittelten Quelle-Ziel-Matrizen für jede Stunde des Prognosetages als Ausgangsdaten. Je nach Verfahren stehen eine unterschiedliche Anzahl Parameter zur Verfügung mit denen das Modell kalibriert werden kann. Um die Prognosequalität zu bestimmen, werden die Verkehrsbelastungen der Prognose mit den tatsächlich an dem jeweiligen Testtag gemessenen Verkehrsbelastungen verglichen.

6.3.3.1 Ermittlung eines geeigneten Umlegungsverfahren

Im ersten Schritt wird das für diese Aufgabe am besten geeignete Umlegungsverfahren identifiziert und das Modell kalibriert. Dazu werden nicht die geclusterten Quelle-Ziel-Matrizen verwendet, sondern die Quelle-Ziel-Matrix, die sich aus den gemessenen Belastungen eines bestimmten Testtages ergibt. Abbildung 152 zeigt die mittleren GEH-Werte der Gesamtverkehrsbelastung für den 23. Juni 2008. Es werden die Abweichungen der Gesamtstreckenbelastung zwischen Modell und Realität aus verschiedenen statischen Umlegungsverfahren verglichen. Für jedes Verfahren werden jeweils die Parameterkombinationen verwendet, die die besten Ergebnisse liefern. Es wird deutlich, dass mit dem stochastischen Verfahren die besten Ergebnisse zu erzielen sind. Die mittleren GEH-Werte sind niedriger, als bei den anderen beiden Umlegungsverfahren. Knapp sieben Prozent aller GEH-Werte überschreiten die Grenze von 5.

Da sich die Umlegung und die Vergleichsbelastungen auf den gleichen Tag bezieht, resultieren Abweichungen hier entweder aus Ungenauigkeiten in der Modellierung oder aus unvorhersehbaren Ereignissen, wie z.B. Unfällen, die im Verkehrsmodell nicht dargestellt werden können.

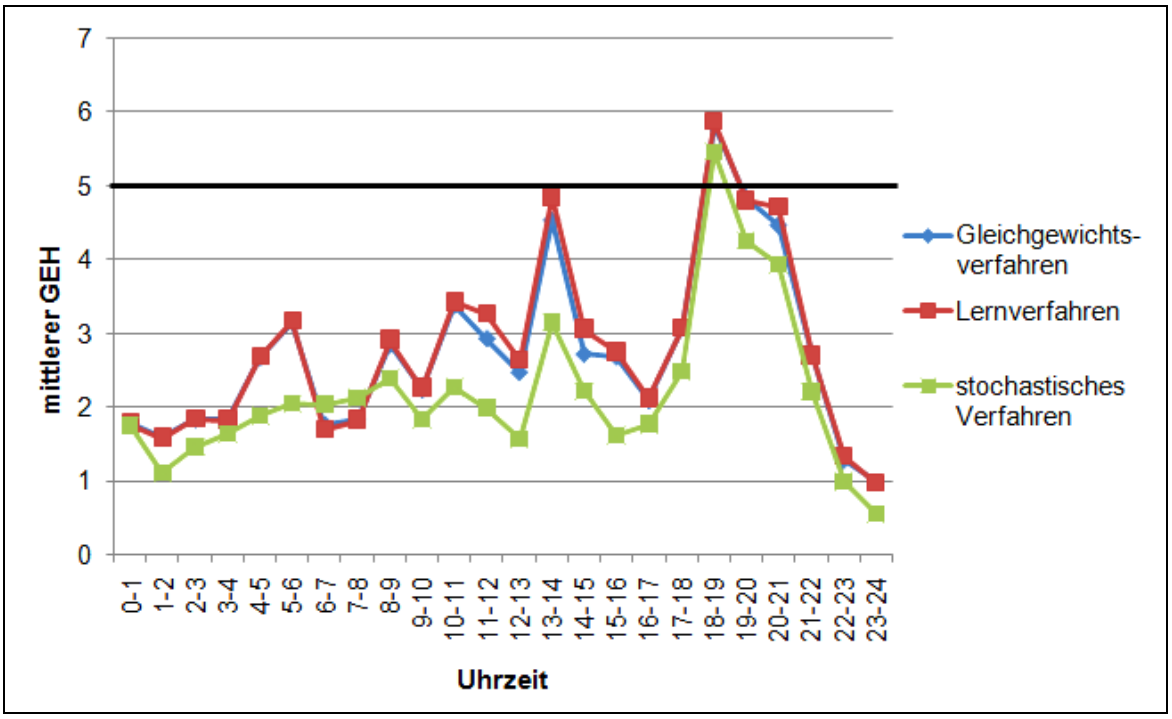


Abbildung 152: Vergleich verschiedener Umlegungsverfahren für den 23.06.2008.

Mittels der statischen Umlegungsverfahren werden für insgesamt 68 Tage in der Vergangenheit Umlegungen erstellt und die Ergebnisse mit den in der Realität gemessenen Verkehrsstärken verglichen. Wie am 23.06.2008 werden auch an den anderen Tagen mit der stochastischen Umlegung die besten Ergebnisse erzielt. Abbildung 153 zeigt für die einzelnen Tage die mittleren GEH-Werte auf den Teststrecken als Maß für die Qualität der Verkehrsprognose. Dabei werden die 24 stündlichen mittleren GEH-Werte eines Tages zu einem Tagesmittel zusammengefasst.

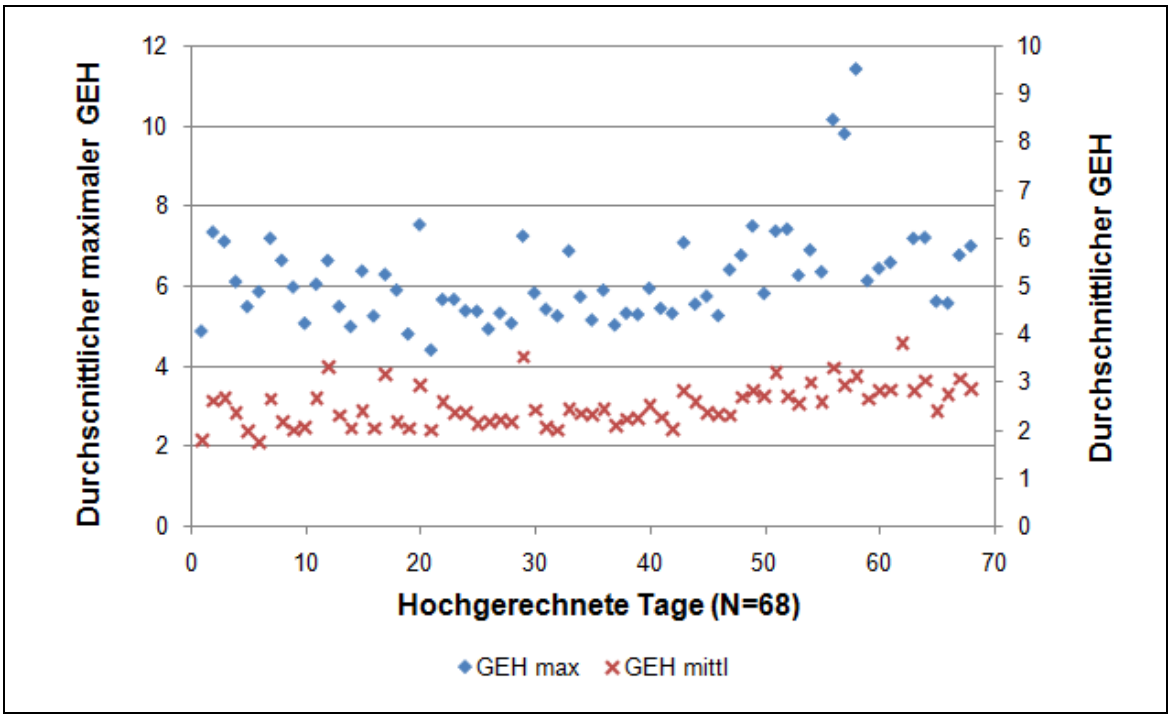


Abbildung 153: Prognosequalität für mehrere Testtage mit stochastischer Umlegung.

An keinem der Tage liegt das Tagesmittel aller an diesem Tag berechneten GEH-Werte über fünf. Ein großer Teil der GEH-Werte bewegt sich zwischen zwei und vier, was für eine gute Umlegungsqualität mit geringen Abweichungen zwischen modellierten und gemessenen Verkehrsstärken spricht.

Über alle Tage betrachtet, beträgt der Anteil der GEH-Werte, die die Grenze von 5 überschreiten, elf Prozent. Damit wird die Empfehlung der Highway Agency, nur zu 15 % GEH-Werte größer fünf zuzulassen, eingehalten.

Die guten Ergebnisse werden nach Anwendung des zweiten Kontrollverfahrens zur Überprüfung der Routenwahl im Modell bestätigt. Abbildung 154 zeigt die an den einzelnen Tagen aufgetretenen maximalen Abweichungen der Verkehrsstärken auf einzelnen Relationen zwischen Tagesprognose und Realität. Auch hier werden die stündlichen Spitzenwerte zu einem Tagesmittel zusammengefasst. An keinem der Testtage übersteigen die maximalen Abweichungen den von der Highway Agency angegebenen Kontrollwert. Bezogen auf die Einzelwerte werden nur auf 0,45% aller Relationen an den betrachteten Tagen größere Abweichungen mit einem GEH von größer fünf zwischen der Realität und der Tagesprognose ermittelt.

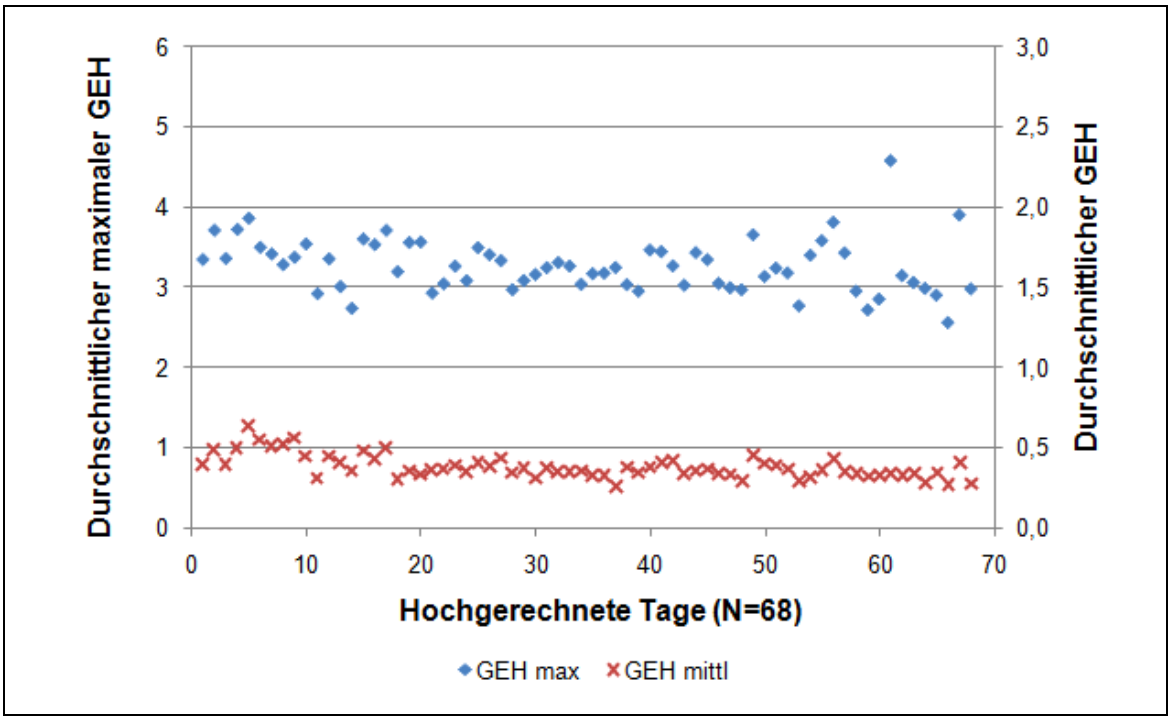


Abbildung 154: Prognosequalität der Routenwahl auf einzelnen Relationen für mehrere Testtage.

Für die beobachteten Abweichungen der modellierten von den gemessenen Verkehrsbelastungen gibt es im Wesentlichen zwei Erklärungen. Zum einen können in der Realität nicht vorhersehbare Verkehrsstörungen wie zum Beispiel Unfälle den Verkehrsablauf beeinflussen. Dieser Einfluss kann aufgrund seiner Unvorhersehbarkeit im Modell allerdings nicht abgebildet werden. Die Effekte auf das Verkehrsgeschehen werden daher im Modell nicht mit berücksichtigt. Zum anderen resultieren Abweichungen der Verkehrsbelastungen aus den Vorgang der Modellierung. Da das Umlegungsverfahren und die gewählten Parameter nicht für jeden Tag individuell angepasst werden können, wurden die entsprechenden Werte so kalibriert, dass für alle Tage sehr geringe Abweichungen auftreten.

6.3.3.2 Tagesprognose aus statischem Umlegungsverfahren

Nachdem das Umlegungsmodell kalibriert wurde, kann im zweiten Schritt die Tagesprognose für die 68 Testtage erstellt werden. Naturgemäß kann für eine Prognose in der Zukunft nun nicht wie im oben beschriebenen Teil auf die für diesen Tag gemessenen Verkehrsbelastungen zurück gegriffen werden. Statt dessen wird die für den jeweiligen Tag berechnete Matrix aus der Clusteranalyse verwendet und umgelegt. Im Ergebnis erhält man die Tagesprognose. Das Verkehrsmodell enthält Informationen zu den Baustellen für den jeweiligen Tag.

Um die Qualität der Tagesprognose zu bestimmen, werden die prognostizierten Verkehrsbelastungen mit den tatsächlich an diesem Tag gemessenen Verkehrsbelastungen verglichen. Abweichungen ergeben sich hier ebenfalls aus nichtvorhersehbaren Ereignissen, wie Unfällen. Außerdem stellt die Quelle-Ziel-Matrix des Clusters, die als Eingangsgröße für die Prognose verwendet wird, nur eine Schätzung dar. Die tatsächlichen Verkehrsbelastungen an dem betrachteten Prognosetag können davon abweichen, was ebenfalls zu einer Verschlechterung der Prognosequalität führt.

Tabelle 65 zeigt die Abweichungen, die sich nur durch die Clusterbildung ergeben, exemplarisch für vier zufällig ausgewählte Tage. Der GEH-Wert in der zweiten Spalte resultiert aus der Differenz zwischen den am jeweiligen Tag auf den 16 Teststrecken gemessenen Verkehrsbelastungen und den gemittelten Verkehrsstärken des jeweiligen Clusters. Demgegenüber stehen die Abweichungen aus der Tagesprognose in der dritten Spalte. Zusätzlich zu dem Fehler aus der Clusterbildung kommen hier Abweichungen aufgrund der Umlegung und aus nichtvorhersehbaren Ereignissen, wie z.B. Unfällen, hinzu. Durch die Einbeziehung der Baustellen erreicht man mit der Tagesprognose aber einen stärkeren Bezug zur Realität, als im anderen Verfahren. Aus diesem Grund sind die Abweichungen aus der Tagesprognose trotz höherer Anzahl möglicher Fehlerquellen nur wenig höher, als die Abweichungen zwischen Cluster- und Zählwerten.

Testtage	**durchschnittlicher Tages-GEH**	
	Zählwerte – Cluster	**Tagesprognose (Umlegung - Cluster)**
21.07.2008	6,36	5,45
22.07.2008	3,89	4,76
09.08.2008	4,84	5,04
23.08.2008	5,67	5,57

Tabelle 65: Vergleich der Abweichungen der Verkehrsstärken durch Clusterbildung und in der Tagesprognose.

Aufgrund der zusätzlichen Abweichungen durch die Clusterbildung sind die GEH-Werte der Tagesprognose als Maß für die Abweichung der Verkehrsbelastungen in der Realität und in der Prognose höher als in Kapitel 6.3.3.1, wo die tagesspezifische Quelle-Ziel-Matrix verwendet wurde. Abbildung 155 zeigt die Tagesmittel des GEH-Wertes für alle Testtage.

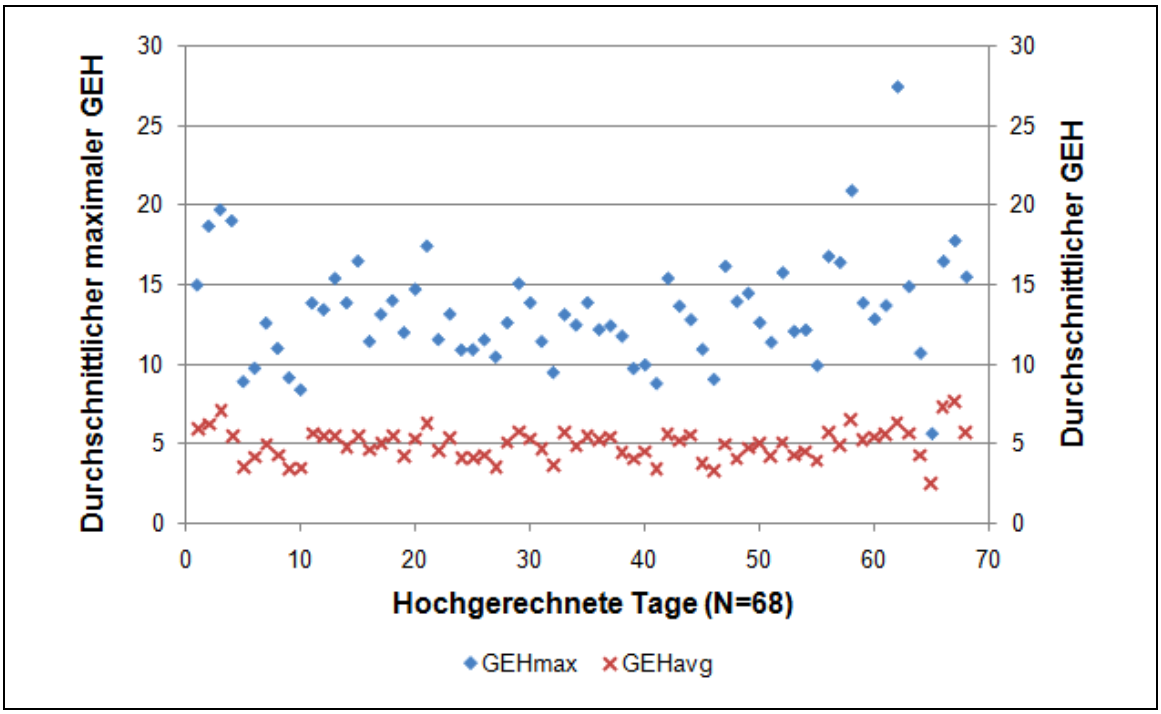


Abbildung 155: Prognosequalität für mehrere Testtage in der Tagesprognose.

6.3.3.3 Ableitung des LOS

Da für den Fahrzeugfahrer als Nutzer der Verkehrsprognose weniger die genaue Verkehrsbelastung sondern die Stausituation von Interesse ist, wird das Ergebnis in LOS-Stufen in Abhängigkeit der Auslastung nach Tabelle 63 dargestellt (Abbildung 156).

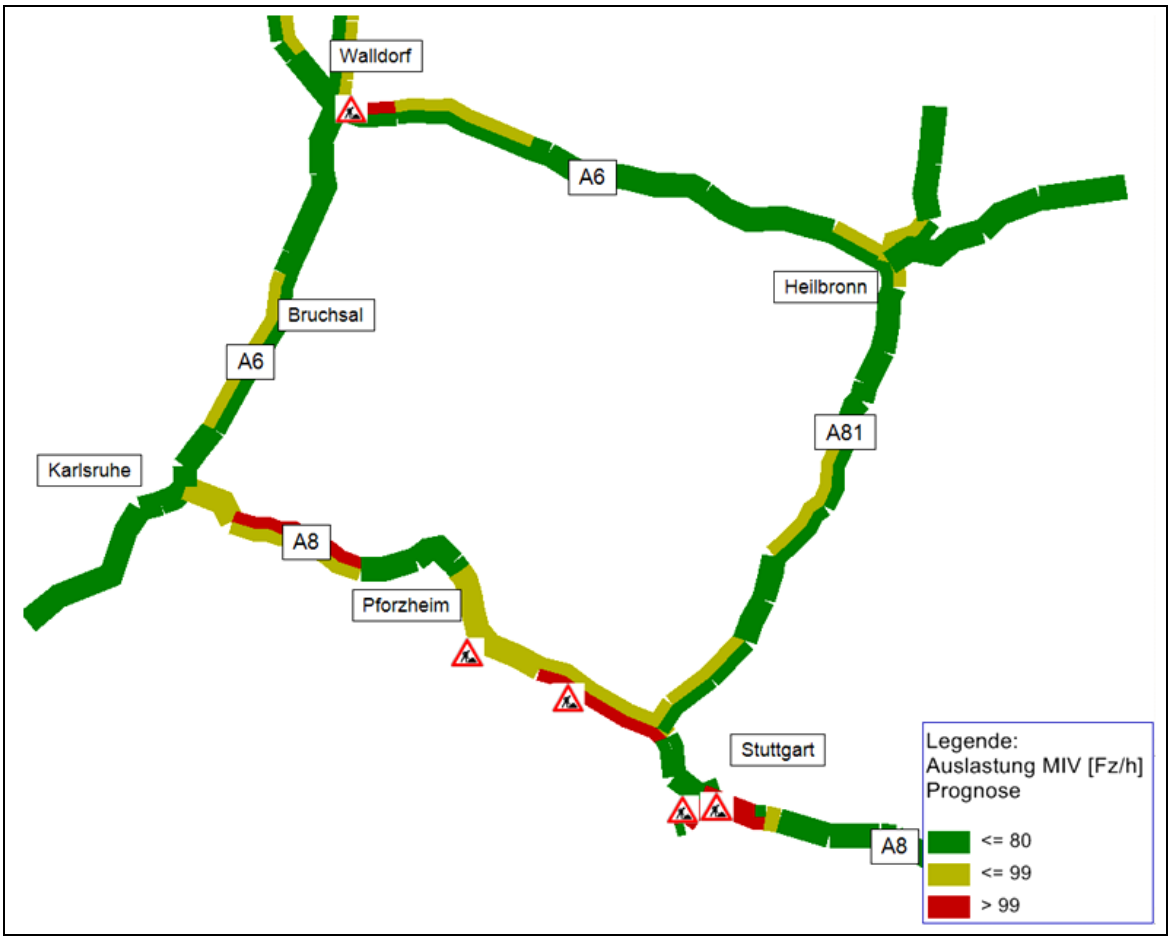


Abbildung 156: Level of Service-Darstellung als Ergebnis einer Tagesprognose am 23.06.08 von 7:00 Uhr bis 8:00 Uhr.

Für die LOS-Darstellung wird das betrachtete Autobahnnetz in 34 Teilabschnitte unterteilt. Diese Straßenabschnitte entsprechen den einzelnen Strecken im VISUM-Netzmodell. Für jede Stunde des Prognosetages werden für alle Straßenabschnitte in beiden Fahrtrichtungen Qualitätsstufen berechnet. Abbildung 157 zeigt die Häufigkeit der Abweichung zwischen prognostizierter und tatsächlich eingetretener LOS-Stufe an 68 Testtagen.

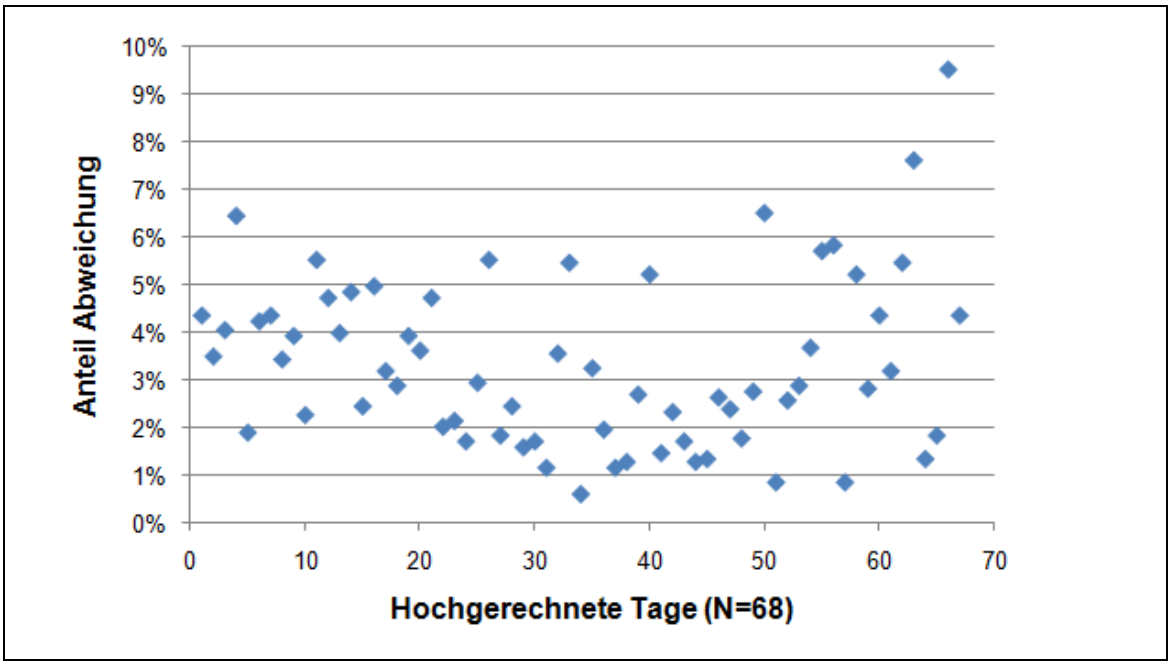


Abbildung 157: Häufigkeit der Abweichung zwischen LOS-Stufen nach Prognose und Realität.

Die Häufigkeit der Abweichung zwischen prognostizierter und gemessener LOS-Stufe ist kein Maß für die Umlegungsqualität. Trotz starker Abweichung der gemessenen und der prognostizierten Verkehrsstärke kann einem Streckenabschnitt die gleiche Qualitätsstufe zugeordnet werden. Andererseits ist es auch möglich, wenn Auslastungen im Grenzbereich zweier Qualitätsstufen ermittelt werden, dass unterschiedliche Qualitätsstufen zugeordnet werden, obwohl die absoluten Abweichungen der Verkehrsbelastungen niedrig sind.

6.3.4 Ergebnisse aus statischen Verfahren mit Rückstauberechnung

Dieses Verfahren unterscheidet sich von den statischen Verfahren nur dann, wenn es zur Staubildung kommt. In diesem Fall kann mittels der Rückstauberechnung festgestellt werden, wie viele Fahrzeuge vom Stau betroffen sind und auf welchen Straßenabschnitten es durch den Rückstau an einem Engpass zu Behinderungen kommt. Aufgrund der Tatsache, dass das vereinfachte Planungsnetz, das für die Erstellung der Tagesprognose verwendet wird, relativ grob ist, bringt die Rückstauberechnung hier wenig neue Erkenntnisse. Die Staulänge übersteigt in den meisten Fällen nicht die Standkapazität des betroffenen Streckenabschnittes.

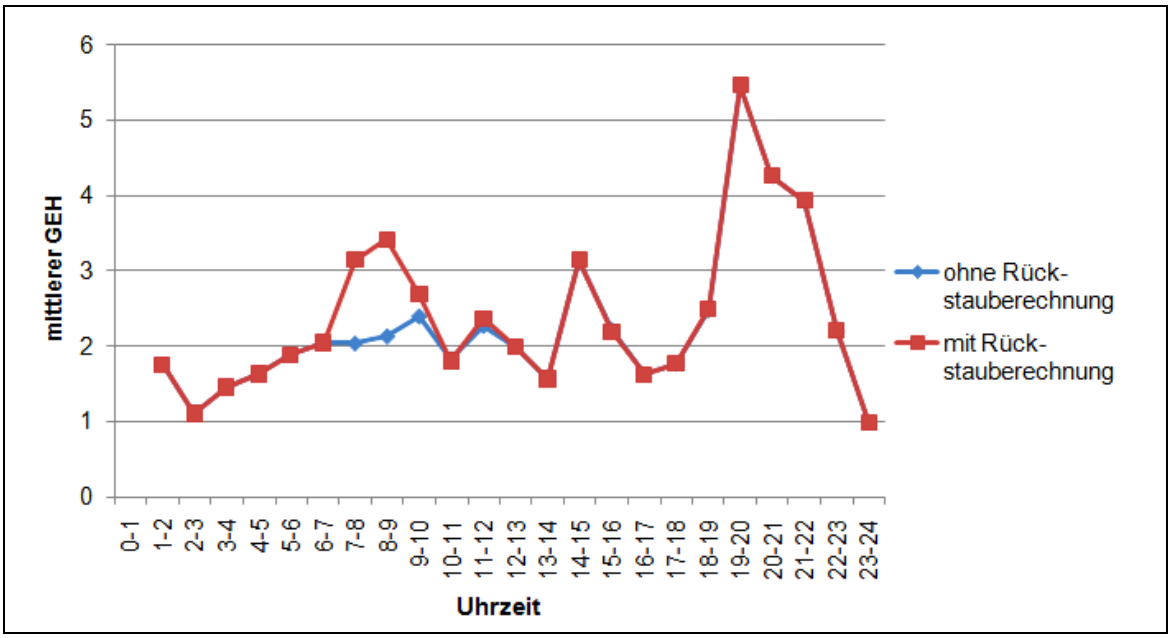


Abbildung 158: Vergleich einer stochastischen Umlegung mit und ohne Rückstauberechnung.

Die Qualität des Verfahrens hängt daher in großem Maße von der Qualität des verwendeten statischen Umlegungsverfahrens ab. Zu einer Verbesserung der Umlegungsqualität kann die Rückstauberechnung kaum beitragen. Allerdings liefert sie zusätzliche Informationen über die Länge eines Staus oder die Wartezeit für die Verkehrsteilnehmer (Abbildung 159).

Da aus den genannten Gründen eine Verbesserung für die Prognose nicht zu erwarten ist, wurden mit diesem Verfahren keine Tagesprognosen erstellt.

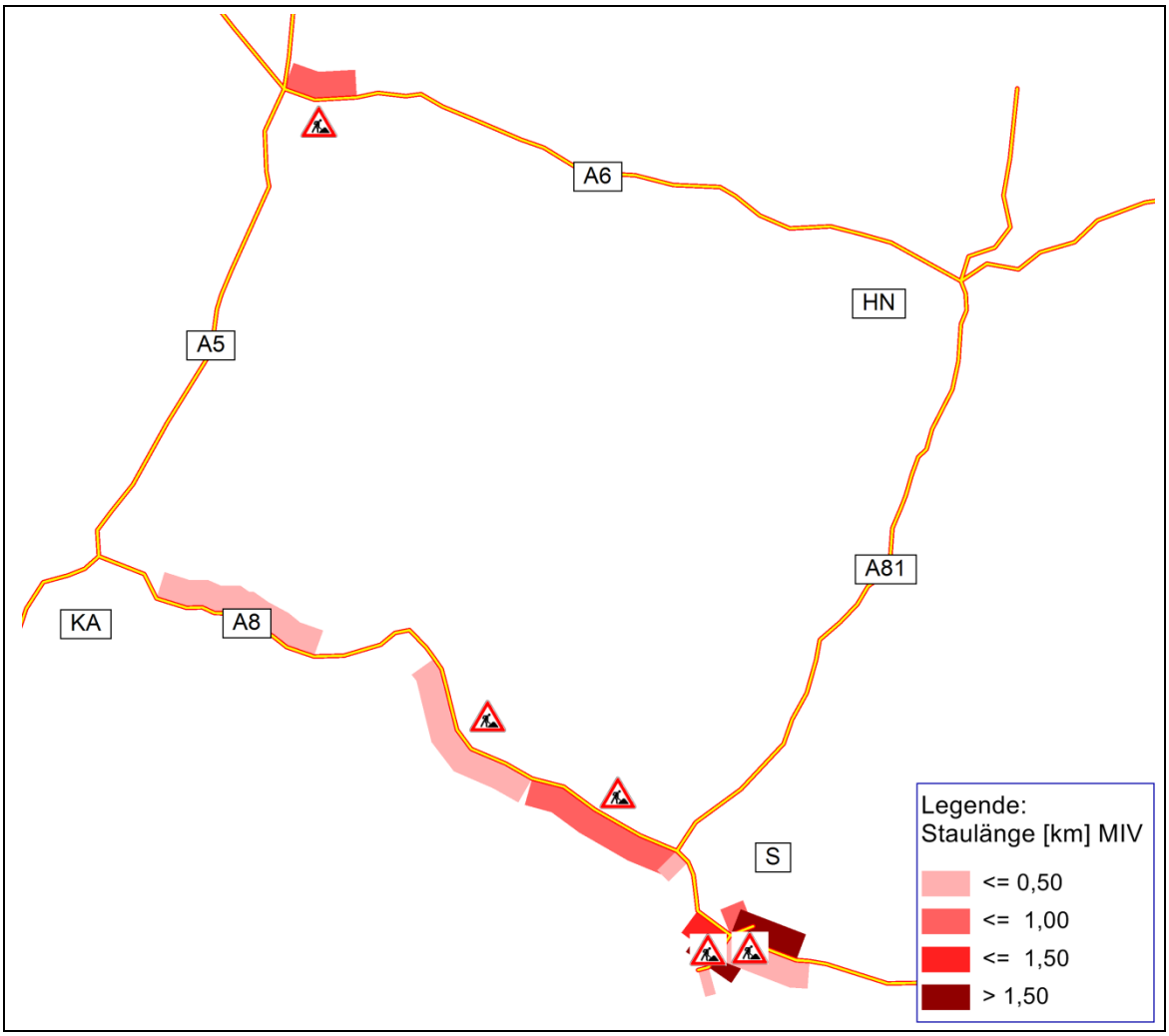


Abbildung 159: Staulängen mit stochastischer Umlegung und Rückstauberechnung am 23.06.08 von 7:00 Uhr bis 8:00 Uhr.

6.3.5 Ergebnisse aus dynamischen Verfahren

Die dynamische Umlegung hat sich für die Berechnung der Tagesprognose nicht bewährt. Die bei ersten Versuchen berechneten Verkehrsstärken weichen stark von den Messwerten ab. Aus diesem Grund und vor dem Hintergrund der guten Ergebnisse mit dem statischen Umlegungsmethoden wird dieses Verfahren nicht weiter verfolgt.

6.4 Grenzen der Tagesprognose

Anders als Behinderungen durch Baustellen können Engpässe, die durch nicht vorhersehbare Ereignisse, wie z. B. Unfälle entstehen und daher im Vorfeld nicht bekannt sind, nicht in der Tagesprognose dargestellt werden. Dies wird an folgendem Beispiel

verdeutlicht. Am 23. Juni 2008 ereignete sich kurz nach 17:00 Uhr ein Unfall auf der A 8 bei Pforzheim. Es folgte eine Vollsperrung mit örtlicher Ausleitung des betroffenen Autobahnabschnittes von 17:45 Uhr bis 21:11 Uhr. In dieser Zeit waren die Verkehrsbelastungen auf der A 8 niedriger als zuvor in der Tagesprognose berechnet. Der Verkehr aus Stuttgart kommend Richtung Karlsruhe und Walldorf wurde auf die A 81 und die A 6 über Heilbronn umgeleitet. Hier wurden nun höhere Verkehrsbelastungen gemessen, als im Modell prognostiziert. In diesem Zeitraum steigen die GEH-Werte aufgrund der Abweichung von Prognose und tatsächlicher Verkehrsbelastung an (Abbildung 160).

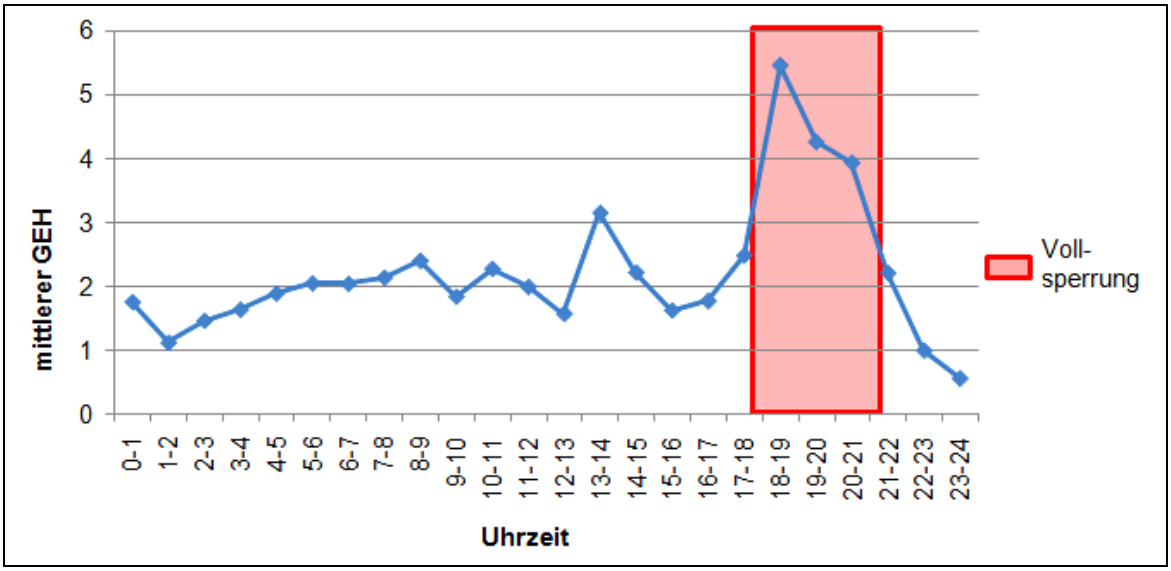


Abbildung 160: Entwicklung der mittleren GEH-Werte bei unvorhergesehener Vollsperrung eines BAB-Abschnittes.

Auch Änderungen der Anzeige der dynamischen Wechselwegweisung können nicht modelliert werden, da diese in der Regel aufgrund der aktuellen Verkehrslage erfolgen und demnach nicht prognostizierbar sind.

7 Monitoring der Angebotsqualität

7.1 Vorgehen

Definition von Angebotsqualität

Die Angebotsqualität in einem Verkehrsnetz beschreibt die Qualität des Verkehrsablaufs aus der Sicht des Nutzers. Das aus Nutzersicht entscheidende Kriterium ist der Zeitaufwand, der durch unterschiedliche Kenngrößen beschrieben werden kann. In diesem Kapitel werden die Kenngrößen Geschwindigkeit, Verlustzeit und Fahrzeitindex verwendet:

- Geschwindigkeit: mittlere Geschwindigkeit in einem Netzabschnitt [km/h],
- Verlustzeit: Summe der Verlustzeiten [Fahrzeugstunden] in einem Netzabschnitt,
- Verkehrszeitindex: Quotient aus der mittleren Fahrzeit und der Soll-Fahrzeit.

Definition von Monitoring

Unter Monitoring ist die kontinuierliche Erfassung der Angebotsqualität zu verstehen. Das Monitoring ermöglicht eine Analyse der Angebotsqualität über einen längeren Zeitraum. Es ermöglicht Trends zu erkennen und die Wirksamkeit verkehrlicher Maßnahmen zu quantifizieren.

Vergleich der Datenquellen

Ein kontinuierliches Monitoring der Angebotsqualität ist derzeit noch nicht üblich. Das liegt u.a. daran, dass die direkte Ermittlung von Fahrzeiten mit lokalen stationären Detektoren nicht möglich ist und die Detektordichte nicht ausreicht. FPD können hier eine Lücke schließen. In diesem Kapitel werden zwei Methoden der Fahrzeitermittlung im Straßennetz verglichen:

- Fahrzeitermittlung aus lokal gemessenen Geschwindigkeiten,
- Fahrzeitermittlung aus Mobilfunktrajektorien.

Datenverfügbarkeit für das Monitoring

Für das Monitoring stehen sowohl Daten aus stationären Detektoren als auch Mobilfunktrajektorien zur Verfügung. Die RMQ-Daten der stationären Detektoren liegen für einen Zeitraum von 25 Monaten vor und umfassen in 15-Minuten-Intervallen die Kenngrößen

- Verkehrsstärke $Q_{SD,RMQ}(i)$ in [Kfz/h] und

- Ist-Geschwindigkeit $v_{ist,RMQ}(i)$ in [km/h] als mittlere Geschwindigkeit eines Intervalls am stationären Detektor.

Für VPN-Strecken ohne stationäre Detektoren liegen keine Messdaten vor. Für einzelne VPN-Strecken mit mehreren referenzierten stationären Detektoren liegen dagegen mehrere Messdaten vor. Um für alle VPN-Strecken einen Messwert zu erhalten ist eine Datenvervollständigung erforderlich, die in Kapitel 7.3 beschrieben wird.

Durch die mit Verkehrsstärken aus RMQ hochgerechneten Mobilfunktrajektorien werden alle Autobahnstrecken im Untersuchungsgebiet belastet. Die hochgerechneten Mobilfunktrajektorien sind eine Fusion aus den Eingangsdaten:

- Verkehrsstärke $Q_{SD,Strecke}(i)$ in [Kfz/h] aus stationären Detektoren und
- FPD-Fahrzeiten $t_{FPD,Strecke}$ in [s] aus FPD-Daten als Fahrzeit einer Trajektorie über eine Strecke. In Verbindung mit der Streckenlänge $l_{Strecke}$ ist die FPD-Geschwindigkeit $v_{FPD,Strecke}$ bekannt.

Mit den Mobilfunktrajektorien werden alle VPN-Strecken mit ihrer Verkehrsstärke belastet, so dass hier keine Datenvervollständigung nötig ist. Jedoch entspricht die FPD-Geschwindigkeit der VPN-Strecken nicht der realen Geschwindigkeit. Wie in Kapitel 3.12 beschrieben, sind nur die langen Trajektorien geeignet um aus ihrer Fahrzeit eine mittlere Geschwindigkeit zu bestimmen. Da Pkw und Lkw unterschiedliche Fahrtweitenverteilungen haben, sind bei der Betrachtung der langen Trajektorien die Lkw überrepräsentiert und die FPD-Geschwindigkeiten erheblich geringer als die realen mittleren Geschwindigkeiten. Diese Geschwindigkeitskorrektur wird in Kapitel 7.4 durchgeführt. Aus diesen zusätzlichen Inputdaten können weitere Kenngrößen bestimmt werden die zum Monitoring der Angebotsqualität beitragen.

7.2 Kenngrößen der Angebotsqualität

Für die Bestimmung der Verlustzeit ist es nötig, eine Soll-Geschwindigkeit v_{soll} bzw. eine Soll-Fahrzeit t_{soll} als Bezugsgröße zu wählen. Es gibt verschiede Möglichkeiten diese Bezugsgröße festzulegen:

- Zulässige Höchstgeschwindigkeit bzw. Richtgeschwindigkeit v_0 in [km/h],
- Definition einer Soll-Geschwindigkeit für jede Straßenkategorie nach der RIN (Richtlinie für integrierte Netzgestaltung*, Kapitel 5),*
- Ableitung einer Soll-Geschwindigkeit aus der Verteilung der gemessenen Geschwindigkeiten. Als Soll-Geschwindigkeit kann z.B. das 5-, 10- oder 15-Percentil der höchsten Geschwindigkeiten verwendet werden.

Im Folgenden wird als Soll-Geschwindigkeit das 15-Percentil der lokal erfassten Geschwindigkeitsverteilungen verwendet.

7.2.1 Kenngrößen für eine Strecke

Verkehrsleistung VL [Fahrzeugkilometer]

Die Verkehrsleistung $VL_{Strecke}$ [Fahrzeugkilometern] einer Strecke im Intervall *i* ist definiert als das Produkt von Streckenlänge $l_{Strecke}$ und der Verkehrsstärke $Q_{Strecke}$ im Intervall *i*:

$$VL_{Strecke}(i) = Q_{Strecke}(i) \cdot l_{Strecke}$$

Soll-Verkehrszeitaufwand VZA_{soll} [Fahrzeugstunden]

Für jede Strecke lässt sich für das Intervall *i* aus der Soll-Geschwindigkeit $v_{soll,Strecke}$, der Streckenlänge $l_{Strecke}$ und der Verkehrsstärke $Q_{Strecke}(i)$ im Intervall der Verkehrszeitaufwand bestimmen, der bei durchgehend freiem Verkehrsfluss mit der aktuellen Verkehrsstärke erreicht worden wäre.

$$VZA_{soll,Strecke}(i) = Q_{Strecke}(i) \cdot \frac{l_{Strecke}}{v_{soll,Strecke}}$$

Wenn Geschwindigkeit und Verkehrsstärke nach Pkw und Lkw differenziert vorliegen, lässt sich der Verkehrszeitaufwand genauer berechnen:

$$VZA_{soll,Strecke}(i) = VZA_{Pkw,soll,Strecke}(i) + VZA_{Lkw,soll,Strecke}(i)$$

Die Soll-Geschwindigkeit v_{soll} ist im Gegensatz zur aktuellen Geschwindigkeit nicht vom betrachteten Intervall *i* abhängig.

Ist-Verkehrszeitaufwand VZA_{ist} [Fahrzeugstunden]

Für jede Strecke lässt sich für das Intervall *i* aus der aktuellen Geschwindigkeit $v_{ist,Strecke}(i)$, der Streckenlänge $l_{Strecke}$ und der Verkehrsstärke $Q_{Strecke}(i)$ im Intervall der Verkehrszeitaufwand bestimmen, der tatsächlich erreicht worden ist.

$$VZA_{ist,Strecke}(i) = Q_{Strecke}(i) \cdot \frac{l_{Strecke}}{v_{ist,Strecke}(i)}$$

Verlustzeitaufwand $VZA_{Verlust,Strecke}$ [Fahrzeugstunden]

Die Verlustzeiten über das Intervall *i* werden aus der Differenz von Ist-Verkehrszeitaufwand und Soll-Verkehrszeitaufwand berechnet.

$$VZA_{Verlust,Strecke}(i) = VZA_{ist,Strecke} - VZA_{soll,Strecke}$$

Fahrzeitindex als Indikator für die Angebotsqualität

Der Fahrzeitindex für ein Intervall *i* ist der Quotient zwischen Soll- und Ist-Geschwindigkeit und repräsentiert die Fahrzeitveränderung aus Sicht des Verkehrsteilnehmers.

$$FZI_{Strecke}(i) = \frac{v_{soll,Strecke}}{v_{ist,Strecke(i)}}$$

7.2.2 Aggregierte Kenngrößen

Die oben definierten Kenngrößen beziehen sich jeweils auf eine VPN-Strecke und ein Intervall. Sie können räumlich auf Streckenzüge (alle Strecken zwischen zwei Autobahnkreuzen) oder auf das ganze Untersuchungsgebiet aggregiert werden. Zeitlich können die Kenngrößen auf Tage oder längere Zeiträume zusammengefasst werden. Die Kenngrößen Verkehrsleistung, Verkehrszeitaufwand und Verlustzeit werden hierbei durch Addition berechnet:

$$\begin{aligned}
VL_{aggregiert} &= \sum_{Intervalle}\sum_{Strecken} VL_{Strecke}(i) \\
VZA_{ist,aggregiert}(i) &= \sum_{Intervalle}\sum_{Strecke} VZA_{ist,Strecke}(i) \\
VZA_{soll,aggregiert}(i) &= \sum_{Intervalle}\sum_{Strecke} VZA_{soll,Strecke}(i) \\
VZA_{Verlust,aggregiert}(i) &= \sum_{Intervalle}\sum_{Strecke} VZA_{Verlust,Strecke}(i)
\end{aligned}$$

Der Fahrzeitindex bestimmt sich aus der gewichteten mittleren Ist- und Soll-Geschwindigkeit des zu aggregierenden Gebietes und Zeitraums.

$$FZI_{aggregiert} = \frac{v_{soll,aggregiert}}{v_{ist,aggregiert}}$$

mit:

$$v_{ist,aggregiert} = \frac{\sum_{Intervalle}\sum_{Strecken} Q_{Strecke}(i) \cdot v_{ist,Strecke}(i)}{\sum_{Intervalle}\sum_{Strecken} Q_{Strecke}(i)}$$

$$v_{soll,aggregiert} = \frac{\sum_{Intervalle} \sum_{Strecken} Q_{Strecke}(i) \cdot v_{soll,Strecke}}{\sum_{Intervalle} \sum_{Strecken} Q_{Strecke}(i)}$$

7.3 Ermittlung der Kenngrößen der Angebotsqualität aus stationären Detektoren

7.3.1 Verfügbare Daten

Im Untersuchungsgebiet wurden auf Autobahnen stationäre Detektoren referenziert. Einige Zählwerte stammen von hochgerechneten Messungen des linken Fahrstreifens und besitzen lediglich eine Gesamtverkehrsstärke und eine Ist-Geschwindigkeit für alle Kfz gemeinsam. Die übrigen stationären Detektoren liefern sowohl Verkehrsstärke, als auch Ist-Geschwindigkeit nach Pkw und Lkw differenziert. Für den Verkehrslagebericht werden Verkehrsstärken und Geschwindigkeiten auf Kfz aggregiert und die Kenngrößen daraus bestimmt. Für das gesamte Untersuchungsgebiet liegen RMQ-Zähldaten von stationären Detektoren aus 25 Monaten vor.

Um die Kenngrößen der Angebotsqualität berechnen zu können, sind für jede untersuchte Strecke Messwerte erforderlich. Da viele Strecken nicht detektiert werden und da Detektoren ausfallen können, müssen fehlende Daten vervollständigt werden. Hierfür wird ein Verfahren angewendet, das aus RMQ-Ganglinien vollständige und konsistente Streckendaten ableitet. Diese besitzen für jeden Tag und jede VPN-Strecke genau eine Ganglinie, die als Grundlage für die Berechnung der benötigten Kenngrößen verwendet wird. .

7.3.2 Aggregierung von Messdaten zu VPN-Streckendaten

In der Clusterung entstehen für jeden Tag und jeden RMQ die folgenden drei Typen von Ganglinien.

- RMQ-Ganglinien: Verkehrsstärken und Geschwindigkeiten aus tagesgenauen Messdaten,
- Verkehrstagtypische Ganglinien: Verkehrsstärken und Geschwindigkeiten, die der mittleren Ganglinie eines Clusters entspricht, dem der Tag zugeordnet ist.
- Wochentagtypische Ganglinie: mittlere Verkehrsstärken und Geschwindigkeiten aus der Vorklassifizierung am Attribut „Wochentag“. Je stationären Detektor gibt es somit genau sieben kalendertagtypische Ganglinie.

Um für den Verkehrslagebericht eine möglichst vollständige Abdeckung von Daten aus stationären Detektoren zu erhalten, werden alle drei Ganglinientypen für die Erzeugung

eines Verkehrslageberichtes berücksichtigt. Im Normalfall werden für einen RMQ und einen gegebenen Tag die tagesgenauen Ganglinien verwendet. Sollten die Daten für diesen Tag nicht vorhanden sein, wird die verkehrstagtypische Ganglinie aus dem den Tag repräsentierenden Cluster verwendet. Wurde der RMQ in den Tagen dieses Clusters nicht versorgt, oder hat das Cluster an dem RMQ weniger als 10 versorgte Tage, wird für den RMQ die wochentagtypische Ganglinie verwendet.

Ganglinienfilter der RMQ-Messdaten

Von allen verwendeten Ganglinien werden mittels eines Qualitätsmerkmals die schlechtesten und ungenauesten Ganglinien herausgefiltert. Die übrigen Ganglinien werden für die Bestimmung der Streckendaten verwendet.

Qualitätsbewertung von Messwerten stationären Detektoren

Für die Bestimmung der Eignung einer Ganglinie werden folgende Qualitätsmerkmale berücksichtigt, deren Erfülltheitsgrad Punkte vergibt, die zur Aggregierung bzw. Filterung der Ganglinien dienen. Das Qualitätsmerkmal richtet sich ausschließlich nach der Verkehrsstärke der Ganglinie und nicht nach der Geschwindigkeit.

1. Vollständigkeit (QM_{vs}): Vollständigkeit der Versorgung im Erfassungszeitraum. Dieser Qualitätswert soll die Zuverlässigkeit im gesamten Zeitraum darstellen und ist abhängig vom Anteil versorgter Tage V_{ant}

$$V_{ant} = \frac{Anzahl\,versorgter\,Tage}{Anzahl\,betrachteter\,Tage} \cdot 100$$

$$QM_{VS} = 0 : \; falls\, V_{ant} < 50\%$$

$$QM_{VS} = 2 \cdot \frac{V_{ant}}{50} - 2 : \; falls\, V_{ant} >= 50\%$$

2. Echtheit (QM_{Echt}): Die Bewertung der Tatsache, dass ein RMQ durch eine echte Ganglinie oder durch eine gemittelte Ersatzganglinie repräsentiert wird
   - QM_{Echt} = 1 falls echte Ganglinie
   - QM_{Echt} = 0 falls Clusterganglinie oder historische Ganglinie
3. Abweichung ($QM_{Abweichung}$): Ist eine gemittelte Ganglinie aus der entsprechenden Klassifikationsgruppe vorhanden, entspricht eine hohe Abweichung von ihr einer auftretenden Anomalie. Diese könnte sowohl verkehrstechnisch als auch messstellentechnisch (Defekt) bedingt sein. Eine verkehrstechnische Abweichung der Zählwerte würde auf allen RMQs derselben Strecke auftreten und würde auf allen RMQ zu einer verminderten Qualität führen.
   - Wird der RMQ an diesem Tag durch eine Ersatzganglinie versorgt, so ergibt sich für die Ganglinie keine Abweichung.

- GEH –Wert:

$$GEH_{RMQ} = \frac{1}{Intervalle\ i} \sum_i \sqrt{\frac{2(G_i - G_{i,avg})^2}{G_i + G_{i,avg}}}$$

- $QM_{Abweichung}$ = Max(0;1-GEH/20)

Aus den genannten Qualitätsmerkmalen wird über eine Bewertungsfunktion ein Qualitätswert $QW_{gl,rmq}$ für die Ganglinie berechnet:

$$QW_{Gl,rmq} = QM_{Vs} + QM_{echt} + QM_{Abweichung}$$

Gibt es für eine Strecke mehrere Ganglinien zur Auswahl, werden alle RMQs mit einem Qualitätswert von $QW < 0{,}5$ aus der Berechnung entfernt. Alle übrigen Ganglinien werden zu einer Streckenganglinie gemittelt.

Schließlich wird jede VPN-Strecke entweder durch keine, durch genau eine oder durch mehrere Messstellen in Form einer gemittelten Streckenganglinie versorgt. Für den Fall, dass keine Versorgung vorliegt, wird die Strecke durch die Verwendung benachbarter versorgter Strecken im folgenden Kapitel interpoliert.

7.3.3 Interpolation von nicht versorgten VPN-Strecken

Für jede nicht versorgte VPN-Strecke wird eine gewichtete Interpolation der stationären Detektordaten durchgeführt. Dabei werden die nicht versorgten Strecken durch eine Ersatzganglinie aus den nächsten versorgten vorhergehenden Strecken und den nächsten versorgten nachfolgenden Strecken interpoliert. Das Untersuchungsgebiet umfasst die acht Streckenzüge der Autobahnen A5, A6, A8 und A81 jeweils in beide Richtungen. Für eine beliebige Strecke ist eine vorhergehende / nachfolgende Strecke diejenige, die in demselben Streckenzug vorkommt und beim Überfahren des Streckenzuges vor / nach der Strecke passiert wird. Gehört eine Strecke zu mehreren Streckenzügen, kann diese mehrere vorhergehende / nachfolgende Strecken besitzen – je Streckenzug eine. Eine schematische Darstellung der Interpolation der Verkehrsstärke ist in Abbildung 161 zu finden.

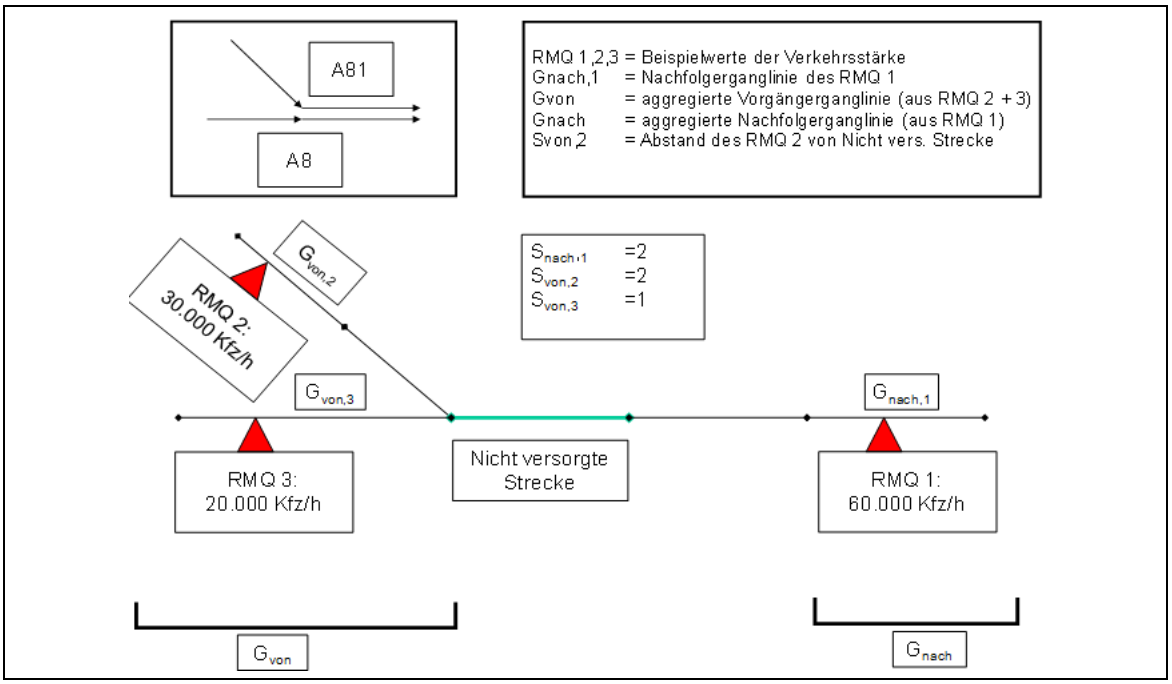


Abbildung 161: Schematische Darstellung von Vorgänger- und Nachfolgerganglinien

Im Folgenden wird die Anzahl der unterschiedlichen gefundenen Vorgänger- und Nachfolgerstrecken als cv und cn bezeichnet. Alle Vorgängerganglinien $G_{von,1}$ bis $G_{von,cv}$ und Nachfolgerganglinien $G_{nach,1}$ bis $G_{nach,cn}$ werden zu jeweils einer gemittelten Vorgängerganglinie G_{von} und G_{nach} aggregiert. Der Abstand der Vorgänger- und Nachfolgerganglinien $S_{von,i}$ bzw. $S_{nach,i}$ zur nicht versorgten Strecke entspricht jeweils der Anzahl der Knoten zwischen den Strecken. Im VPN-Netz entsprechen die Knotenpunkte auf Autobahnen den Zu- und Abfahrten.

Zusammenfassung von Vorgänger- und Nachfolgerganglinien

Ist $c_{von}>1$ bzw. $c_{nach}>1$ werden alle ungleichen Vorgänger- bzw. Nachfolgerganglinien zu einer zusammengefassten Vorgänger- bzw. Nachfolgerganglinie durch Summierung der Verkehrsstärke aggregiert.

$$G_{vor} = \sum_{i} G_{vor,i}$$

Der Streckenabstand S_{von} bzw. S_{nach} von der Strecke zur aggregierten Vorgänger- bzw. Nachfolgerganglinie wird durch das arithmetische Mittel der einzelnen Streckenabstände $S_{von,i}$ bzw. $S_{nach,i}$ bestimmt.

Die Geschwindigkeiten aus Vorgänger- und Nachfolgerganglinien werden aus den entsprechenden Ganglinien durch Mittelung gewichtet nach ihrer Verkehrsstärke zur jeweils aggregierten Vorgänger- bzw. Nachfolgerganglinie berechnet.

Interpolation einer Strecke

Die (aggregierte) Vorgängerganglinie G_{vor} und die (aggregierte) Nachfolgerganglinie G_{nach} werden durch entsprechende Mittelung zu einer Ersatzganglinie der nicht versorgten Strecke zusammengefasst.

$$G_{ersatz,t}(ID) = \frac{S_{nach} \cdot G_{von,t} + S_{von} \cdot G_{nach,t}}{S_{von} + S_{nach}}$$

Sowohl Verkehrsstärke als auch Geschwindigkeit werden hier entsprechend der Entfernung gewichtet. Der Interpolationsvorgang für die Verkehrsstärke ist in Abbildung 162 verdeutlicht.

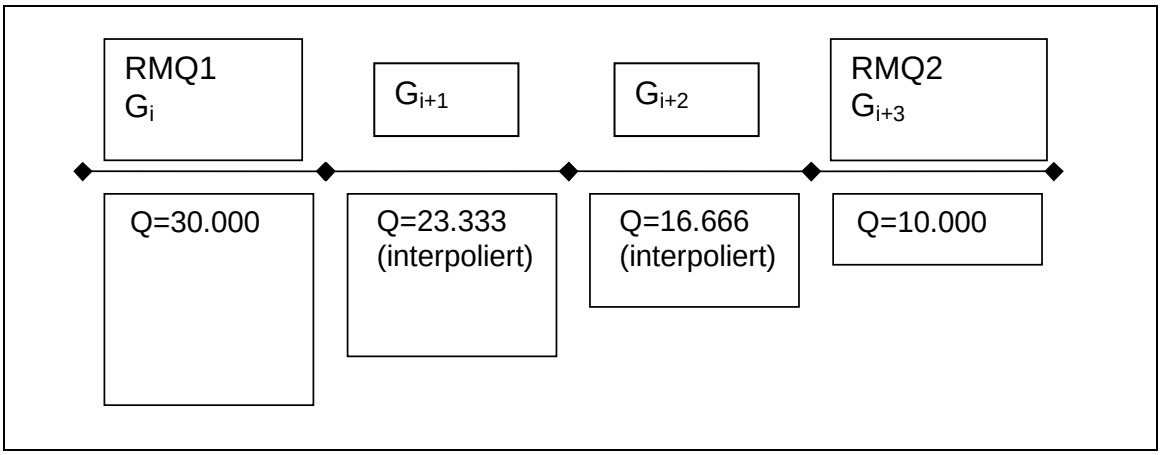


Abbildung 162: Schematische Darstellung der Interpolation an einem Streckenzug für Strecken.

Extrapolation von Lücken

Hat eine Strecke in der am Anfang oder am Ende eines Streckenzugs nur noch eine Vorgängerstrecke bzw. Nachfolgerstrecke, jedoch aber keine Nachfolgerstrecke bzw. Vorgängerstrecke, so können die Streckenwerte nicht interpoliert werden.

Die Ersatzganglinie der Strecke wird direkt aus der Vorgängerganglinie bzw. Nachfolgerganglinie ohne Änderung versorgt

$$G_j(ID) = G_{vor,i}$$

Diese Extrapolation tritt nur am Anfang bzw. Ende eines Streckenzuges auf, wenn die ersten bzw. letzten Strecken nicht versorgt wurden. Dies tritt nur an einer Stelle innerhalb des relevanten Autobahnvierecks auf. Am Autobahndreieck Karlsruhe endet der Streckenzug der A8 ohne weiterführende Strecken. Da die erste/letzte Strecke nicht durch einen RMQ versorgt ist, und die Streckenzuginterpolation nicht streckenzugübergreifend durchgeführt wird, gibt es nur die Möglichkeit die nächstgelegenen Streckenganglinien auf diese zu extrapolieren.

7.4 Ermittlung der Kenngrößen der Angebotsqualität aus stationären Detektoren und Mobilfunkdaten

7.4.1 Verfügbare Daten

Wie in der Zusammenfassung von Kapitel 3.12 beschrieben, ist die Qualität der Mobilfunkdaten je nach Länge der Route und der LAC-Folgen sehr unterschiedlich. Für die Generierung von FPD-Daten auf dem VPN-Netz werden für die Verkehrsstärken die hochgerechneten Trajektorien verwendet, die sowohl auf Strecken mit stationären Detektoren als auch auf Strecken ohne stationäre Detektoren eine zeitgenaue Verkehrsbelastung erzeugen. Dadurch dass die Streckenbelastung aufgrund von Trajektorien keine Lücken aufweist, ist auf den nicht durch RMQ-Daten versorgten Strecken weder Inter- noch Extrapolation notwendig. Die Belastungsdaten können als qualitativ besser betrachtet werden. Die FPD-Geschwindigkeiten werden aus den Geschwindigkeiten der Trajektorien berechnet. Da diese bei kurzen Trajektorien sehr stark schwankt, werden für die Berechnung der FPD-Geschwindigkeit auf den Autobahnstrecken ausschließlich lange Trajektorien verwendet, die über mindestens einen ganzen Streckenzug führen.

7.4.2 Anpassung der FPD-Geschwindigkeit

Bestimmung von Pkw und Lkw-Geschwindigkeiten

Wie bereits in Kapitel 3.9.1 erläutert, ist der Lkw-Anteil bei den FPD-Trajektorien überproportional hoch, so dass eine arithmetische Mittelung der FPD-Geschwindigkeiten zu einer zu niedrigen Streckengeschwindigkeit führen würde. Eine Unterscheidung der Trajektorien in Pkw und Lkw anhand ihrer Fahrzeit wird durch den in Kapitel 4.3.1 vorgestellten Algorithmus durchgeführt. Da die Identifizierung des Pkw- und Lkw-Verkehrs in der Routenwahl aufgrund der Fahrzeit über zwei Streckenzüge (Diagonale) bestimmt wird, werden die Parameter entsprechend angepasst um für die kürzeren Fahrzeiten über nur einen Streckenzug eine äquivalente Pkw-Lkw-Trennung zu erhalten. Das Verhältnis der Trajektorien, die zur Bestimmung von Pkw- und Lkw-Geschwindigkeiten verwendet werden (Lkw-Anteil) befindet sich auf den Streckenzügen im Bereich von 50-60%.

Bestimmung der FPD-Geschwindigkeiten für Strecken

Die Geschwindigkeiten aus den getrennten Pkw- und Lkw-Fahrzeiten werden für die FPD-Geschwindigkeit mit realen Lkw-Anteilen gewichtet. In Tabelle 66 sind Lkw-Anteile aus einer manuellen Straßenverkehrszählung (BASt, 2007), sowie gemittelte Lkw-Anteile aus den RMQ-Ganglinien auf den einzelnen Streckenzügen dargestellt.

Sowohl die manuellen Messdaten der BASt, als auch die Daten der Detektoren sind verwendbar und die erkennbaren Unterschiede sind geringfügig, wenn man berücksichtigt, dass diese nur für die Bestimmung der mittleren Geschwindigkeit verwendet werden und es sich hier um einen leicht abweichenden Gewichtungsfaktor der Lkw-Geschwindigkeiten handelt.

Quelle	A5	A6	A8	A81
RMQ-Daten	16,9%	22,5%	18,0%	14,7%
Manuelle Zählung	14,3%	22,0%	14,5%	11,0%

Tabelle 66: Lkw-Anteile aus Mobilfunkdaten, RMQ-Daten und BASt-Zählwerten.

Da die RMQ-Messwerte in 15-Minuten-Intervallen und nach Wochentagen getrennt vorliegen, werden sie hier für die Berechnung der mittleren Streckengeschwindigkeiten verwendet. Dabei werden Lkw- und Pkw-Geschwindigkeiten abhängig vom Wochentag und der Uhrzeit mit den mittleren Lkw-Anteilen der Streckenzüge berechnet, wie sie in Abbildung 163 dargestellt sind.

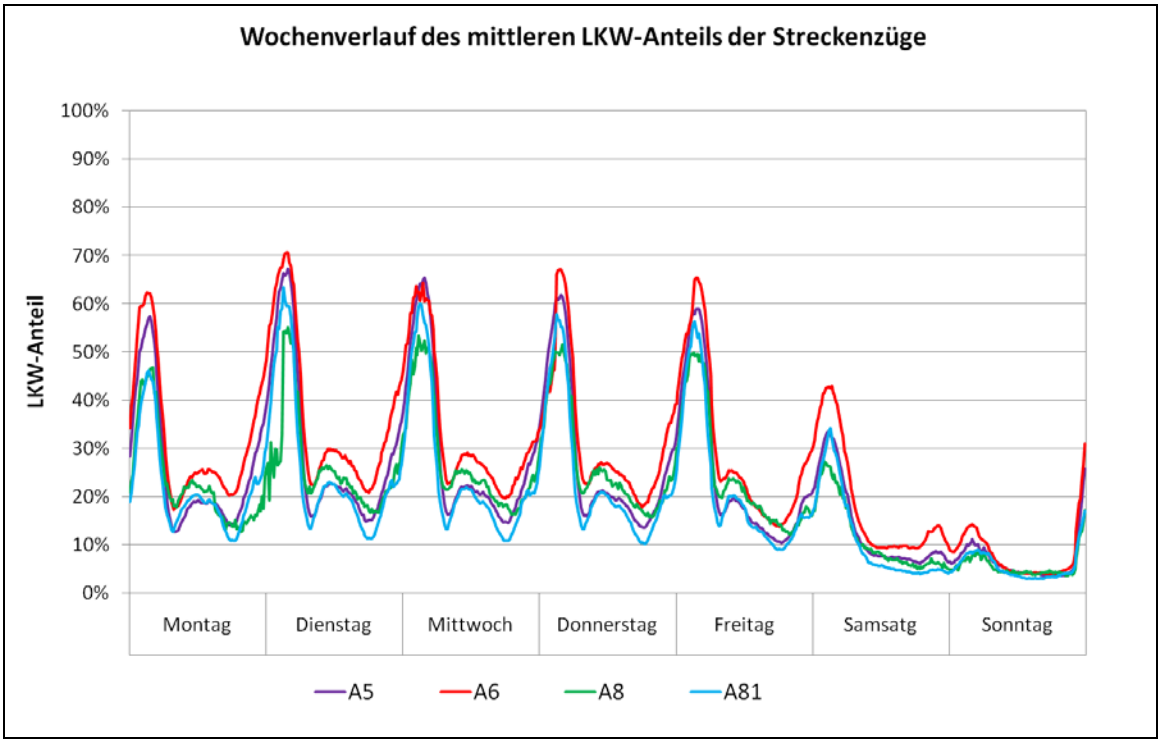


Abbildung 163: Zeitabhängiger Lkw-Anteil aller Detektoren unterschieden nach Wochentagen und nach den Streckenzügen A5, A6, A8 und A81.

Die sehr hohen Spitzen des Lkw-Anteils entstehen in der Nachtzeit an Werktagen, an denen im Vergleich zur Tageszeit nur etwa 10% der Verkehrsstärke der Pkw, jedoch immer noch etwa 50% Verkehrsstärke der Lkw vorliegt, wie in Abbildung 164 zu sehen ist.

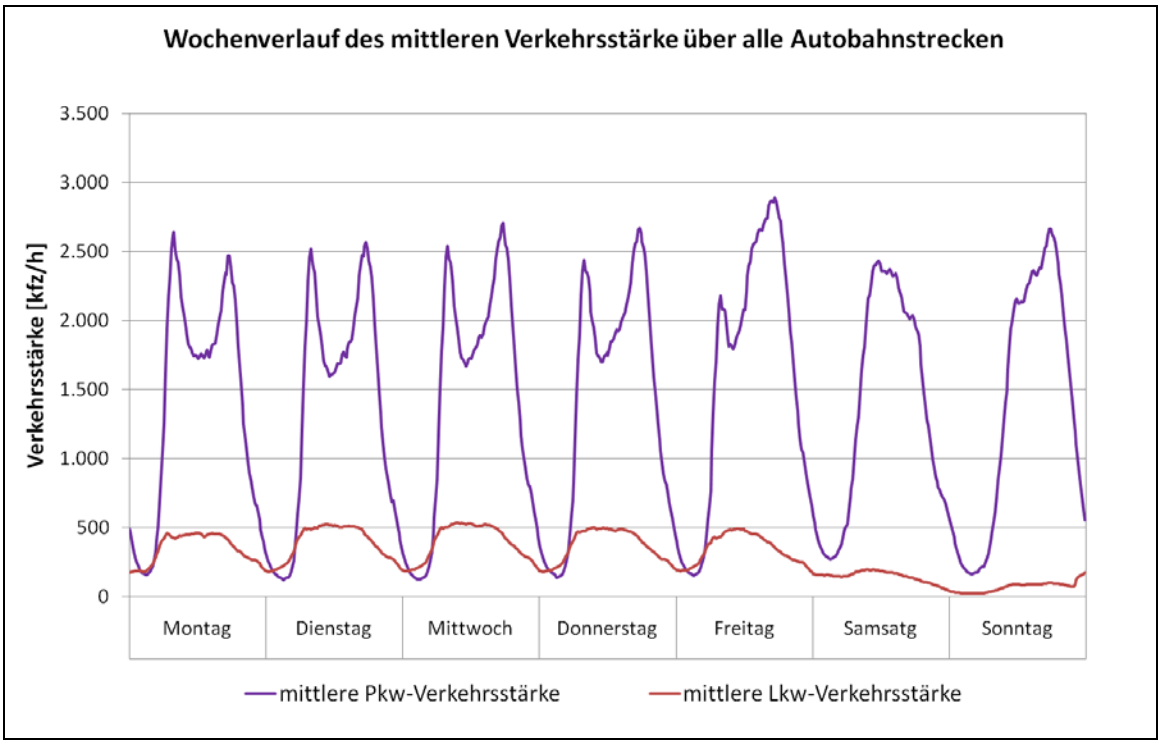


Abbildung 164: Mittlere Verkehrsstärke der Fahrzeugtypen Pkw und Lkw über alle Detektoren auf Autobahnen.

Die Geschwindigkeiten für Lkw und Pkw werden zeitabhängig mit dem ermittelten Lkw-Anteil des Streckenzugs gewichtet. Diese FPD-Geschwindigkeit wird schließlich in den 15-Minuten-Intervallen der Ganglinie aus Mobilfunkdaten zusammen mit der hochgerechneten Verkehrsstärke abgelegt. Die Auswirkungen der Korrektur auf die Geschwindigkeit ist in der Darstellung der Summenhäufigkeiten in Abbildung 165 verdeutlicht.

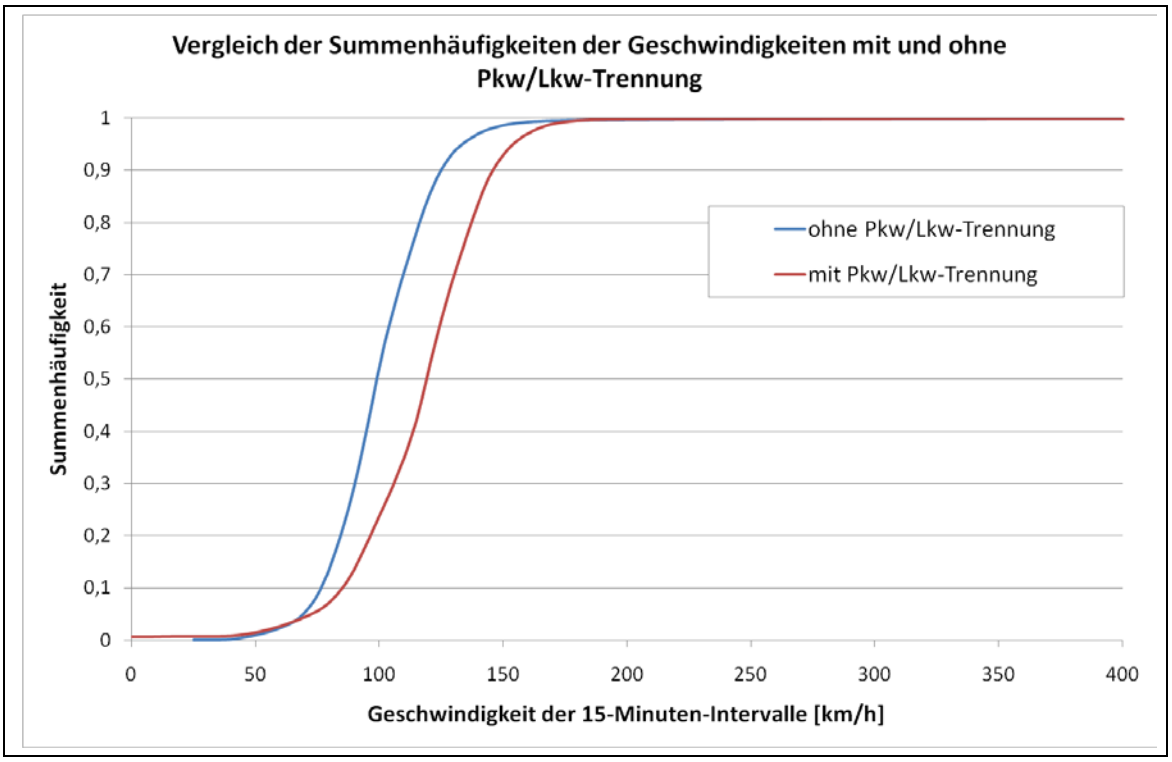


Abbildung 165: Summenhäufigkeit der Geschwindigkeiten aller Autobahnstrecken mit und ohne korrigierende Gewichtung nach Lkw-Anteilen.

7.5 Verkehrslagebericht aus RMQ-Daten

Für das VPN-Netz der Bundesautobahn-Strecken wird ein Verkehrslagebericht erzeugt. Der Berichtszeitraum umfasst entweder einen Tag oder einen Monat. Für Verkehrsdaten, die einen längeren Zeitraum umfassen ist auch ein Jahresbericht denkbar.

7.5.1 Tagesbericht

Der Tagesbericht dient zur Darstellung der tagesaktuellen Angebotsqualität. Hierbei werden für jeden Streckenzug die Kenngrößen Verkehrsleistung und mittlere Geschwindigkeit zeitabhängig dargestellt und mit den verkehrstagtypischen Werten aus der Clusterung verglichen. In den Abbildungen 166 und 167 sind beispielhaft die Verläufe von Kenngrößen über einen Tag über jeweils einen Streckenzug dargestellt.

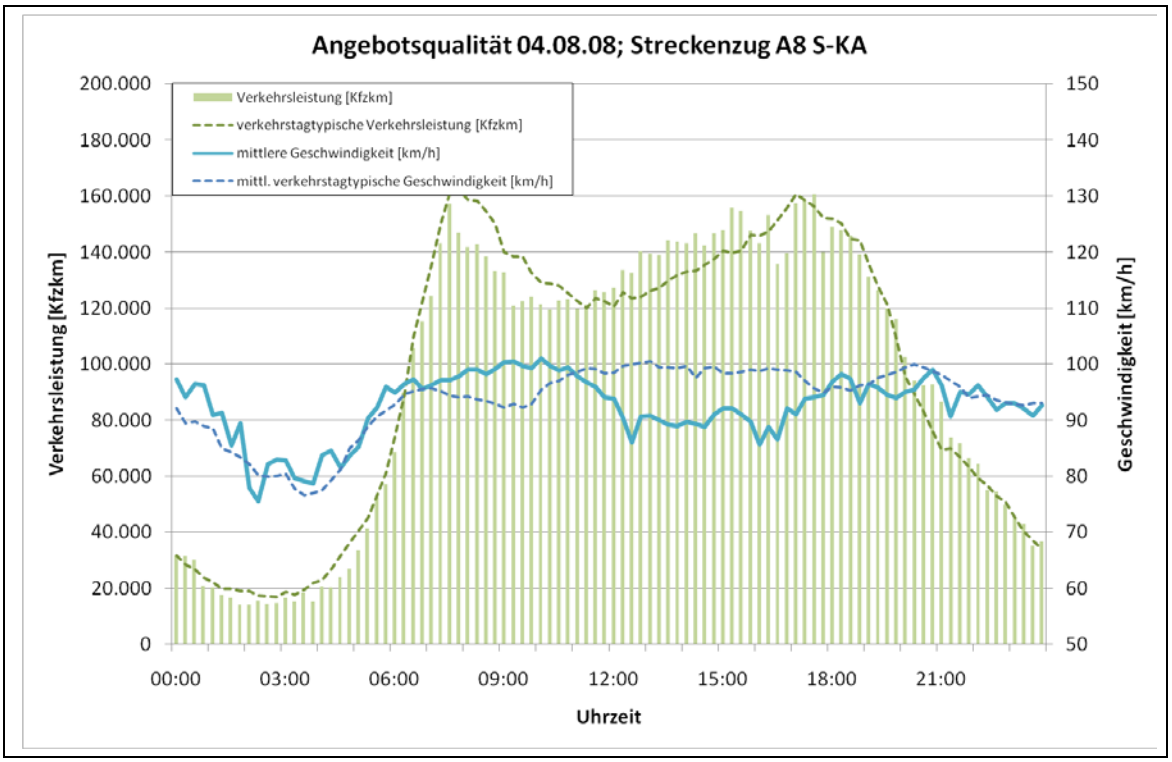


Abbildung 166: Tagesübersicht vom 04.08.2008

Der Kalendertag 04.08.2008 ist dem Cluster „Mo-Do (Ferien)“ zugeordnet. Der in Abbildung 166 gezeigte Tagesbericht für die A8 von Stuttgart nach Karlsruhe zeigt im Vergleich mit den verkehrstagestypischen Werten dieselbe Tendenz sowohl in der Verkehrsleistung, als auch in der Ist-Geschwindigkeit aus stationären Detektoren. In der Morgenspitze von 7:00 bis 11:00 Uhr beträgt die Verkehrsleistung etwa 20.000 Fahrzeugkilometer weniger je 15-Minuten-Intervall. Gleichzeitig ist die Ist-Geschwindigkeit um etwa 5 bis 10 km/h höher als die verkehrstagtypische Referenz. Im Gegensatz dazu ist am Nachmittag des Tages zwischen 12:00 und 17:00 die Verkehrsleistung deutlich erhöht, sowie die Ist-Geschwindigkeit verringert. Die grundsätzlich geringere mittlere Geschwindigkeit in der Nacht entspricht aufgrund des übermäßig hohen Lkw-Anteils den Erwartungen.

In Abbildung 167 ist der Tagesbericht des 09.08.2008 auf der A6 von Walldorf nach Heilbronn aus dem „Samstag“-Cluster dargestellt. In der Zeit zwischen 13 und 14 Uhr ist eine Veränderung sowohl bei den stationären Detektorgeschwindigkeiten als auch an der Verkehrsleistung klar erkennbar. Die Verkehrsstörung ist durch einen Abgleich mit den Verkehrsmeldungen auf einen Unfall zurückzuführen wodurch 2 von 3 Fahrstreifen gesperrt wurden. Die verkehrstagtypische Geschwindigkeit repräsentiert die für diesen Kalendertag typische Geschwindigkeit aus dem „Samstag“-Cluster. Der Vergleich zwischen dieser und der tatsächlichen Geschwindigkeit zeigt, dass es sich hierbei um ein einmaliges Ereignis handelt, das nicht im Tagesverlauf der typischen Geschwindigkeit vorkommt. Auch an diesem Tag sind im Allgemeinen etwas niedrigere

Verkehrsleistungen sowie höhere Geschwindigkeiten gemessen worden als bei den berechneten verkehrstagtypischen Werten.

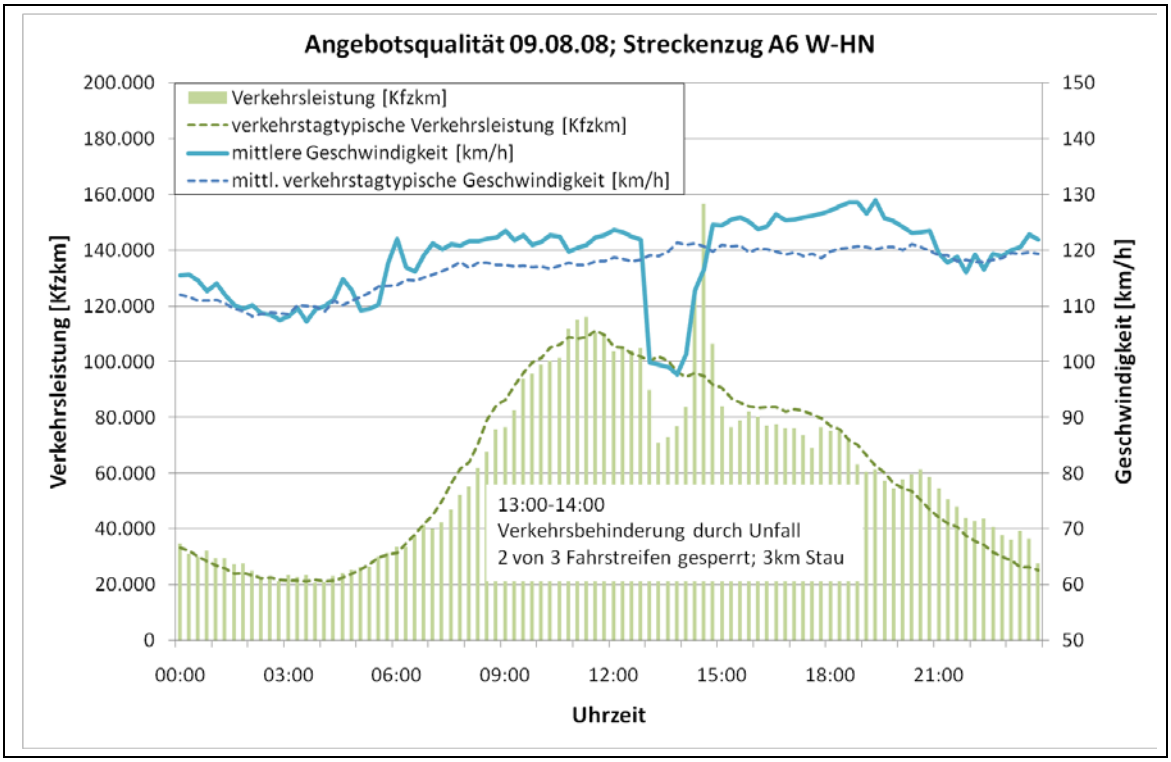


Abbildung 167: Tagesübersicht vom 09.08.2008 auf der A6 mit Verkehrsstörung durch Unfall

7.5.2 Monatsbericht

Für jeden Monat wird ein Übersichtsdiagramm erstellt, in dem der tägliche Fahrzeitindex und der damit verbundene Verkehrszeitaufwand im Netz sowie die mittlere Verlustzeit pro Streckenkilometer dargestellt sind. In den Abbildungen 168 und 169 sind die deutlich niedrigere Verkehrsleistung der Wochenendtage, sowie die daraus resultierende geringere Verlustzeit zu erkennen. Ebenso ist in Abbildung 168 der erhöhte Freitagsverkehr am 14. März zu Beginn der Osterferien zu erkennen.

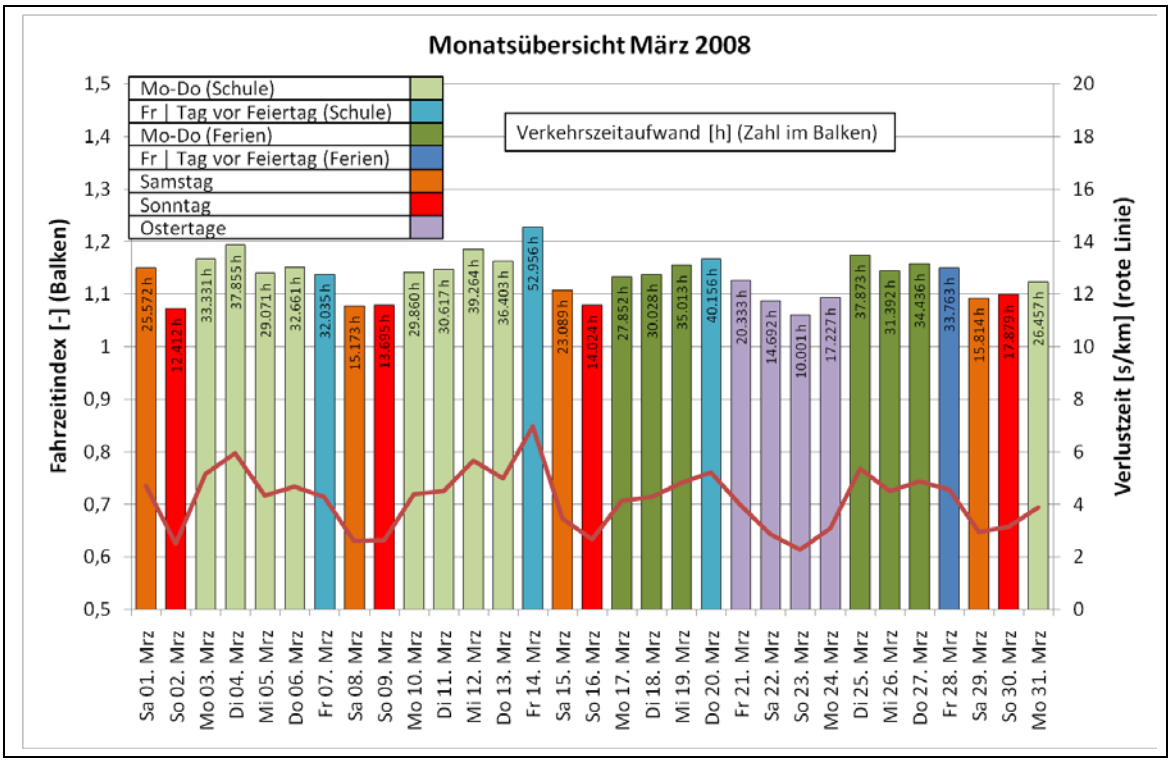


Abbildung 168: Verlauf der Angebotsqualität im Monat März 2008.

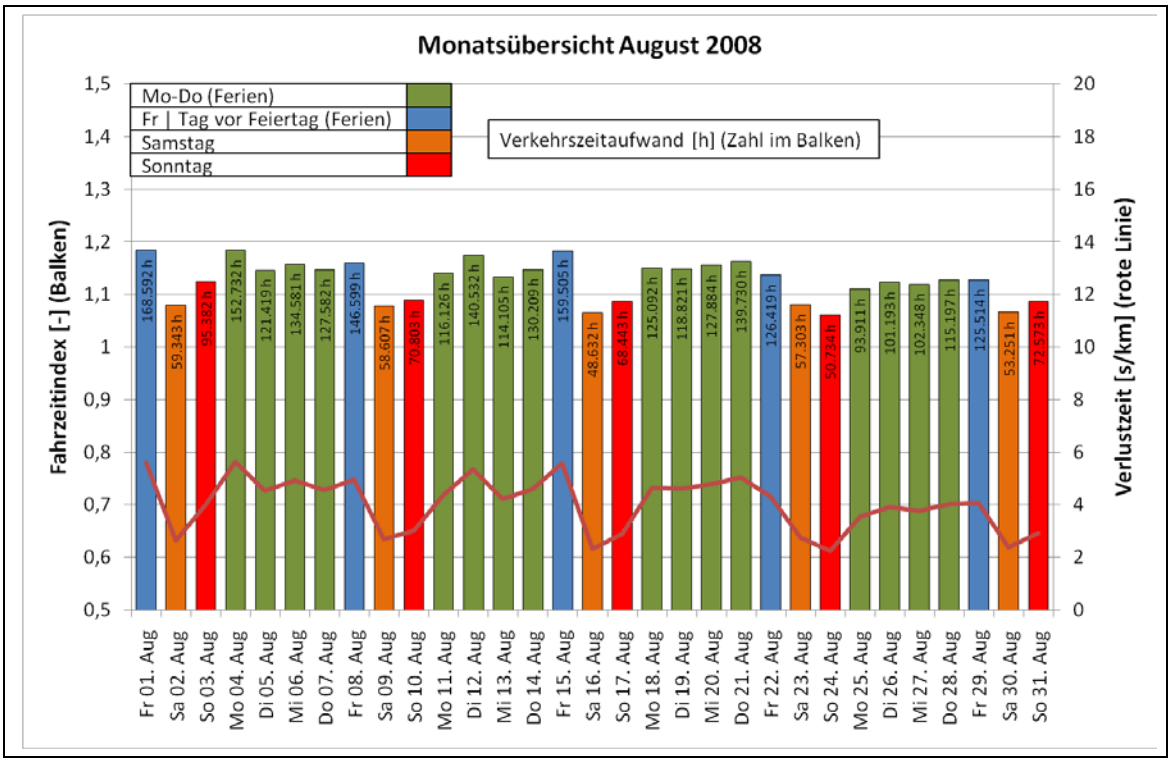


Abbildung 169: Verlauf der Angebotsqualität im Monat August 2008.

7.5.3 Jahresbericht

Liegen Messdaten über einen längeren Zeitraum vor sind Zeitreihenanalysen möglich. Durch einen Vergleich gemittelter Kenngrößen über ganze Monate hinweg kann man z.B. eine langzeitbedingte Veränderung in der Angebotsqualität erkennen. Im Jahresvergleich der Fahrzeitindizes in Abbildung 170 und der Verkehrsleistung in Abbildung 171 ist der ähnliche Jahresverlauf der Jahre 2007 und 2008 zu sehen. Auffällig sind die Monate Juli und August der Jahre 2007 und 2008. Im Juli beider Jahre hat sich der Fahrzeitindex kaum geändert, obwohl Verkehrsleistung und Verkehrszeitaufwand um etwa 6% zugenommen haben. Im August 2008 hat sich der Fahrzeitindex um etwa 2% verringert. Sowohl Verkehrsleistung, als auch Verkehrszeitaufwand reduzierten sich in diesem Monat um etwa 6%.

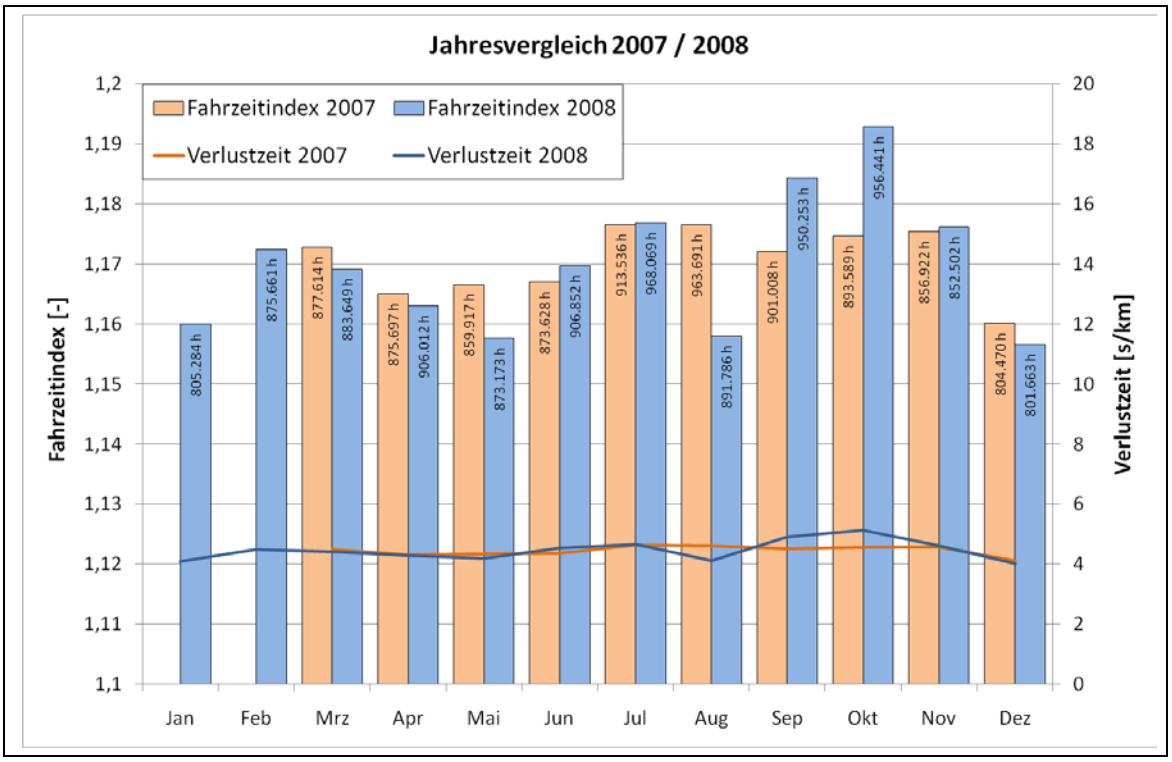


Abbildung 170: Vergleich der Fahrzeitindizes, durchschnittlichen täglichen Verkehrszeitaufwände und die damit verbundenen Verlustzeitaufwände der Jahre 2007 und 2008 auf dem Autobahnviereck.

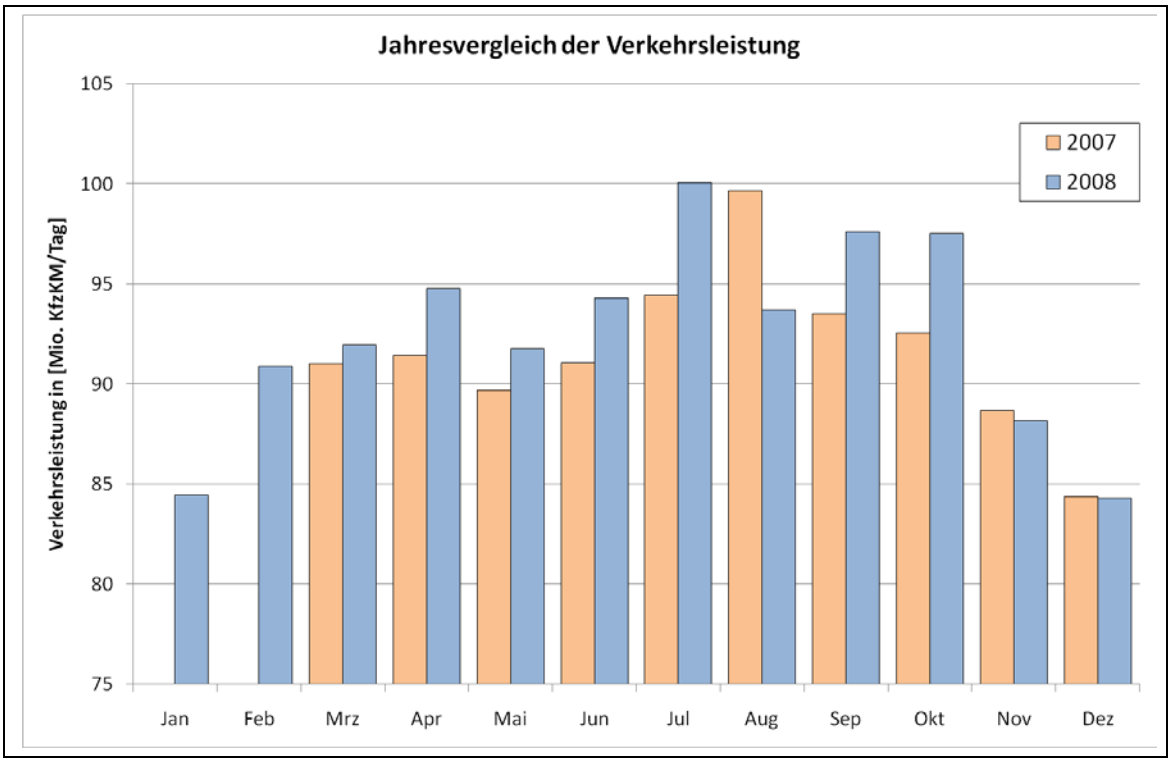


Abbildung 171: Vergleich der mittleren täglichen Verkehrsleistung im Monat der Jahre 2007 und 2008 auf dem Autobahnviereck.

Zu einem geringen Teil können die Differenzen der monatlichen Verkehrsleistungen in 2007 und 2008 auf die unterschiedliche Anzahl von Wochenenden in den jeweiligen Monaten zurückgeführt werden. Eine Normierung der Kenngrößen auf einen durchschnittlichen Wochentag zeigt jedoch keine nennenswerten Änderungen der Kenngrößen. Eine Verschiebung der Verkehrsleistung vom August in den Juli könnte durch eine Verschiebung der Sommerferien begründet werden, jedoch beginnen die Sommerferien nur im benachbarten Bundesland Hessen um etwa zwei Wochen früher am 23.06.2008 und in Teilen des Nachbarlandes Niederlande um etwa eine Woche früher. Der Sommerferienanfang der Bundesländer Nordrhein-Westfalen und Baden-Württemberg und der übrigen Regionen der Niederlande sowie des Nachbarlandes befand sich in den Jahren 2007 und 2008 in demselben Zeitraum. Die Ferien weiterer Nachbarländer und Bundesländer werden für das Untersuchungsgebiet nicht als verkehrsrelevant betrachtet.

Eine detaillierte Analyse der Verkehrsleistungszunahme in den Monaten Juli und August in Abbildung 172 zeigt, dass die stärkere Verkehrsleistung im Juli 2008 eher am Wochenende und insbesondere am Freitag aufgebracht wurde. Die schwache Verkehrsleistung im August 2008 kann jedoch den Werktagen Montag bis Donnerstag zugeordnet werden.

Inwieweit die Veränderungen auf eine fehlerhafte Erfassung der Verkehrsstärken, z.B. bedingt durch Baustellen, zurückzuführen ist, kann aus den RMQ-Rohdaten nicht

eindeutig festgestellt werden. Die RMQ-Daten liefern keine fahrstreifenfeine, sondern bereits zu Verkehrsquerschnitten aggregierte Werte.

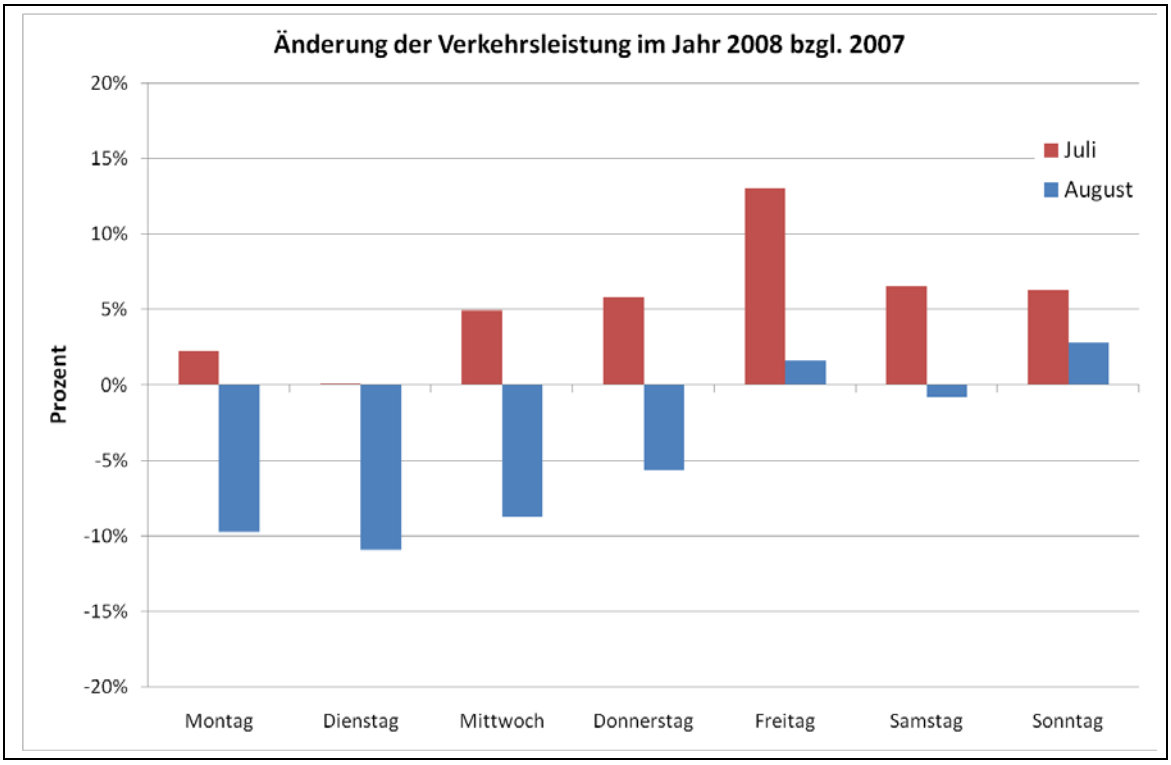


Abbildung 172: Veränderung der mittleren täglichen Verkehrsleistung in den Monaten Juli und August der Jahre 2007 und 2008 auf dem Autobahnviereck.

Eine allgemeine Tendenz lässt sich bei einem Erfassungsumfang von nur zwei Jahren nicht ableiten, jedoch würde man bei einem langjährigen Erfassungszeitraum zwischen statistischer und tendenzieller Veränderung unterscheiden sowie weitere kalendarische Eigenschaften definieren können.

7.6 Verkehrslagebericht aus RMQ- und Mobilfunkdaten.

Zusätzlich zum Verkehrslagebericht aus Messwerten stationärer Detektoren liegen für einige Tage zusätzlich die angepassten FPD-Geschwindigkeiten aus den Mobilfunktrajektorien vor. Der in Abbildung 173 gezeigte Tagesbericht entspricht dem aus Abbildung 167 jedoch ist hier zusätzlich die mittlere FPD-Geschwindigkeit zum Vergleich angegeben.

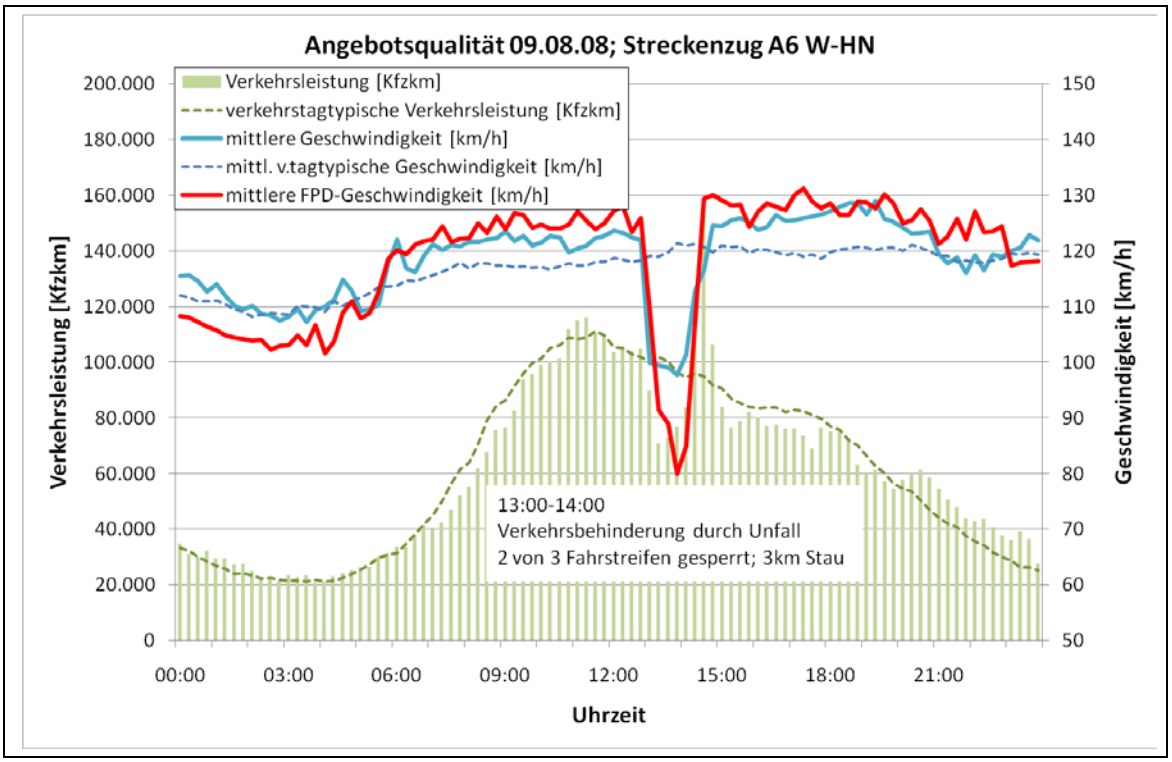


Abbildung 173: Tagesübersicht vom 09.08.2008 aus stationären Detektoren und Mobilfunkdaten mit Verkehrsstörung durch Unfall.

Der Geschwindigkeitsverlauf beider Quellen verhält sich ähnlich. Der Geschwindigkeitseinbruch zwischen 13 und 14 Uhr von über 40 km/h ist sowohl bei den Trajektorien als auch bei den Messdaten zu erkennen, jedoch in unterschiedlich stark ausgeprägter Form. Die Daten der FPD-Geschwindigkeit sind in diesem Abschnitt besser, da sie den Geschwindigkeitseinschnitt der Verkehrsbehinderung vollständig erkennen. Die Messdaten von stationären Detektoren haben zwar zeitgleich eine Reduzierung der Ist-Geschwindigkeit festgestellt, jedoch nur auf etwa 100 km/h.

Andere Streckenzüge weisen dagegen stark unterschiedliche Verläufe zwischen FPD- und stationären Detektorgeschwindigkeiten auf. Der in Abbildung 174 dargestellte Verkehrslagebericht vom 25.08.2008 zeigt bei den FPD-Geschwindigkeiten über den Streckenzug A8 verlaufenden Trajektorien eine erhebliche Störung, die mit einer gleichzeitigen Verkehrsmeldung über Instandhaltungsarbeiten und einem damit verbundenen Stau korrespondiert. Die Geschwindigkeiten der stationären Detektoren jedoch bleiben von diesem Ereignis unbeeinflusst.

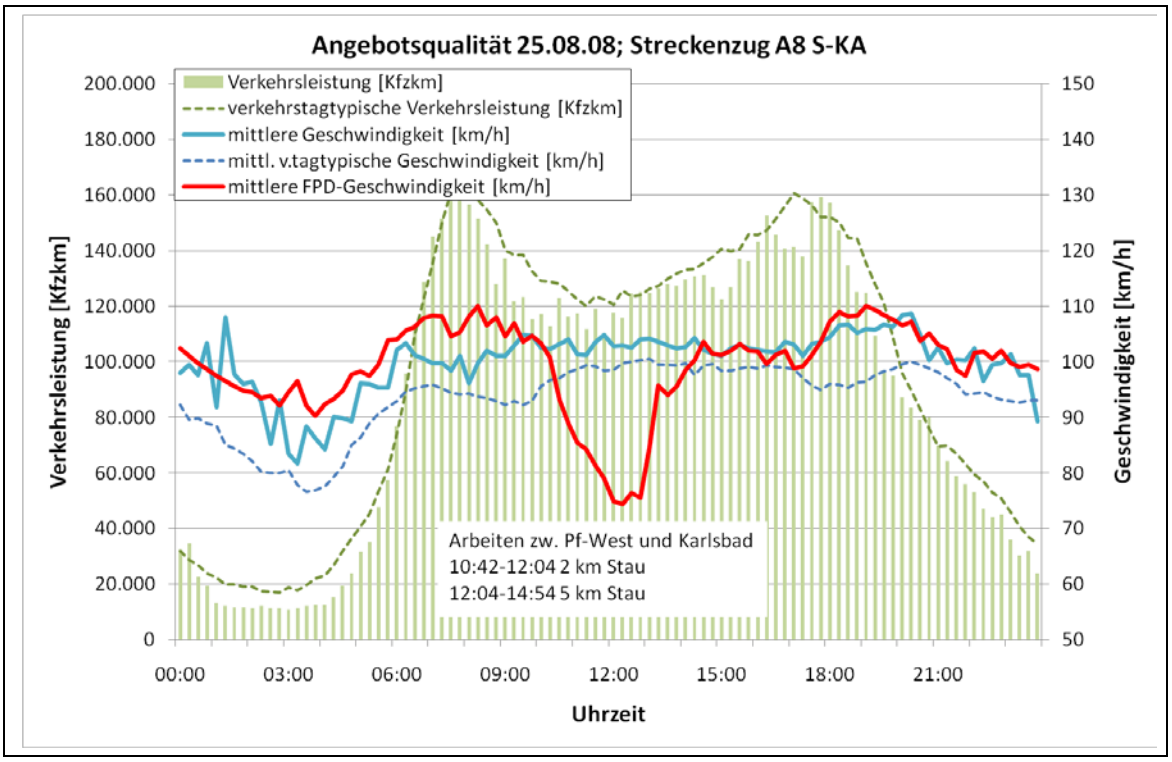


Abbildung 174: Tagesübersicht vom 25.08.2008 mit abweichender FPD- und stationärer Detektorgeschwindigkeit.

Diese Tatsache lässt sich durch die ungünstige Messstellenverteilung auf dem Streckenzug der A8 erklären, auf dem es zwischen dem Dreieck Stuttgart und Karlsbad einen 35 km langen Streckenabschnitt gibt, für den keine verwendbare Versorgung durch stationäre Detektoren vorliegt. Räumlich begrenzte Störungen in diesem Abschnitt können nicht erkannt werden, wenn diese sich nicht bis zur Position der stationären Detektoren auswirken. Die für den Streckenzug A8 von Stuttgart nach Karlsruhe verwendbaren stationären Detektorgeschwindigkeiten für den 25. August 2008 sind in Abbildung 175, deren räumliche Verteilung in Abbildung 176 dargestellt.

Die Interpolation der Messwerte aus den stationären Detektoren bei Karlsbad (RMQ 3410 und 8870) und Stuttgart (RMQ 3442, 3450, 3454 und 3458) für die Erzeugung des Verkehrslageberichts bewirkt dass sich die mittlere Geschwindigkeit über den gesamten Streckenzug nicht verändert. Die FPD-Geschwindigkeiten hingegen umfassen alle Störungen entlang des Streckenzuges. Im Gegensatz zu den stationären Detektoren können bei den Mobilfunkdaten Störungen erkannt, jedoch nicht exakt lokalisiert werden, da die Fahrzeiten und daraus resultierenden FPD-Geschwindigkeiten ausschließlich über den gesamten Streckenzug bekannt sind.

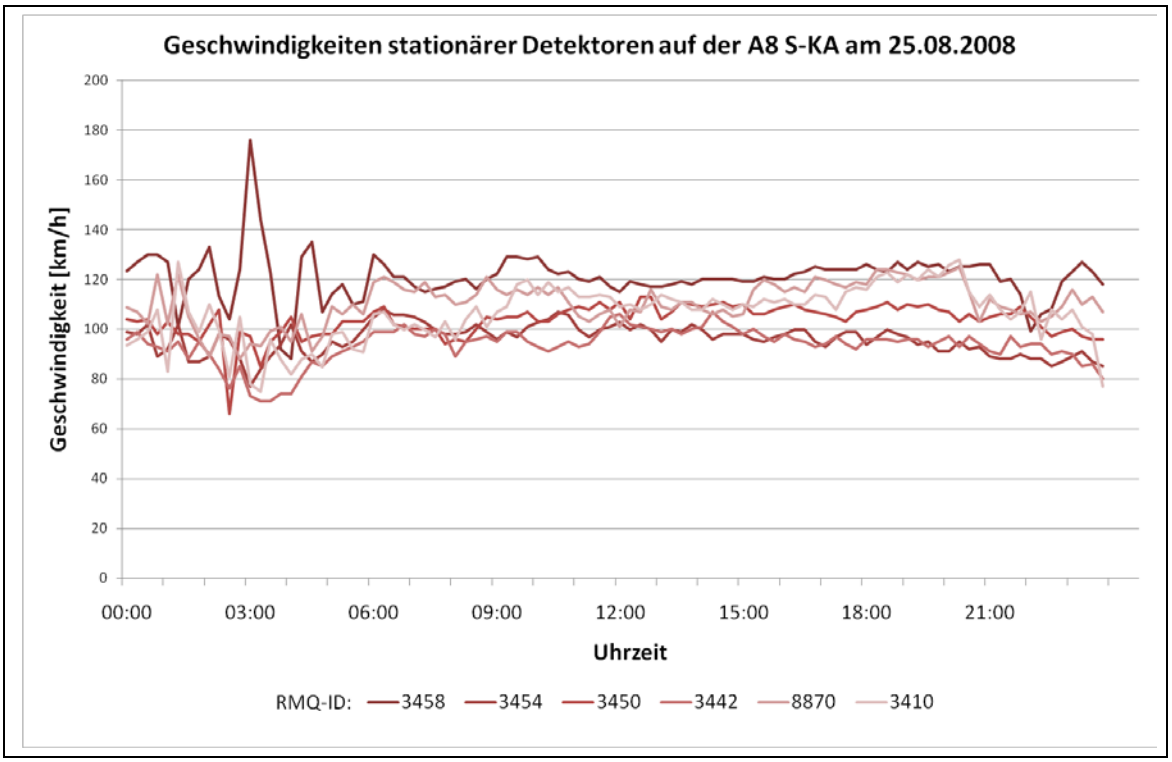


Abbildung 175: Vergleich der mittleren Geschwindigkeiten aller stationären Detektoren auf der A8 von Stuttgart nach Karlsruhe am 25.08.2008.

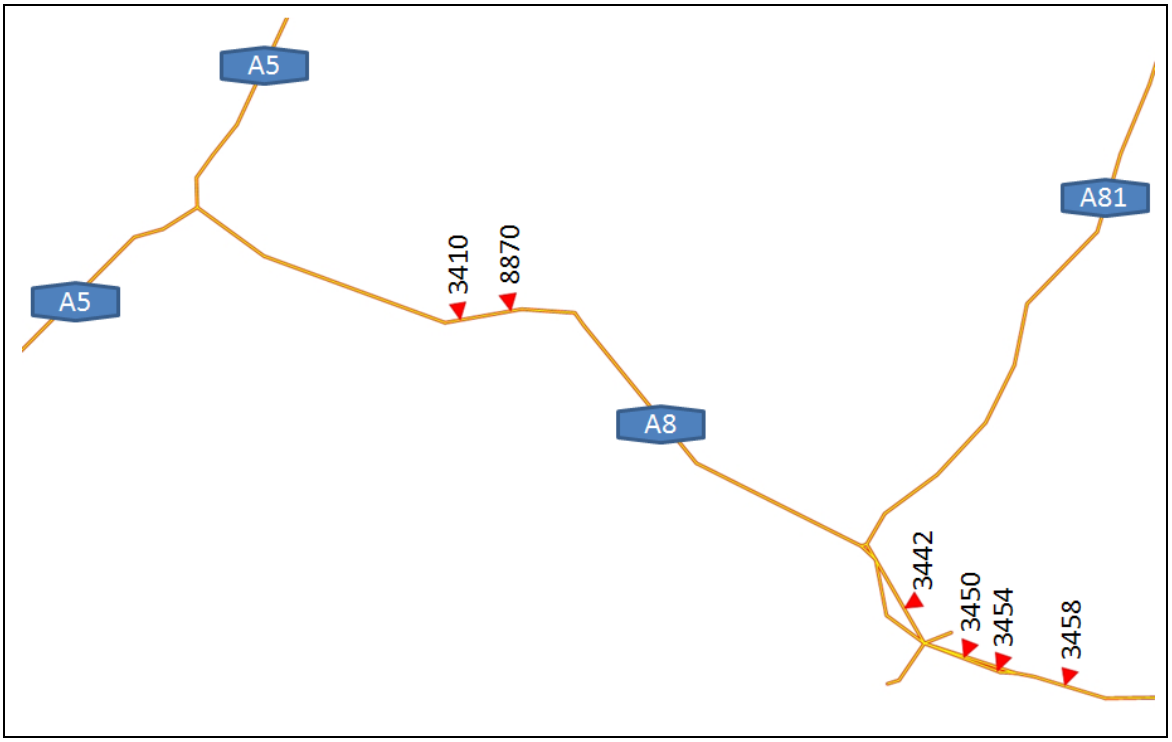


Abbildung 176: Räumliche Verteilung der stationären Detektoren die auf dem Streckenzug A8 von Stuttgart nach Karlsruhe verwendbare Messwerte liefern.

8 Glossar

A-Ebene bzw. A-Interface

- Obere Ebene im Mobilnetz. Am A-Interface werden Location Area Updates sowie Handover erfasst.

Abis-Ebene bzw. Abis-Interface

- Untere Ebene im Mobilnetz. Am Abis-Interface werden u.a. die aktuelle Cell-ID und Signalstärken während eines Telefonates im 480ms-Takt erfasst.

Abstandfunktion

- Funktion zur Bestimmung der Ähnlichkeit bzw. der Differenz zweier Ganglinien. Definierte Funktionen sind: Summe der Abstandsquadrate (SAQ), der GEH-Wert und die Prozentfunktion.
- Formelzeichen: z.B. $D_{Typ}(G(RMQ_i, Tag_j).Q, G(RMQ_k, Tag_l).Q)$ mit $Typ \in [SAQ, GEH, \%]$

ANPR

- Abkürzung für Automatic Number Plate Recognition (Automatische Kennzeichenerfassungssysteme, vgl. Kapitel 4.2.1).

BAB

- Abkürzung für Bundesautobahn.

BAB-Netzmasche

- Netzmasche der Bundesautobahnen A5, A6, A8 und A81 zwischen Stuttgart (S), Heilbronn (HN), Walldorf (W) und Karlsruhe (KA), siehe Abbildung 57.

Befolgungsgrad

- Anteil des Durchgangsverkehrs der Hauptroute, der durch eine aktive Wechselwegweisung auf die Alternativroute umgelenkt wird (s. Kapitel 4.5.3)

Best-Server-Plot

- Der Best-Server-Plot gibt für jeden Ort die Cell-ID an, der nach Modellrechnungen an diesem Ort die höchste Sendeleistung hat (s. Abbildung 55).

Cluster

- Anhand eines Clusterverfahrens und einer Abstandsfunktion bestimmte Gruppe von Objekten (Ganglinie, Tagesganglinie, Netzganglinie …) die ähnliche Eigenschaften haben.
- Formelzeichen : C_i

Dynamische Hochrechnung

- Orts- und zeitabhängige Hochrechnung von Trajektorien anhand von Messdaten.

dWiSta

- Abkürzung für Dynamische Wegweiser mit integrierten Stauinformationen.

FPD-Geschwindigkeit

- Mittlere Geschwindigkeit der auf das Projektnetz umgelegten, nach Pkw und Lkw sortierten und mit erfassten zeitgenauen Lkw-Anteilen gewichteten Trajektorien.

FPD-Trajektorie

- Weg-Zeit-Verlauf eines Mobilfunkteilnehmers.

Ganglinie

- Menge von Datensätzen zu je 96 Messwerten über 15 Minuten-Intervalle eines Tages.
- Die Datensätze repräsentieren die von stationären Detektoren oder FPD-Mobilfunktrajektorien versorgten Messwerte. Diese sind (Verkehrsstärke im Intervall, mittlere Geschwindigkeit des Intervalls und optional Lkw-Anteil sowie Geschwindigkeiten getrennt nach Pkw und Lkw).

Ganglinie, gemittelt

- Durchschnittliche Ganglinie als Mittelwert mehrerer Ganglinien eines RMQ.

Ganglinie, kalendertagtypisch

- Gemittelte Ganglinie der Netzganglinien aller Tage, die derselben TagTyp-Gruppe zugeordnet sind.
- Die zugeordneten Tage haben als Ergebnis der Vorklassifizierung alle dieselben kalendarischen Eigenschaften.

Ganglinie, verkehrstagtypisch

- Gemittelte Ganglinie der Netzganglinien aller Tage, die demselben Clusters zugeordnet sind.
- Die zugeordneten Tage haben als Ergebnis der Clusterung ähnliche Netzganglinien.

Ganglinie, Formelzeichen

- $G(l,\hat{t}).Wert(T)$ mit

  l als Ort (location), der einer RMQ-ID entspricht. Ist $\vec{l} = \begin{pmatrix} RMQ_1 \\ \vdots \\ RMQ_n \end{pmatrix}$ ein Vektor, dann repräsentiert G der Dimension von $\vec{l}$ entsprechend viele Ganglinien. In diesem Fall ist die Rede von einer Netz-Ganglinie.

  $\hat{t}$ als Datum. Ist $\vec{\hat{t}} = \begin{pmatrix} Tag_1 \\ \vdots \\ Tag_n \end{pmatrix}$ ein Vektor, dann repräsentiert G der Dimension von $\vec{\hat{t}}$ entsprechend viele Ganglinien. In diesem Fall ist die Rede von einem Cluster von

Ganglinien.
$Wert \in \left[Q_{Kfz}, Q_{Pkw}, Q_{Lkw}, V_{Kfz}, V_{Pkw}, V_{Lkw}\right]$
$T = Zeitindex \in [1,2,\ldots,95,96]$ für jedes 15-Minuten-Intervall eines (Cluster)Tages.

GEH

- Formelzeichen: $D_{GEH}\left(G(j,d), G(k,d)\right)$, Wert zum Vergleich von Verkehrsstärken (vgl. Kapitel 5.1.2.3).

Geschwindigkeit

- Formelzeichen: V_{Typ} in [km/h] mit $Typ \in [Kfz, Pkw, Lkw]$

Geschwindigkeit, kalendertagtypisch

- Geschwindigkeit oder –verlauf aus einer kalendertagtypischen Ganglinie.

Geschwindigkeit, verkehrstagtypisch

- Geschwindigkeit oder –verlauf aus einer verkehrstagtypischen Ganglinie.

Handover

- Übergabe einer Mobilstation zwischen zwei Funkzellen während eines Telefonates.

Hochrechnungszeitintervall

- Zeitliche Auflösung der dynamischen Hochrechnung.

Hochrechnungszeitraum

- Zeitraum, in dem Trajektorien und Zähldaten von der dynamischen Hochrechnung berücksichtig werden.

Ist-Geschwindigkeit

- Die Ist-Geschwindigkeit ist die mittlere Geschwindigkeit in einer Ganglinie aus FPD-Mobilfunkdaten oder aus Daten stationärer Detektoren in einem 15-Minuten-Intervall. Sie entspricht der tatsächlichen ermittelten oder gemessenen Geschwindigkeit (vgl. Soll-Geschwindigkeit).
- Formelzeichen: v_{Ist}

Korrekturfaktor

- $K_i = \frac{1}{n}\sum_{j=1}^{n} K_{i,j}$ Korrekturfaktor einer Trajektorie i als Mittelwert aller Korrekturfaktoren $K_{i,j}$ die für diese Trajektorie existieren. n entspricht dabei der Anzahl Zählstellen, die die Trajektorie passiert. Ein einzelner Korrekturfaktor ergibt sich, wie folgt aus dem Zählwert $Z(l,t)$ und der Verkehrsstärke $Q(l,t)$ zum Zeitpunkt t am Ort l:
$K_{i,j} = \frac{Z(l,t)}{Q(l,t)}$

Location Area (LA)
- Eine Location Area umfasst mehrere Funkzellen („cells“).

Location Area Code (LAC)
- Nummer bzw. Zeichen einer Location Area.

Location Area Update (LAU)
- Die Übergabe einer Mobilstation zwischen zwei Location Areas im Standby-Modus.

Lokalisierungsnetz (LN)
- Verkehrsnetz das auf Roadelementen (Navteq) basiert.

Messwert
- Auf ein quellenspezifisches Intervall aggregierter Wert über den Verkehrszustand der an stationären Detektoren erfasst wird. (Q, V, LKW-Anteil,…)
- Einheiten: Kfz/h; Km/h

MIV
- Motorisierter Individualverkehr

NBA
- Abkürzung für Netzbeeinflussungsanlage, in dieser Arbeit die Netzbeeinflussungsanlage in Stuttgart (s.a. WWW).

Netzganglinie
- Die erfassten Ganglinien aller RMQ eines Tages als zusammengefasstes Objekt.

PN-Trajektorie
- Weg-Zeit-Verlauf eines Mobilfunkteilnehmers auf RoadElementen.

Projektnetz (PN)
- Der für Do-iT relevante Teil des Lokalisierungsnetzes.

Projektnetzstrecke
- Logische Strecke auf dem Projektnetz, definiert durch Streckennamen (A8, A81, A5,…)

QZ-Matrix (Kfz)
- Quelle-Ziel-Matrix von Kraftfahrzeugfahrten (i. A. Ergebnis der Hochrechnung)
- Einheiten: Kfz /h, Fz/d, Kfz /t

QZ-Matrix (FPD)
- Quelle-Ziel-Matrix von Mobilfunktrajektorien
- Einheiten: Fahrten/h, Fahrten/d, Fahrten/t

QZ-Matrix, Formelzeichen

- $M_{Typ,\Delta t}(t)$ in $[Typ / \Delta t]$ mit
  $Typ \in [Kfz, Pkw, Lkw, FPD]$ wobei FPD in [Mobilfunkortsveränderungen/ Δt].
  t als Datum mit Uhrzeit. Ist $\vec{t} = \begin{pmatrix} Tag_1 \\ \vdots \\ Tag_n \end{pmatrix}$ ein Vektor dann repräsentiert M die mittlere Nachfrage.

Richtungsmessquerschnitte (RMQ)

- Stelle auf einer Strecke, an der durch einen stationären Detektor Messwerte erfasst werden.

RMQ-ID

- Die eindeutige Nummer eines RMQ.

RMQ-Geschwindigkeit

- Die gemittelte Geschwindigkeit eines 15-Minuten-Intervalls in einer Ganglinie, die aus den Messdaten stationärer Detektoren erzeugt wurde.

Roadelement

- Elementare Strecke im Lokalisierungsnetz (basiert auf einem Streckenelement des Navteq-Netzes)

Soll-Geschwindigkeit

- Die Soll-Geschwindigkeit ist die definierte mögliche mittlere Geschwindigkeit an einem RMQ oder einer Strecke, die bei freiem Verkehrsfluss erreicht werden kann. Sie entspricht dem Geschwindigkeitswert, der von 15% der 15-Minuten-Intervalle aller Tage und Intervalle erreicht oder überschritten wird. (vgl. Soll-Geschwindigkeit)
- Formelzeichen: v_{soll}

Stationärer Detektor

- Ein System, welches fest an oder in der Fahrbahn installiert ist, lokale Messwerte erfasst und diese Systemintern zu Intervallen aggregiert (z.B. Induktionsschleifen, Kennzeichenerfassungssysteme).

Strecke

- Eine Kante im Verkehrsmodell des Vereinfachten Projektnetzes zwischen zwei TMC-Locations. Die Kannte besitzt die relevanten Attribute:
  - Länge
  - Anzahl Fahrstreifen
  - Erlaubte Höchstgeschwindigkeit

Streckenzug

- Eine Verkettung von hintereinander liegenden Strecken mit jeweils gemeinsamen Anfangs- und Endpunkt. Alle Strecken haben dieselbe Streckenbezeichnung (z.B. A8) und Richtung (z.B. Stuttgart-Karlsruhe).

Streckenzählwert

- Zählwert an einer Strecke l im Zeitintervall t (siehe auch Zählwert)
- Einheiten: Kfz/h, Kfz/d, Kfz/t
- Formelzeichen: z

TagTyp

- Eine Vorklassifizierung des Datums anhand von Kalenderdaten (Wochentag, Werktag, Feiertag, Ferientag,...)

TMC-Location

- Codierte Position eines POIs auf dem Straßennetz.

TMC-Locationcodeabschnitts

- Streckenabschnitt (vonKM, bisKM) auf einer Strecke des Projektnetz auf dem eine TMC-Location definiert ist.

Trajektorienbelastung

- Belastung einer Trajektorie, kann vom Wert 1 abweichen.
- Einheit: Fahrten

Vereinfachtes Projektnetz (VPN)

- Abstraktes Verkehrsnetz, das auf TMC-Locations basiert (vgl. Kapitel 2.1).

Verkehrsleistung

- P in [Personenstunden/ Δt]

Verkehrsstärke

- Anzahl der Fahrzeuge, die an einer lokalen Position innerhalb einer bestimmten Zeit durchfahren.
- Formelzeichen: $Q_{Typ}(l,t)$ mit l als Ort (location), der einer Strecken-ID oder RMQ-ID entspricht. t als Datum mit Uhrzeit, ist $\vec{t}$ ein Vektor, dann repräsentiert Q die mittlere Verkehrsstärke dieser Zeitpunkte. Auch als z um zu verdeutlichen, dass es sich um eine gezählte Verkehrsstärke handelt.
- Einheit: in [$Typ/\Delta t$] mit $Typ \in [Kfz, Pkw, Lkw, FPD]$ wobei FPD in [Mobilfunkorts-veränderungen/ Δt]

Vollständigkeit

- Prozentuelle Angabe über den Anteil der versorgten Tage einer Ganglinie bezüglich aller gemessenen Tage.

VPN-Trajektorie

- Weg-Zeit-Verlauf eines Mobilfunkteilnehmers im VPN.

WWW

- Abkürzung für Wechselwegweisung, in dieser Arbeit die Wechselwegweisung in der BAB-Netzmasche (s.a. NBA).

Zeitindex

- $T = Zeitindex \in [1,2,\ldots,95,96]$ für jedes 15-Minuten-Intervall eines (Cluster)Tages.

Zeitintervall

- Δt in Sekunden, Minuten, Stunden, Tagen, Wochen, Monaten oder Jahren.

Zeitpunkt

- $_t$ als Datum mit Uhrzeit auch als $\hat{t}$ ein Datum ohne Uhrzeit, z.B. 5. Februar 2008 (00:00 Uhr).

9 Literatur

Backhaus, K., Erichson, B., Plinke, W., Weiber, R. (2006), *Multivariante Analysemethoden*, 11. Auflage, Springer Verlag, Berlin/Heidelberg, Deutschland.

de la Barra, T., Perez, B., Anez, J. (1993), Multidimensional Path Search and Assignment, *Proceedings of the 21st PTRC Summer Meeting*, Manchester, UK

Bastian, M., Boltze, M., Bobinger, R., Friedrich, B., Friedrich, M. Frommer, A., Heck, H.M., Papageorgiou, M., Kirschfink, H. (2008): *OptiV – Erschließung von Entscheidungs- und Optimierungsmethoden für die Anwendung im Verkehr*, Wissenspapier der FGSV, CD bzw. Internet Veröffentlichung unter www.optiv.de, FGSV Verlag, Köln.

Bierlaire, M. (2003), BIOGEME: A free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Schweiz.

Bierlaire, M. (2009), *An introduction to BIOGEME Version 1.8*, online verfügbar unter http://biogeme.epfl.ch, heruntergeladen am 08.06.2009

Bundesanstalt für Straßenwesen (BASt, 2007), *Manuelle Straßenverkehrszählung – Ergebnisse auf Bundesautobahnen*, online verfügbar unter http://www.bast.de/cln_007/nn_42256/DE/Statistik/Verkehrsdaten/Downloads/zaehlung-2005-BAB-strassen.html, heruntergeladen am 21.03.2009

Bundesministerium der Justiz (BMJ, 2009a), Verordnung zur Erleichterung des Ferienreiseverkehrs auf der Straße (Ferienreiseverordnung), zuletzt geändert durch V v. 13.6.2008, online verfügbar unter www.gesetze-im-internet.de/ferreisev_1985/, heruntergeladen am 17.04.2009

Bundesministerium der Justiz (BMJ, 2009b), Straßenverkehrs-Ordnung (StVO), zuletzt geändert durch durch Art. 1 V v. 26.3.2009 I 734, online verfügbar unter http://www.gesetze-im-internet.de/stvo/, heruntergeladen am 17.04.2009

Ben-Akiva, M., Bergman, M., Daly, A., Ramaswamy, R. (1984), Modelling Inter Urban Route Choice Behavior, *Ninth International Symposium on Transportation and Traffic Theory*, VNU Science Press, Utrecht, Niederlande

Ben-Akiva, M., Lerman, S. (1985), *Discrete Choice Analysis - Theory and Application to Travel Demand*, MIT Press, MA, USA.

Cascetta, E. (2001), *Transportation Systems Engineering: Theory and Methods,* Kluver Academic Publishers, Dordrecht, Niederlande.

Cowan, G. (2003), *Statistical data analysis*, Clarendon Press; Oxford University Press, Oxford, New York, USA.

Domencich, T., McFadden, D.L. (1975), *Urban Travel Demand: A Behavioral Analysis*, North-Holland Publishing Co., Niederlande, online verfügbar unter http://www.econ.berkeley.edu/~mcfadden/travel.html, heruntergeladen am 15.03.2009.

Everts, K. (1978), *Einsatz eines Steuerungsmodells zur Verkehrsstromführung mit Hilfe von Wechselwegweisern*, Forschung Straßenbau und Straßenverkehrstechnik, Heft 251, Hrsg.: Bundesministerium für Verkehr, Bonn.

Fastenrath, U. (2009), *TMCpro: Presence and Future of Real Time Traffic Information*, online verfügbar bei SlideShare Inc., online verfügbar unter www.slideshare.net/TMCpro/tmcpro-ii-presentation, heruntergeladen am 25.05.2009,

Felderhoff, M. (2005), *persönliches Gespräch mit Martin Felderhoff*, RAN Service Design, T-Mobile International AG & Co KG, Bonn am 19.05.2005.

Friedrich, B., Engelmann, F. (2003), *Verkehrsablauf an Fahrstreifenreduktionen*, Institut für Verkehrswirtschaft, Straßenwesen und Städtebau, Universität Hannover

Friedrich, M., Hofsäß, I., Nökel, K., Vortisch, P. (2000), Umlegung zeitlich differenzierter Nachfragematrizen: ein dynamisches Verfahren für Verkehrsplanung und Telematik. *Tagungsband zum 1. Aachener Kolloquium Mobilität und Stadt*, Schriftenreihe Stadt Region Land, Heft 69, Institut für Stadtbauwesen RWTH Aachen

Friedrich, M., Jehlicka, P., Schlaich, J. (2009), Erfassung von Verkehrsdaten mit automatischen Kennzeichenerfassungssystemen, *Straßenverkehrstechnik*, Heft 12, im Druck.

Friedrich, M., Jehlicka, P., Wacker, M. (2007): *Durchgangsverkehr Stuttgart: Erhebung des Durchgangsverkehrs durch die Landeshauptstadt Stuttgart auf der Bundesstraße B10 und den Bundesstraßen B27 (Nord) und B14 (Ost),* unveröffentlichter Projektbericht.

Forschungsgesellschaft für Straßen- und Verkehrswesen (FGSV, 2007), Hinweise zur Wirksamkeitsschätzung und Wirksamkeitsberechnung von Verkehrsbeeinflussungsanlagen, FGSV-Nr. 311, Ausgabe 2007, FGSV-Verlag, Köln.

Google Inc. (2007), *Google Maps*, CA 94043, USA, abgerufen am 09. November 2008, http://maps.google.de/maps

Huelsenbeck, J., Crandall, K. (1997), Phylogeny estimation and hypothesis testing using maximum likelihood, *Annual Review of Ecology and Systematics* 1997 28, S. 437-466.

Kayser, H.J., Krause, S. (1986), *Verkehrslenkungssytem Dernbacher Dreieck und Autobahnkreu Koblenz – Überprüfung und Bewertung*, Forschung Straßenbau und Straßenverkehrstechnik, Heft 491, Hrsg.: Bundesministerium für Verkehr, Bonn.

Louviere, J.J., D.A. Hensher. Swait, J. (2000), Stated Choice Methods: Analysis and Application in Marketing, *Transportation and Environmental Valuation*, Cambridge University Press, Cambridge.

McFadden, D.L. (2003), Economic Choices, S. 330-365, in: Nobel Lectures in Economic Sciences, 1996-2000, Hrsg.: T. Persson, World Scientific Publishing, Singapur.

Miller, E., Mahmassani, H., Ziliaskopoulos, A. (1994), Path search techniques for transportation networks with time-dependent, stochastic arc costs, *IEEE International Conference on Systems, Man, and Cybernetics. Humans, Information and Technology*, New York, USA

Mobilität in Deutschland (MID 2002), *Ergebnisbericht im Auftrag des Bundesministeriums für Verkehr*, Bau- und Wohnungswesen, Projekt 70.0736/2003, April 2004

Needleman, S.P., Wunsch, C.D. (1970), A general method applicable to the search for similarities in the amino acid sequence of two proteins, *Journal of Molecular Biology*, 48:443–453, 1970.

PTV Planung Transport Verkehr AG (PTV, 2006), *VISUM Benutzerhandbuch Version 9.4*, Karlsruhe, Kap. 5.5

Schlaich, J., Friedrich, M. (2008), Staumeldungen und Routenwahl in Autobahnnetzen – Teil 1: Analyse von Staumeldungen, *Straßenverkehrstechnik*, Heft 10/08, S. 621-627, Kirschbaum Verlag, Köln.

Schollmeyer, R., Wiltschko, T. (2007), Classification of public transport vehicles using cellular mobile radio data. *Proceedings on 6th European Congress and Exhibition on Intelligent Transport Systems and Services,* Aalborg, Dänemark.

Schreckenberg, M., Selten, R., Chmura, T., Pitz, T. und Wahle, J. (2001), Experiments on Day-to-Day Route Choice, *Tagungsband zu den 18. Verkehrswissenschaftlichen Tagen*, Dresden, Deutschland.

Sheffi, Y., Powell, W. (1982), An Algorithm for the Equilibrium Assignment Problem with Random Link Times, *Networks*, 12, 2, 191-207.

Toll Collect Gmbh (2008), *Vier Jahre Lkw-Maut in Deutschland*, Pressemitteilung vom 19.12.2008, online verfügbar unter www.toll-collect.de/frontend/press/PressEntryVP.do?pressId=601, heruntergeladen am 12.04.2009

Trapp, R., Feldges, M. (2009), Ermittlung der Wirksamkeit von Verkehrsbeeinflussungsanlagen, *Straßenverkehrstechnik*, Heft 10/08, S. 375-379, Kirschbaum Verlag, Köln.

Trapp, R., (2009), E-Mail-Kontakt mit Roland Trapp, Ing.-Büro Dr. Roland Trapp, Leiter des FGSV-Arbeitsausschusses „3.2 Verkehrsbeeinflussung ausserorts“, 23.07.2009.

UK Highway Agency (2008), *Design Manual for Roads and Bridges (DMRB)*, Volume 12, Section 2, Part 1, Office of Official Publications. Richmond, Surrey, Großbritannien.

Voeske, A. (2008), *persönliches Gespräch mit Achim Voeske*, Geschäftsführer der Antenne Radion GmbH & Co. KG, Stuttgart am 04.06.2008, siehe auch Präsentation unter http://www.ktmc.de/pdfs/080603_Staupilotenclub.pdf, abgerufen am 10.06.2009.

Vortisch, P. (2005), *Modellunterstützte Messwertpropagierung zur Verkehrslage-schätzung in Stadtstraßennetzen*, Dissertation an der TH Karlsruhe.

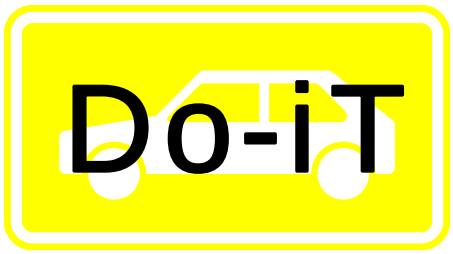
Wermuth, M. (2007), Personen und Personenwirtschaftsverkehr, in: *Handbuch Verkehrspolitik* von Schöller, O., Canzler, W. & Knie, A., S. 323-347. Springer Verlag, Heidelberg

Wermuth, M., Sommer, C. und Wulff, S. (2004), *Erhebung der individuellen Routenwahl zur Weiterentwicklung von Umlegungsmodellen*, Forschungsbericht FE-Nr. 01.158/2002/CGB im Auftrag des Bundesministeriums für Verkehr, Bau- und Wohnungswesen, Braunschweig/Berlin.

Arbeitspaket 4 / 10 Bericht

Datenoptimierung für integrierte Telematik

Do-iT



Förderkennzeichen: 19 B 5007

Zusammenfassender Bericht

Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren auf Abis und A-Ebene (AP4.6) sowie Evaluierung der Positionsbestimmung durch höherwertige Ortungsverfahren (AP10.3)

Alexander Beetz, Renate Czommer, Ralf Laufer, Volker Schwieger

September 2009

Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen

Dokument-Informationen

Version	Datum	Bearbeiter	Bemerkung
0.1	**02.03.09**	**Schwieger**	• Gliederung und Struktur
0.2	**19.06.09**	**Laufer**	• Ergänzung um Kapitel 3 und 4.2 sowie 5.2
0.3	**23.06.09**	**Czommer**	• Kapitel 2 und Teile 5
0.4	**07.07.09**	**Laufer**	• Kapitel 3, 4.2 vervollständigt und 5.2 erweitert
0.5	**09.07.09**	**Schwieger**	• Überarbeitung und Kommentare zum aktuellen Zustand
0.6	**10.08.09**	**Beetz / Czommer**	• Ergänzungen zu Kapitel 2 und 4 sowie Einarbeitung Kommentare von Schwieger
0.7	**13.08.09**	**Schwieger**	• Überarbeitung (pre-final) für Kapitel 2, 4.1, 5.1 Anfang 3
0.8	**31.08.09**	**Laufer**	• Ergänzung und Überarbeitung der Kapitel 3, 4.2 und 5.2
1.0	**11.09.09**	**Schwieger**	• Abschlusskorrektur und Zusammenfassung

Gliederung

1 Einleitung 4

2 Generierung von FPD-Trajektorien 5
2.1 Abis-Ebene 5
2.1.1 Priorisierung von Mobilfunkdaten für die Online-Auswertung 5
2.1.2 Positionsbestimmung 7
2.1.3 Kartengrundlage 22
2.1.4 Routengenerierung 25
2.1.5 Weg-Zeit-Diagramme 31
2.2 A-Ebene 35
2.2.1 Priorisierung 35
2.2.2 Kartengrundlage 37
2.2.3 Mobilfunknetz 39
2.2.4 Routengenerierung 42
2.2.5 Weg-Zeit-Diagramme 51

3 Qualitätskriterien 54
3.1 Qualitätskriterien: GPS als Referenz 56
3.2 Qualitätskriterien: SES als Referenz 59

4 Beschreibung der Referenzdaten 62
4.1 GPS-Trajektorien 62
4.2 Stationäre Erfassungssysteme 68

5 Qualitätsbeurteilung 78
5.1 Vergleich mit GPS-Trajektorien 78
5.1.1 Abis-Ebene 78
5.1.2 A-Ebene 86
5.1.3 Abis- versus A-Ergebnisse 93
5.2 Vergleich mit SES 97
5.2.1 Abis-Ebene 97
5.2.2 A-Ebene 99
5.2.3 A-Ebene versus Abis 101

6 Zusammenfassende Bewertung 103

7 Literatur 105

© Do-iT 2009

1 Einleitung

Dieser Bericht fasst die Ergebnisse der Arbeitspakete 4.6 „Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren auf Abis und A-Ebene“ und 10.3 „Evaluierung der Positionsbestimmung durch höherwertige Ortungsverfahren“ zusammen. Eine Trennung dieser beiden Berichte ist aus fachlichen Gesichtspunkten nicht sinnvoll, da sich die Aussagen beider Arbeitspakete sehr gut ergänzen.

Der Bericht umfasst dieser Einleitung folgend in Kapitel 2 die Vorgehensweise zur Generierung der FPD-Trajektorien. Hier ist, wie auch in den nachfolgenden Kapiteln zwischen den Datenquellen A-Schnittstelle und Abis-Schnittstelle zu unterscheiden. Die Entwicklung der Methodik für die Daten der Abis-Schnittstelle sowie deren nachfolgende Bewertung erfolgte teilweise bereits im Bericht zum AP4.5 [AP4-5 2008], so dass dieser Teil hier nur in seinen wesentlichen Auszügen wiedergegeben wird.

Kapitel 3 umfasst in Anlehnung an den Bericht zu AP 10.1 die Darstellung der einzelnen Qualitätskriterien und -parameter. Des Weiteren wird die Berechnung der Parameter behandelt. Hier muss zwischen Parameteren, die nur mit Referenzdaten (absolute Parameter) und denen die auch ohne Referenzdaten bestimmt werden können (relative Parameter), unterschieden werden. Als Referenzdaten kommen parallel aufgezeichnete Global Positioning System (GPS) Messungen und Stationäre Erfassungssysteme (SES) in Frage. Diese werden in diesem Kapitel erstmalig eingeführt.

Im nachfolgenden Kapitel werden die Referenzdaten ausführlich beschrieben. Dabei werden sowohl die GPS Trajektorien als auch die SES Standorte beschrieben, die jeweils zur Evaluierung herangezogen wurden.

Das zentrale Kapitel 5 fasst die Ergebnisse unter Nutzung der Methoden aus Kapitel 2 und der Referenzdaten aus Kapitel 4 zusammen. Die Ergebnisse werden in Form der Qualitätsparameter des Kapitels 3 ausgedrückt und beurteilt. Es wird sowohl zwischen A- und Abis-Schnittstelle als auch zwischen GPS und SES Evaluierung unterschieden. Den Abschluss bildet die vergleichende Wertung der A- und Abis-Ergebnisse.

Das abschließende Kapitel 6 fasst alle Ergebnisse zusammen und bietet einen Ausblick auf eventuelle zukünftige weitere Entwicklungen.

Alle Ausführungen beziehen sich, wenn nicht ausdrücklich abweichend erwähnt, auf die Online-Variante der FPD-Server. Parallel wurden zum Teil leicht abweichende Offline-Varianten aufgebaut, die hier nicht dokumentiert sind.

© Do-iT 2009

2 Generierung von FPD-Trajektorien

Zur Generierung von Trajektorien aus Mobilfunkdaten müssen zwei prinzipiell verschiedene Ortungsverfahren unterschieden werden. Am A-Interface stehen bei passivem Mobilfunkgerät, also außerhalb von Telefonaten, nur einzelne, zeitlich und räumlich weit voneinander entfernte Ortungsereignisse (Location Area Updates) zur Verfügung. Bei aktivem Mobilfunkgerät werden zusätzlich alle Zellwechsel (Handover) erfasst. Diese Einzelpositionen müssen mittels geeigneter logischer Zuordnungsverfahren mit dem digitalen Straßennetz in Verbindung gebracht werden.

Dagegen kann am Abis-Interface ein Mobiltelefon während eines Telefonats zusätzlich von einer BSC (Base Station Controller) fortlaufend in kurzen Zeitabständen (etwa alle 0.5 Sekunden) unter Nutzung von gemessenen Signalstärken geortet werden.

2.1 Abis-Ebene

An der Abis-Schnittstelle können die Zell-ID und der Timing-Advance-Wert (TA-Wert) sowie die Signalstärken zur bedienenden Zelle und zu den sechs Nachbarzellen abgegriffen werden. Zur Positionsbestimmung mit Signalstärken-Matching werden als Referenz die Signalstärkenkarten verwendet, die als Planungsdaten beim Mobilfunkbetreiber vorliegen und dort u. a. zur Konfiguration und Optimierung des GSM-Netzes genutzt werden.

2.1.1 Priorisierung von Mobilfunkdaten für die Online-Auswertung

Ein großes Problem bei der Online-Auswertung ist die Menge der im Online-Betrieb anfallenden Mobilfunkdaten, so dass im Sinne einer effizienten Verkehrslageerfassung die Rechenressourcen möglichst auf die relevanten Mobilfunkdaten zu richten sind. Hierfür ist eine Priorisierung der Mobilfunkdaten erforderlich. Die Einordnung der Priorisierung in den Ablauf der Online-Auswertung zeigt die folgende Abbildung 2-1.
Im Online-Prozess wird ein Datenpaket von Abis-Daten priorisiert. Danach werden die priorisierten Abis-Daten dem RX-Matching zur Auswertung übermittelt. Die berechneten Positionen werden anschließend als Verkehrsteilnehmer (Fußgänger, Individualverkehr, öffentlicher Verkehr usw.) identifiziert und am Schluss der FPD-Trajektoriengenerierung weitergeleitet. Alle hier aufgeführten Prozesse laufen parallel.

© Do-iT 2009

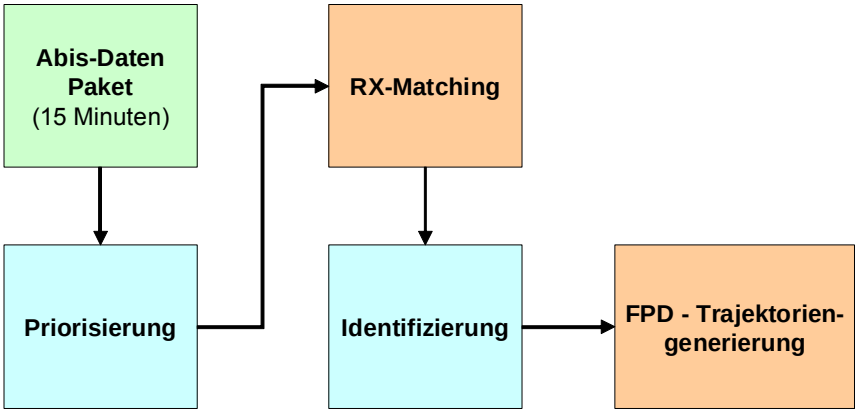


Abbildung 2-1: Grobablauf FPD-Generierung aus Abis-Daten

Während der Online-Auswertung werden 15 Minuten lang Daten gesammelt und anschließend ausgewertet. Es hat sich bei empirischen Untersuchungen gezeigt, dass eine Länge von 15 Minuten ein guter Kompromiss ist, zwischen dem Zerschneiden einzelner Telefongespräche und der Forderung zeitnahe Verkehrsinformationen zu liefern.
Im Weiteren soll nun auf die eigentliche Priorisierung eingegangen werden. Priorisiert wird z.B. aus Erfahrungen zurückliegender Analysen (Historie) oder aus den Mobilfunkdaten selbst. Eine Historieninformation ist bspw., ob ein Teilnehmer in einer der letzten Berechnungen bereits als „aktiv“ eingestuft worden ist. Kommen weitere Mobilfunkdaten dieses Teilnehmers hinzu, könnten diese eher von Interesse sein als von einem bislang „passiven“ Teilnehmer. Die Kriterien für die Priorisierung sind folgend aufgelistet.

1. H_{ap}: Grad der IV-Aktivität (**H**istorie): Teilnehmer war **a**ktiv = 1, **p**assiv = 0
2. Grad der Verkehrsrelevanz:
   - TAC_{Vr}: **TAC** Identifizierung für **V**erkehrs**r**elevanz (0...1)
   - P_{Sektor}: arithmetisches Mittel der Wahrscheinlichkeit (**P**) der Verkehrsrelevanz bestimmt mittels der erweiterten **Sektor**en (0...1)
3. Grad der Bewegung:
   - Anzahl unterschiedlicher Sektoren (0,1,2 ... $n_{Sektoren}$)
   - Anzahl der Messelemente (0,1,2 … $n_{Messelemente}$)

Zur effizienten Abarbeitung der Rangliste wird aus den Kriterien eine Priorisierungszahl (PZ) mit der folgenden Formel abgeleitet. Je höher diese Zahl ist, desto höher ist die Priorität des Teilnehmers für die Auswertung.

$$PZ = \mathrm{int}\left(\begin{array}{lll} & H_{ap} & \cdot 100\,000 \\ + & TAC_{Vr} & \cdot 1\,000\,000 \\ + & P_{Sektor} & \cdot 10\,000 \\ + & \dfrac{n_{Sektoren}}{n^{max}_{Sektoren}} & \cdot 1000 \\ + & \dfrac{n_{Messelemente}}{n^{max}_{Messelemente}} & \cdot 100 \end{array} \right)$$

mit : $n^{max}_{Sektoren}$ bislang maximale Anzahl Sektoren je Trajektorie

$n^{max}_{Messelemente}$ bislang maximale Anzahl Messwerte je Trajektorie

Für detaillierte Informationen sei hier auf den AP3.5-Bericht [AP3-5 2008] verwiesen.

2.1.2 Positionsbestimmung

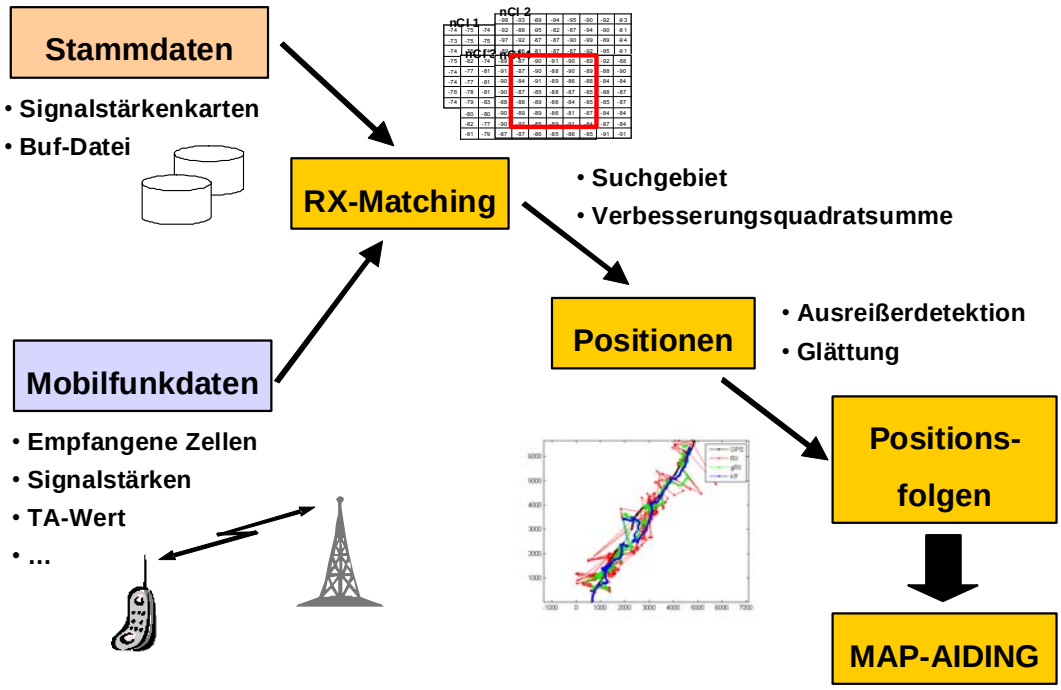


Abbildung 2-2: Ablauf der Mobilfunkortung aus Abis-Daten

Abbildung 2-2 zeigt den grundsätzlichen Ablauf der Mobilfunkortung bis zur Weitergabe an das Map-Aiding. An der Abis-Schnittstelle können die Zell-ID und der Timing-Advance-Wert (TA-Wert) sowie die Signalstärken zur bedienenden Zelle und zu sechs Nachbarzellen abgegriffen werden. Zur Positionsbestimmung mit Signalstärken-Matching werden als Referenz die Signalstärkenkarten genutzt, die als Planungsdaten

beim Mobilfunkbetreiber vorliegen und dort u. a. zur Konfiguration und Optimierung des GSM-Netzes genutzt werden.
Über die aus den Planungsdaten bekannte Ausrichtung der Antenne (BTS=Base Transceiver Station) und dem TA-Wert, der einen Hinweis zur Laufzeit des Signals und somit zur Entfernung zwischen Mobilstation und Antenne liefert, erfolgt eine erste grobe Bestimmung der Position. Anhand dieser Grobposition und den Messunsicherheiten wird der zu berücksichtigende Bereich aus den Signalstärkenkarten eingegrenzt. In diesem Bereich werden für alle 150 m x 150 m großen Rasterbereiche für die jeweils vorliegenden (maximal sieben) gemessenen Signalstärken eine bestmögliche Anpassung an Referenzdaten mit Hilfe der Verbesserungsquadratsumme über alle Suchbereiche bestimmt (vergleiche [Ramm & Schwieger 2007], [Schwieger et al. 2008]).

Auf diese Weise ergibt sich für jeden Zeitpunkt eine Position. Das vorgestellte Ortungsverfahren kann dabei nur während eines Telefonats oder einer Datenkommunikation des Mobilfunkteilnehmers erfolgen, da nur in diesem Fall die entsprechenden Daten auf der Abis-Schnittstelle zur Verfügung stehen.

Die ermittelte Positionsfolge des Mobilfunkteilnehmers wird im Anschluss über einen Kalman Filter geglättet. Im Ergebnis liegen die Positionen des anonymisierten Mobilfunkteilnehmers in zeitäquidistanten Abständen vor. Die Positionen werden zu einer Positionsfolge zusammengefasst, die eine zeitabhängige Beschreibung des zurückgelegten Wegs repräsentiert. Ein logischer Bezug zum Verkehrsnetz ist an dieser Stelle noch nicht gegeben.

Für die weitere Auswertung, insbesondere für die Eingrenzung des Straßennetzes und für die Qualitätsbewertung, wird auch im Online-Betrieb die Genauigkeit der Einzelpositionen benötigt. Beim reinen Signalstärken-Matching ist die Genauigkeit nur grob, anhand des durch die Messunsicherheiten des TA-Wertes eingegrenzten Bereichs, abzuschätzen und liegt im Mittel bei 2 km. Im Kalmanfilter wird die Positionsgenauigkeit auf Werte zwischen 2.5 m und 400 m geschätzt.

2.1.2.1 Stammdaten

Wie bereits oben erwähnt werden zur Mobilfunkortung Stammdaten der Mobilfunkortung benötigt. Diese beschränken sich hier auf Signalstärkenkarten in einem Rasterformat von 150 m x 150 m und die sogenannte Buf-Datei, die Informationen über die jeweiligen Standorte, Bauarten, Abstrahlrichtungen und weitere Antenneneigenschaften enthält. Diese werden nun im weiteren Verlauf detaillierter erklärt.

Signalstärkenkarten

Signalstärkenkarten werden als ASCII-Files genutzt und ca. alle 2 Monate aktualisiert. In diesen ASCII-Files wird die Signalausbreitung der Antenne rasterförmig in dBm (Dezibel Milliwatt) dargestellt. Die Rastergröße beträgt 150 m x 150 m bzw. 5 Bogensekunden. Wie in [AP4-5 2008] gezeigt werden konnte, ergaben sich für die

endgültigen Positionen keine deutlichen Verbesserungen wenn Karten mit einer höheren Auflösung gewählt wurden (untersucht wurden Karten mit einer Auflösung von einer Bogensekunde und 5 Bogensekunden). Da sich der Datenumfang innerhalb der Signalstärkenkarten jedoch deutlich reduziert, konnte bei der Berechnung ein deutlicher Performancegewinn erzielt werden.

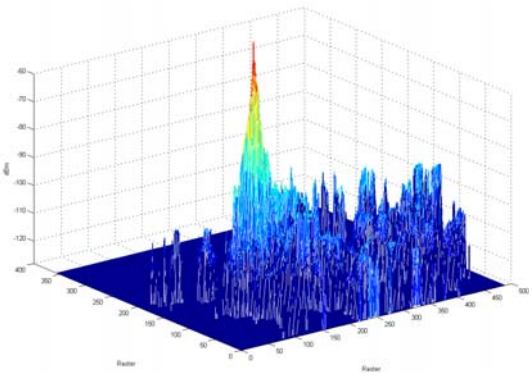


Abbildung 2-3: Verteilung der Signalstärkeninformation der Zelle 54121

Als Beispiel zeigt Abbildung 2-3 die Verteilung der Signalstärken von Zelle 54121. Sehr schön ist die Antenne als Peak mit der höchsten Signalstärke zu erkennen.

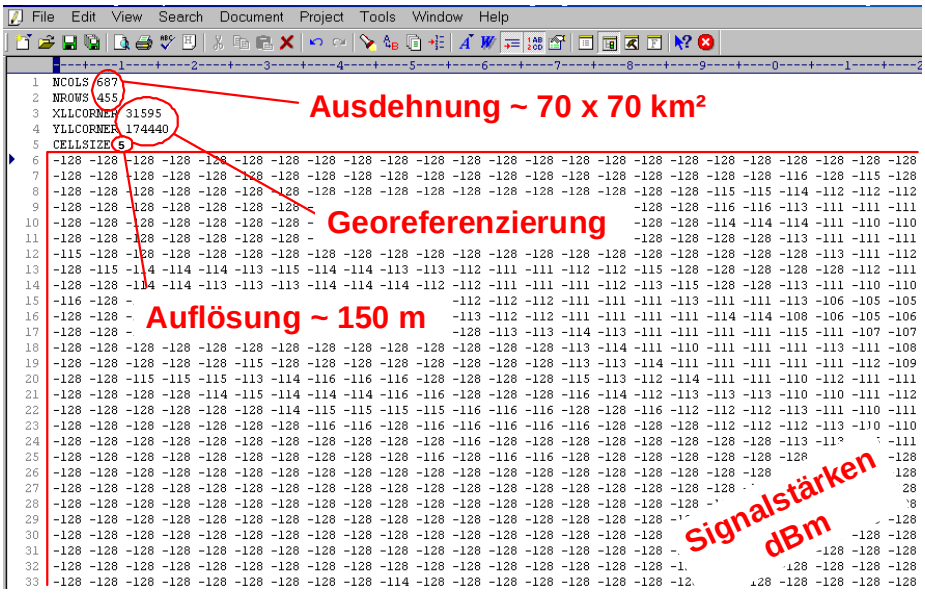


Abbildung 2-4: Darstellung der Signalstärkenkarte als ASCII-File

Abbildung 2-4 zeigt den Ausschnitt einer solchen Signalstärkenkarte. Links oben erkennt man den Header in dem die Anzahl der Zeilen und Spalten gegeben werden. Des Weiteren erfolgt die Längen- und Breitenangabe der linken unteren Ecke in Bo-

gensekunden. Als letzte Information erfolgt die Auflösung in Sekunden (hier 5 Bogensekunden). Durch diese Angaben ist es theoretisch möglich, für jeden Wert in der Karte ein Koordinatenpaar zu berechnen, was für die spätere Positionsbestimmung wichtig ist.

Für die spätere Auswertung werden Lookup-Tabellen erstellt, die zum einen den Dateinamen der jeweiligen Signalstärkenkarte und zum anderen die 4 Eckpunkte der Karte in Gauß-Krüger-Koordinaten (projiziert auf das WGS84 Ellipsoid) enthalten. Dadurch wird eine schnellere Berechnung der gesuchten Position erreicht, da dann für jeden Wert in der Karte die entsprechende Koordinate interpoliert werden kann. Andernfalls müsste jeder Rasterpunkt in der Signalstärkenkarte gesondert berechnet werden.

Buf-Datei (BSSDATA.BUF)

Der vollständige Namen der Datei heißt BSSDATA.BUF. BSS bedeutet hierbei Base Station Subsystem und umfasst alle Base Transceiver Stations und Base Station Controller. Wie schon eingangs erwähnt, gibt die Buf-Datei Informationen über die Antennen und Zellen im Projektgebiet. Buf-Dateien werden tagesaktuell vom Mobilfunkbetreiber generiert und umfassen ca. 6800 Antennen, welche in ganz Baden-Württemberg verteilt sind. Da die Buf-Datei eine Vielzahl an Informationen enthält, die nicht alle für das RX-Matching benötigt werden, wird hierfür eine vereinfachte Buf-Datei erstellt.

Die enthaltenen Informationen sind im Einzelnen:

- Zellen-Nummer
- LAC-Nummer
- Standpunktkoordinaten
- Abstrahlrichtung der Antenne
- Öffnungswinkel der Antenne
- Repeater Ja/Nein
- Index

Neben der Position und dem Abstrahlwinkel der Antenne ist beim RX-Matching auch der Öffnungswinkel der Antenne für die Berechnung des Suchgebiets wichtig. Da in der Buf-Datei des Mobilfunkbetreibers nur die codierten Antennenbezeichnungen vorkommen, müssen diese zunächst über Kennzahlen in die entsprechenden Öffnungswinkel übersetzt werden. Die Information des Repeaters konnten nicht verwendet werden, da aus den Mobilfunkdaten nicht eindeutig hervorgeht, ob die empfangene Zelle einen Repeater darstellt oder nicht. Zukünftig kann diese Information dennoch als Parameter dienen, um Positionen, die in Zellen mit Repeatern fallen, herunter zu gewichten, um so eine Verfälschung der Positionsfolge zu minimieren. Der Index erleichtert in der spä-

teren Berechnung die Suche nach Information in der vereinfachten Buf-Datei. Im Echtzeitbetrieb wird die vereinfachte Buf-Datei beim täglichen Systemneustart generiert.

2.1.2.2 Das RX-Matching

Im Folgenden wird das Verfahren des RX-Matchings erläutert. Zunächst soll das Verfahren grundsätzlich erklärt werden. Danach werden die einzelnen Schritte behandelt, welche für den Echtzeitbetrieb notwendig waren, um eine Beschleunigung der Berechnungsgeschwindigkeit und eine höhere Stabilität der Rechenprozesse zu erreichen.

Grundsätzliches Verfahren

Wichtige Daten für das RX-Matching sind sowohl die bereits oben genannten Stammdaten, als auch die Signalstärken der bedienenden Zelle sowie aller Nachbarzellen, die am Mobiltelefon empfangen werden. Des Weiteren wird der TA-Wert (Timing-Advance-Wert) der bedienenden Zelle, welcher die Entfernung zur Antenne wiedergibt, zur Bestimmung des Suchgebiets für das RX-Matching benötigt. Das Suchgebiet entspricht einem Ausschnitt aus allen Signalstärkenkarten, in dem sich der Teilnehmer befindet. Würde dieses Suchgebiet nicht erstellt, müsste über den gesamten Bereich der Signalstärkenkarte ein RX-Matching durchgeführt werden, was einen enormen Rechenaufwand bedingen würde. Grundsätzlich können beim RX-Matching nur Daten verwendet werden, die während des Telefonats im zwei Hz Takt an der Abis-Schnittstelle aufgezeichnet werden. Dabei werden unter anderem folgende, für das RX-Matching wichtige Daten ausgegeben:

- Teilnehmer ID (anonymisiert)
- ID Location Area (Location Area in der sich die bedienende Zelle befindet)
- ID bedienende Zelle (Zelle in der Mobilfunkteilnehmer angemeldet ist)
- TA-Wert (Timing Advance – Abstand zur Antenne in 550 m - Schritten)
- Signalstärke bedienende Zelle (dBm)
- IDs aller empfangenen Nachbarzellen (bis zu 6)
- Signalstärken aller empfangen Nachbarzellen (dBm)

Dem RX-Matching selbst liegt folgende Idee zu Grunde. Es wird die Position gesucht, in die sich alle gemessenen Signalstärken im Vergleich zu den Soll-Signalstärken der Signalstärkenkarten optimal einpassen lassen. Das Minimum dieser Quadratsumme ergibt, bezogen auf die koordinierten Rasterwerte der Signalstärkenkarten, die gesuchte Position. Dies lässt sich zunächst durch die aus [Ramm & Schwieger 2007] bekannte Formel ausdrücken.

$$\min_{y,x\in A}\sum_{i=1}^{n}(P_{Soll}(y,x)_i - P_{Mess,i})^2$$

A: Suchgebiet
n: Anzahl der gemessenen Zellen
P_{Soll}: Signalstärke aus Signalstärkenkarten auf Position (y,x) bezogen
P_{Mess}: Signalstärke aus Mobilfunkgespräch bezogen

Das RX-Matching wurde inzwischen etwas erweitert. Da gemessene Daten mit Planungsdaten verglichen werden, kann es zu systematischen Fehlern kommen. Die Planungsdaten entsprechen nicht ganz der tatsächlichen Ausbreitung der Antennensignale. Dieser Unterschied ist jedoch bei allen Planungsdaten gleich, da hinter allen Berechnungen für die Signalstärkenkarten das gleiche Dämpfungsmodell steckt. Dieser systematische Fehler drückt sich näherungsweise durch einen Offset aus, der sich durch Bildung von Differenzen beseitigen lässt und zu einer deutlichen Positionsverbesserung beiträgt. Die Differenzen werden dabei zwischen einer Signalstärkenkarte einer bestimmten empfangenen Zelle und den Signalstärkenkarten aller anderen empfangenen Zellen gebildet. Dabei werden die Signalstärken gleicher Koordinaten voneinander abgezogen. Des Weiteren werden die empfangenen Signalstärken der einzelnen Zellen des Mobilfunkgerätes in gleicher Weise voneinander abgezogen. Anschließend werden aus den gebildeten Differenzen die Abweichungen gebildet und die Verbesserungsquadratsumme berechnet. Das Minimum ergibt dann wieder die gesuchte Position. Die oben genannte Formel ändert sich dadurch folgendermaßen ab:

$$\min_{y,x\in A}\sum_{i=1}^{n}((P_{Zell,BED}(y,x) - P_{Zell,NACH}(y,x)_i) - (P_{Mess,BED} - P_{Mess,NACH,i}))^2$$

A: Suchgebiet
n: Anzahl der gemessenen Zellen
P_{Zell}: Signalstärke aus Signalstärkenkarten auf Position (y,x) bezogen
P_{Mess}: Signalstärke aus Mobilfunkgespräch
BED: Bedienende Zelle
NACH: Nachbarzelle

Die folgende Tabelle 2-1 zeigt exemplarisch die deutlichen Verbesserungen für die Standardabweichung der Positionsbestimmung an Hand dreier Beispiele:

Tabelle 2-1: Vergleich RX-Matching mit RX-Matching aus Differenzen. Die Messfahrten wurden am 13.09.2009 durchgeführt und werden in [AP3-6 2009] beschrieben.

	RX-Matching	RX-Matching mit Differenzenbildung
Bundesstraße 3	919	607
Bundesstraße 4	603	418
Kreisfahrt 2	1150	760

Im Weiteren werden nun alle für das RX-Matching notwendigen Schritte aufgeführt und detailliert beschrieben. Zunächst soll das nachfolgende Ablaufdiagramm einen groben Überblick verschaffen (vgl. Abbildung 2-5).

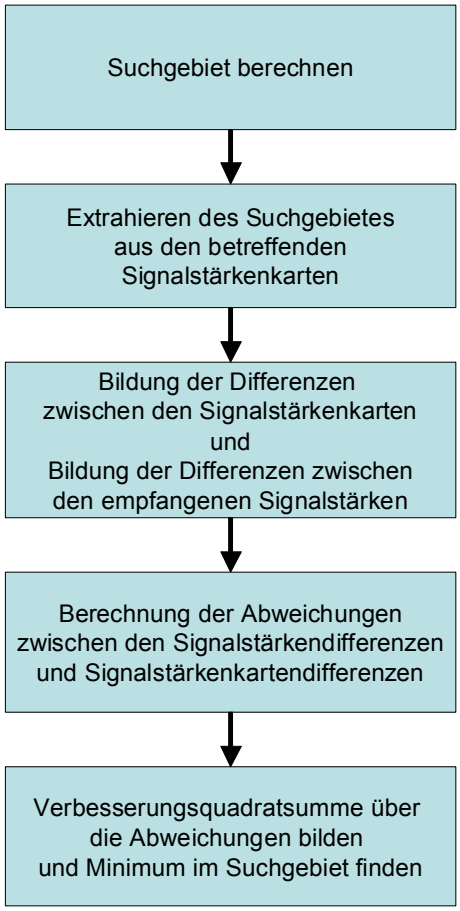
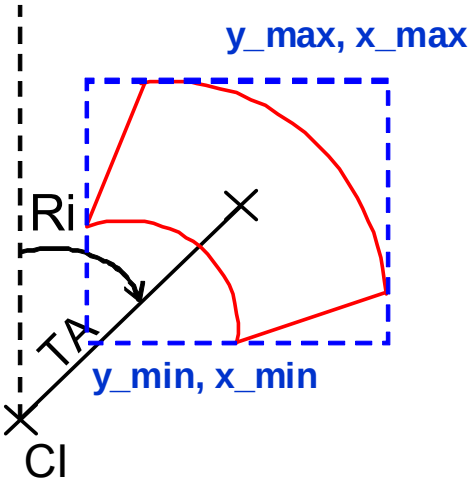


Abbildung 2-5: Ablauf RX-Matching

1. Berechnung Suchgebiet

Im ersten Schritt wird aus den Informationen „Bedienende Zelle (CI)“, TA-Wert sowie die Abstrahlrichtung und dem Öffnungswinkel der Antenne ein Suchgebiet in Rechteckform berechnet. Vom Standpunkt der Antenne wird durch polares Anhängen das Zentrum des Suchgebietes ermittelt und, je nach Ausdehnung, die Eckpunkte berechnet. Die Ausdehnung ergibt sich aus den Grenzen des Öffnungswinkels der Antenne und jeweils einem halben TA-Wert in Richtung der Antenne und in die entgegen gesetzte Richtung vom Zentrum des Suchgebietes aus gesehen (vgl. Abbildung 2-6).

Suchgebiet:



Sektormaske:

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Abbildung 2-6 : Berechnung des Zielgebiets und der Sektormaske

Im zweiten Schritt wird eine Sektormaske generiert, die der Größe des Suchgebietes entspricht. Die Sektormaske wird am Ende der Quadratsummenbildung über das Ergebnisarray gelegt um das Minimum nur innerhalb des Sektors zu suchen. Das heißt alle Werte außerhalb des Sektors werden ausgeschaltet. In der Sektormaske werden alle Werte, die innerhalb des rot dargestellten Sektors (Abbildung 2-6 links) sind, auf 1 gesetzt. Alle anderen Werte werden sehr hoch gewichtet. In Abbildung 2-6 (rechts) sind diese hoch gewichteten Werte, auf Grund der Übersichtlichkeit, mit einer 0 dargestellt. Die hohe Gewichtung ergibt sich aus der Suche nach dem Minimum der Quadratsumme, die nur im Sektorbereich zu suchen ist.

2. Extrahieren des Suchgebiets aus den Signalstärkenkarten

Beim Extrahieren werden mit Hilfe der Eckpunkte des Suchgebiets die entsprechenden Gebiete aus den Signalstärkenkarten herausgeschnitten (vgl. Abbildung 2-7).

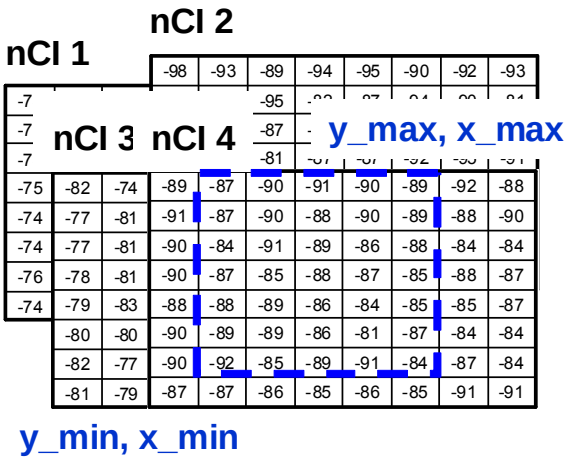


Abbildung 2-7: Extrahieren des Suchgebiets

Es entstehen dadurch, je nach Anzahl der empfangenen Mobilfunkzellen bis zu 7 Kartenauszüge.

3. Bildung der Differenzen

Ablauf der Differenzenbildung:

a) Im ersten Schritt werden die Differenzen zwischen den Signalstärkenkarten (rx_{Zell}) gebildet. Dabei werden die aus den Signalstärkenkarten extrahierten Daten (Ausschnitt Suchgebiet) elementweise voneinander abgezogen. Elementweise bedeutet, dass die Werte jeweils gleicher Koordinatenpaare voneinander abgezogen werden. Da die bedienende Zelle den zuverlässigsten Signalwert bietet, wird mit dieser die Differenz mit allen anderen Signalstärkenkarten gebildet. Das folgende Beispiel soll dies verdeutlichen:

   Ausschnitt bedienende Zelle – Ausschnitt 1. Nachbarzelle
   Ausschnitt bedienende Zelle – Ausschnitt 2. Nachbarzelle
   Ausschnitt bedienende Zelle – Ausschnitt 3. Nachbarzelle
   usw.

   Bei maximal 7 Kartenausschnitten entstehen so 6 Matrizen mit Signalstärkendifferenzen.

b) Im zweiten Schritt werden nun in derselben Weise die empfangenen Signalstärken (rx_{Mess}) voneinander abgezogen.

   empf. Signalstärke bedienende Zelle – empf. Signalstärke 1. Nachbarzelle
   empf. Signalstärke bedienende Zelle – empf. Signalstärke 2. Nachbarzelle
   empf. Signalstärke bedienende Zelle – empf. Signalstärke 3. Nachbarzelle
   usw.

   Bei maximal 7 empfangenen Signalstärken entstehen so 6 Signalstärkendifferenzen.

4. Abweichung bilden zwischen Soll-Signalstärken und empfangener Signalstärke

Hier werden bei allen Werten innerhalb der extrahierten Kartenauszüge die Differenz zwischen den Empfangenen Signalstärken der einzelnen Zellen und den in den dazugehörigen Karten befindlichen Sollsignalstärken gebildet (vgl. Abbildung 2-8).

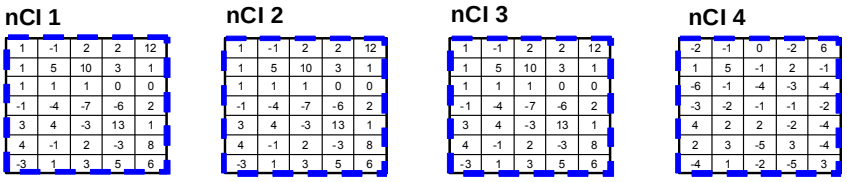
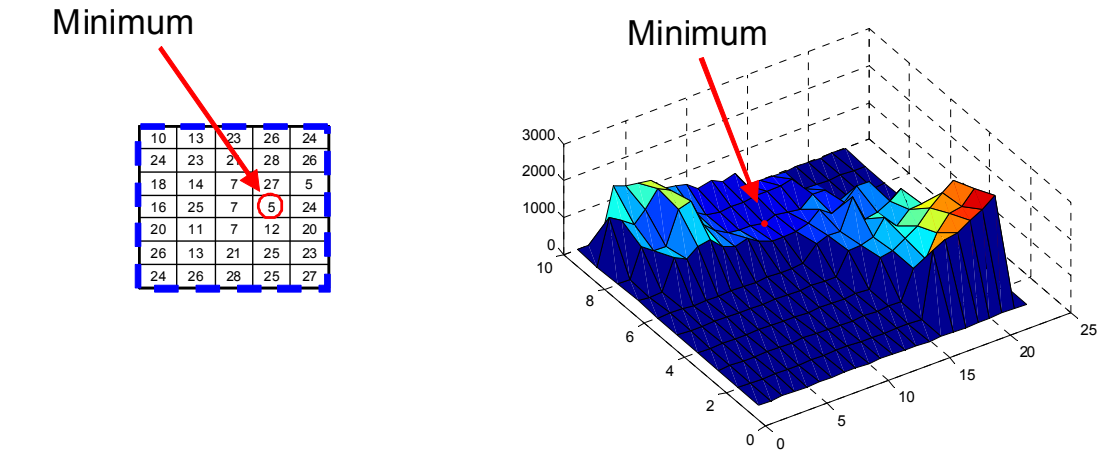


Abbildung 2-8: Kartenauszüge mit Differenzen

4. Bildung der Verbesserungsquadratsumme

In diesem Schritt wird über die gebildeten Differenzen die Quadratsumme gebildet. Dabei werden alle Differenzen derselben Koordinaten quadratisch aufaddiert. Das Ergebnis ist eine Matrix der Quadratsummen, wobei das Minimum der gesuchten Position entspricht. Im letzten Schritt wird nun über die Ergebnismatrix, die im 1. Schritt berechnete Sektormaske gelegt, sodass nur das Minimum innerhalb des Suchsektors betrachtet wird. Hierzu muss gesagt werden, dass es beim Berechnen der Verbesserungsquadratsumme vorkommen kann, dass für den gesuchten Punkt mehrere Minima gefunden werden. Eine Möglichkeit zur Lösung dieses Problems ist hier die Berechnung des Mittels über alle Minima, die andere Möglichkeit ist die Berechnung der Strecken von der letzten Position zu allen neuen Minima, wobei unterstellt wird, dass die neue Position mit hoher Wahrscheinlichkeit sehr nah an der vorherigen liegt. Im Algorithmus wurde es so umgesetzt, dass im ersten Durchlauf unterschieden wird, ob nur ein Minimum oder mehrere Minima vorliegen. Im ersten Fall wird das Ergebnis übernommen, im zweiten Fall wird ein Mittel über alle Minima gebildet. Da die Minima zum Teil sehr stark gestreut sind, werden für alle weiteren Durchläufe die Strecken zu den Minima berechnet, sollte es keine eindeutige Lösung geben. Danach wird der Punkt als neue Position gewählt, welcher den kürzesten Abstand zur vorherigen Position hat.



Ergebnis der Quadratsummenbildung 3D-Darstellung eines Ergebnissektors

Abbildung 2-9: Ergebnis der Quadratsummenbildung

In Abbildung 2-9 wird zum einen eine Ergebnismatrix der Quadratsummenbildung, zum anderen ein 3D-Bild eines Ergebnissektors gezeigt. Zur Verdeutlichung wurden auch hier alle hoch gewichteten Werte auf Null gesetzt.

Ausreißerdetektion und Glättung

Nach der Berechnung der Positionen mit dem RX-Matching sind die Positionsfolgen stark verrauscht. Um beim anschließenden Map-Aiding bessere Ergebnisse zu erhalten, werden die Positionsfolgen geglättet. Es wurden zwei verschieden Glättungsverfahren untersucht. Zum einen eine Glättung durch ein gleitendes Mittel, zum anderen eine Glättung durch einen einfachen Kalman Filter. Betrachtet man die Genauigkeit allein, sind beide etwa gleich. Auf Grund der besseren Glättungseigenschaften und Formtreue fiel die Wahl zu Gunsten des Kalman Filters aus. Mit Hilfe des Kalman Filters können die Positionen und Geschwindigkeiten zusätzlich geschätzt und etwaige Ausreißer eliminiert werden. Der Kalman Filter selbst ist auf Grund von Performancegründen recht einfach gehalten und gestattet eine sehr schnelle Berechnung der gesuchten Werte. Im Folgenden ist die mathematische Beschreibung des Kalman Filters aufgezeigt [vgl. Ramm & Schwieger 2007].

Im Allgemeinen gilt:

$$\tilde{x}_{k+1} = T \times \tilde{x}_k + C \times w$$

$\boldsymbol{T}$ ist die Transitionsmatrix, $\boldsymbol{C}$ ist die Störgrößenmatrix und in $\boldsymbol{w}$ ist die Gewichtung der Störgrößen. Da $\boldsymbol{w}$ nicht bekannt ist ergibt sich für den Erwartungswert $\boldsymbol{E(w)=0}$:

$$\overline{x}_{k+1} = T \cdot \hat{x}_k$$

Das konkrete mathematische Modell lässt sich wie folgt darstellen:

$$\begin{vmatrix} x_{k+1} \\ y_{k+1} \\ v_{x,k+1} \\ v_{y,k+1} \end{vmatrix} = \begin{vmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} x_k \\ y_k \\ v_x \\ v_y \end{vmatrix},$$

mit v_x und v_y als Geschwindigkeiten in x- und y-Richtung.

Durch Kovarianzfortpflanzung lässt sich die Kovarianzmatrix der Prädiktion mit der Störmatrix $\boldsymbol{C}$ schätzen:

$$\Sigma_{\bar{x}_{k+1}\bar{x}_{k+1}} = T \cdot \Sigma_{\hat{x}_k\hat{x}_k} \cdot T^T + C \cdot \Sigma_{ww} \cdot C^T$$

wobei:

$$C = \begin{vmatrix} 0.5\Delta t^2 & 0 \\ 0 & 0.5\Delta t^2 \\ \Delta t & 0 \\ 0 & \Delta t \end{vmatrix}$$

Die Gleichung für die gemessenen Werte ergibt sich aus:

$$l_{k+1} = A \cdot \bar{x}_{k+1} + \varepsilon_{k+1}$$

Der Beobachtungsvektor $\boldsymbol{l}_{k+1}$ sind die geschätzten Trajektorien als Positionen l_x und l_y, wobei sich die Designmatrix $\boldsymbol{A}$ folgendermaßen gestaltet:

$$A = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix}$$

Die Innovation berechnet sich zu:

$$d_{k+1} = l_{k+1} - A \cdot \bar{x}_{k+1}$$

$$D = \Sigma_{l_{k+1}l_{k+1}} + A \times \Sigma_{\bar{x}_{k+1}\bar{x}_{k+1}} \times A^T$$

Die Kalmanverstärkungmatrix ergibt sich mittels:

$$K = \Sigma_{\bar{x}_{k+1}\bar{x}_{k+1}} A \cdot D^{-1}$$

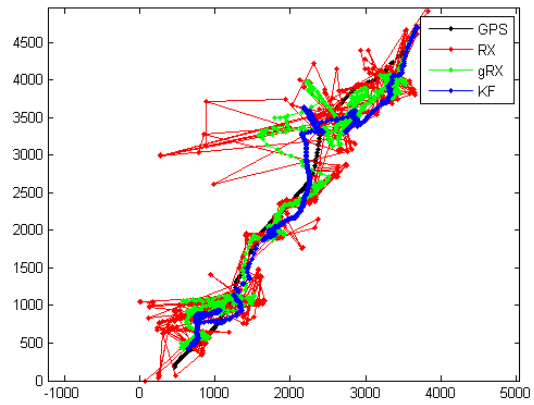
Für den Zustandsvektor und dessen Kovarianzmatrix ergeben sich folgende Gleichungen:

$$\hat{x}_{k+1} = x_{k+1} + K \cdot d_{k+1}$$

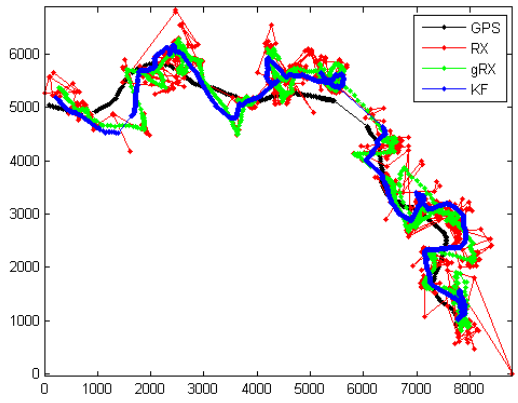
$$\Sigma_{\hat{x}_{k+1}\hat{x}_{k+1}} = \Sigma_{\bar{x}_{k+1}\bar{x}_{k+1}} - K \cdot D \cdot K^T$$

In den nachfolgenden Abbildungen sind einige Messfahrten, die die Ergebnisse der unterschiedlichen Auswertemethoden aufzeigen, als Beispiele dargestellt. Schwarz ist

dabei die GPS-Trajektorie als Referenz (GPS). Rot sind die ungefilterten Positionen (RX), grün sind die mit einem gleitenden Mittel gefilterten (gRX) und letztendlich blau, die mit einem Kalman Filter gefilterten Positionen (KF). Des Weiteren sind für jede Messfahrt die Standardabweichungen der einzelnen Auswertestrategien bezogen auf die GPS-Solltrajektorie angegeben.

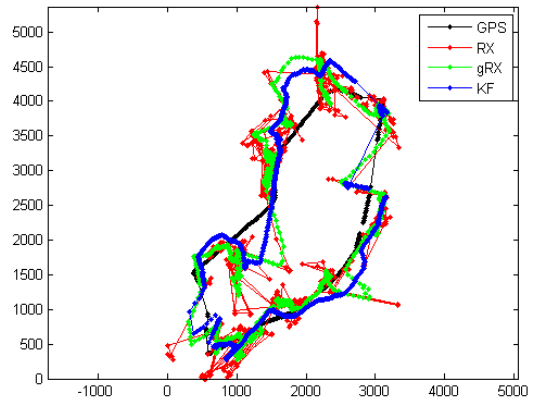


s_{RX}= 496 m; s_{gRX}= 316 m; s_{Kal}= 317 m

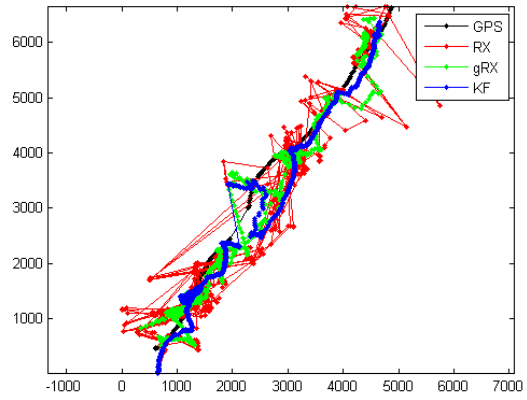


s_{RX}= 479 m; sg_{RX}= 409 m; s_{Kal}= 373 m

Abbildung 2-10: Versuchsfahrten Bundestrasse_B3_4_H2, Nr.72 (links) und Handy_an_Handy_1_H1, Nr.241 (rechts) (vgl. Messfahrten in [AP3-6 2009])



s_{RX}= 395 m; sg_{RX}= 333 m; s_{Kal}= 333 m



s_{RX}= 527 m; sg_{RX}= 384 m; s_{Kal}= 374 m

Abbildung 2-11: Versuchsfahrten Kurve_gr_4_H1, Nr.232 (links) und Bundesstrasse_B3_1_H2, Nr.42 (vgl. Messfahrten in [AP3-6 2009])

Anhand der gezeigten Beispiele lässt sich sehr gut erkennen, dass eine bessere Formanpassung an die gefahrene Strecke beim Kalman Filter zu finden ist. Weiter wird deutlich, dass die Genauigkeit der Positionen durch die Filteralgorithmen gesteigert werden kann, was für die weitere Datenprozessierung sehr wichtig ist.

Erweiterungen für die Robustheit

Im Folgenden werden zwei notwendige Erweiterungen aufgezeigt die zur Robustheit des Auswerteprozesses beitragen. Im Einzelnen überprüfen nachfolgende Funktionen die Plausibilität der empfangenen Zellen.

1. Überprüfung empfangener Zellen auf das Vorhandensein im Projektgebiet

Ein Problem, das im Echtzeitbetrieb vorkommt, ist die Aufzeichnung von Zellen, die nicht im Projektgebiet existieren. Dies kann mehrere Gründe haben. Zum Einen können diese Fehlzuordnungen bei der Aufbereitung der Rohdaten entstehen, zum Anderen kann es am Rand des Projektgebietes vorkommen, dass in Zellen gewechselt wird, die außerhalb des Projektgebiets liegen. Da diese Inkonsistenz jedoch den Algorithmus zum Absturz bringen kann, müssen solche Zellen vorab aus den Daten gelöscht werden. Zu diesem Zweck wird jedes auszuwertende Telefonat zunächst darauf überprüft, für welche Zellen Signalstärkenkarten in den Stammdaten vorhanden sind und für welche nicht. Sind für bestimmte Zellen keine Signalstärkenkarten in den Stammdaten vorhanden, werden diese aus dem Datensatz entfernt.

2. Überprüfung empfangener Zellen auf Korrektheit

Ein weiteres Problem sind falsch zugeordnete Zellen, die als Signalstärkenkarten in den Stammdaten vorhanden sind. Es wird zunächst kein Fehler festgestellt, jedoch bringen die meisten falsch zugeordneten Zellen den Algorithmus zum Absturz, wenn das Suchgebiet außerhalb der Signalstärkenkarte liegt. Gelöst wird dieses Problem indem vor der Positionsberechnung überprüft wird, ob alle Nachbarzellen innerhalb des Suchgebietes liegen. Ist dies bei einzelnen Zellen nicht der Fall, werden diese aus der Berechnung herausgenommen. Ist die bedienende Zelle selbst falsch zugeordnet, ist dies nicht unmittelbar festzustellen. Im günstigen Fall sind alle Nachbarzellen außerhalb des Suchgebiets und es kommt keine Berechnung zu Stande, da wenigstens zwei richtig empfangene Zellen für das RX-Matching notwendig sind. Im ungünstigen Fall kommt es zur Positionsberechnung und es wird eine falsche Positionsbestimmung durchgeführt. Dieses Problem wurde im Testbetrieb bisher vernachlässigt, da die Wahrscheinlichkeit für das Eintreten dieses Falles als gering einzuschätzen ist. Dies müsste jedoch noch durch genauere Untersuchungen verifiziert werden.

Erweiterungen für den Echtzeitbetrieb

In diesem Unterkapitel werden kurz einige Maßnahmen beschrieben, welche die Performance des Systems verbessern. Darunter fällt auch die Verwendung von 5-Sekunden Signalstärkenkarten anstatt 1-Sekunden Signalstärkenkarten. Dies wird hier jedoch nur der Vollständigkeit halber genannt. Für weitere Informationen diesbezüglich sei hier auf den AP 4.5-Bericht verwiesen [AP4-5 2008]. Es folgen nun weitere Maßnahmen die zur Beschleunigung des Auswerteprozesses beigetragen haben.

1. Vorhalten der Kartendaten im Arbeitsspeicher

Einen sehr großen Performancegewinn bringt das Vorhalten der Signalstärkenkarten im Arbeitsspeicher. Dabei wird das auszuwertende Telefonat vor Beginn der Berechnung hinsichtlich der vorkommenden Zellen überprüft, welche dann vor den eigentlichen Positionsberechnungen in den Arbeitsspeicher geladen wird. Bei den eigentlichen Berechnungen müssen nun nicht jedes Mal die Karten einzeln geladen werden, sondern es kann jederzeit auf die Daten im Arbeitsspeicher zugegriffen werden. In unserem Fall haben sich 3 dimensionale Matrizen zum Vorhalten der Kartendaten bewährt.

2. Änderung der Abtastrate

Ein weiterer Performancegewinn kann durch das Ändern der Abtastrate erzielt werden. Dabei wird nicht jede, sonder eventuell nur jede fünfte oder jede zehnte gemessene Datenzeile (Signalstärken, Zellen, usw.) für Auswertung herangezogen. Bei unterschiedlichen Szenarien konnten bis zu 9 Datenzeilen weggelassen werden, wobei sich das Ergebnis nicht signifikant änderte. Ein Problem stellt jedoch die in der Identifikation (vgl. [AP3-5 2008]) benutzte Schwerpunktberechnung zur Bewegungsanalyse dar, die auf eine große Anzahl von Positionen angewiesen ist, um brauchbare Ergebnisse zu liefern. Aus diesem Grund wurde die originale Abtastrate beibehalten. Dieser Ansatz sei hier dennoch genannt, sollte bei späteren Anwendungen eine große Anzahl von Positionen keine Rolle mehr spielt.

3. Beibehalten der Sektormaske bei gleichem TA-Wert und gleicher Zelle

Bei der Berechnung der Sektormaske kann etwas Zeit gespart werden, wenn bei jeder Positionsberechnung zuerst überprüft wird, ob eine Neuberechung der Sektormaske nötig ist. Da der TA-Wert sich erst alle 550 Meter ändert und auch ein Zellwechsel nicht bei jeder neuen Datenzeile eintritt, kann eine Sektormaske für mehrere Durchgänge benutzt werden. Zukünftig ist es sinnvoll, Sektormasken für die verschiedenen TA-Werte der einzelnen Zellen vorab zu berechnen und vorzuhalten. Hier ist sicher ein weiterer Performancegewinn zu erreichen.

Die folgende Tabelle 2-2 gibt einen groben Überblick über die Performanceverbesserungen der einzelnen Maßnahmen. Die Berechnungszeiten wurden während der Entwicklungsphase nur sporadisch protokolliert und geben somit keine exakte Angabe für die Berechnungszeiten. Diese würden sowieso, auf Grund unterschiedlicher Rechnerleistungen, von System zu System variieren. Sie geben aber dennoch Aufschluss über erreichbare Performancegewinne. Das hier verwendete Testsystem hatte einen Quad-Core-Prozessor mit jeweils 2,66 GHz und einen Arbeitsspeicher von 4 Gigabyte. Der verwendete Testdatensatz entstammt den Messfahrten vom 13.09.2007. Insgesamt wurden 48 von 52 Messfahrten ausgewertet. Dies entspricht einer Gesamtgesprächsdauer von ca. 8 Stunden (vgl. Messfahrten in [AP3-6 2009]). In der Tabelle wird die zeitliche Verbesserung der jeweiligen zugeschalteten Maßnahme gezeigt. Jede Maßnahme beinhaltet dabei auch die zuvor aufgeführte(n) Maßnahme(n).

Tabelle 2-2: Übersicht der Rechenzeit von 48 Messfahrten vom 13.09.2007 mit einer Gesamtgesprächsdauer von ca. 8 Stunden

Maßnahme	**Zeitaufwand (Rechenzeit)**
Auswertung ohne performancesteigernde Maßnahmen	ca. 2 Wochen an mehreren PCs
+ Vorhalten der Kartendaten im Arbeitsspeicher	ca. 8 Stunden an einem PC
+ Verwenden von 5-Sekunden Karten anstatt 1-Sekunden Karten	ca. 1,5 Stunden an einem PC
+ Abtastrate jeder 10. Wert	ca. 6 Minuten an einem PC

2.1.3 Kartengrundlage

Im Projekt Do-iT werden alle für Ortung, Identifikation und Map-Matching erforderlichen Kartendaten aus dem Lokalisierungsnetz (Do-iT Datenmodell V05) abgeleitet, das auf GDF-Daten der Version NavTeq 38 Q3/2005 basiert. Im Gegensatz zum Projektnetz, das nur die für die Weiterverarbeitung interessanten Straßen enthält, beinhaltet das Lokalisierungsnetz alle Straßen im Projektgebiet, so dass auch Routen abseits des Projektnetzes gefunden werden können.

Im Lokalisierungsnetz ist das Verkehrsnetz als kantenorientierter Graph dargestellt, wobei für jede Fahrtrichtung eine eigene Kante angegeben wird und der geometrische Verlauf einer Kante durch fortlaufend gespeicherte Formpunkte gegeben ist. Das Datenmodell des Verkehrsnetzes ist in Abbildung 2-12 dargestellt und beinhaltet weitere Attribute sowie Verknüpfungen. Zur Erläuterung der Bezeichnungen sei auf den Bericht [Schollmeyer 2007] verwiesen.

Da im Lokalisierungsnetz die Verkehrsnetze für schienen- und straßengebundene Verkehrsmittel gemeinsam gespeichert werden, ist zur Unterscheidung der Verkehrswege für jede Kante die Funktionsklasse gespeichert (vgl. Datenmodellierung der Abbildung LN_Form_Of_Way in [Schollmeyer 2007]). Beim Aufbau des Straßennetzes als knoten- und kantenorientierter Graph werden für den öffentlichen Verkehr nicht erlaubte Kanten, die nur für Fußgänger, Busse und LKW (Lieferverkehr) zulässig sind, nicht verwendet, um die Anzahl möglicher Straßensegmente zu reduzieren.

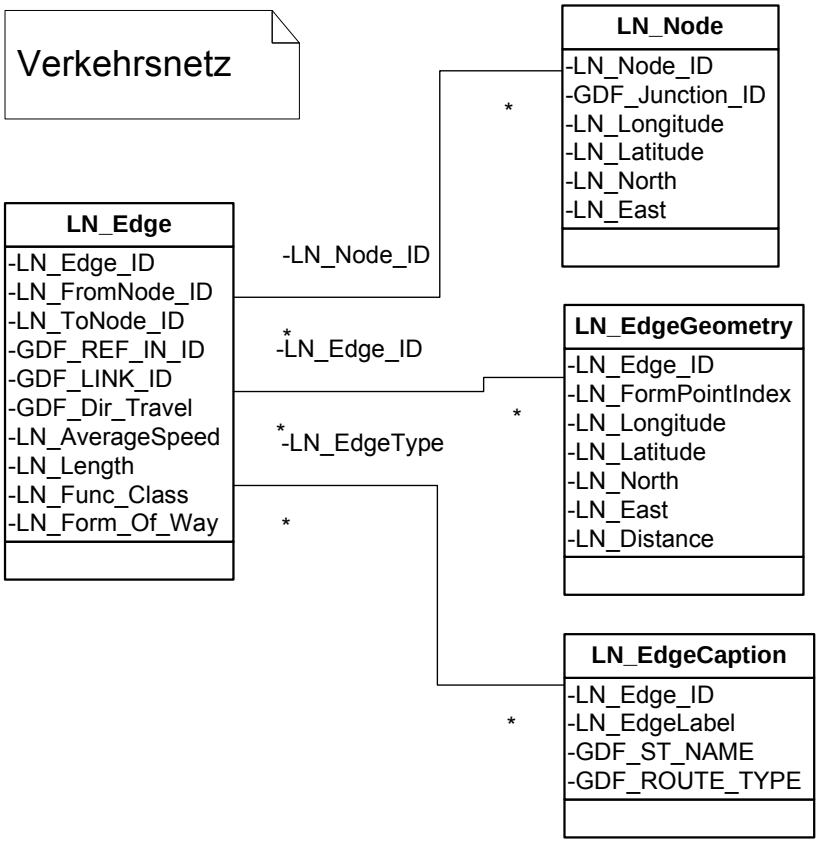


Abbildung 2-12: Datenmodellierung des Verkehrsnetzes im Lokalisierungsnetz

Um die für ein bestimmtes Gebiet relevanten Inhalte (Elemente) des Lokalisierungsnetzes schnell auffinden zu können, wurde ein regelmäßig angeordnetes Raster (Kachelgröße 100 m x 100 m) über das gesamte Projektgebiet gelegt und für jede Kachel die beteiligten Knoten, Kanten und POIs in weiteren Tabellen der Datenbank abgelegt. (vgl. Abbildung 2-13 und [Schollmeyer 2007])

Außerdem kann zur Reduktion der Antwortzeiten für die Auswertung von Abis-Daten das Lokalisierungsnetz auf die in einer MSC enthaltenen Location Areas begrenzt werden, z.B. mit dem umschließenden Rechteck plus einem empirisch festgelegten Sicherheitsrand.

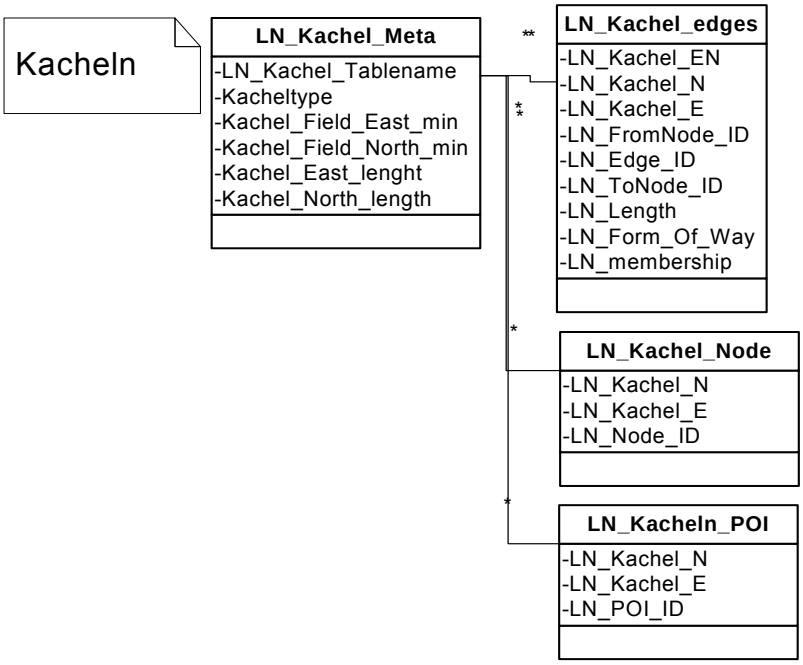


Abbildung 2-13: Datenmodellierung der Kachelung

Um für die gemessenen und damit fehlerbehafteten Positionen der Mobilstationen (MS) aus dem Lokalisierungsnetz wahrscheinliche Trajektorien extrahieren zu können, wird zunächst ein Suchgebiet festgelegt, in dem sich das Fahrzeug während der auszuwertenden Messzeit (z.B. während des Telefonats) mit Sicherheit befand. Die Größe des Suchgebiets wird entsprechend der Ortungsgenauigkeit (3σ) gewählt, wobei zur schnelleren Suche die Konfidenzellipse auf ihr umschließendes Rechteck erweitert wird.

Alle Knoten innerhalb des Suchgebiets werden aus dem Lokalisierungsnetz, das in einer MySQL- Datenbank gespeichert ist, herausgesucht und in einer doppelt verketteten Liste abgelegt. Alle für den öffentlichen Verkehr zulässigen Kanten, die von den Knoten abgehen, werden ebenfalls als doppelt verkettete Liste an ihrem Startknoten abgelegt. Verkettete Listen ermöglichen im Gegensatz zu starren Matrixdarstellungen mit vorgegebener Maximalgröße eine dynamische Speicherplatzverwaltung, die ein schnelles Einfügen und Löschen von Elementen erlaubt. Doppelt verkettete Listen können, indem jeweils die Speicheradresse des Vorgängerknotens und des Nachfolgeknotens gespeichert werden, in beiden Richtungen durchsucht werden. Vorgänger und Nachfolger beziehen sich dabei nur auf die Reihenfolge der Speicherung. Dieser knoten- und kantenorientierte Graph dient als Grundlage für die kartenbasierten Ortungsverfahren.

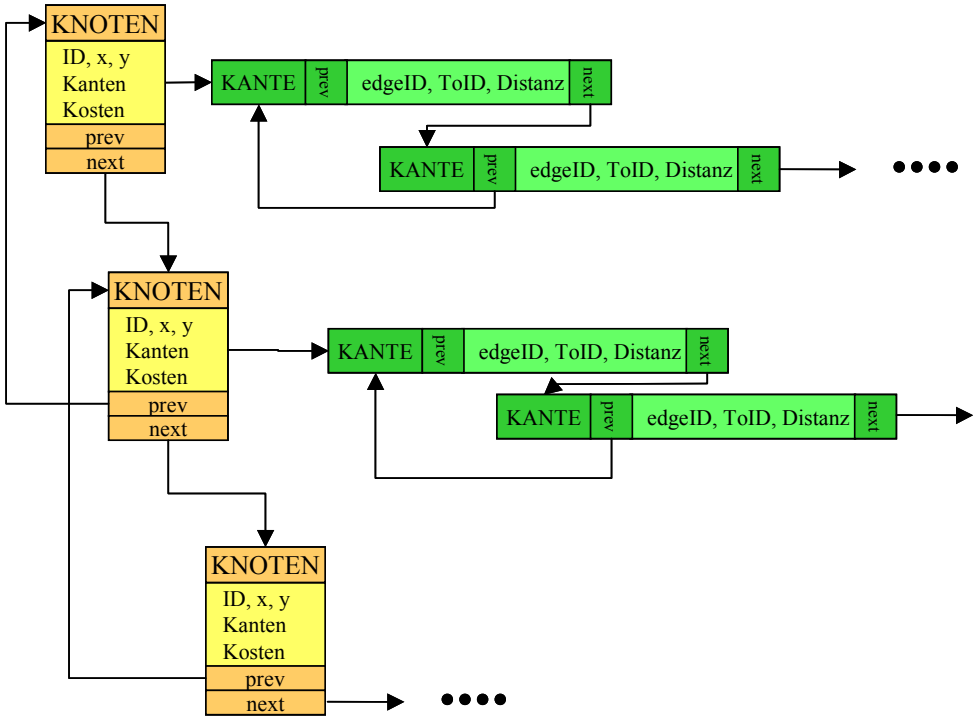


Abbildung 2-14: Speichermodell für Knoten und Kantenlisten

2.1.4 Routengenerierung

Zur Bestimmung der Verkehrslage sollen die Bewegungen der Mobilfunkteilnehmer auf dem Verkehrsnetz in ihrem geometrischen und zeitlichen Verlauf als Trajektorien dargestellt werden. Der geometrische Verlauf dieser Bewegung, der im Folgenden als Route bezeichnet wird, lässt sich als gerichtete Folge von Straßensegmenten beschreiben und bildet im Datenmodell des Lokalisierungsnetzes einen Kantenzug. Der zeitliche Verlauf der Bewegung des Mobilfunkteilnehmers wird in einem Weg-Zeit-Diagramm bereitgestellt, das den Zusammenhang zwischen der Zeit und dem auf der Route zurückgelegten Weg beschreibt.

Die Genauigkeiten für die in 2.1.2 beschriebenen Positionen werden bei reinem Signalstärken-Matching und bei der Glättung als achsenparallele Rechtecke abgeschätzt. Bei der Kalmanfilterung werden die Konfidenzellipsen berechnet, die zur Einsparung von Rechenzeit auf umschließende Rechtecke vereinfacht werden.

Die Positionen mit ihren Genauigkeitsangaben werden dazu benutzt, das Straßennetz auf einen Korridor einzugrenzen, in dem sich das Mobiltelefon bewegt haben kann. Theoretisch liegt bei angenommener Normalverteilung die wahre Position mit einer Wahrscheinlichkeit von 99,7 % innerhalb des Gebiets mit dreifacher Standardabweichung. Die ungünstige Schätzung der Positionsgenauigkeit im Signalstärken-

Matching mit und ohne Glättung, die grob mit einem mittleren Wert von 2000 m angegeben ist, kann vor allem im dichten Straßennetz als zu grob zur Festlegung der Suchfenster angesehen werden. Dagegen führen die zum Teil zu optimistisch geschätzten Genauigkeiten der Kalman-Filterung zu einem extrem schmalen Korridor, so dass der verbleibende Graph des Verkehrsnetzes nicht komplett befahrbar ist, bzw. offensichtliche Alternativtrassen durch den Korridor unterbrochen sind (vgl. Abbildung 2-15).

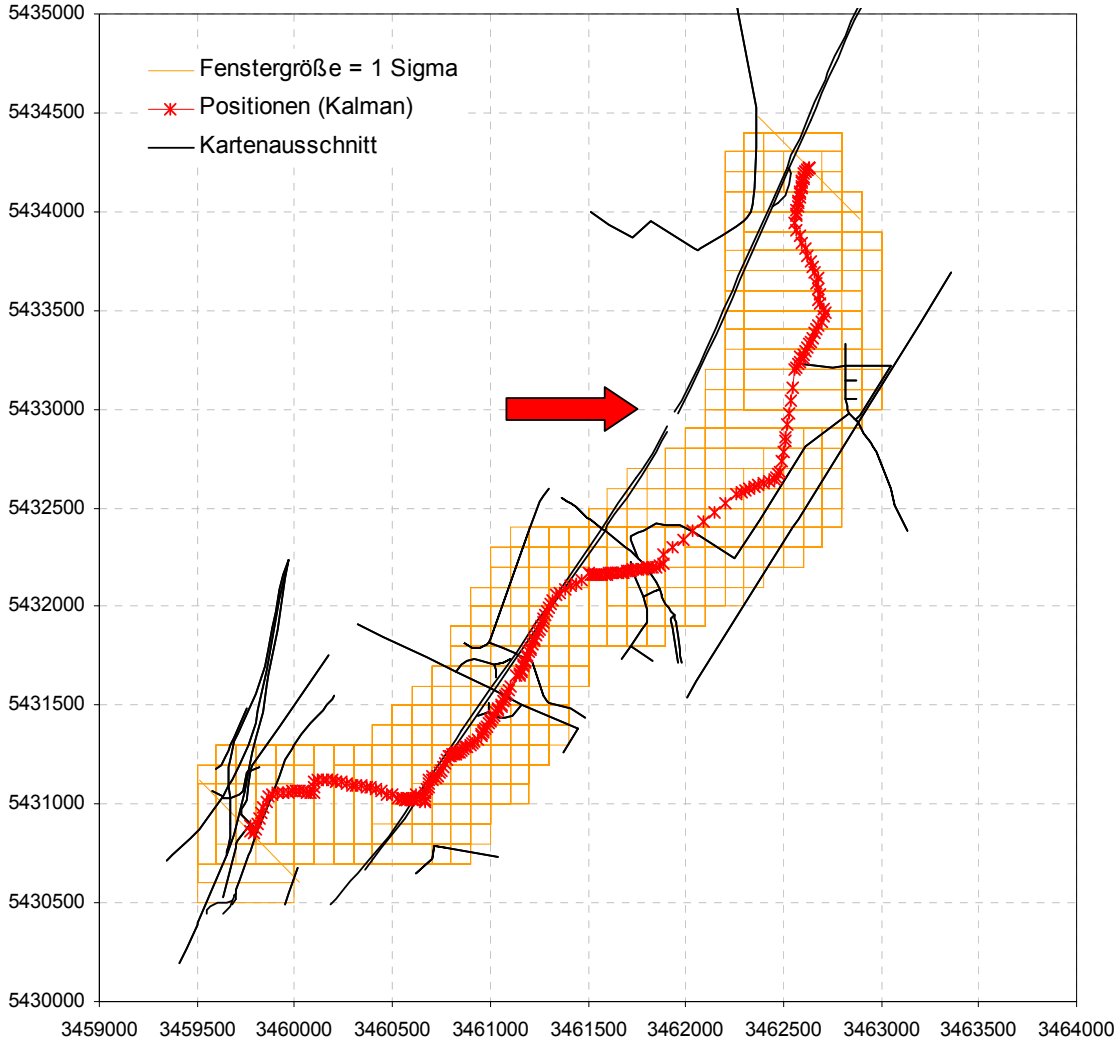


Abbildung 2-15: Suchfensterdefinition; Beispiel für ein zu klein gewähltes Suchfenster

Deshalb werden für Positionen mit zu groß oder zu klein geschätzten Genauigkeitsparametern die Suchfenster auf empirisch bestimmte Grenzwerte zwischen F_{min} und F_{max} beschränkt. Abbildung 2-16 zeigt die Positionen einer Testfahrt und ihre aus den Genauigkeiten berechneten Suchfenster, die auf $F_{min} = 300$ m und $F_{max} = 600$ m begrenzt wurden, für die Glättung in blau und für die Kalmanfilterung in orange.

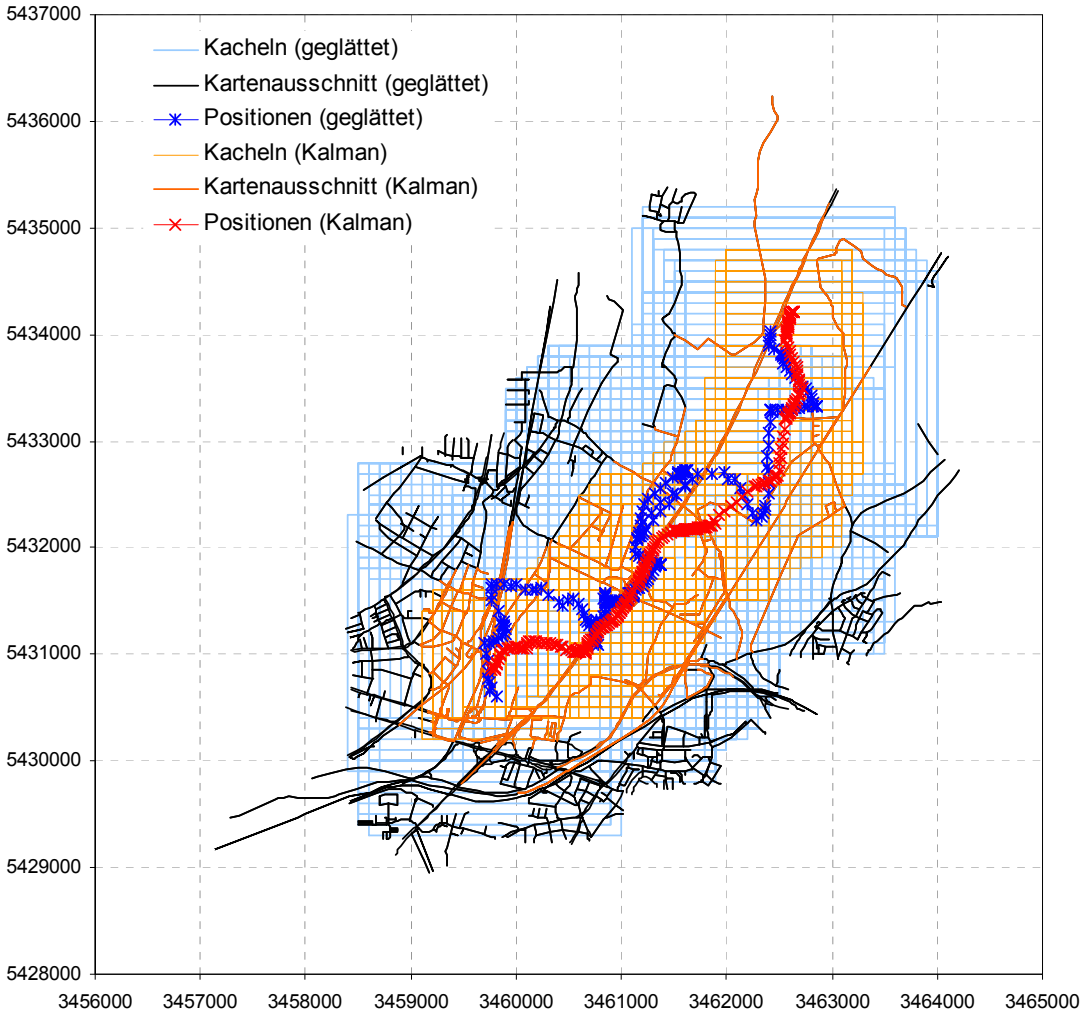


Abbildung 2-16: Suchgebiet in Abhängigkeit von der Genauigkeit

Bei Lücken in der Datenaufzeichnung werden aufeinander folgende Suchfenster, die sich nicht überlappen, auf das umschließende Rechteck erweitert, so dass sich ein geschlossenes Suchgebiet ergibt. Außerdem wird sichergestellt, dass das Straßennetz im Korridor zusammenhängend ist, indem das Suchfenster sukzessiv um 100 m vergrößert wird, solange keine Wege im Korridor gefunden werden.

Nur innerhalb des so definierten Korridors wird der Graph des Straßennetzes aufgebaut, und es sollen alle Wege bestimmt werden, auf welchen das Mobiltelefon bewegt werden konnte. Da alle Algorithmen zur Wegesuche jeweils einen Start- und einen Zielknoten benötigen, müssen dafür plausible Punkte ausgewählt werden. Dazu wird eine grobe Fahrtrichtung aus den ersten Positionen bestimmt und von den beiden Diagonalen des ersten Suchfensters diejenige ausgewählt, deren Normale mit der Fahrtrichtung den kleineren Winkel bildet. Die Anfangsknoten der Kanten, die diese Diagonale in gleicher Richtung wie die Fahrtrichtung schneiden, werden als Startknoten betrachtet. Analog können Endpunkte festgelegt werden, indem die Endknoten der

Kanten gewählt werden, die die Diagonale des letzten Suchfensters in Richtung der Fahrtrichtung schneiden.

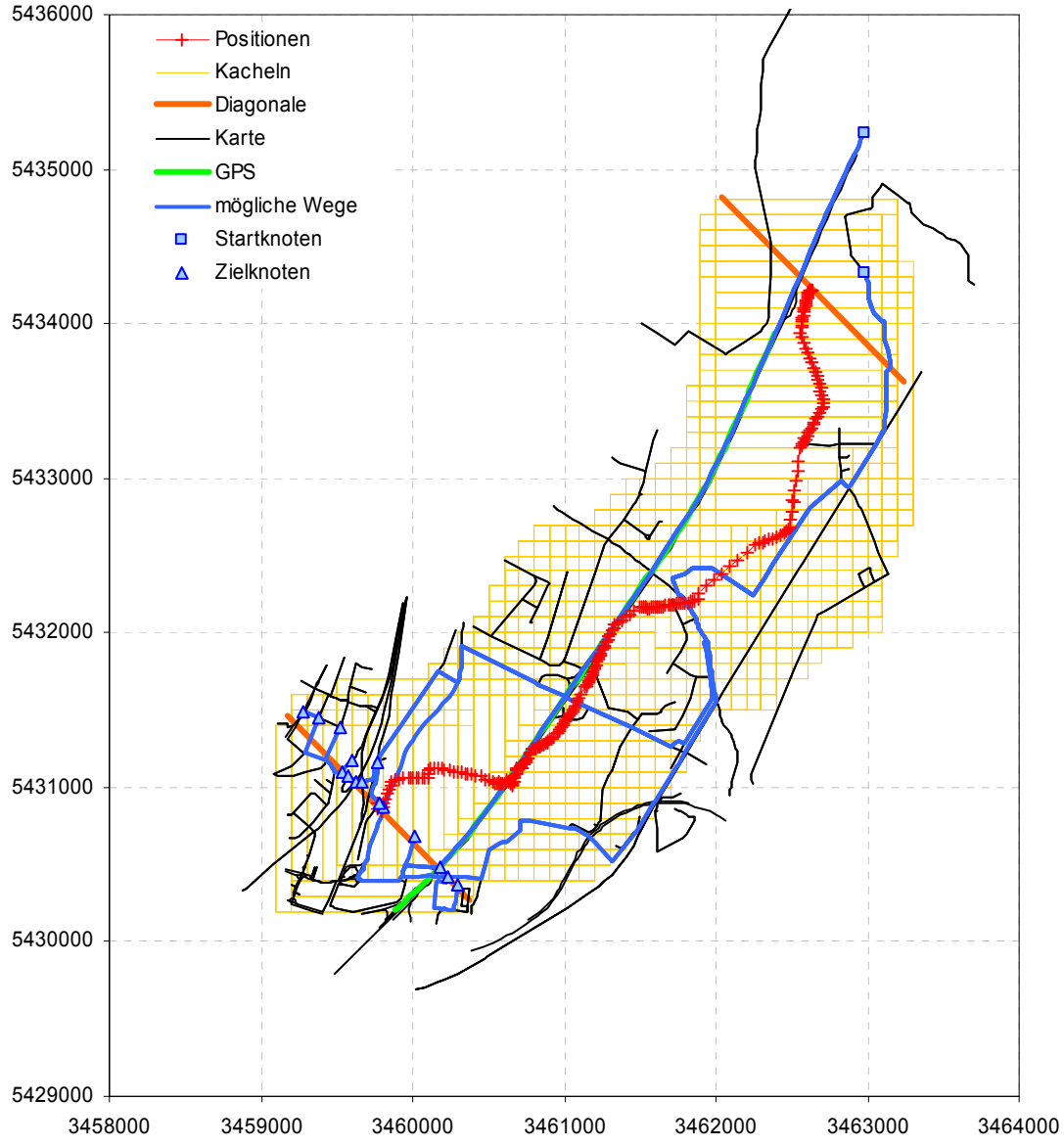


Abbildung 2-17: Wahl der Start- und Zielknoten

Die kürzesten Verbindungen von allen n_a Startknoten zu allen n_e Zielknoten, die mit dem Kürzeste-Wege-Algorithmus nach Dijkstra berechnet werden, ergeben $n_a \cdot n_e$ mögliche Routen, wenn alle Start- und Zielknoten über den Graphen zu erreichen sind.

Die Minimierung im Kürzeste-Wege-Algorithmus kann über den Fahrweg als Summe der Kantenlängen oder über die Fahrzeit mit der für jede Kante gespeicherten mittleren Geschwindigkeit erfolgen. In der Regel führen beide Varianten zu plausiblen und oft auch identischen Routen, aber in Einzelfällen kommt es vor, dass bei minimaler Stre-

cke die Routen durch die Gemeinden führen, bei minimaler Fahrzeit dagegen Routen über schneller zu befahrende Umgehungen gewählt werden.

Ein Beispiel dafür sind die Testfahrten auf der B3 zwischen Durlach und Grötzingen (vgl. Abbildung 2-18). Die etwa 400 m lange Strecke im Westen von Grötzingen, die in der Karte mit einer Geschwindigkeit von 31 km/h ausgewiesen ist, führt dazu, dass bei Minimierung der Fahrzeit alle Alternativen über die kurvigere, mit 71 km/h gespeicherte Teilstrecke führen.

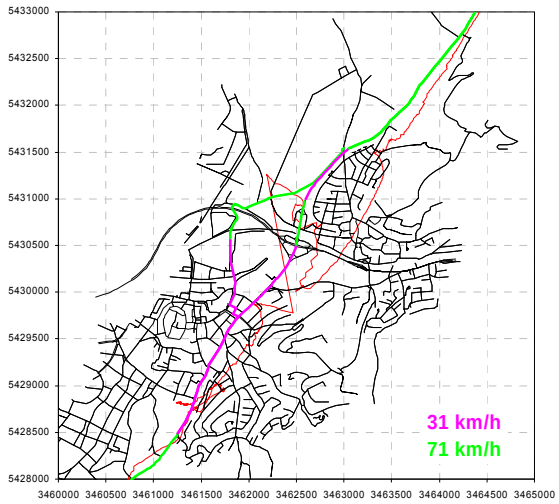


Abbildung 2-18: Vergleich kürzester und schnellster Weg

Abbildung 2-19: Umgehungsstraße od. Feldweg? (Quelle: Google Maps)

Da aufgrund der schlechten Positionsgenauigkeiten der Mobilfunkortung die Suchfenster nicht verkleinert werden können, werden sowohl kürzeste als auch schnellste Wege als Alternativrouten verwendet.

Diese Suche kürzester oder schnellster Wege in allen Kombinationen findet häufig Routen mit nur geringen Unterschieden, da manche Wege über andere Startknoten und / oder Zielknoten verlaufen. Diese fast identischen Lösungen werden zusammengefasst, indem alle Wege, die vollständig in einem anderen Weg enthalten sind, gelöscht werden.

Im Beispiel der Abbildung 2-17 ergeben sich für 2 Start- und 15 Zielknoten statt der theoretisch möglichen 30 Wege nur 24, da von einem Startknoten aus nicht alle Ziele auf dem Straßennetz innerhalb des Korridors zu erreichen sind. Da einige Wege über andere Zielknoten laufen, werden die längeren Wege verwendet, so dass sich die Anzahl der Wege auf 13 reduziert.

Um die Route zu bestimmen, auf der sich das Mobiltelefon mit größter Wahrscheinlichkeit bewegt hat, müssen die verbleibenden Möglichkeiten beurteilt werden. Da bei der Auswertung der Abis-Daten Positionsfolgen bestimmt werden, wird die Route gewählt,

die den geringsten Abstand zu den gemessenen Positionen hat. Dabei ist es nicht erheblich, ob es sich hierbei um einen kürzesten oder schnellsten Weg handelt.

Zur Bestimmung der Lotfußpunkte wird für jede gemessene Position P_i aus allen Knotenpunkten und den dazwischen liegenden Formpunkten der Trassenpunkt T_i mit der geringsten Entfernung gesucht. Von der gemessenen Position P_i wird das Lot auf die beiden zum Trassenpunkt T_i benachbarten Polygonseiten $(VT_i,\ T_iN)$ gefällt.

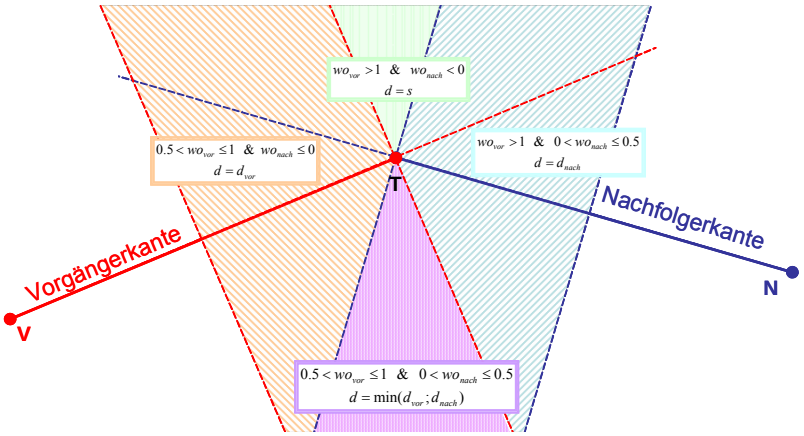


Abbildung 2-20: Berechnung der Lotfußpunkte

Welcher Lotfußpunkt der gemessenen Position zugeordnet wird, wird anhand der Streckenrelationen

$$wo_{vor} = \frac{s(V, L_v)}{s(V, T_i)}; \qquad wo_{nach} = \frac{s(T_i, L_n)}{s(T_i, N)} \quad \text{mit} \quad \begin{matrix} L_V & \text{Lotfußpunkt auf Vorgängerkante} \\ L_N & \text{Lotfußpunkt auf Nachfolgerkante} \end{matrix}$$

festgelegt (vgl. Abbildung 2-20). Werte mit $wo_{vor} < 0.5$ und $wo_{nach} > 0.5$ können nicht vorkommen, da sonst nicht der nächstgelegene Trassenpunkt T_i gewählt wurde. Fallen beide Lote auf die zu T_i benachbarten Kantenhälften $(0.5 < wo_{vor} < 1;\ 0 < wo_{nach} < 0.5;)$, wird der Lotfußpunkt mit kleinerem Abstand gewählt. Trifft ein Lot zwischen Anfangs- und Endpunkt der Kante, das andere aber außerhalb auf die Kante, wird der Lotfußpunkt gewählt, der die Kante trifft $(0.5 < wo_{vor} \le 1;\ wo_{nach} \le 0$ ⇨ Vorgängerkante; $wo_{vor} > 1 \ \& \ 0 < wo_{nach} \le 0.5$ ⇨ Nachfolgerkante). Wenn beide Lotfußpunkte außerhalb der Kanten liegen $(wo_{vor} > 1 \ \& \ wo_{nach} < 0)$ wird die Strecke zwischen Position P_i und nächstgelegenem Trassenpunkt T_i als Abstand definiert. Die so bestimmten Lotfußpunkte der Positionsfolge werden auch zur Festlegung des auf dem Kantenzug zurückgelegten Wegs verwendet, der zusammen mit den Messzeitpunkten Informationen über die Geschwindigkeit liefert (vgl. Kapitel Weg-Zeit-Diagramm).

Bei Positionsmessungen mit Genauigkeiten besser als die halbe Kantenlänge (wie z. B. GPS) lassen sich Abweichungen zu einer richtigen Trasse problemlos berechnen. Bei sehr schlechten Ortungsgenauigkeiten oder beim Vergleich mit grob falschen Tras-

sen, fällt eine gemessene Position jedoch nicht in die Bereiche innerhalb der halben Kantenlängen. Der Abstand wird dann definiert als Strecke zwischen gemessener Position und nächstgelegenem Trassenpunkt. Allerdings werden am Anfang alle Positionen P_i eliminiert, die vor die Anfangskante fallen (i<ia), und analog alle Punkte am Ende, die hinter die Endkante (i >ie) einer Route fallen.

Um zu beurteilen, wie gut eine Alternativtrasse zu der gemessenen Positionsfolge passt, wird aus den orthogonalen Abweichungen der verbleibenden Punkte eine Standardabweichung berechnet.

$$\sigma = \sqrt{\frac{\sum_{ia}^{ie} d^2}{ie - ia + 1}}$$

Bei der Auswahl der Route wird neben der Standardabweichung auch berücksichtigt, wie viele Positionen am Anfang und Ende abgeschnitten wurden, indem für jede Alternative der Quotient Q aus Standardabweichung und prozentualem Anteil der verwendeten Positionen berechnet wird. Die Route mit minimalem Q wird ausgewählt.

$$P = \frac{ie - ia + 1}{n} \qquad Q = \frac{\sigma}{P}$$

Zur Qualitätsbewertung der gewählten Route wird die Standardabweichung σ als „mittlere Querabweichung QA_{FPD}“ und der prozentuale Anteil P der verwendeten Positionen einer Positionsfolge als „Trajektorienvollständigkeit VO_{FPD}“ im Online-Betrieb zur Verfügung gestellt.

2.1.5 Weg-Zeit-Diagramme

Zur Generierung von Stauinformationen wird neben der befahrenen Route, die den geometrischen Verlauf in Form eines Kantenzugs angibt, auch der zeitliche Verlauf auf der Route benötigt. Dazu ist im Datenmodell „FPD-Trajectory“ vorgesehen, ein aus der gemessenen Positionsfolge zu bestimmendes Weg-Zeit-Diagramm in „FPD_PathTime“ abzulegen (vgl. [AP10-1 2008]).

Für jeden Messzeitpunkt wird von der gemessenen Position das Lot auf die nächstgelegene Kante der Route gefällt. Der Lotabstand dient zur Bestimmung der Querabweichung von der Route und dient zum Auswählen der FPD-Route. Außerdem geben die Lotfußpunkte an, wann und wo sich die Mobilstation auf der Route befunden hat, so dass für jeden Messzeitpunkt der zurückgelegte Weg berechnet werden kann. In den Weg-Zeit-Diagrammen bezieht sich der zurückgelegte Weg nicht auf den Gesprächsanfang sondern auf den ersten Knoten der Route.

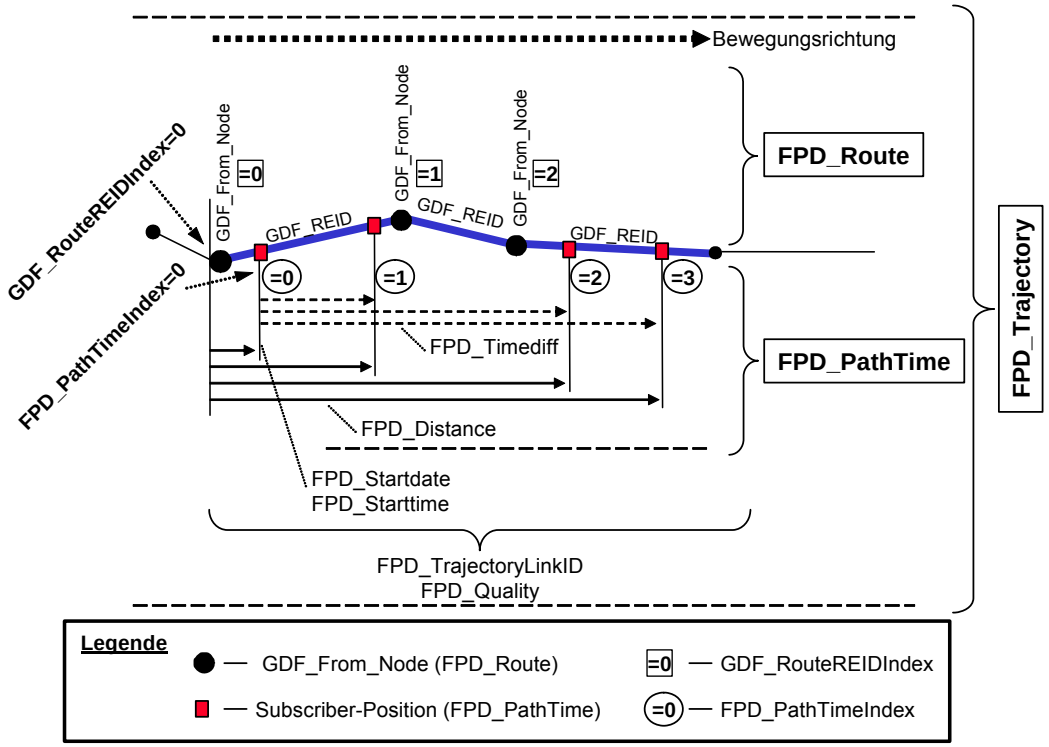


Abbildung 2-21: Datenmodell „FPD-Trajectory“

Liegen die gemessenen Positionen auf der Kurveninnenseite, fallen die Lote benachbarter Punkte auf verschiedene Kanten und die Abstände zwischen den Lotfußpunkten sind größer als die Abstände zwischen den Positionen. Dies führt zu einer Vergrößerung der daraus abgeleiteten Geschwindigkeiten. Liegen dagegen die gemessenen Positionen auf der Kurvenaußenseite werden die nicht eindeutig berechenbaren Lotfußpunkte auf den Knickpunkt (Knoten bzw. Formpunkt) der Route gelegt. Dies verursacht dort zu klein geschätzten Geschwindigkeiten.

Diese durch den Algorithmus bedingten Probleme sind aber gering im Vergleich zu den Einflüssen bei groben Positionsfehlern, die zu einem nicht monotonen Verlauf im Weg-Zeit-Diagramm führen und damit Rückwärtsfahrten vortäuschen können.

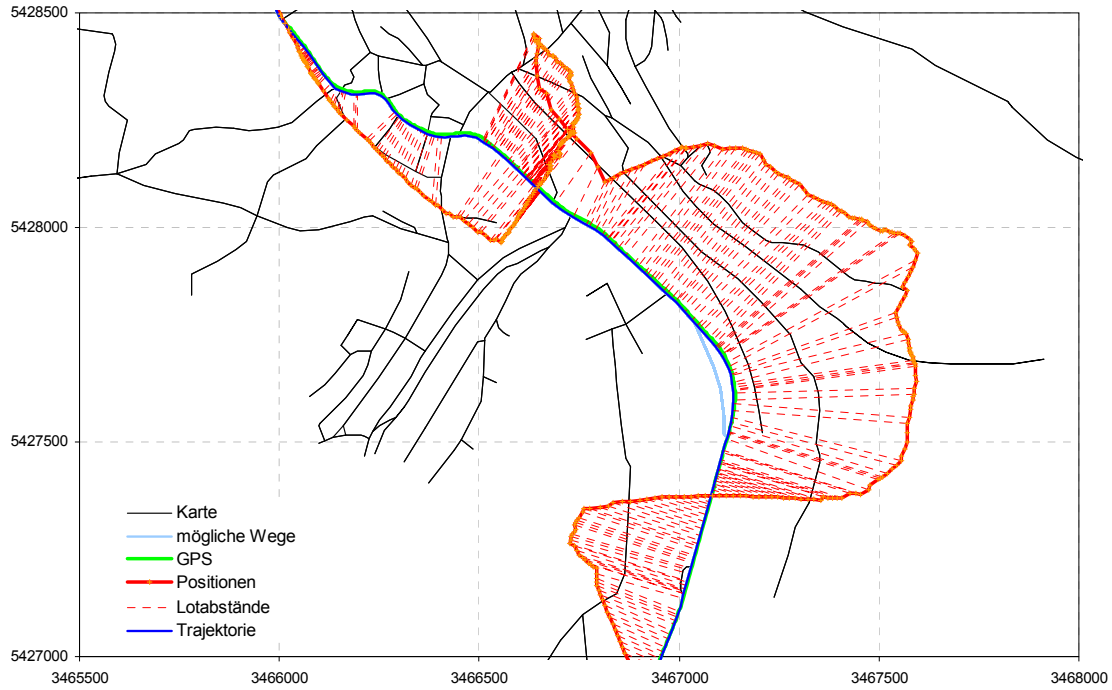


Abbildung 2-22: Lotfußpunkte

Um diese durch die Ungenauigkeiten der Positionsbestimmung verursachten Streckungen und Stauchungen im Weg-Zeit-Diagramm abzumildern, werden die mit ca. 2 Hz getasteten Strecken über ein 30 s Intervall gemittelt. Die Mittelwerte der Zeiten und Strecken ergeben dann ein ausgedünntes Weg-Zeit-Diagramm, wobei die Monotonie im Weg-Zeit-Diagramm gewährleistet wird, indem die Zeitpunkte, die zu negativen Streckeninkrementen führen, weggelassen werden.

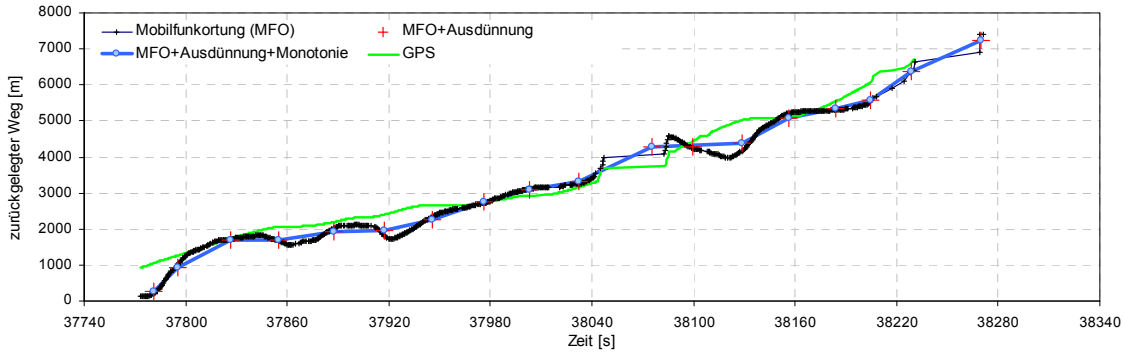


Abbildung 2-23: Weg-Zeit-Diagramm

Die wichtigsten Berechnungsschritte zur Generierung von FPD-Trajektorien inklusive der Weg-Zeit-Diagramme auf Abis-Ebene sind in Abbildung 2-24 noch einmal zusammengefasst.

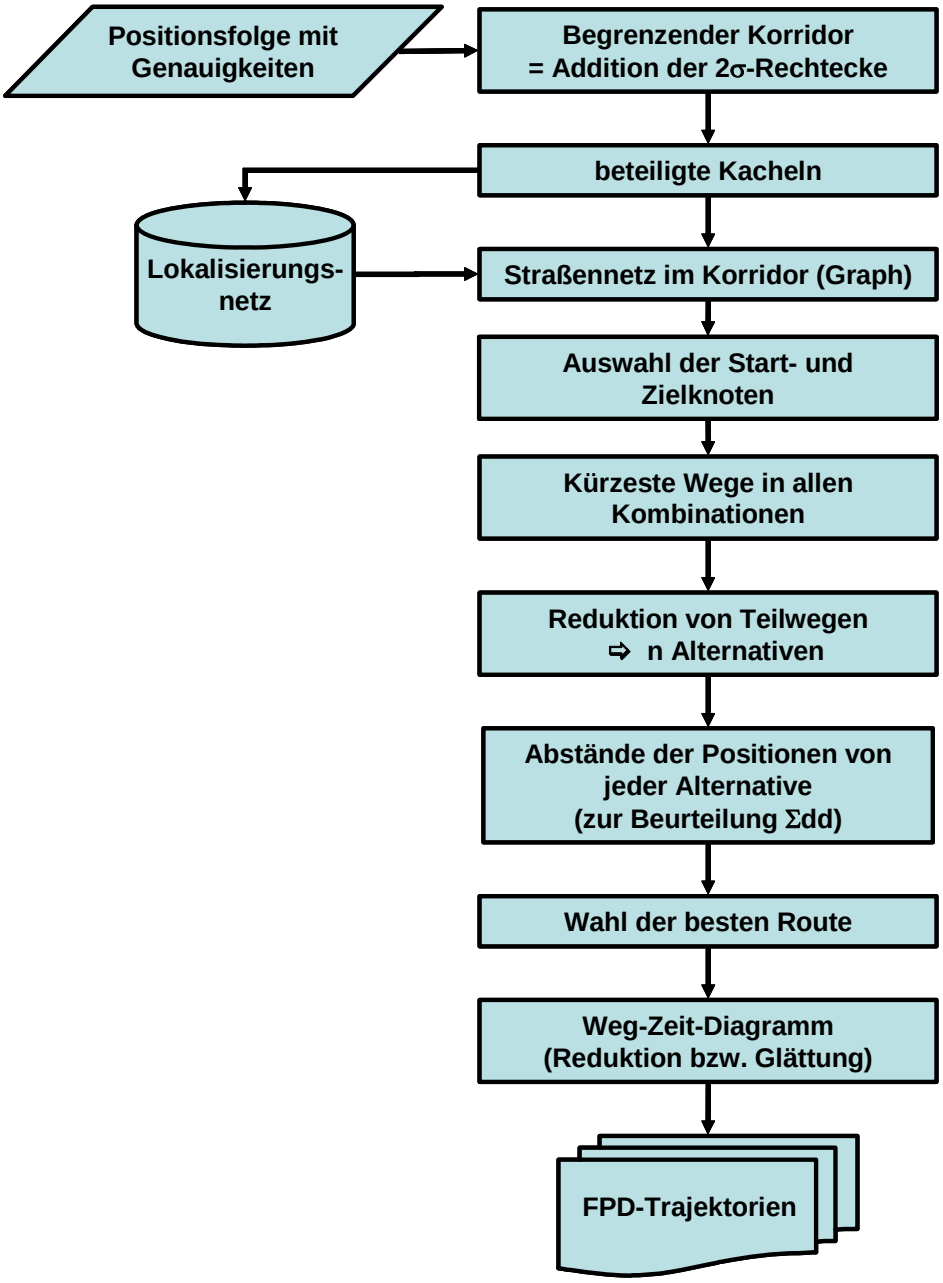


Abbildung 2-24: Schematischer Ablauf der Trajektoriengenerierung auf Abis-Ebene

2.2 A-Ebene

Die A-Daten werden durch so genannte Network Probes aus dem normalen Datenverkehr des Mobilfunknetzes über das bundesweite A-Monitoring-System der TMD und die Nexus-XML-Schnittstelle abgegriffen und stehen so für das gesamte Projektgebiet zur Verfügung.

Während eines Anrufs liegen Informationen über alle Wechsel der Funkzellen vor und es werden unabhängig vom Grund des Handovers und des Erfolgs (complete/failure) alle Ereignisse protokolliert, die die Ortungsinformationen Location Area Code (LAC), Cell Identity (CI) und Timing Advance Werte (TA) enthalten. Wenn nicht telefoniert wird, sind nur die alte LAC (ohne CI) und die neue LAC mit CI bekannt.

2.2.1 Priorisierung

Wie bei den Abis-Daten ist eine Priorisierung der eingehenden Mobilfunkdaten auf A-Ebene notwendig. Dies ergibt sich aus der Tatsache, dass zwar weniger Berechnungsschritte aber dafür ein deutlich größeres Gebiet abgedeckt werden muss und daher ein großes Aufkommen an Mobilfunkdaten zu verarbeiten ist. Wie bei den Abis-Daten ist es auf Grund der Rechnerperformance nicht möglich alle A-Daten in Echtzeit zu berechnen. Deshalb bietet sich auch hier an, Trajektorien besonders interessanter Mobilfunkteilnehmer zuerst zu berechnen.

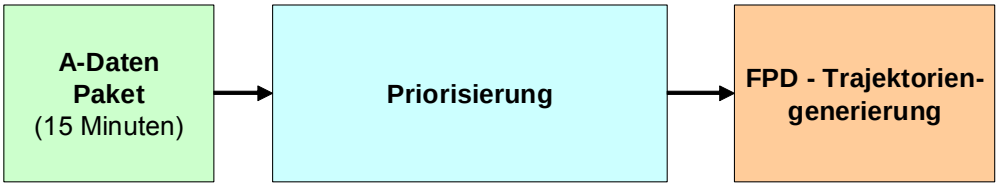


Abbildung 2-25: Grobablauf FPD-Generierung aus A-Daten

Wie bei den Abis-Daten werden im Echtzeitbetrieb immer 15-Minuten-Blöcke gebildet. Das bedeutet, dass 15 Minuten lang Daten gesammelt werden, die anschließend priorisiert und dann FPD-Trajektorien generiert werden (vgl. Abbildung 2-25). Daraus folgt ein Offset zur Echtzeit von maximal 30 Minuten.

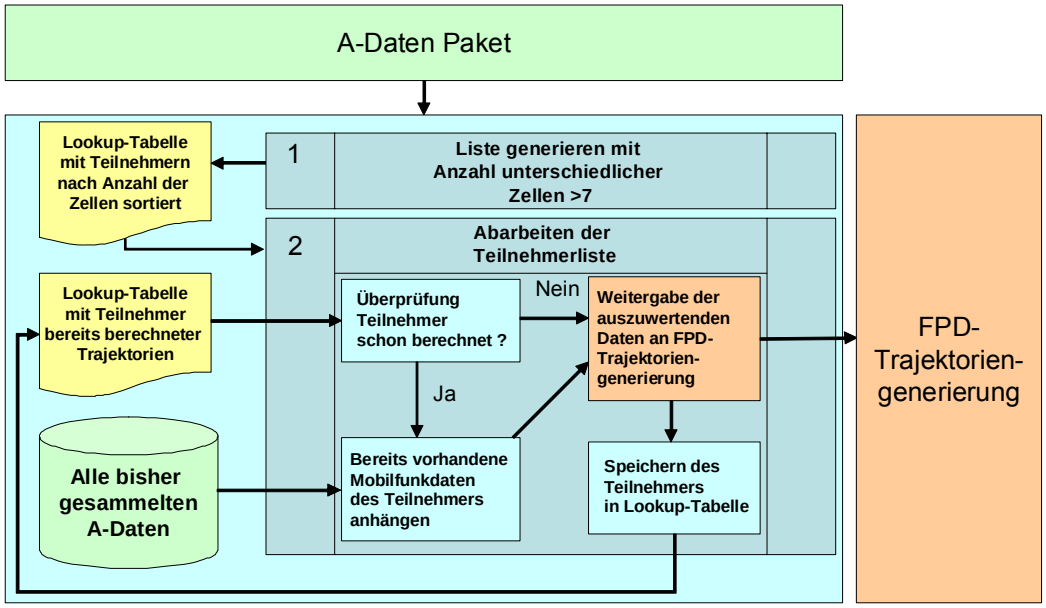


Abbildung 2-26: Schematische Darstellung der Priorisierung

Die eigentliche Priorisierung ist auf Grund der kurzen Entwicklungszeit sehr einfach gehalten. Abbildung 2-26 zeigt die schematische Darstellung der A-Daten-Priorisierung. Im ersten Schritt wird der neue A-Daten-Block nach Anzahl der benutzten Mobilfunkzellen der einzelnen Teilnehmer sortiert. Dabei werden im Echtzeitbetrieb alle Teilnehmer sortiert, die eine höhere Anzahl als 7 Zellen während eines Gespräches oder insgesamt mehr als 7 Location Areas durchfahren haben. Es hat sich empirisch gezeigt, dass Teilnehmer sich erst ab 7 Zellen eindeutig bewegen. Bei weniger als 7 Zellwechseln kann es sein, dass bei einer ungünstigen Position des Teilnehmers die Zellwechsel nur durch netzinterne Weiterleitung zu begründen sind. Dies ergibt sich aus der Tatsache, dass neben der bedienenden Zelle 6 weitere Nachbarzellen empfangen werden und in alle Zellen netzseitig gewechselt werden kann. Bei Anzahl der Location Areas ist eine eindeutige Bewegung schon ab 4Location Area Wechsel zu unterstellen. Dies hätte die Anzahl an auswertenden Teilnehmern jedoch verzehnfacht, was zum jetzigen Zeitpunkt noch nicht verarbeitbar gewesen wäre. In der Offline-Auswertung werden diese Location Area Wechsel aber berücksichtigt, worauf weiter unten noch einmal eingegangen wird. Im nächsten Schritt findet eine Überprüfung statt, ob priorisierte Mobilfunkteilnehmer schon ein- oder mehrmals berechnet wurden. Die bereits berechneten Teilnehmer finden sich in einer Lookup-Tabelle. Falls für einen Teilnehmer schon Trajektorien berechnet wurden, werden die bisher gesammelten Mobilfunkdaten mit den Neuen vereint und an die FPD-Trajektoriengenerierung weitergeschickt. Falls für einen Mobilfunkteilnehmer noch keine Trajektorien berechnet wurden, werden nur die aktuellen Mobilfunkdaten verarbeitet. Abschließend wird die neue Berechnung in die Lookup-Tabelle von bereits berechneten Trajektorien aufgenommen.

Bei der Offline-Auswertung können alle Daten verarbeitet werden, wodurch die Priorisierung etwas abgewandelt werden kann. Es werden alle Daten eines Tages sortiert und Mobilfunkdaten zur Auswertung weitergegeben, die nicht nur mehr als 7 Zellen während eines Gespräches durchfahren haben, sondern auch mehr als 3 Location Areas ohne Telefonat. Zur Verdeutlichung für die gewählte Anzahl der Location Area Wechsel soll die Abbildung 2-27 dienen. Sie zeigt alle Location Areas, die im Projektgebiet vorkommen und entstand auf Grundlage des Best-Server-Plots. Die roten Kreise zeigen die Stellen, an denen sich 3 Location Areas berühren. Das bedeutet, dass es an diesen Stellen theoretisch möglich ist, dass Teilnehmer Location Area Wechsel über drei Location Areas durchführen, ohne dass sie sich bewegen. Aus diesem Grund wurden nur Teilnehmer in die Auswertung mit einbezogen, die mindestens 4 Location Area Wechsel durchgeführt haben.

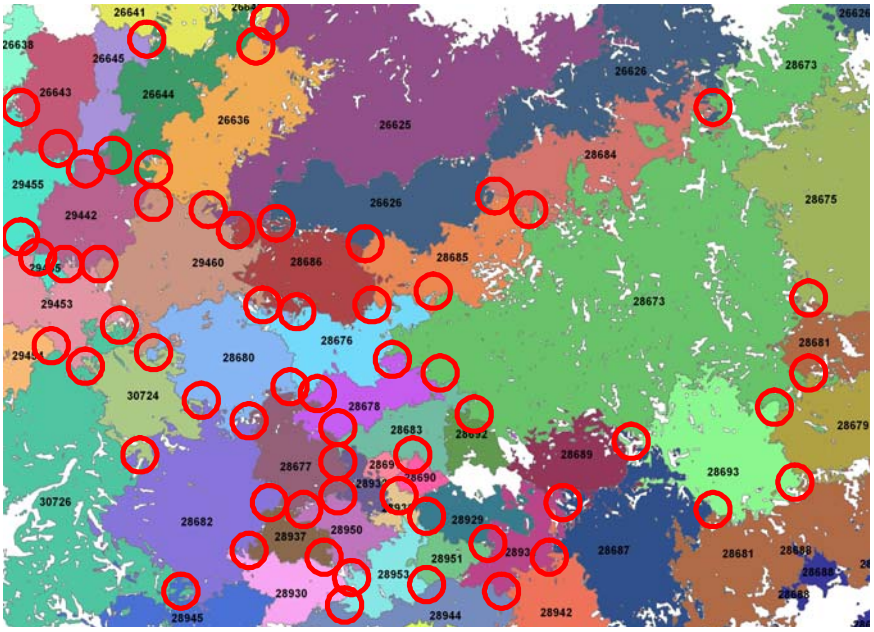


Abbildung 2-27: Location Areas im Projektgebiet

2.2.2 Kartengrundlage

Auch für die Auswertung der A-Daten bildet das aus NavTec Daten (38 Q3 /2005) abgeleitete Lokalisierungsnetz die Kartengrundlage. Im Gegensatz zum Projektnetz, das nur die für die Weiterverarbeitung interessanten Straßen enthält, beinhaltet das Lokalisierungsnetz alle Straßen im Projektgebiet, so dass auch Routen abseits des Projektnetzes gefunden werden können.

Allerdings führt die Verwendung aller Straßen (über 1 Mio. Kanten, vgl. Tabelle unten) zu langen Such- und Antwortzeiten. Da eine Gebietseinschränkung wie bei den Abis-Daten hier nicht möglich ist, kann die Rechenzeit nur verbessert werden, indem der Datenumfang durch eine Beschränkung auf die Durchgangsstraßen reduziert wird.

Anhand der im NavTec GDF definierten „Functional Road Class“ [1] erscheint ein Reduzieren auf die Klassen FC 1 bis FC 3 sinnvoll, so dass nur noch 12,5 % der Kanten im Lokalisierungsnetz zu berücksichtigen sind.

Nach den Testfahrten in Stuttgart mit ortskundigen Fahrern zeigte sich in seltenen Fällen, dass stark frequentierte Straßen im GDF nur mit FC 4 bewertet sind. Um zu niedrig bewertete Straßen nicht zu vernachlässigen und diese Testfahrten auswerten zu können, wird zur Bearbeitung der A-Daten das Lokalisierungsnetz nur um die Straßen mit FC 5 reduziert. Anhand der Tabelle 2-3 unten kann die Reduktion der Datenmengen abgeschätzt werden.

Tabelle 2-3: Reduktion der Datenmenge durch Ausschluss einzelner FC

	Gesamtes Netz	Function-Class<=4		Function-Class<=3	
Knoten	416891	101445	24,3%	68234	16,3%
Kanten	1018558	195389	19,2%	126002	12,4%
Geometrie	4418019	1030356	23,3%	657192	14,9%
Kachelung	2964125	271147	9,1%	176384	6,0%

Abbildung 2-28 zeigt die dadurch verbleibenden Straßen in Rot im Vergleich zu allen Straßen (grau) für das Do-iT Projektgebiet.

[1] http://www.santa-eufemia.it/cms/data/pages/files/000033_resource2_orig.pdf:
Functional Class 1 (FC 1) roads allow for high volume, maximum speed traffic movement between and through major metropolitan areas. Access to the road is usually controlled. FC 1 roads comprise approximately 2% of total road network.
Functional Class 2 (FC 2) roads are used to channel traffic FC 1 roads for travel between and through cities in the shortest amount of time. FC 2 roads comprise approximately 2- 3% of total road network.
Functional Class 3 (FC 3) is applied to roads that interconnect FC 2 roads and provide a high volume of traffic movement at a lower level of mobility than FC 2 roads. FC 3 roads comprise approximately 5% of total road network.
Functional Class 4 (FC 4) is applied to roads that provide for a high volume of traffic movement at moderate speeds between neighborhoods. These roads connect with higher functional class roads to collect and distribute traffic between neighborhoods. FC 4 roads comprise approximately 12-15% of total road network.
Functional Class 5 (FC 5) is applied to roads whose volume and traffic movement are below the level of any functional class. FC 5 roads comprise approximately 77-80% of total road network.

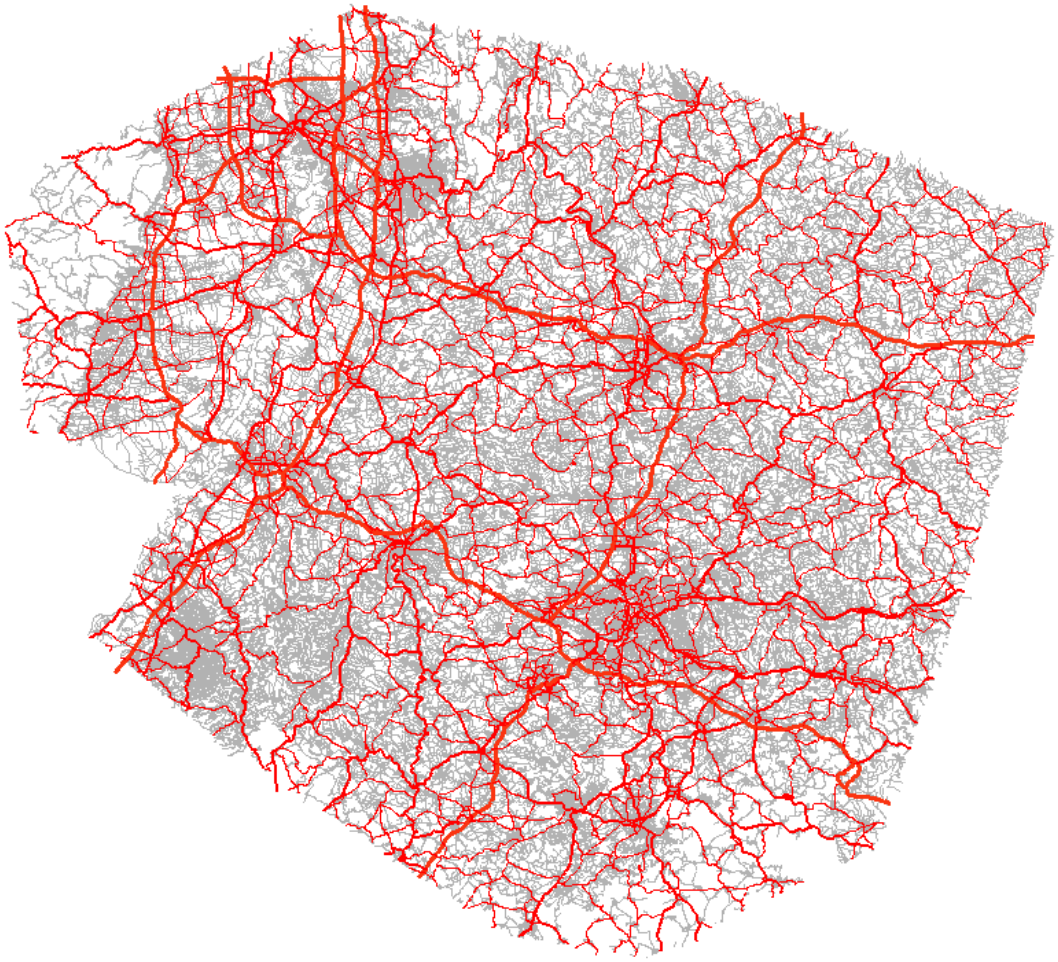


Abbildung 2-28: Lokalisierungsnetz im Projektgebiet

2.2.3 Mobilfunknetz

Informationen zum Mobilfunknetz liegen im Projekt Do-iT in Form so genannter Best-Server-Plots (BSP) vor, die von T-Mobile aus theoretisch bestimmten Signalstärken und einem digitalen Geländemodell für eigene Anwendungen berechnet wurden und die für jeden Ort (Auflösung von 32 m x 32 m) angeben, welche Antenne mit der besten Qualität zu empfangen ist. Abbildung 2-29 zeigt den BSP für einen Innenstadtbereich, wobei die Antennen durch die Farben und die Angabe der CI zu unterscheiden sind. Durch Neu- und Umbau von Antennen unterliegt der BSP ständigen Veränderungen, die durch die monatliche Aktualisierung von T-Mobile berücksichtigt werden können.

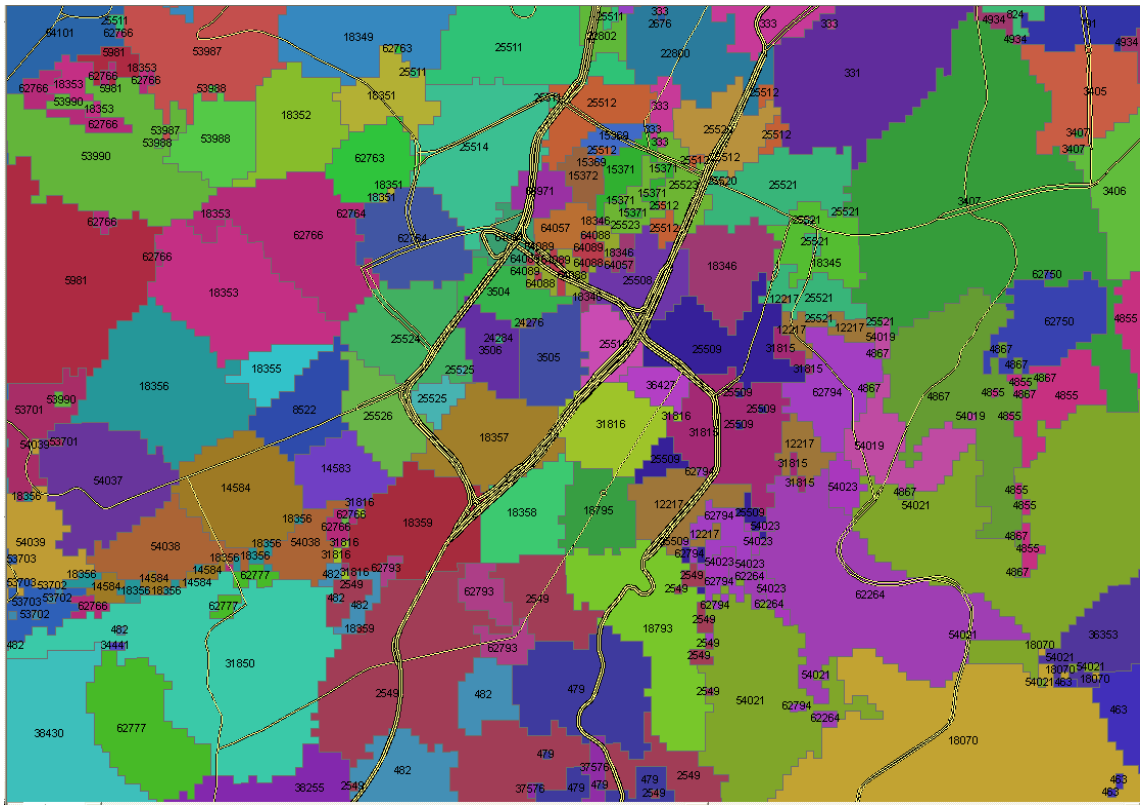


Abbildung 2-29: Original Best Server Plot (Ausschnitt)

Aufgrund von Topographie und Abschattungen an Gebäuden werden die Mobilfunkzellen im Best Server Plot nicht als einheitliche Flächen dargestellt, sondern sind in mehrere Teilflächen unterteilt und enthalten häufig Fremdflächen. Da Mobiltelefone beim Durchqueren kleiner Teilflächen in der Regel nicht auf die zugehörige Antenne umgeschaltet werden, muss der Best Server Plot generalisiert werden. Dies erfolgt in ArcGIS mit folgenden Berechnungsschritten:

- Fläche jedes Umringpolygons
- Gesamtfläche einer Mobilfunkzelle als Summe der zugehörigen Teilflächen
- Flächen kleiner 5000 m^2 (<6 Pixel) oder kleiner 10 % der Gesamtfläche einer Mobilfunkzelle werden der Nachbarfläche zugeordnet, die die längste gemeinsame Grenze besitzt. Dieser Schritt muss aufgrund hoher Fremdpixeldichte mehrmals durchgeführt werden.
- Kleine Löcher im Best Server Plot (= Gebiete die nicht von Antennen im Projektgebiet versorgt werden) werden in gleicher Weise den Nachbarzellen zugeordnet.

Diese Generalisierung reduziert die Anzahl der Umringpolygone um ca. 80 % und ist in nachfolgender Abbildung für den gleichen Ausschnitt wie oben dargestellt.

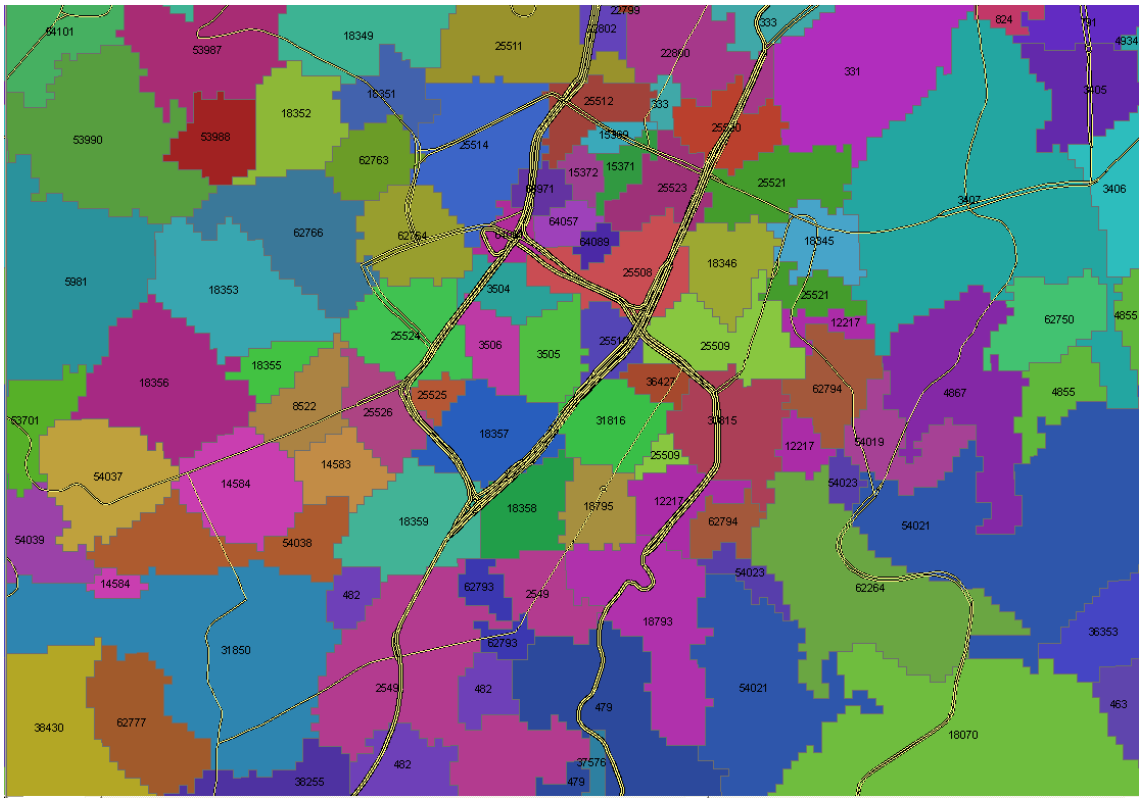
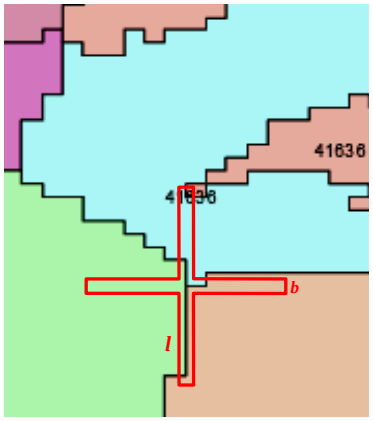


Abbildung 2-30: Generalisierter Best Server Plot

Um aus diesen Umringpolygonen einen Graphen des Mobilfunknetzes mit korrekten Nachbarschaftsbeziehungen herzuleiten, wurden folgende Schritte durchgeführt:

- Die Umringpolygone werden in einer Datenbank abgelegt, wobei die Koordinaten der Punkte auf ganze Meter gerundet werden. Zur Vergabe eindeutiger Punktnummern wird eine weitere Tabelle erzeugt, die die Polygonpunkte zuerst nach Rechtswert, dann nach Hochwert geordnet enthält. In dieser Ordnung werden eindeutige Punktnummern vergeben.
- Zugehörigkeit
  Für jeden Punkt wird bestimmt, zu welchen Polygonen er gehört. Da die Polygonseiten (bis auf Digitalisierungsfehler) achsenparallel verlaufen kann ein Punkt zu maximal 4 Polygonen gehören.
- Topologie
  Da die Umringpolygone nur an Knicken Punkte enthalten, müssen zum Aufbau der Topologie dort zusätzliche Polygonpunkte eingeführt werden, wo Nachbarpolygone abzweigen. Deshalb muss für jeden Punkt geprüft werden, ob er auf einer weiteren Polygonseite liegt. Um in dieser Doppelschleife über alle Punkte und alle Polygonseiten Rechenzeit einzusparen, werden nur die Polygonseiten verwendet, die innerhalb eines Fadenkreuzes mit Länge l (= 2km ≅ 2* Maximum der Polygonseiten) und Breite b (= 60m ≅ 2* Pixelgröße) liegen.



- Punktreduktion
  Punkte, die zu 3 oder mehr Umringpolygonen (= Flächen) gehören, entsprechen in der Graphentheorie einem Knoten und sollen zur Erhaltung der Topologie bei einer Punktreduktion erhalten bleiben. Eine Folge von Punkten, die alle zu den gleichen zwei Polygonen gehören, entspricht einer Kante, deren geometrischer Verlauf durch Formpunkte gegeben ist. Um die Anzahl dieser Formpunkte und damit den Speicherplatz und die Antwortzeiten zu reduzieren, werden diese Formpunkte einer Kante nach dem Verfahren von Douglas-Peucker ausgedünnt. Das Ergebnis dieser Punktreduktion hängt davon ab, ob bei der Berechnung die Punktfolge vorwärts oder rückwärts durchlaufen wird. Um eindeutige Ergebnisse zu erhalten, wird die Punktreduktion auf Kanten benachbarter Zellen immer im Polygon mit kleinerem Zählindex durchgeführt.

Der Best Server Plot mit über 2 Millionen Punkten in über 9000 Polygonen kann bei einer Punktreduktion nach Douglas-Peucker mit einem 75 m breiten Toleranzband als Graph mit 128000 Knoten und 135000 Kanten dargestellt werden.

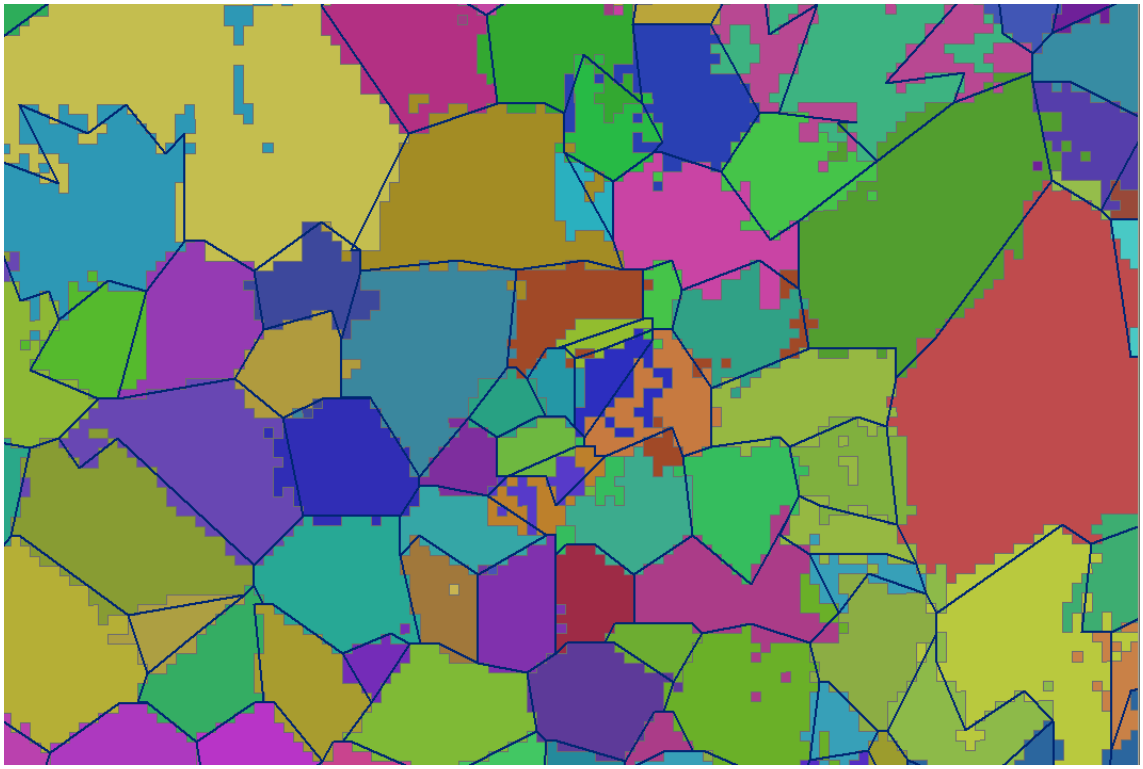


Abbildung 2-31: Zellgrenzen aus Best Server Plot

2.2.4 Routengenerierung

Im GSM-Netz (GSM = Global System for Mobile Communications) sind jeweils mehrere Funkzellen zu einer Location Area zusammengefasst. Bei einem eingeschalteten Mobilfunkgerät (Standby-Mode) wird netzseitig im Wesentlichen nur das Wechseln einer Location Area über einen so genannte Location Area Update (LUP) erfasst. Erst bei einem Verbindungsaufbau wird die aktive Funkzelle (Serving Cell) ermittelt, über die dann die weitere Kommunikation abläuft. Wechselt das Mobilfunkgerät die Funkzelle während einer aktiven Funkverbindung wird dies über ein Handover (HO) erfasst, bei

dem Informationen über den Zeitpunkt des Zellwechsels sowie die beiden Cell Identities vorliegen.

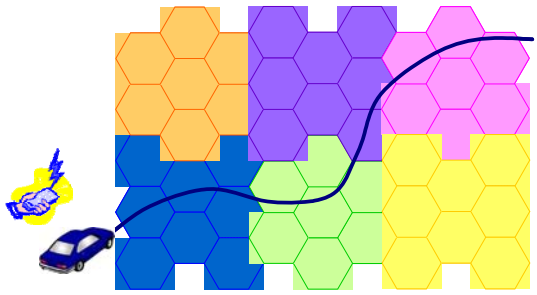


Abbildung 2-32: Handover

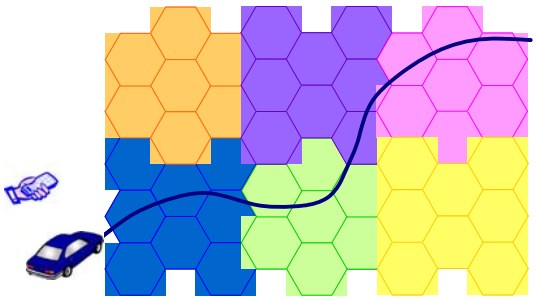


Abbildung 2-33: Location Area Update

Unter der Annahme, dass ein Mobilfunkgerät immer mit der Antenne verbunden ist, die den besten Empfang liefert, lassen sich aus der Zellinformation und dem BSP Ortsinformationen ableiten. Bei einem Ereignis (HO oder LUP) müsste sich das Mobiltelefon auf einem der Schnittpunkte zwischen dem Verkehrsnetz und den Grenzen der beteiligten Mobilfunkzellen befinden. Dazu werden aus dem generalisierten Best Server Plot die Zellgrenzen zwischen allen an einem Ereignis beteiligten Zellen wie folgt festgelegt:

a. Bei einem HO sind beide CI bekannt, so dass die Zellgrenze aus allen Kanten besteht, die zu beiden Zellen gehören.

b. Findet man im Fall a. keine gemeinsamen Kanten, so sind alte und neue Zelle nicht direkt benachbart. Die Grenze für dieses Ereignis ergibt sich aus allen Kanten der neuen Zelle.

c. Bei einem LUP sind die alte LAC und die neue Zelle bekannt. Die Grenze ergibt sich aus allen Kanten, die zur alten LAC und zur neuen CI gehören.

d. Findet man im Fall c. keine gemeinsamen Kanten, so liegt die neue Zelle nicht am Rand zur alten LAC. Die Grenze für dieses Ereignis ergibt sich aus allen Kanten der neuen Zelle.

Anhand realer Daten zeigte sich, dass auch unlogische Zellwechsel vorkommen können, die eliminiert werden:

- Zellen, die nicht im Best-Server-Plot enthalten sind (z.B. Funkzellen fremder Anbieter)
- Umschalten in die gleiche Zelle (z.B. neuer Gesprächsanfang in der gleichen Zelle)
- mehrfaches Hin- und Herschalten zwischen zwei Zellen.

Zum Zeitpunkt des Umschaltens befinden sich als bewegt identifizierte Mobilstationen

auf den Zellgrenzen bzw. bei Umschaltverzögerungen in der Nähe der Zellgrenzen und auf dem Straßennetz, also auf den Schnittpunkten zwischen den aus dem BSP abgeleiteten Zellgrenzen und dem Straßennetz. Da die Grenzen des Mobilfunknetzes aus dem BSP nicht fehlerfrei bestimmt werden können und die Umschaltung von einer Zelle zur Nächsten zeitlich verzögert sein kann, werden auch die Punkte außerhalb der Zellgrenze berücksichtigt, wenn die maximal 1,5 km lange Verlängerung der Zellgrenze mit dem Straßennetz geschnitten wird. Durch die vorherige Bestimmung aller Schnittpunkte und deren Abspeicherung konnte die Rechenzeit der Online-Algorithmen erheblich reduziert werden. Zu einem Umschaltzeitpunkt können sich mehrere Schnittpunkte ergeben, die Position ist also nicht eindeutig.

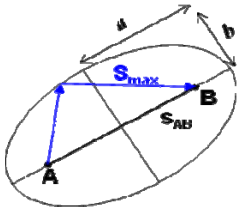
Im Zeitraum zwischen zwei Umschaltungen kann die Mobilstation höchstens einen Weg s_{max} zurückgelegt haben, der sich aus der höchsten in diesem Gebiet erlaubten und im Lokalisierungsnetz für jede Kante gespeicherten Geschwindigkeit v_{max} abschätzen lässt. Um sicherzustellen, dass dieser Weg nicht zu kurz geschätzt wird, wird er um einen empirisch ermittelten Maßstabsfaktor (m=1.1) und einen gleichfalls aus dem Datenmaterial abgeleiteten Zuschlag für eventuelle Umschaltverzögerungen (c=500m) vergrößert.

$$s_{\max} = \Delta t \cdot v_{\max} \cdot m + c$$

Wenn sich die MS zum Zeitpunkt des Zellwechsels auf dem Schnittpunkt A zwischen Best Server Plot und Straßennetz befunden hat und zum Zeitpunkt des nächsten Zellwechsels auf dem Schnittpunkt B und in dem Zeitraum dazwischen höchstens den Weg s_{max} zurücklegen konnte, kann die MS sich nur innerhalb einer Ellipse bewegt haben, deren Brennpunkte mit den Schnittpunkten A und B zusammenfallen und deren Halbachsen sich aus dem Abstand s_{AB} zwischen den Brennpunkten und dem maximal zurückgelegten Weg berechnen lassen.

Große Halbachse $a = 0.5 \cdot s_{\max};$

Kleine Halbachse $b = \frac{1}{2}\sqrt{s_{\max}^2 - s_{AB}^2};$



Wenn die beiden Schnittpunkte A und B den richtigen MS-Positionen entsprächen, müsste das Straßennetz nur innerhalb dieser Ellipse aufgebaut werden. Da sich die Grenzkanten eines Zellwechsels aber mehrmals mit dem Straßennetz schneiden können, ergibt sich der im Zeitraum zwischen zwei Zellwechseln mögliche Aufenthaltsbereich aus der Summe aller Ellipsen, die sich aus allen Kombinationen der Schnittpunkte ergeben.

Da für krummlinige Begrenzungen das Auslesen der Straßenkarte sehr zeitintensiv ist, wird der mögliche Aufenthaltsbereich durch ein geschlossenes Polygon angenähert, indem die Hüllkurve um alle an zwei Zellwechseln beteiligten Schnittpunkte gelegt wird und zusätzlich um das Maximum der kleinen Halbachsen b durch Parallelverschiebung

nach außen vergrößert wird.

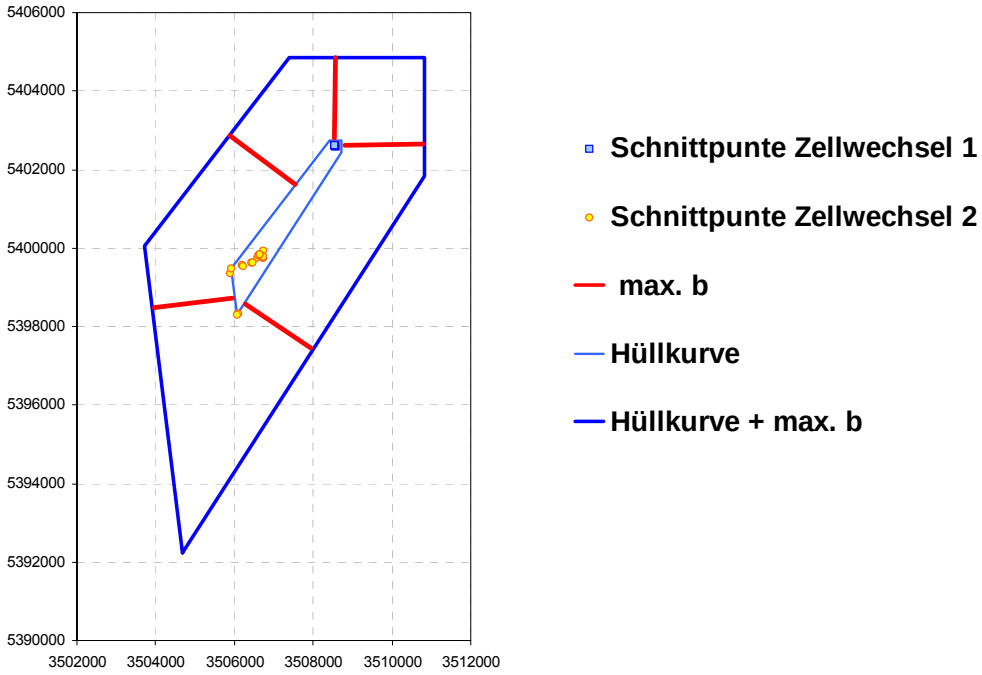


Abbildung 2-34: Hüllkurve und möglicher Aufenthaltsbereich

Innerhalb des möglichen Aufenthaltsbereichs können mit Hilfe der Kachelung des Lokalisierungsnetzes alle darin enthaltenen Knoten und Kanten des Straßennetzes zu einem Teilgraph aufgebaut werden. Diese Aufenthaltsbereiche zwischen zwei aufeinander folgenden Zellwechseln können auf verschiedene Arten zusammengesetzt werden.

Unter der Annahme, dass die protokollierten Zellwechsel immer durch die Positionsänderungen der Mobilstation verursacht wurden, muss die FPD-Route durch jeden Zellwechsel führen. Um in jedem Teilgraph eines Aufenthaltsbereichs mögliche Wege zwischen aufeinander folgenden Zellwechseln zu bestimmen, werden zwischen den Anfangsknoten der Schnittkanten des Straßennetzes mit den Grenzen des vorherigen Zellwechsels und den Endknoten der Schnittkanten des Straßennetzes mit den Grenzen des nachfolgenden Zellwechsels kürzeste Wege in allen Kombinationen berechnet. Alle Knoten und Kanten dieser kürzesten Wege werden zu einem Ersatzgraph zusammengefasst. In gleicher Weise ergeben sich für die weiteren Aufenthaltsbereiche mögliche Wege, um von einem Zellwechsel zum nächsten zu gelangen. Die Knoten und Kanten dieser Wege werden dem Ersatzgraph hinzugefügt, wenn sie noch nicht enthalten sind. Anschließend werden in diesem Ersatzgraph mögliche Routen als kürzeste Wege zwischen den Startknoten des ersten Zellwechsels und den Endknoten des letzen Zellwechsels bestimmt und die wahrscheinlichste Route daraus ausgewählt.

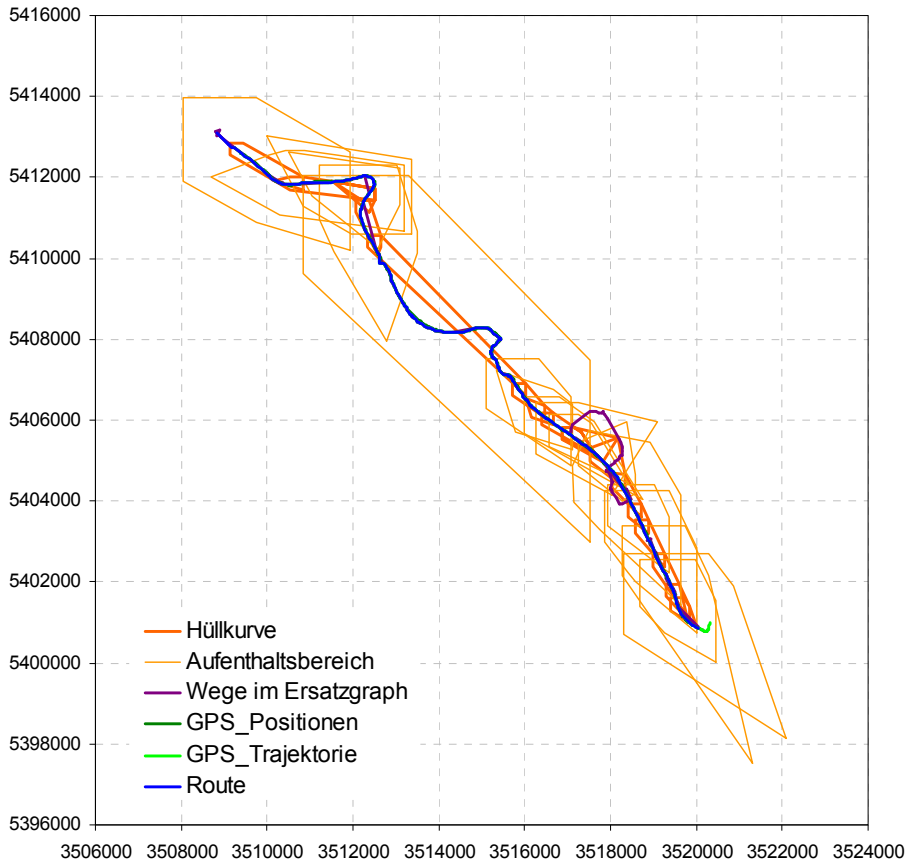


Abbildung 2-35: B10_Zuffenhausen_NS (Variante 1)

Als positives Beispiel zeigt Abbildung 2-35 das Ergebnis dieser Variante für eine Testfahrt auf der B10 im Stadtgebiet Stuttgart. Der aus den einzelnen Aufenthaltsbereichen zusammengesetzte Ersatzgraph besteht nur aus wenigen Straßenkanten, die aber trotz des Umschaltens in eine neben der Route liegende Funkzelle zusammenhängen und nur eine Alternative, die der richtigen Fahrstrecke entspricht (vgl. GPS-Referenz), zulassen. Leider enthalten die A-Daten häufig Umschaltungen in Zellen, die sich nicht aus der Position der Mobilstation erklären lassen. Das nächste Beispiel (Abbildung 2-36) zeigt anhand einer Autobahnfahrt von Stuttgart nach Karlsruhe das typische Auftreten solcher Fehlschaltungen, die nicht nur in den kleinen Funkzellen im Innenstadtbereich vorkommen. Der Ersatzgraph weist an vielen Stellen Lücken auf, so dass kein zusammenhängender Weg gefunden werden kann.

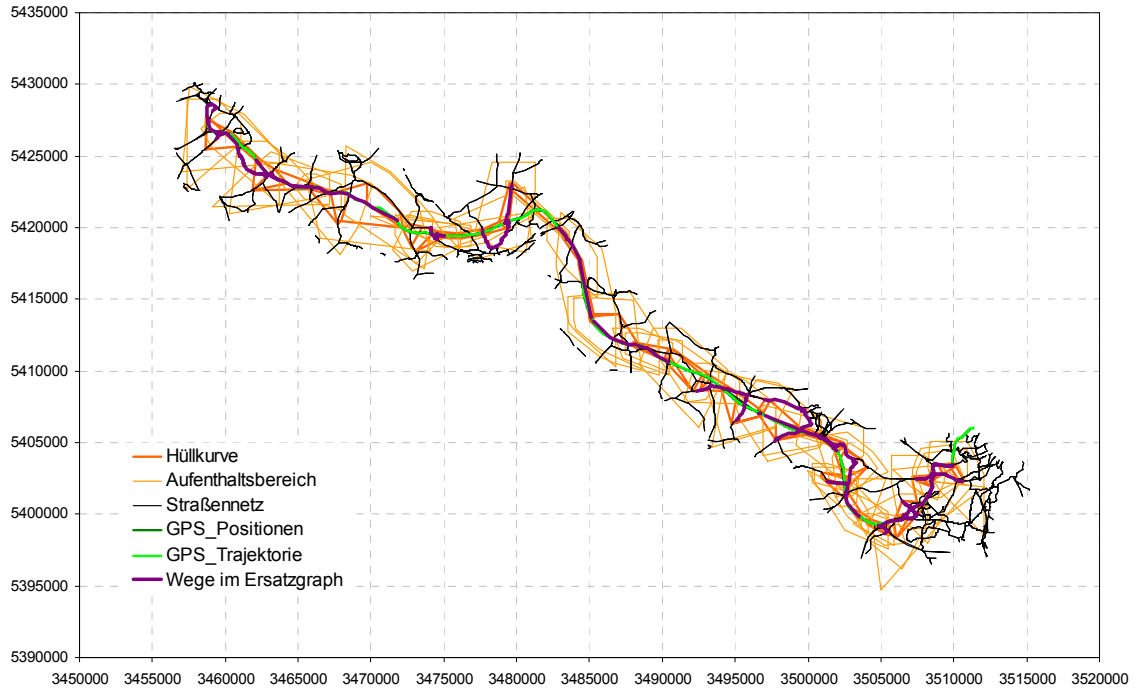


Abbildung 2-36: Testfahrt A8 Stuttgart ⇨ Karlsruhe (Variante jeder ZW)

Für eine gegenüber solchen Fehlschaltungen toleranteren Variante wird das Straßennetz innerhalb der Aufenthaltsbereiche aller Zellwechsel aufgebaut. Die möglichen Wege ergeben sich auch hier aus den kürzesten Wegen in allen Kombinationen der Startknoten des ersten Zellwechsels zu den Endknoten des letzten Zellwechsels.

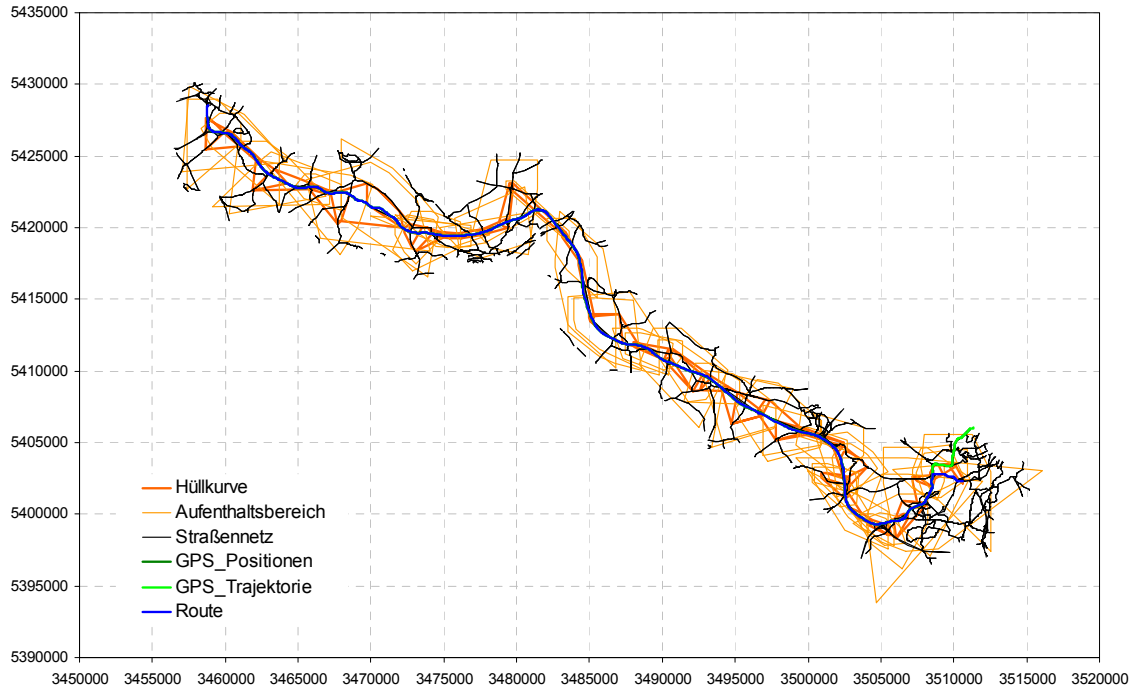


Abbildung 2-37: Testfahrt A8 Stuttgart ⇨ Karlsruhe (Variante mit Korridor)

Durch das Zusammenfassen der Aufenthaltsbereiche zu einem Korridor können die Funkzellen der nicht geometrisch bedingten Umschaltungen umgangen werden Bei der Autobahnfahrt (vgl. Abbildung 2-36 und Abbildung 2-37) zeigt sich dies z.B. bei der Funkzelle in Pforzheim. Im Vergleich zur GPS-Referenz werden 91.5 % der Strecke mit diesem Verfahren richtig zugeordnet.

Die Routenberechnung im Korridor führt aber dazu, dass die zeitliche Reihenfolge der Zellwechsel bei der Routengenerierung unberücksichtigt bleibt. Im folgenden Beispiel der Umgehungsstrecke U3 zur Autobahn A81 werden durch die langsame Fahrt die Aufenthaltsbereiche so breit, dass sie sich im umfahrenen Autobahnabschnitt überlappen und der Korridor in diesem Bereich geschlossen wird. Als schnellster Weg wird die Autobahn gefunden.

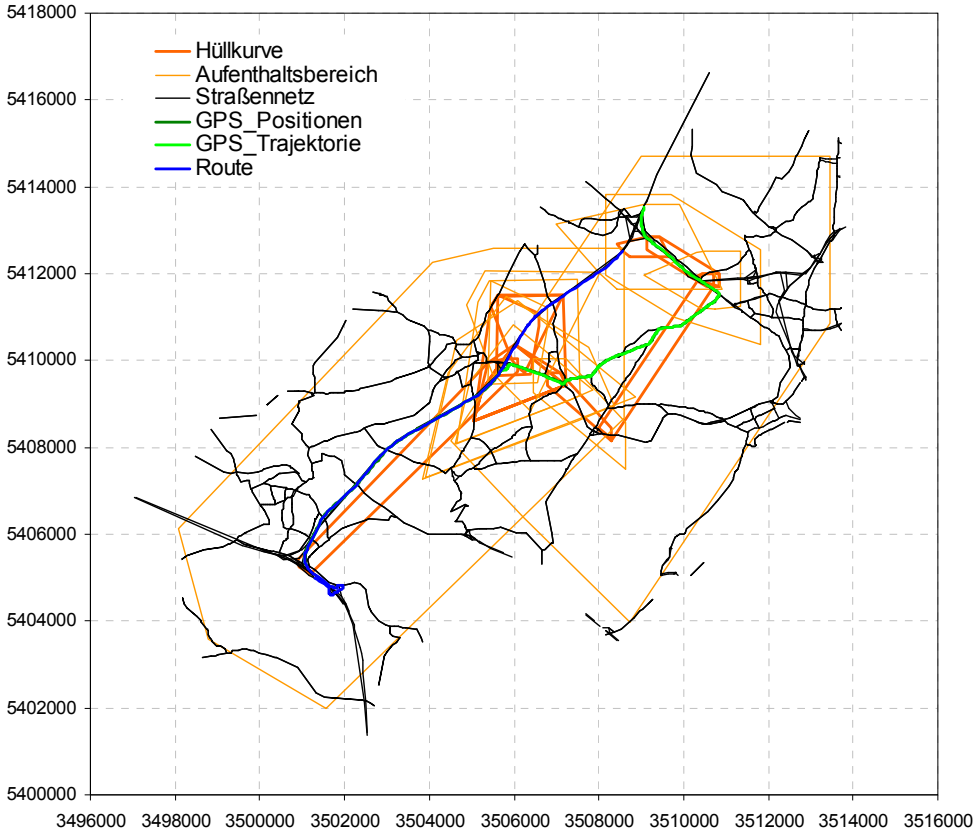


Abbildung 2-38: Testfahrt A81 U3 (Variante mit Korridor)

Die Variante, die zur Routenberechnung das Straßennetz im Korridor aller Aufenthaltsbereiche benutzt, ist gegenüber Fehlschaltungen sehr tolerant, findet aber bei Umgehungen, die in so geringem Abstand verlaufen, dass sie im Korridor liegen, immer die kürzeste Route. Wird die Route dagegen durch jeden Zellwechsel gezwungen, führen die nicht geometriebedingten Umschaltungen zu Lücken im Ersatzgraph, so dass keine Route gefunden wird. Um die Nachteile beider Varianten abzumildern, wurden weitere Varianten ausgetestet.

Um den Zwang durch alle Zellwechsel abzumildern, wird die Variante 1 dadurch erweitert, dass der Ersatzgraph nicht nur aus den Wegen zwischen aufeinander folgenden Zellwechseln besteht, sondern auch aus den Wegen zu den nächsten *n* Zellwechseln. Abbildung 2-39 zeigt eine Testfahrt auf der A8 von Karlsruhe zur A81 mit Umgehung des Leonberger Dreiecks, die mit einem Ersatzgraph aus den Wegen zu den jeweils nächsten 3 Zellwechseln über weite Teile die richtige Route findet.

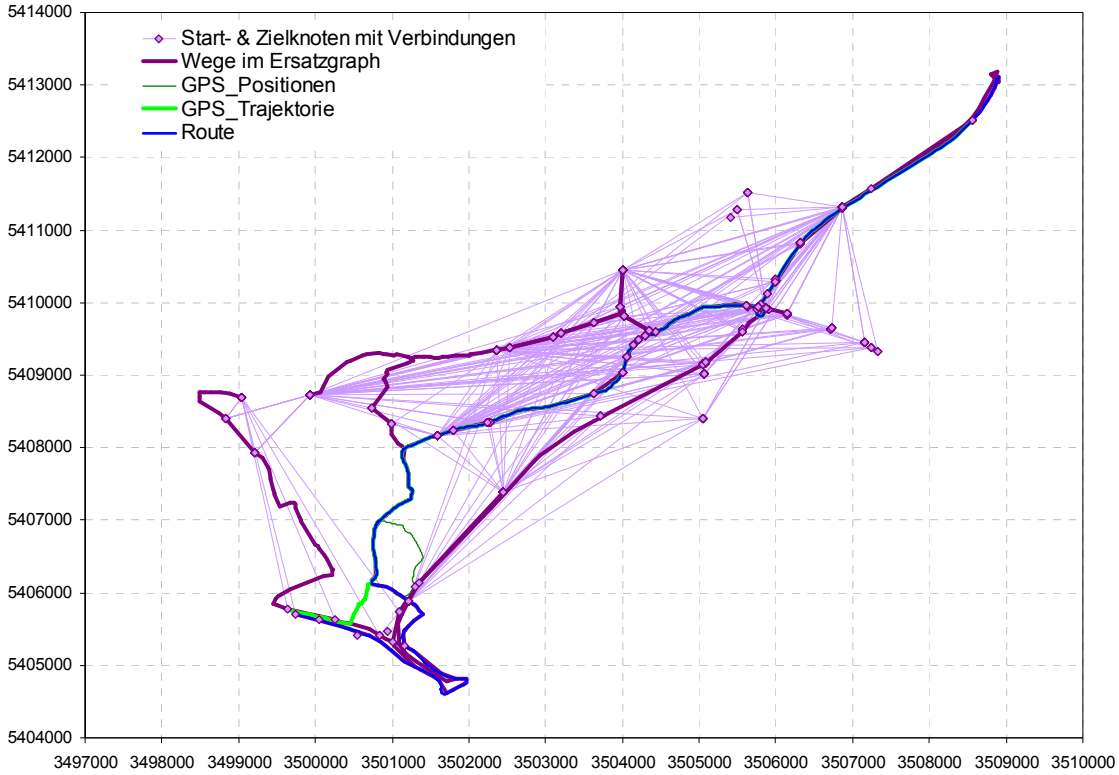


Abbildung 2-39: Testfahrt A8 von Karlsruhe mit Umgehung zur A81 (Variante mit 3 ZW)

Diese Variante, die die Reihenfolge der Zellwechsel berücksichtigt, aber tolerant ist gegenüber nicht geometriebedingten Umschaltungen, kann ähnlich wie die Variante mit dem Korridor aller Aufenthaltsbereiche nicht aus Alternativen mit geringem Abstand die richtige identifizieren, wenn die Aufenthaltsbereiche aufgrund zu hoch geschätzter Geschwindigkeiten zu breit sind. Vor allem bei den Umgehungen werden geringere Geschwindigkeiten als die für den Aufenthaltsbereich maximalen Geschwindigkeiten, die für die Autobahnen gespeichert sind, gefahren. Deshalb wurde empirisch versucht, die Aufenthaltsbereiche auf eine maximale Breite zu begrenzen (siehe nachfolgende Seiten).

Zur objektiven Beurteilung der verschiedenen Auswertevarianten und Breitenbegrenzungen wurde aus den bis Mitte März durchgeführten Testfahrten (ca. 960 km Länge) jeweils das Handy im Dauerbetrieb ausgewertet, da bei reinen Location Area Updates bzw. mehrmaligen kurzen Telefonaten eine schlechtere Auflösung eng benachbarter Routen zu erwarten ist.

Tabelle 2-4: Länge und Anteil der richtigen Zuordnungen

Begrenzung [m]	Korridor		4 ZW		3 ZW		2 ZW		jeder ZW	
	km	%	km	%	km	%	km	%	km	%
500	801	80	183	18	158	16	53	5	12	1
1000	889	89	458	46	223	22	118	12	12	1
1500	891	89	496	49	276	28	113	11	12	1
2000	876	87	493	49	290	29	147	15	12	1
2500	876	87	666	66	287	29	142	14	9	1
3000	863	86	660	66	287	29	145	14	9	1

Wie aus Tabelle 2-4 und Abbildung 2-40 ersichtlich führt die Auswertung mit dem Korridor aller Aufenthaltsbereiche zu den besten Ergebnissen, wobei die Breitenbegrenzung der Aufenthaltsbereiche nur geringen Einfluss hat. Diese Variante wird in der Online-Auswertung verwendet.

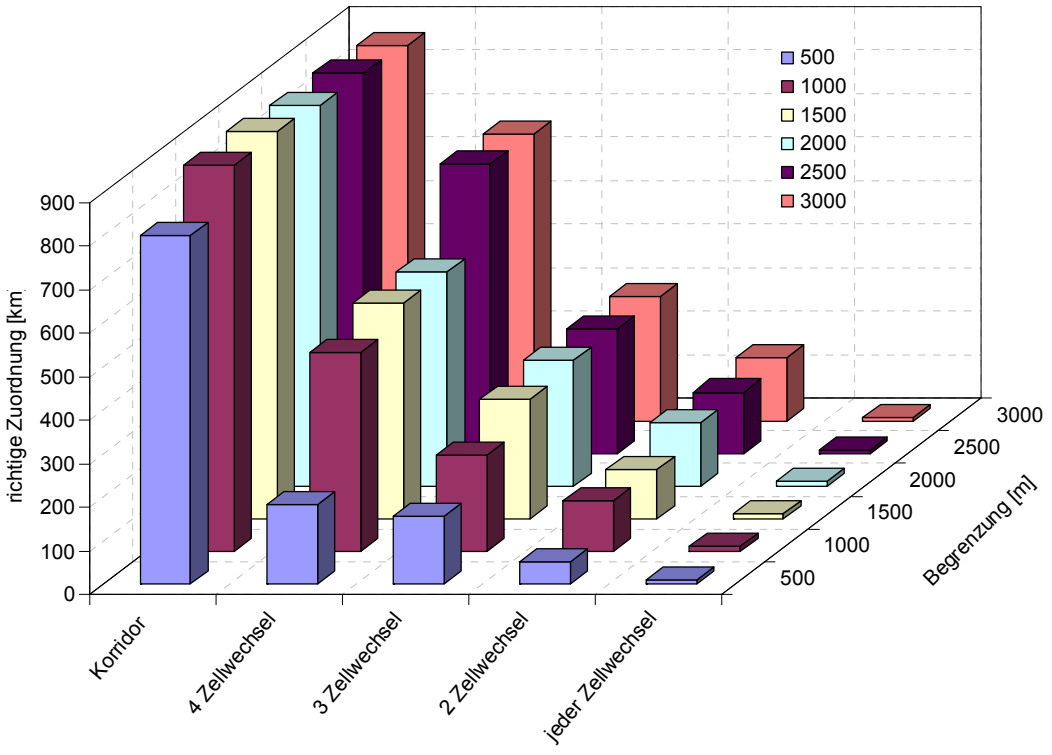


Abbildung 2-40: Vergleich der Varianten anhand der Länge der richtigen Zuordnungen

2.2.5 Weg-Zeit-Diagramme

Zur Weitergabe der Trajektorie eines Teilnehmers wird neben der befahrenen Route auch der zeitliche Verlauf auf der Route benötigt, der wie bei Abis-Daten entsprechend dem Datenmodell (vgl. [AP10-1 2008]) als Weg-Zeit-Diagramm zu bestimmen ist. Bei A-Daten muss das Weg-Zeit-Diagramm aus den wenigen Zeitpunkten der Umschaltungen von einer Funkzelle zur nächsten abgeleitet werden. Zum Zeitpunkt einer Umschaltung muss sich die Mobilstation sowohl auf der Route als auch auf den Grenzen der am Zellwechsel beteiligten Funkzellen befinden. Für einen Umschaltzeitpunkt können sich aufgrund geteilter Funkzellen mehrere Schnittpunkte ergeben. Bei Schaltungen in Zellen abseits der Route, die aufgrund der Kapazität häufig vorkommen, ergeben sich keine echten Schnittpunkte zwischen den Straßenkanten und den Kanten der Zellgrenzen. Um auch die Zeitinformationen dieser Zellwechsel nutzen zu können, werden Schnittpunkte auch noch verwendet, wenn der Schnittpunkt auf eine maximal 1.5 km lange Verlängerung der Zellkante fällt. Ergeben sich dadurch für einen Umschaltzeitpunkt mehrere Schnittpunkte werden sie nach einer dieser Verlängerung umgekehrt proportionalen Gewichtung gemittelt

$$q = \begin{cases} 10 & \text{bei echten Schnitten} \\ 1/S_{Verlängerung} & \text{sonst} \end{cases}$$

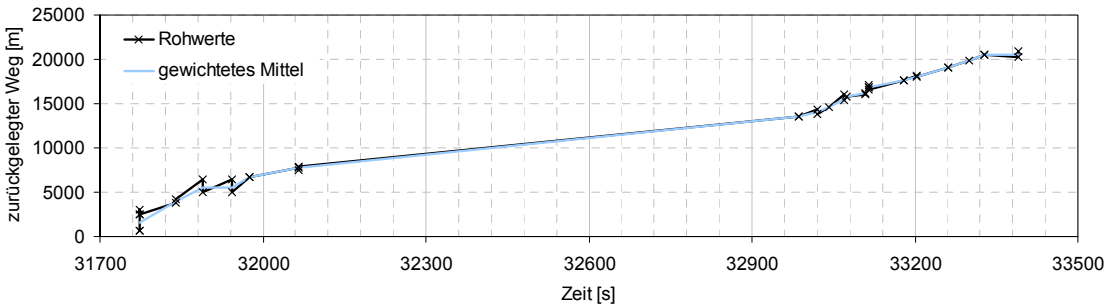


Abbildung 2-41: Rohwerte und gewichtetes Mittel des Weg-Zeit-Diagramms

Für die Testfahrt auf der B10 (Zuffenhausen), bei der zweimal mit einer Pause von 15 Minuten telefoniert wurde, zeigt Abbildung 2-41 die aus allen Schnittpunkten ermittelten Rohwerte des Weg-Zeit-Diagramms, die für einen Zeitpunkt Differenzen von bis zu 2 km aufweisen können. Die gewichtete Mittelwertbildung führt zu einer deutlichen Glättung, kann aber noch extrem hohe Geschwindigkeiten oder Rückwärtsfahrten beinhalten. Deshalb werden die Zeitpunkte, die zu diesen nicht-plausiblen Geschwindigkeiten führen, aus dem Weg-Zeit-Diagramm eliminiert. In Abbildung 2-42 kann diese Korrektur mit der GPS-Referenz verglichen werden, wobei das aus der GPS-Referenz abgeleitete Weg-Zeit-Diagramm (in dunkelgrün) auf die Umschaltzeiten der Mobilfunkortung (hellgrün) umgerechnet wurde, um die Geschwindigkeiten besser beurteilen zu können (GPS-Trajektorie in Abbildung 2-42). Bei dieser Testfahrt ergab sich eine Stan-

dardabweichung der Geschwindigkeit von 6,4 km/h, wobei maximale Abweichungen von -15 bis 25 km/h auftraten. Die aus A-Daten ermittelte Durchschnittsgeschwindigkeit weicht von der mit GPS bestimmten Durchschnittsgeschwindigkeit um 2,5 km/h ab.

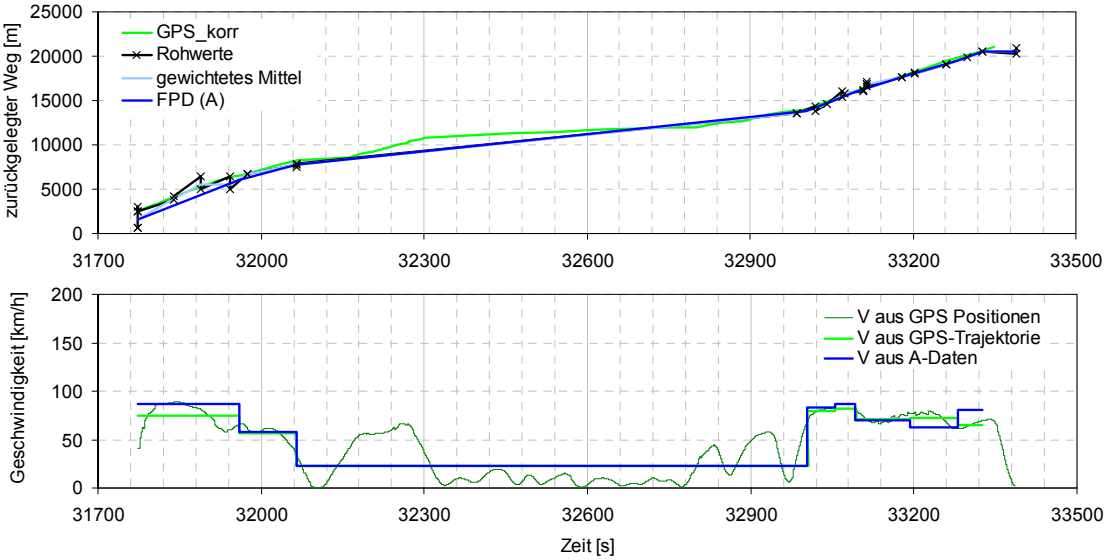


Abbildung 2-42: Weg-Zeit- und Geschwindigkeits-Zeit-Diagramm

Die Ergebnisse der für die anderen Testfahrten durchgeführten Genauigkeitsuntersuchungen sind im Kapitel 5.1.2 (Qualitätsbeurteilung A-Ebene) dargestellt.

Die wichtigsten Berechnungsschritte zur Generierung von FPD-Trajektorien inklusive der Weg-Zeit-Diagramme auf A-Ebene sind in Abbildung 2-43 noch einmal zusammengefasst.

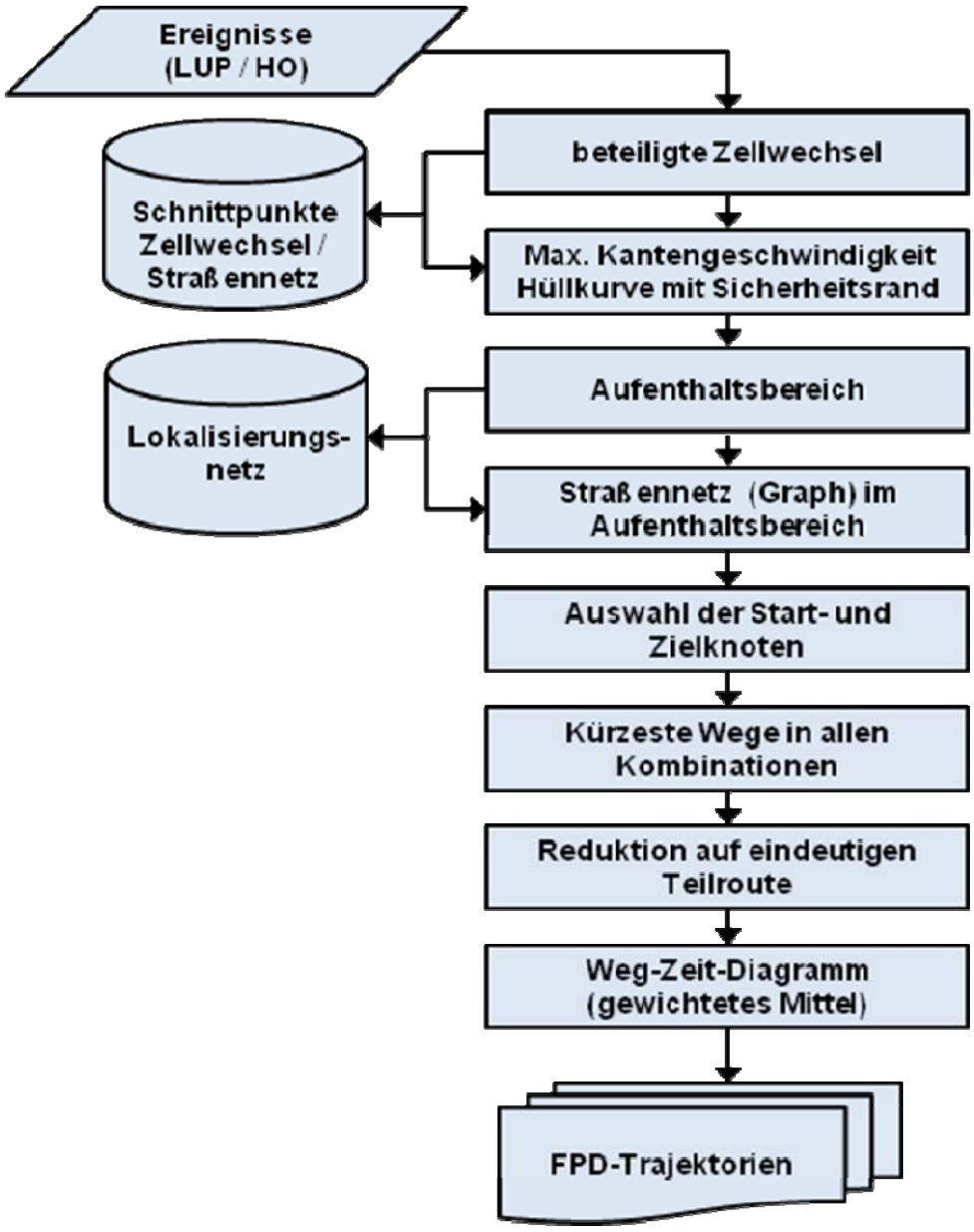


Abbildung 2-43: Schematischer Ablauf der Trajektoriengenerierung auf A-Ebene

3 Qualitätskriterien

Zur Beschreibung der Verkehrsdatenqualität werden im Qualitätsmodell für Do-iT (vgl. [AP10-1 2008]) die zwei Zuverlässigkeitsmerkmale Verfügbarkeit und Aktualität, die zeitbezogene Aspekte der Daten beschreiben, die drei Integritätsmerkmale Vollständigkeit, Konsistenz und Korrektheit, die die Gebrauchstauglichkeit der Daten hinsichtlich ihrer physikalischen und semantischen Eigenschaften beschreiben, und das Merkmal Genauigkeit verwendet. Diese Merkmale sind durch messbare Qualitätsparameter zu konkretisieren. Im Einzelnen werden die in Tabelle 3-1 dargestellten Qualitätsparameter zur abschließenden Beurteilung der Qualität der FPD-Trajektorien herangezogen. Die Parameter gehen aus dem Qualitätsmodell für FPD-Trajektorien hervor, wie es im Evaluierungskonzept (vgl. [AP10-1 2008]) dargestellt ist bzw. im Abschlussbericht zum Nachweis der Wirksamkeit der Qualitätssicherungsmaßnahmen (vgl. [AP8-7 2009]) für A-Daten erweitert wurde.

Der Zusammenhang der beurteilbaren Aspekte der FPD mit den konkret bestimmbaren Qualitätsparametern wird ebenfalls erläutert.

Tabelle 3-1: Zur Evaluierung herangezogenes Qualitätsmodell für FPD-Trajektorien

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
P1.1	Verfügbarkeit VE	Globale Ausfallrate (offline+relativ)	AUR_{global} [%]	Beschreibt die Verfügbarkeit des FPD-Servers in einem best. Zeitraum
P1.2		Trajektoriendichte (online+relativ)	ρ_{Tra} [1/s]	Summe(Routenlängen)/L/T mit L = Länge des betrachteten Streckenabschnitts und T = Dauer
P2.1	Aktualität AK	Berechnungszeit (online+relativ)	Δt_{ber} [s]	Berechnungszeit für eine FPD-Trajektorie aus der Differenz *FPD_CalculationTime – FPD_Starttime*
P2.2		Mittlere Berechnungszeit (offline+relativ)	$\overline{\Delta t}_{ber}$ [s]	Mittlere Berechnungszeit (Stunde, Tag) aus P2.1
P2.3		Alter der Trajektorie (online+relativ)	Δt_{FPD} [min]	Alter einer Trajektorie, berechnet aus der Differenz zwischen der aktuellen Zeit und dem Zeitstempel der ersten Position *aktuelle Zeit –FPD_Starttime*
P2.4		Mittleres Trajektorienalter (offline+relativ)	$\overline{\Delta t}_{FPD}$ [min]	Mittleres Alter der Daten (aus P2.3)
P3.1 Abis	Vollständigkeit VO	Trajektorienvollständigkeit (online+relativ)	VO_{FPD} [%]	**Abis:** Anteil der verwendeten Positionen einer Positionsfolge eines Teilnehmers, der für die Trajektorie verwendet wird *FPD_Quality (2)*
P3.1 A				**A:** Anteil der verwendeten Zellwechsel eines Teilnehmers der für die Trajektorie verwendet wird *FPD_Quality (2)*

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
P3.2		Abdeckungsrate (offline+absolut)	ABR_{FPD} [%]	Grad der Abdeckung des betrachteten Straßennetzes mit FPD (Begriff: „vollständige Abdeckung“ ist zu definieren)
P3.3		Durchdringung mit FPD (offline+absolut)	d [%]	Punktueller Anteil der mit FPD detektierten Verkehrsstärke im Vergleich zur tatsächlichen (jeweils an SES-Standorten)
P4.1	Konsistenz KO	Wird durch Einhaltung des Datenmodells gewährleistet		
P5.1	Korrektheit KR	Klassifizierungskorrektheit (online+absolut)	KR_{Kl} [%] oder [0...3]	Wahrscheinlichkeit, mit der die Teilnehmerklassen identifiziert werden (Aggregation auf einen QP "aktiver IVT" geplant) *FPD_Identification*
P5.2		Mittlere Klassifizierungskorrektheit (offline+absolut)	$\overline{KR}_{Kl}$ [%] oder [0...3]	Mittlere Klassifizierungskorrektheit der Klassen z. B. über einen Tag berechnet (aus P5.1)
P5.3		Trajektorienlänge (online+relativ)	L_{FPD} [km]	Länge der FPD-Trajektorie (*FPD_RouteLength*)
P5.4		Mittlere Trajektorienlänge (offline+relativ)	$\overline{L}_{FPD}$ [km]	Mittlere Länge der FPD-Trajektorie (aus P5.3)
P5.5		Zuordnungskorrektheit Typ A (offline+absolut)	KR_{ZuA} [%]	Korrekter Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, bezogen auf die Länge der GPS-Route (GPS als Referenz erforderlich)
P5.6		Zuordnungskorrektheit Typ B (offline+absolut)	KR_{ZuB} [%]	Korrekter Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, bezogen auf die Länge der ermittelten FPD-Route (GPS erforderlich)
P5.7		Rangkorrelation nach Spearman (offline+absolut)	r	Punktuelle Korrelation des Tagesgangs der detektierten Verkehrsstärke mit der tatsächlichen mit SES ermittelten (jeweils an SES-Standorten)
P6.1 ABis	Genauigkeit GE	Mittlere Querabweichung (online+relativ)	QA_{FPD} [m]	Mittel der orthogonalen Abweichungen der verwendeten Positionen von der wahrscheinlichsten Route *FPD_Quality (1)* **nur Abis-Daten**
P6.1 A		Anzahl der Zellwechsel eines Teilnehmers (online+absolut)	Anz_{ZW} [-]	Anzahl der Zellwechsel, die einem Teilnehmer zugeordnet werden konnten *FPD_Quality (1)* **nur A-Daten**
P6.2		Standardabweichung der Geschwindigkeit (offline+absolut)	S_V [km/h]	Gibt die Genauigkeit der aus den Trajektorien ermittelbaren Geschwindigkeit an. Referenz erforderlich (Messfahrt oder TEU)
P6.3		Differenz der Durchschnittsgeschwindigkeiten (offline+absolut)	δ [km/h]	Vorzeichenfreie Differenz zwischen Durchschnittsgeschwindigkeit aus GPS und aus MFD

Einerseits werden intern im Online-Betrieb bestimmbare Parameter zum Vergleich der Ergebnisse verschiedener Zeiträume herangezogen. Andererseits ist eine absolute Beurteilung der Mobilfunktrajektorien durch Vergleich mit Referenzdaten erfolgt.

Zu den im Online-Betrieb bestimmbaren Qualitätsparametern, die sich auf einzelne Trajektorien beziehen und in der Tabelle mit „online+relativ“ bezeichnet sind, gehören die Parameter Alter der Trajektorie, Trajektorienvollständigkeit, Trajektorienlänge, und die Genauigkeit. Das Alter einer Trajektorie, das sich aus der Differenz zwischen der aktuellen Zeit und dem Zeitstempel der ersten Position berechnet, kann im Online-Betrieb jederzeit berechnet werden, und die Trajektorienlänge wird im Online-Betrieb für jede FPD-Trajektorie zur Verfügung gestellt. Die beiden anderen Qualitätsparameter (P3.1 und P6.1) können aufgrund der unterschiedlichen Auswertealgorithmen für Abis- und A-Daten nicht einheitlich definiert werden.

Bei Abis-Daten liegt nach Signalstärken-Matching und Kalman-Filterung für jeden aktiven Mobilfunkteilnehmer eine Positionsfolge vor, die Grundlage für das Festlegen der plausibelsten Route aus mehreren Alternativen bildet. Es wird die Route gewählt, auf die möglichst viele Positionen projiziert werden können und dabei möglichst geringe Abstände haben (vgl. Kapitel 2.1.4). Zur Qualitätsbeurteilung im Online-Betrieb werden die gleichen Parameter verwendet. Aus den orthogonalen Abweichungen der Abis-Positionen zur plausibelsten Route wird die „mittlere Querabweichung QA_{FPD}“ berechnet. Da nicht für alle Positionen Lotfußpunkte auf der Route gefunden werden, wird der prozentuale Anteil *P* der verwendeten Positionen einer Positionsfolge als „Trajektorienvollständigkeit VO_{FPD}“ berechnet.

Bei A-Daten können diese Qualitätsparameter nicht bestimmt werden, da keine Positionsinformationen vorliegen, sondern nur wenige Ereignisse wie Location Area Updates und Handover. Anhand realer Daten bestätigte sich die Vermutung, dass aus A-Daten abgeleitete Trajektorien den Referenzwerten näher kommen, je mehr Zellwechsel zur Verfügung stehen. Da auch Zellwechsel auftreten können, die nicht durch eine Positionsänderung sondern z.B. durch eine Kapazitätsüberschreitung der Antenne (BTS) verursacht sind, und im Algorithmus eliminiert werden, werden im Online-Betrieb zwei Parameter zur Verfügung gestellt. Der erste gibt die Anzahl der Zellwechsel einer MS reduziert um die Zellen, die nicht im Best-Server-Plot enthalten sind (Fremdanbieter), an; der zweite die Anzahl der zur Trajektorienbestimmung verwendeten Zellwechsel.

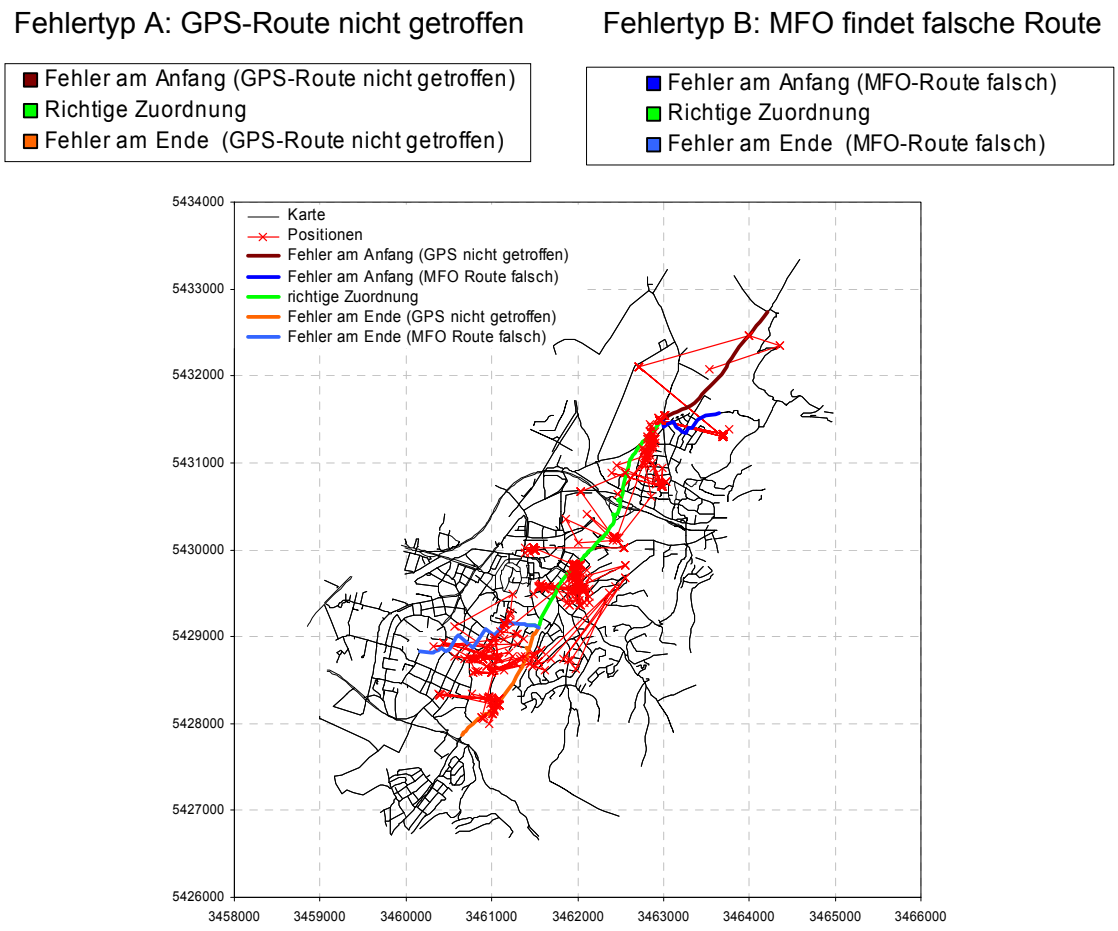
3.1 Qualitätskriterien: GPS als Referenz

Eine absolute Beurteilung einer aus Abis- oder A-Daten generierten Trajektorie kann nur durch Vergleich mit Referenzdaten erfolgen. Hierfür wurden eine Reihe eigener Testfahrten mit GPS durchgeführt.

Mit den aus Abis-Daten bestimmten Positionsfolgen liefert die Zuordnung auf das Straßennetz im mittleren Bereich meist die richtige Route (in der folgenden Abbildung in

grün dargestellt). Am Anfang und Ende treten jedoch Abweichungen auf, die in zwei Kategorien eingeteilt werden können:

Der Fehlertyp A „GPS-Route nicht getroffen“ verursacht keine Fehlzuordnungen und führt lediglich zu kürzeren Trajektorien und damit zu weniger Informationen für die Schätzung der Verkehrslage. Fehlertyp A beeinträchtigt also nur die Trajektorienvollständigkeit, aber nicht die Korrektheit. Beim Fehlertyp B „MFO findet falsche Route“ werden dagegen Gesprächsteile falschen Straßensegmenten zugeordnet, so dass falsche Informationen zur Verfügung gestellt werden. Fehlertyp B ist daher als deutlich kritischer einzustufen.



3-1: Fehlerkategorien

Zur Unterscheidung dieser beiden Fehlertypen wurden im Qualitätsmodell zwei Parameter definiert.

- P5.5 „Zuordnungskorrektheit Typ A KR_{ZuA} [%]“ setzt den korrekten Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, mit der Gesamtlänge der GPS-Route in Relation.

- P5.6 „Zuordnungskorrektheit Typ B KR_{ZuB} [%]“ bezieht den korrekten Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, auf die Länge der mit Mobilfunkortung ermittelten Route.

Im Gegensatz zu den aus Abis-Daten bestimmten Trajektorien liefert die Routenbestimmung mit A-Daten nicht immer im mittleren Bereich die richtige Route. Um zu mit Abis-Daten vergleichbaren Parametern zur Beurteilung der absoluten Qualität zu gelangen, werden die Längen der richtigen und falschen Teilrouten durch Summation der entsprechenden Kantenlängen bestimmt, so dass die Trajektorie wieder anhand der beiden Qualitätsparameter „Zuordnungskorrektheit Typ A“, der den Anteil der korrekten Routen an der mit GPS bestimmten Route angibt, und der „Zuordnungskorrektheit Typ B“, der den prozentualen Anteil der korrekten Route an der aus A-Daten generierten Route angibt.

Anhand der Zuordnungskorrektheiten Typ A oder Typ B wird nur die Qualität der Route beurteilt, nicht das Weg-Zeit-Diagramm. Zur Beurteilung des Weg-Zeit-Diagramms werden die aus Mobilfunkortung abgeleiteten Geschwindigkeiten mit denen der GPS-Referenz verglichen.

Bei Abis-Daten werden die Weg-Zeit-Diagramme vor der Datenübertragung durch Mittelwertbildung über 30 s Intervalle Δt_k ausgedünnt, so dass sich daraus intervallweise konstante Geschwindigkeiten v_{MFO} ableiten lassen. Zum leichteren Vergleich werden für die gleichen Diskretisierungsintervalle Δt_k aus den mit GPS bestimmten Weg-Zeit-Diagrammen die Geschwindigkeiten v_{GPS} berechnet.

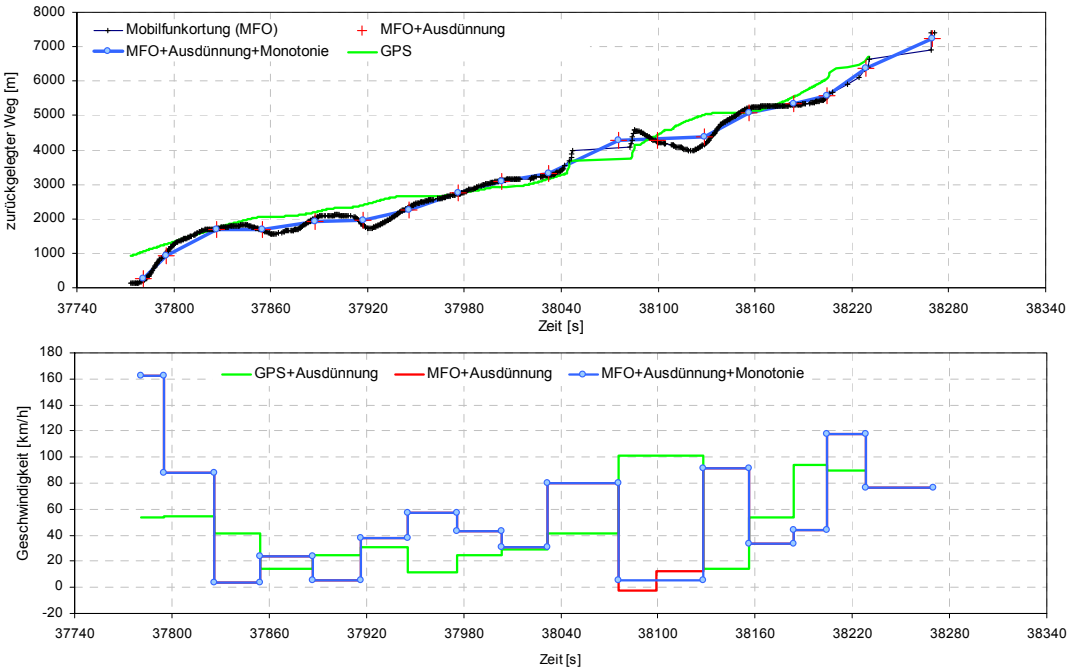


Abb. 3-2: Weg-Zeit- und Geschwindigkeits-Zeit-Diagramm

Zur Genauigkeitsbeurteilung der Geschwindigkeit wird eine Standardabweichung anhand der Differenzen $\varepsilon = v_{GPS} - v_{MFO}$ und der Gesamtzeit *T* (Summe der Zeitintervalle) abgeschätzt (vgl. Qualitätsparameter P6.2):

$$\sigma_v = \sqrt{\frac{1}{T}\sum_{k=0}^{m} \Delta t_k \varepsilon_k^2}$$

Da die Abis-Positionen, die quer zur Fahrtrichtung einige hundert Meter von der Route abweichen können, auch in Fahrtrichtung in dieser Größenordnung falsch sein können, treten in kurzen Zeitintervallen große Geschwindigkeitsdifferenzen auf, die sich aber über längere Zeiträume ausmitteln. Deshalb wurde neben dieser Standardabweichung zur Genauigkeitsbeurteilung der Geschwindigkeit auch die vorzeichenfreie Differenz der Durchschnittsgeschwindigkeiten (P6.3 in Tabelle 3-1)

$$\delta = \left\| \bar{v}_{GPS} - \bar{v}_{MFO} \right\|; \quad \text{mit } \bar{v} = \frac{s(t_E) - s(t_A)}{t_E - t_A}$$

verwendet.

Anhand dieser beiden Beurteilungskriterien für die Geschwindigkeit wurde aus den Testfahrten mit GPS-Referenz festgelegt, dass die Diskretisierung des Weg-Zeit-Diagramms bei Abis-Daten am günstigsten mit einem Mittelungsintervall von 30 s erfolgt. Für dieses Diskretisierungsintervall liegen die Beträge der Differenzen der Durchschnittsgeschwindigkeiten im Mittel über alle Testfahrten bei 12.5 km/h (≅3.5 m/s).

Zur Beurteilung der aus A-Daten bestimmten Weg-Zeit-Diagramme wird ebenfalls die Standardabweichung der Geschwindigkeit und die vorzeichenfreie Differenz der Durchschnittsgeschwindigkeiten verwendet, die hier in gleicher Weise definiert sind wie bei Abis-Daten. Allerdings liegen die A-Daten nicht mit konstanter Abtastung vor, so dass sich die Zeitintervalle auf die Dauer zwischen zwei Umschaltzeiten der Mobilfunkortung beziehen und zum Vergleich das GPS bestimmte Weg-Zeit-Diagramm auf diese Zeitpunkte umgerechnet wird.

Die Untersuchungsergebnisse zur Beurteilung der absoluten Qualität anhand der GPS-Testfahrten finden sich im Abschnitt 5.1 dieses Berichts.

3.2 Qualitätskriterien: SES als Referenz

Zur absoluten Qualitätsbeurteilung mit Referenzdaten standen zusätzlich zu den Testfahrten mit GPS während der gesamten Evaluierungsphase nahezu lückenlos Messdaten ausgewählter stationärer Verkehrsdatenerfassungssysteme (SES) der Projektpartner zur Verfügung. Die herangezogenen Auswahlkriterien sowie die Beschreibung und Lage der geeigneten SES sind in Kapitel 4.2 detailliert dargestellt, die Vergleichsergebnisse finden sich im Abschnitt 5.2.

Die Anforderungen an die Qualität der Referenzdaten werden weitestgehend erfüllt. Die SES lieferten die genaue **Verkehrsstärke q** sowie teilweise auch Vergleichswerte für die **Durchschnittsgeschwindigkeit v** an definierten Stellen des gesamten Projektnetzes. Aufgrund der ungleich besseren Qualität der Referenzgeschwindigkeiten aus GPS-Messungen, wurde jedoch auf den Vergleich mit den punktuellen Geschwindigkeiten aus den SES verzichtet.

Zusätzlich zu den im Qualitätsmodell für FPD-Trajektorien, wie es für Do-iT entwickelt (vgl. [AP8-2, 2006]) bzw. erweitert worden ist (vgl. [AP10-1 2008] und [AP8-7 2009]), wurden die **Durchdringung** (P3.3) mit FPD sowie die **Rangkorrelation nach Spearman** (P5.7) zum Vergleich mit den SES als zusätzliche Parameter hinzugefügt (vgl. Tabelle 3-1). Die beiden neuen Parameter sind wie folgt definiert:

Tabelle 3-2: Definition der erforderlichen Qualitätsparameter zur Evaluierung der FPD mit SES

Parameter	Berechnung	Beschreibung
Durchdringung (diskret)	$d = \frac{q_{FPD}}{q_{SES}}[\%]$	Anteil des Verkehrs, der mit FPD erfasst werden kann; Verkehrsstärke in KFZ/h aus FPD bezogen auf die Verkehrsstärke aus den SES
Rangkorrelation nach Spearman (diskret)	$r = 1 - \frac{6 \cdot \sum D^2}{n \cdot (n^2 - 1)}$	Robustes Maß zur Beurteilung des Zusammenhangs der Tagesganglinien der Verkehrsstärke aus FPD und aus SES; zunächst werden die Reihen in Rangreihen transformiert, dann werden die Differenzen D der n Rangpaare gebildet

Die zur Evaluierung ausgewählten Qualitätsparameter sowie die ggf. erforderlichen Messgrößen sind im Folgenden näher erläutert.

Tabelle 3-3: Referenzmessgrößen aus SES und beurteilbare Parameter

Referenzmessgröße	beurteilbarer Q-Parameter	Erläuterung
Verkehrsstärke aus SES ***q*** **[KFZ/h]**	Abdeckungsrate ABR_{FPD} [%] (P 3.2) Durchdringung d [%] (P 3.3) Korrelationskoeffizient r [%] (P5.7)	Punktuelle Hinweise, wo KFZ überhaupt detektiert werden können Punktuelle Hinweise, welcher Anteil des Verkehrs mit FPD erfasst wird Punktuelle Hinweise ob FPD zur Hochrechnung der Verkehrsstärke geeignet
Durchschnittsgeschwindigkeit am MQ[2] aus SES: v_{SES} **(km/h)**	Standardabweichung der Geschwindigkeit S_v [km/h]	Ableitbare Verkehrsgeschwindigkeit, ableitbarer LOS[3]

[2] Messquerschnitt: Zusammenfassung aller Spuren einer Fahrtrichtung

Die tageweise ausgewertete Durchdringung gibt Aufschluss über den mit FPD detektierbaren Verkehrsanteil. Durch die genaue Kenntnis der Lage der Referenzsensoren ist auch eine getrennte Beurteilung der Eignung von FPD zur Bestimmung der Verkehrsstärke auf Autobahnen, auf Umgehungsstraßen und innerorts sowie die Einbeziehung der umgebenden Infrastrukturdichte möglich. Zur besseren Beurteilung der Durchdringung, werden neben dem Durchschnittswert über alle ausgewerteten Tage auch die Streuung als Differenz zwischen dem größten und kleinsten vorkommenden Wert sowie die empirische Standardabweichung angegeben.

Neben der tageweisen Beurteilung der Verkehrsstärke dient die Korrelation der Beurteilung der sich im Tagesverlauf ändernden Verkehrsstärke. Der ermittelte Tagesgang – in Stunden diskretisiert – wird direkt mit dem Tagesgang, der von dem entsprechenden SES erfasst wurde, verglichen. Auch hier wurde die Aussagekraft der mittleren Korrelation durch die Streuung über alle ausgewerteten Tage sowie durch die empirische Standardabweichung ergänzt.

Die Referenzmessgröße Durchschnittsgeschwindigkeit (vgl. Tabelle 3-3) wurde nicht bei der Evaluierung mittels SES-Daten verwendet. Für einen flächenhaften Vergleich der aus FPD-Trajektorien ableitbaren Verkehrslage über die Weg-Zeit-Informationen ist die Generierung von Reisezeiten und der LOS erforderlich. Dieser Prozessschritt konnte nur vom Projektpartner DDG durchgeführt werden. Ein Vergleich der ermittelbaren Geschwindigkeiten einzelner Trajektorien wurde daher nur mit GPS als Referenz durchgeführt.

Mit Hilfe der vorgestellten Parameter gelang eine Evaluierung der berechenbaren FPD-Trajektorien im Projektnetz. Im Folgenden werden die verwendeten Referenzdaten und Sensoren dargestellt (Kapitel 4.2) und es erfolgt eine abschließende Evaluierung der erzielten Ergebnisse (Kapitel 5.2).

[3] Level of Service: dabei handelt es sich um eine stufenweise Einteilung zur Angabe der Verkehrslage; hier werden die 3+1 Stufen *frei, stockend, Stau, nicht ermittelbar* verwendet.

4 Beschreibung der Referenzdaten

4.1 GPS-Trajektorien

Die Genauigkeit der GPS-Tracks lag durchgängig bei unter 10 m, da die Positionsbestimmung in den kurzen Tunnels, die Bestandteil der Testfahrten waren, durch einen integrierten Kreisel sowie das Tachosignal gestützt wurden. Oberirdisch wird laut Hersteller eine Genauigkeit von 2,5 m CEP[4] erreicht. Diese Genauigkeit ist damit eine Zehnerpotenz höher als die bestenfalls zu erwartenden Einzelpositionen aus Mobilfunkdaten, daher ist der verwendete GPS-Empfänger als Referenz geeignet. Die Verfügbarkeit lag ebenfalls auf Grund der Unterstützung durch Kreisel und Tacho bei nahezu 100 % und die exakte zeitliche Zuordnung der Messreihen konnte über den GPS-Zeitstempel gewährleistet werden.

Nachfolgend werden die Versuchsfahrten zur Evaluierung der A-Daten aufgeführt. Die entsprechende Fahrten zur Abis-Evaluierung wurden bereits im Bericht 4.5 (vgl [AP4-5 2008]) angegeben und werden daher in diesem Bericht nicht mehr erläutert.

Tabelle 4-1: Autobahnfahrten vom 25.11.2008, 18.02.2009 und 18.03.2009

Nr	Bezeichnung	Szenario	Datum/Uhrzeit	Dauer
1	A5_NS	A5 Höhe Bruchsal in Nord-Südrichtung	2009.03.18 / 16:34	22 min
2	A5_SN	A5 Höhe Bruchsal in Süd-Nordrichtung	2009.03.18 / 16:06	20 min
3	A5_U45_West	Umleitungsstrecke U45 Höhe Bruchsal in Nord-Südrichtung	2009.03.18 / 15:35	25 min
4	A6_U59_1	Umleitungsstrecke U59 Höhe Wiesloch Rauenberg	2009.03.18 / 13:15	35 min
5	A6_68_4	Umleitungsstrecke U68 Höhe Wiesloch Rauenberg	2009.03.18 / 13:53	27 min
6	A6A81	Autobahnviereck: Karlsruhe-Mannheim-Heilbronn-Stuttgart	2008.11.25 / 15:51	78 min
7	A8SK	A8 Stuttgart -Karlsruhe	2008.11.25 / 7:13	64 min
8	AB4_L	Autobahnviereck: Stuttgart-Heilbronn-Mannheim- Karlsruhe-Stuttgart	2009.02.18 / 15:06	120 min
9	AB4_R1	Autobahnviereck: Stuttgart-Karlsruhe	2009.02.18 / 9:20	60 min
10	AB4_R2	Autobahnviereck: Karlsruhe-Mannheim-Heilbronn-Stuttgart	2009.02.18 / 13:45	75 min

[4] CEP: Circular Error Probable; d. h. 50 % aller Messungen liegen innerhalb dieses Radius, bzw. 95 % liegen innerhalb des doppelten Radius

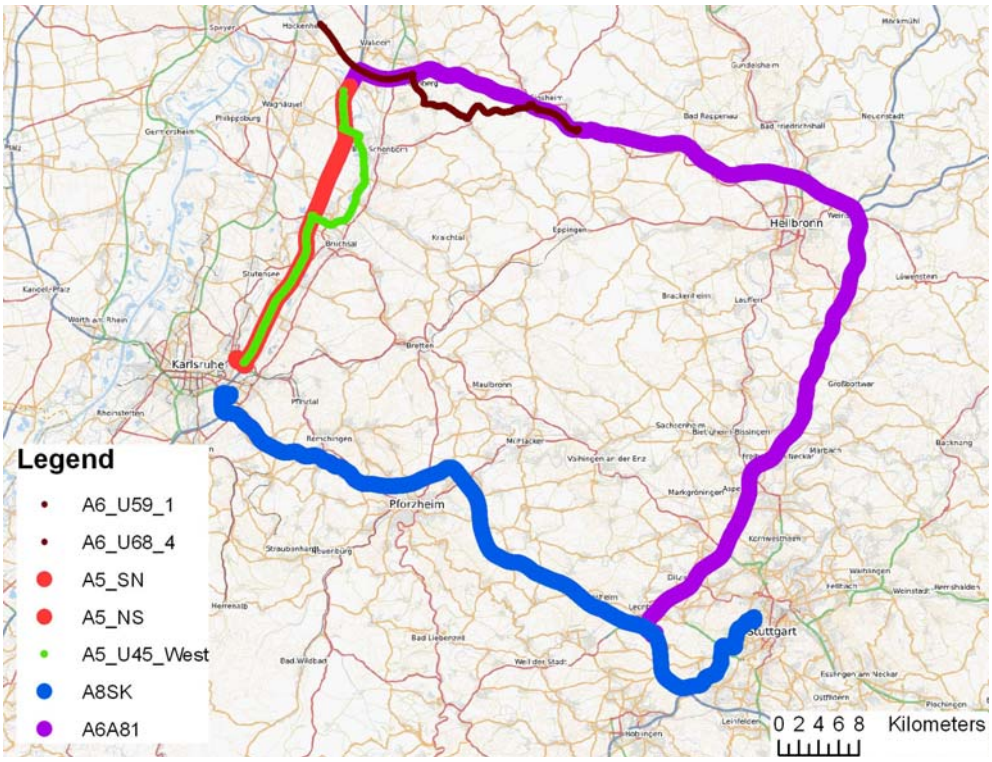


Abbildung 4-1: Autobahnviereck Teilstrecken und Umleitungsstrecken
(Kartenquelle: Open Street Maps)

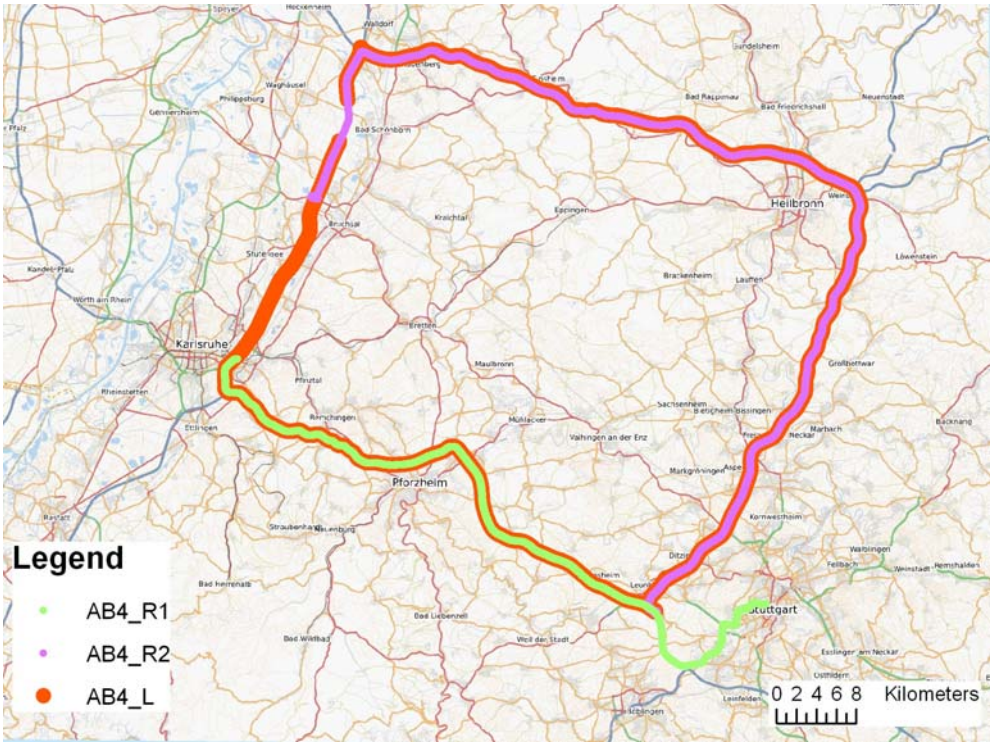


Abbildung 4-2: Autobahnviereck gesamt
(Kartenquelle: Open Street Maps)

Tabelle 4-2: Umleitungsstrecken Leonberger Kreuz vom 09.02.2009 und 03.02.2009

11	A81A8KA	A81-A8 über Leonberger Kreuz (Autobahn)	2009:02.03 / 12:10	20 min
12	A8KA81_h	A8-A81 über Leonberger Kreuz (Autobahn)	2009:02.09 / 14:18	22 min
13	A8KA81_r	A81-A8 über Leonberger Kreuz (Autobahn)	2009:02.09 / 14:43	16 min
14	A8KA81_U_h	A8-A81 über Leonberger Kreuz (Umleitungstrecke)	2009:02.09 / 15:00	29 min
15	A8KA81_U_r	A81-A8 über Leonberger Kreuz (Umleitungstrecke)	2009:02.09 / 15:30	30 min

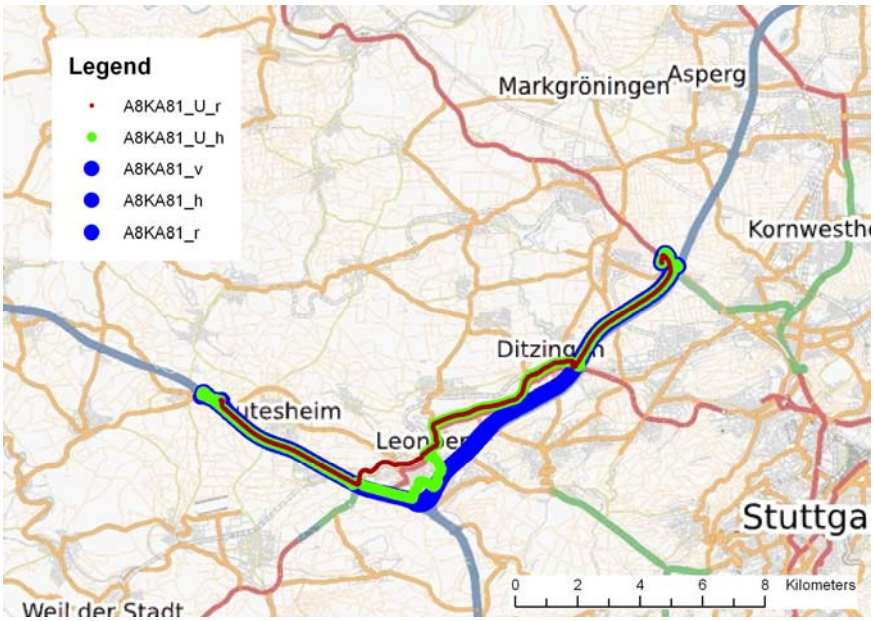


Abbildung 4-3: Umleitungsstrecken Leonberger Kreuz
(Kartenquelle: Open Street Maps)

Tabelle 4-3: Umleitungsstrecken A81 vom 03.02.2009 und 18.03.2009

16	A81_SN	A81 Höhe Feuerbach in Süd – Nord-richtung (Autobahn)	2009:02.03 / 12:30	20 min
17	A81_U13_1	Umleitungsstrecke U13 Höhe Ilsfeld – Mundelsheim in Nord-Südrichtung	2009.03.18 / 11:19	11 min
18	A81_U13_2	A81 Höhe Ilsfeld – Mundelsheim in Nord-Südrichtung	2009.03.18 / 10:30	31 min
19	A81_U13_3	A81 Höhe Ilsfeld – Mundelsheim in Süd-Nordrichtung	2009.03.18 / 11:05	12 min
20	A81_U13_4	Umleitungsstrecke U13 Höhe Ilsfeld – Mundelsheim in Süd-Nordrichtung	2009.03.18 / 9:34	31 min
21	A81_U16_NS	A81-Umleitungstrecke U16	2009:02.03 / 11:11	31 min
22	A81_U3_SN	A81-Umleitungstrecke U3	2009:02.03 / 11:44	26 min

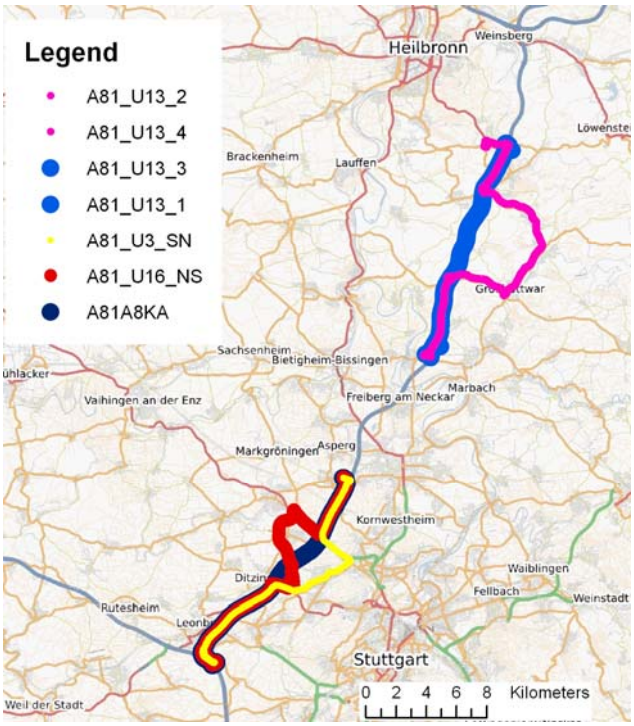


Abbildung 4-4: Umleitungsstrecken Feuerbach und Ilsfeld
(Kartenquelle: Open Street Maps)

Tabelle 4-4: Alternativstrecken Stuttgart vom 03.02.2009 und 09.02.2009

23	B10Feu_NS	Alternativstrecke A81-B10 über AS Feuerbach in Nord-Südrichtung	2009:02.03 / 10:36	33 min
24	B10Feu_SN	Alternativstrecke A81- B10 über AS Feuerbach in Süd-Nordrichtung	2009:02.03 / 9:57	34 min
25	B10Zuff_NS	Alternativstrecke A81- B10 über AS Zuffenhausen in Nord-Südrichtung	2009:02.03 / 9:24	29 min
26	B10Zuff_SN	Alternativstrecke A81- B10 über AS Zuffenhausen in Süd-Nordrichtung	2009:02.03 / 8:43	36 min
27	B14_OW	AS Vaihingen bis Kappelbergtunnel in Ost-Westrichtung	2009:02.09 / 11:01	25 min
28	B14_WO	AS Vaihingen bis Kappelbergtunnel in West-Ostrichtung	2009:02.09 / 10:26	33 min
29	B27_NS	B27 Echterdinger Ei (AS-Flughafen) bis Ludwigsburg in Nord-Südrichtung	2009:02.09 / 12:23	37 min
30	B27_SN	B27 Echterdinger Ei (AS-Flughafen) bis Ludwigsburg in Süd-Nordrichtung	2009:02.09 / 11:40	38 min

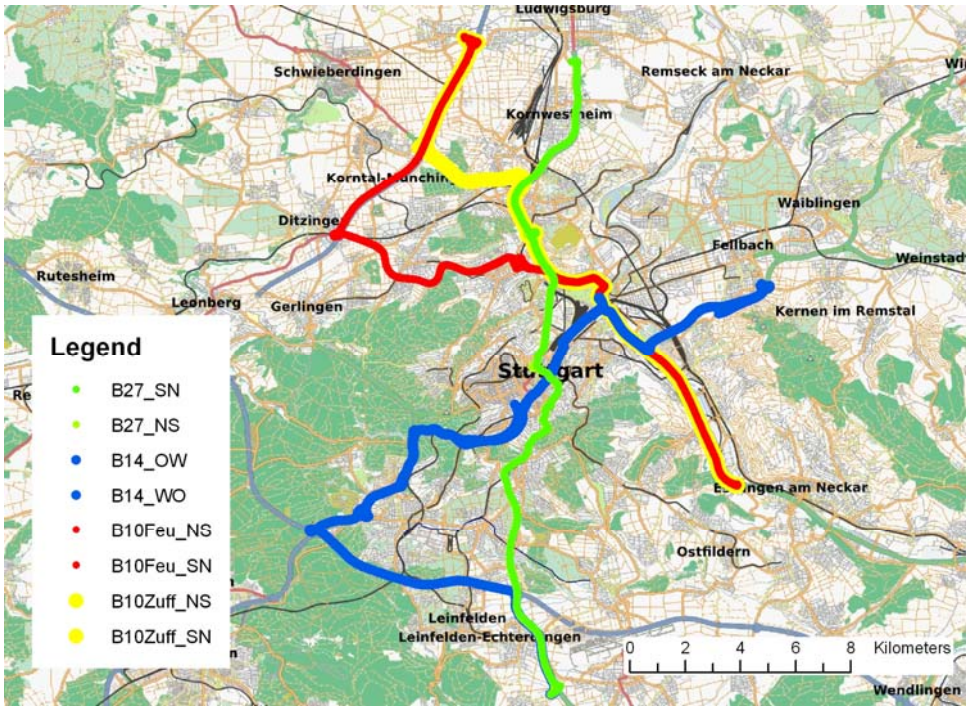


Abbildung 4-5: Alternativstrecken durch Stuttgart (Kartenquelle: Open Street Maps)

Hier schließen sich die Versuchsfahrten zum Vergleich von A-Daten und Abis-Daten an.

Tabelle 4-5: Vergleich Abis-Daten vs. A-daten vom 25.11.2008

Nr	Bezeichnung	Szenario	Datum/Uhrzeit	Dauer
31	2a2	Karlsruhe Südtangente B10 West-Ostrichtung	2008.11.25 / 10:33	7 min
32	2b	Karlsruhe Südtangente B10 Ost-Westrichtung	2008.11.25 / 10:43	9 min
33	2c	Karlsruhe Südtangente B10 West-Ostrichtung	2008.11.25 / 11:10	7 min
34	2e	Karlsruhe Südtangente B10 Ost-Westrichtung	2008.11.25 / 11:20	7 min
35	2f	Karlsruhe Südtangente B10 West-Ostrichtung	2008.11.25 / 11:30	9 min
36	2g	Karlsruhe Südtangente B10 Ost-Westrichtung	2008.11.25 / 11:41	7 min
37	3a	A5 AS-Ettlingen bis AS-Bruchsal in Süd-Nordrichtung	2008.11.25 / 13:45	16 min
38	3b	A5 AS-Ettlingen bis AS-Bruchsal in Nord-Südrichtung	2008.11.25 / 14:10	20 min
39	3d	A5 AS-Ettlingen bis AS-Bruchsal in Nord-Südrichtung	2008.11.25 / 14:56	10 min
40	4a	Parallel zur Straßenbahn (Innenstadt)	2008.11.25 / 13:09	14 min
41	5a	Karlsruhe Kriegstraße in Nord-Südrichtung	2008.11.25 / 15:11	14 min
42	5b	Karlsruhe Kriegstraße in Süd-Nordrichtung	2008.11.25 / 15:29	16 min

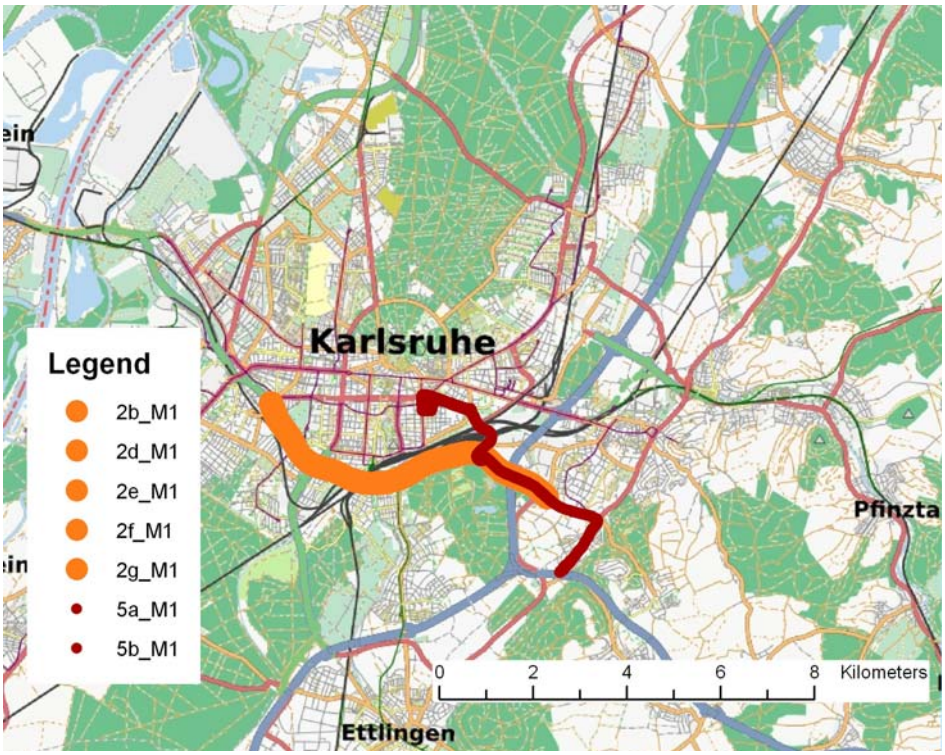


Abbildung 4-6: Vergleich Abis-Daten und A-Daten (Kartenquelle: Open Street Maps)

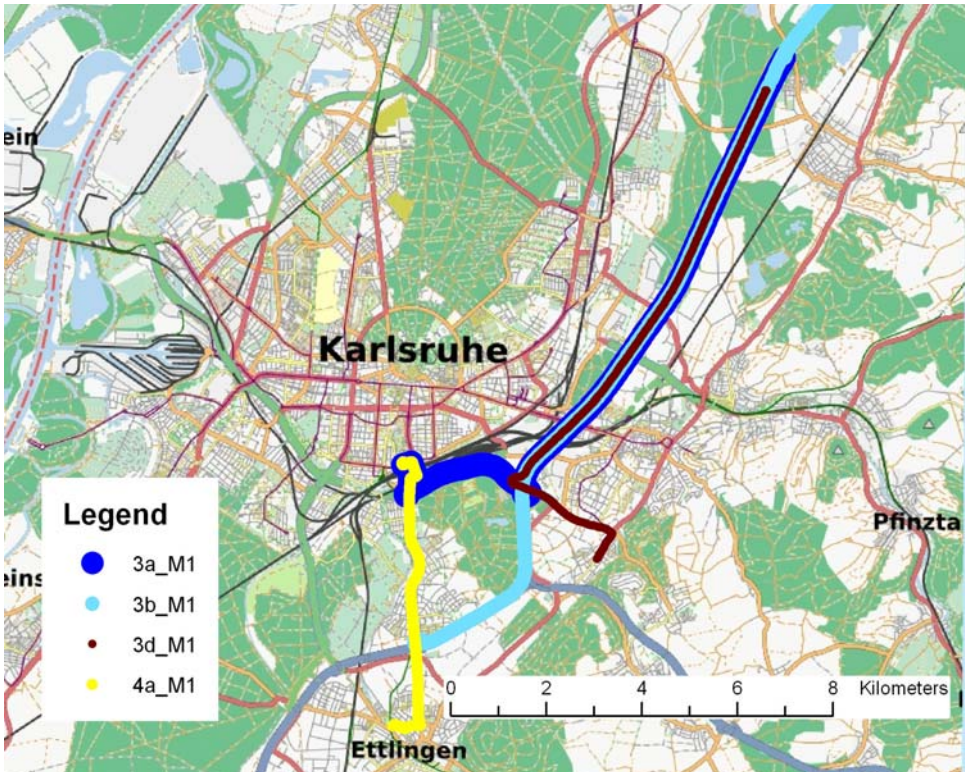


Abbildung 4-7: Vergleich Abis-Daten und A-Daten (Kartenquelle: Open Street Maps)

4.2 Stationäre Erfassungssysteme

Die Projektpartner haben eine große Zahl an Stationären Erfassungssystemen als Referenz zur Verfügung gestellt. Es handelt sich dabei sowohl um Doppelschleifen, wie sie insbesondere auf den Autobahnen betrieben werden, als auch um Einfachschleifen und passive Infrarotsensoren (TEU). Letztere werden vorwiegen innerstädtisch sowie auf Durchgangsstraßen eingesetzt. Aus technischer Sicht eignen sich grundsätzlich alle Sensoren im Projektnetz als Referenz für die Verkehrsstärke, allerdings scheiden einige innerstädtische Sensoren aufgrund der Lage und/oder Anordnung des Sensors aus. Als Referenz für die durchschnittliche Verkehrsgeschwindigkeit hingegen sind die innerstädtischen Schleifensensoren nicht geeignet im Gegensatz zu den Doppelschleifen sowie den TEU.

Im Folgenden werden die verwendeten Sensoren der Projektpartner sowie deren Lage im Projektnetz dargestellt. Es handelt sich dabei sowohl um Doppelschleifen, wie sie insbesondere auf den Autobahnen betrieben werden, als auch um Einfachschleifen und passive Infrarotsensoren (TEU[5]). Letztere werden vorwiegend innerstädtisch sowie auf Durchgangsstraßen eingesetzt.

Die Qualität der verwendeten stationären Erfassungssysteme war ausreichend. Die Korrektheit der Verkehrsstärke q, die mit Doppelschleifen auf der A81 erfasst wird, konnte bereits im Rahmen einer Studienarbeit exemplarisch nachgewiesen werden. Es wurde anhand von punktuellen Referenzzählungen eine maximale Abweichung von 1 % zur tatsächlichen Verkehrsstärke ermittelt [Karrer 2008]. Diese Ergebnisse können aufgrund der ähnlichen Bau- und Funktionsweise aller Sensoren auf den Autobahnen im Projektnetz auf diese übertragen werden. Wie eine Studie von [Sachse 2002] beispielhaft an zwei TEU-Sensoren nachweist, weichen die ermittelten Verkehrsstärken in ca. 95 % der Testintervalle nicht mehr als 1Fz/5Min von der tatsächlichen Verkehrsstärke ab und erfüllen damit die Anforderungen der Bundesanstalt für Straßenwesen [TLS 1997]. Die Geschwindigkeitsmessungen der Sensoren im Autobahnnetz sind aufgrund der Bauweise als Doppelschleifen als ausreichend genau anzusehen. Für die stadtnah installierten TEU liefert die Untersuchung von [Sachse 2002] eine erreichbare Genauigkeit von ±10 km/h für die mittlere Fahrzeuggeschwindigkeit. Wie sich jedoch bereits in den Untersuchungen mit GPS-Referenz herausgestellt hat, wird diese Genauigkeit zur Beurteilung der aus FPD ermittelbaren Geschwindigkeiten in Stadtnähe ausreichen, da sich für die FPD Werte im günstigsten Fall (lange Trajektorien, außerstädtisch) Genauigkeiten von ±12,5 km/h ergeben haben. Innerstädtisch wurden jedoch wesentlich schlechtere Ergebnisse erzielt.

Damit können aus funktionaler Sicht grundsätzlich alle in Do-iT zur Verfügung stehenden Verkehrssensoren zur Evaluierung herangezogen werden. Einige Sensoren insbe-

[5] TEU: Traffic Eye Universal - Es handelt sich dabei um passive Infrarotsensoren der Firma Siemens

sondere innerstädtisch müssen jedoch aufgrund der ungünstigen Lage oder nur teilweiser Abdeckung des Fahrbahnquerschnitts im Bereich von Kreuzungen zunächst ausgesondert werden oder eignen sich nur als Referenz für die Verkehrsstärke und nicht für die mittlere Geschwindigkeit.

Zur groben geografischen Orientierung ist auf der folgenden Abbildung 4-8 zunächst das gesamte Projektnetz dargestellt. Sämtliche Straßenelemente (REIDs) auf denen während des Untersuchungszeitraumes stationäre Sensoren Referenzdaten geliefert haben, sind rot hervorgehoben. Die grau hinterlegte Fläche markiert die beiden LACs KA und Ettlingen, in denen Abis-Daten zur Verfügung standen. In diesem Bereich war sowohl die Auswertung von A- als auch Abis-Daten möglich. Im restlichen Projektgebiet standen aus technischen Gründen nur die A-Daten zur Verfügung.

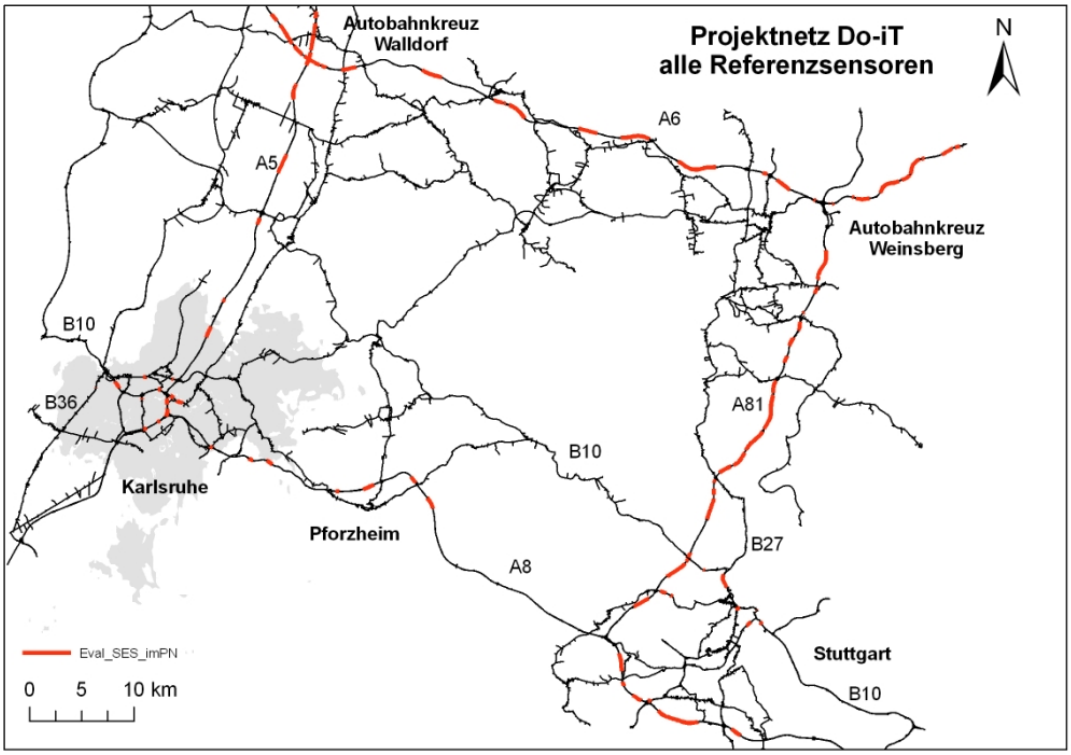


Abbildung 4-8: Übersicht über das gesamte Projektstraßennetz mit rot hervorgehobenen Road Elements (RE) der digitalen Straßenkarte auf denen geeignete SES installiert sind

In den folgenden Darstellungen der Sensorstandorte wird zur Unterscheidung jeweils die RMQ-ID (ID des Regelmessquerschnitts[6]) der Sensoren mit angegeben. Diese projektinterne Bezeichnung wurde eingeführt, um sämtliche Sensoren der beteiligten Partner eindeutig identifizieren zu können. Daher werden im Weiteren nur noch diese Bezeichnungen verwendet.

[6] Als Regelmessquerschnitt (RMQ) wird die Zusammenfassung aller Spuren in einer Fahrtrichtung bezeichnet

Sensoren der DDG (Autobahnnetz)

Die DDG stellte insgesamt 122 Sensoren zur Verfügung, von denen 92 im Untersuchungszeitraum Daten geliefert haben. Es handelt sich vorwiegend um Doppelschleifen auf dem Autobahnnetz, wobei Verkehrsstärke und Durchschnittsgeschwindigkeit i. d. R. aus Messungen auf der Überholspur abgeleitet werden. Wie Langzeiterfahrungen gezeigt haben, stimmen die hochgerechneten Werte jedoch sehr gut mit den realen Verhältnissen überein, weshalb die daraus extrapolierte Verkehrsstärke und der generierte LOS (Level of Service) die tatsächlichen Straßenverhältnisse auf allen Spuren des Messquerschnitts (MQ) sehr gut wiedergibt. Die Verteilung der 92 Sensoren auf dem Projektnetz ist in Abbildung 4-9 dargestellt, die Straßenelemente (Road Elements, RE) auf denen Sensoren installiert sind, wurden rot hervorgehoben. Aus Platzgründen konnten in der Übersicht nicht alle Regelmessquerschnitte mit Ihrer RMQ-ID dargestellt werden. Auffällig ist jedoch die teilweise ungleichmäßige Verteilung mit einer hohen Sensordichte auf der A81 sowie in der Nähe des Autobahnkreuz Weinsberg im Nordosten sowie des Autobahnkreuzes Walldorf im Nordwesten des Projektgebietes. Auf der A5 sowie auf der A8 ist die Sensordichte eher klein. Dies liegt teilweise an dem Ausschluss von Sensoren, die im Untersuchungszeitraum keine Daten geliefert haben. Auf der A8 zwischen PF und KA sowie Auf der A6 zwischen Steinsfurt und Neckarsulm befinden sich jedoch echte Sensorlücken. Beide sind vermutlich auf aktuell noch bestehende und erst kürzlich abgeschlossene Bautätigkeiten zurückzuführen.

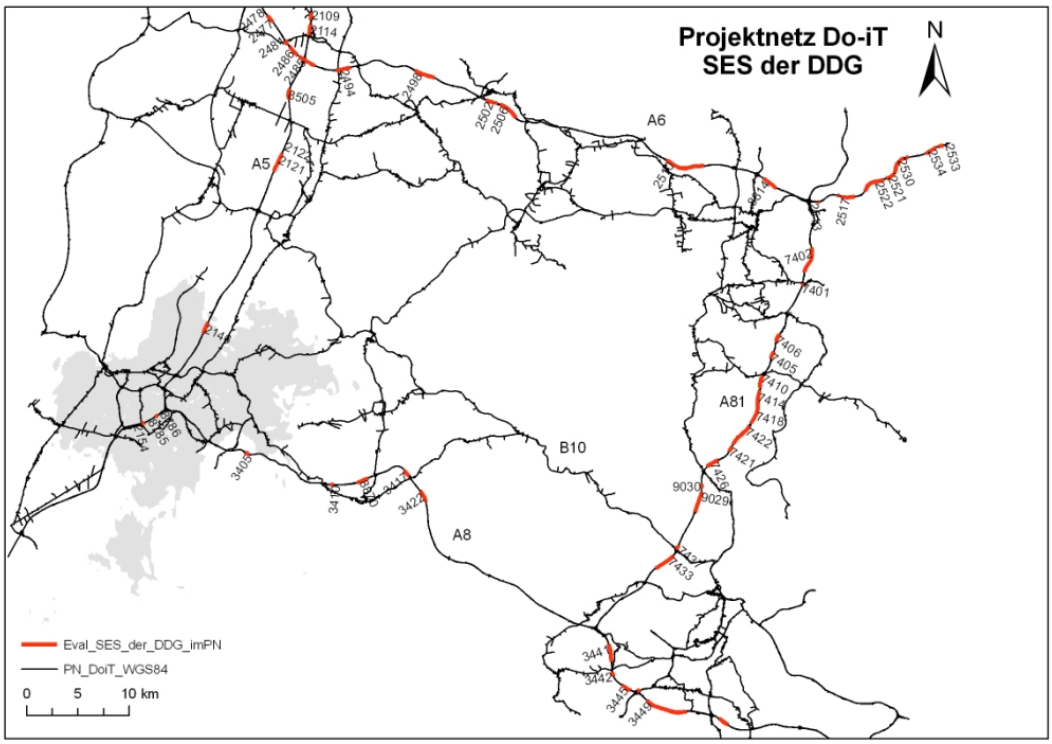


Abbildung 4-9: Übersicht aller verwendeten SES der DDG mit ihrer RMQ-ID auf dem Projektnetz (aus Platzgründen sind nicht alle RMQ_IDs angegeben)

Nach dem technischen Defekt am 24.03., der zum Ausfall von 20 LACs geführt hat, können die Sensoren auf der A81 nicht mehr zur Evaluierung herangezogen werden. Dasselbe gilt für die Sensoren auf der A8 im Raum Pforzheim.

Sensoren der Landesstelle für Straßentechnik BW (Autobahnnetz)

Wie in Abbildung 4-10 ersichtlich, wurden Sensoren der Landesstelle für Straßentechnik ebenfalls aus dem gesamten Projektnetz zur Verfügung gestellt. Insgesamt standen 91 RMQ zur Verfügung. Grundsätzlich waren alle Sensoren der Landesstelle als Referenz geeignet, allerdings standen nur von 56 Sensoren im Evaluierungszeitraum Ende März Messdaten zur Verfügung. Hier fielen einerseits einige Sensoren auf der A8 vermutlich aufgrund des bereits erwähnten Autobahnausbaus sowie einige im Bereich des Kreuz Weinsberg aus.

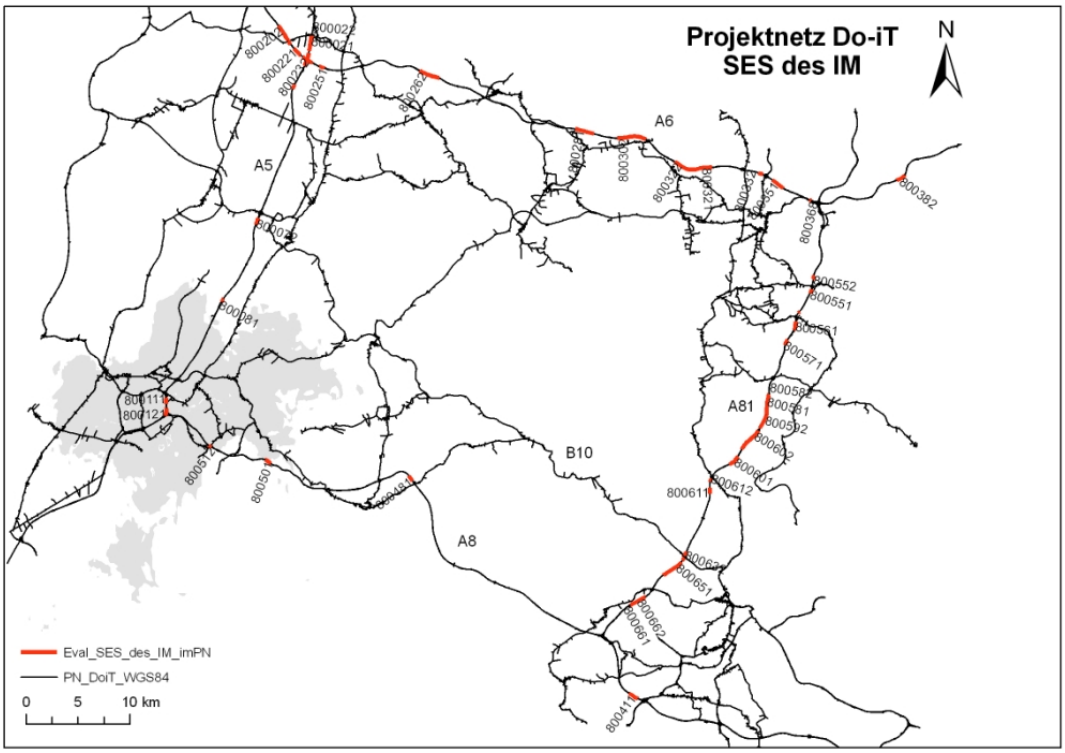


Abbildung 4-10: Projektnetz mit allen MQ der Landesstelle für Straßentechnik BW in Do-iT (aus Platzgründen sind nicht alle RMQ_ID angegeben).

Da die Messquerschnitte vorwiegend der Steuerung einer Netzbeeinflussungsanlage im Autobahnviereck sowie statistischen Zwecken dienen, wird an den einzelnen Messquerschnitten jeweils richtungsscharf der gesamte Verkehr auf allen Spuren erfasst. Es handelt sich ausschließlich um Doppelschleifen, die neben der Verkehrsstärke auch gute Durchschnittsgeschwindigkeiten liefern.

Das gesamte Autobahnviereck im Projektnetz wird somit mit insgesamt 153 Messquerschnitten sehr dicht abgedeckt, auch wenn einzelne Abschnitte aus verschiedenen Gründen im Untersuchungszeitraum nicht überwacht werden konnten. Damit wird eine sehr gute Referenz zur Evaluierung der aus A-Daten generierten FPD-Trajektorien auf dem Autobahnnetz zur Verfügung gestellt. Die Qualitätsaussagen können aufgrund der vergleichbaren Infrastrukurdichte auf die weniger dicht abgedeckten Autobahnabschnitte übertragen werden und es ist daher möglich, allgemeine Aussagen zur Eignung von FPD zur Bestimmung der Verkehrsstärke auf Autobahnen zu treffen.

Auch hier konnten die Sensoren auf der A81 auf Grund des technischen Defekts bei der T-Mobile ab dem 24.03. nicht mehr zur Evaluierung verwendet werden, da die Berechnung von FPD-Trajektorien nicht mehr möglich war.

Sensoren der Stadt Karlsruhe

In Karlsruhe wurden nur TEU Sensoren ausgewählt, da die Induktionsschleifen meist nur einzelne Fahrspuren abdecken und sich die Sensoren i. d. R. unmittelbar vor oder nach Kreuzungen befinden. Daher sind sie weder als Referenz für die Verkehrsstärke des gesamten Richtungsmessquerschnitts noch für die Durchschnittsgeschwindigkeit geeignet. Abbildung 4-11 zeigt in rot hervorgehoben die mit einem geeigneten Sensor besetzen Straßenelemente.

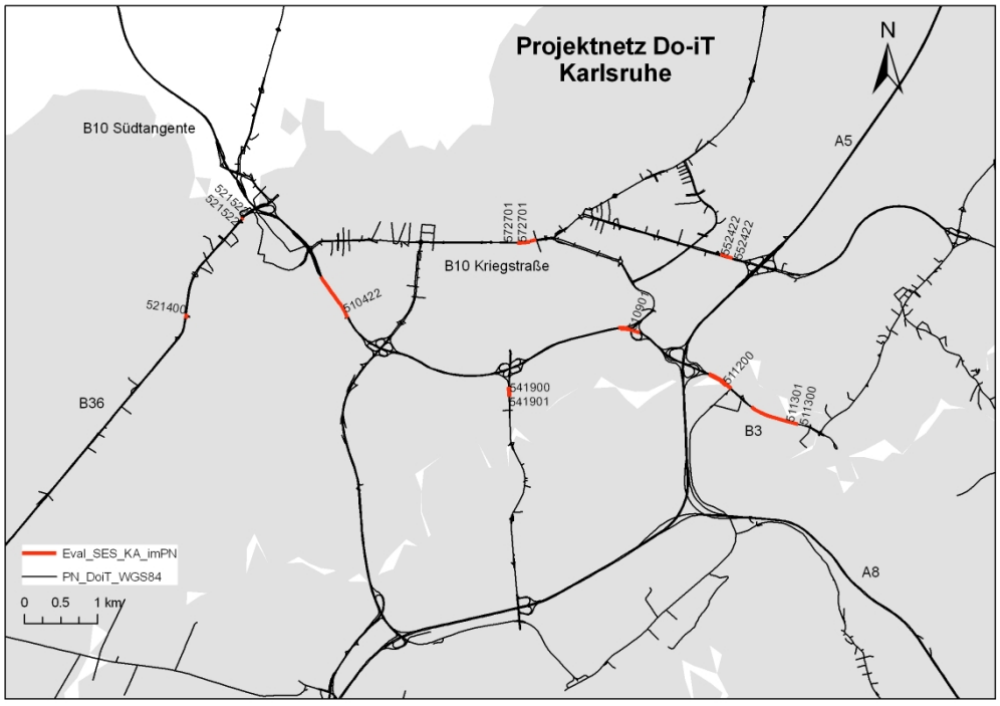


Abbildung 4-11: Projektnetz mit allen verwendeten TEU sowie ihrer RMQ-ID in Karlsruhe

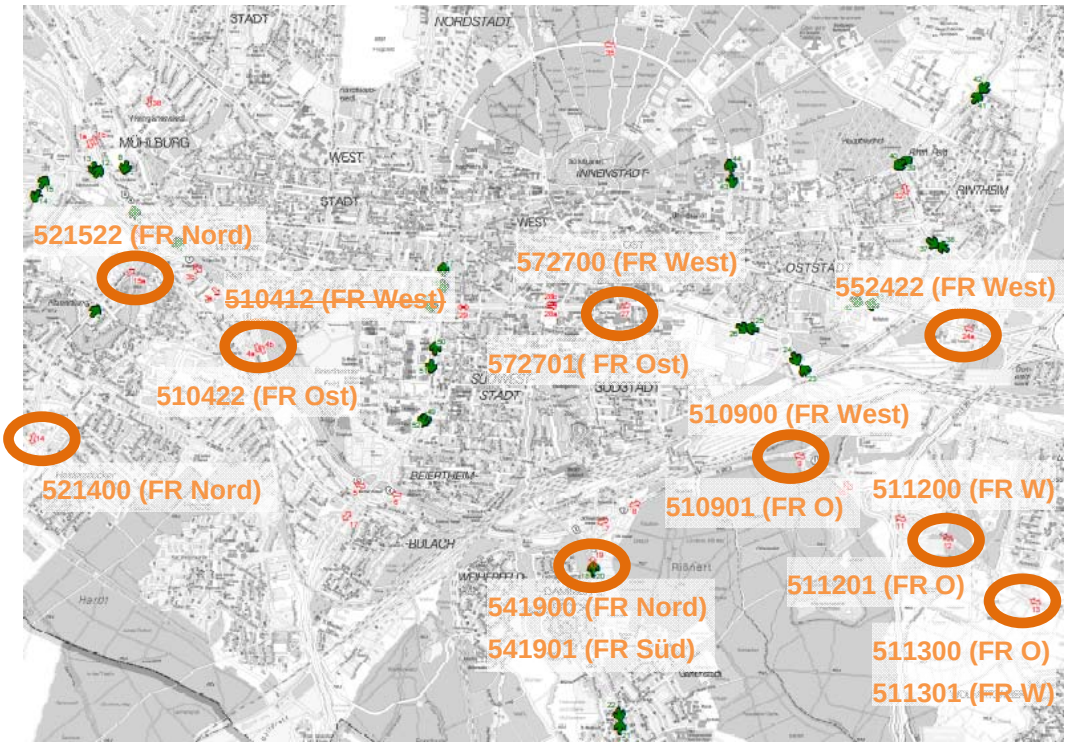


Abbildung 4-12: Übersicht der ausgewählten TEU-Sensoren in Karlsruhe mit Angabe der Fahrtrichtungen (Quelle: Stadt Karlsruhe)

Aufgrund der Lage als besonders geeignet eingestuft wurden die in der Abbildung 4-12 orange hervorgehobenen und mit ihrer projektinternen RMQ-ID dargestellten Infrarotsensoren. Alle Sensoren liegen auf den Hauptein- und -ausfallstraßen sowie auf der Südtangente als verkehrstechnisch wichtige Ost-West Verbindung zwischen den Ländern Baden-Württemberg und Rheinland-Pfalz bzw. als Verknüpfung der Autobahn A8 mit der A65. Es handelt sich um 4 Messquerschnitte auf der kreuzungsfreien Südtangente B10/K9657 (RMQ-ID: 510421/422, 510900/901, 511200/01 und 511300/01) sowie um Messstellen auf den Einfallsstraßen B36 aus Südwest stadteinwärts (521400 und 521522), auf der L561 im Süden stadtein- und -auswärts (541900/01) sowie der Durlacher Allee aus Osten kommend (552422). Zusätzlich wurde eine Messstelle auf der Kriegsbergstraße in der Stadtmitte hinzugenommen (572700/01). Der Sensor mit der RMQ-ID 510412 auf der Südtangete hat im Untersuchungszeitraum keine Daten geliefert. Damit standen insgesamt 14 RMQ zur Verfügung.

Sensoren der Stadt Stuttgart

In Stuttgart (LHS) standen nur Schleifensensoren zur Verfügung. Allerdings sind – anders als in Karlsruhe – viele Schleifensensoren kreuzungsfern installiert und als Doppelschleifen ausgeführt und bilden daher gut den fließenden Verkehr ab. Insgesamt wurden zunächst 20 Richtungsmessquerschnitte mit insgesamt 46 einzelnen Spursensoren auf 34 einzelnen Road Elementen der digitalen Karte ausgewählt. Von diesen grundsätzlich zur Evaluierung geeigneten Sensoren standen jedoch viele im Untersu-

chungszeitraum nicht zur Verfügung. Daher beschränkt sich die Auswahl auf die in den folgenden Abbildungen dargestellten 11 Richtungsmessquerschnitte.

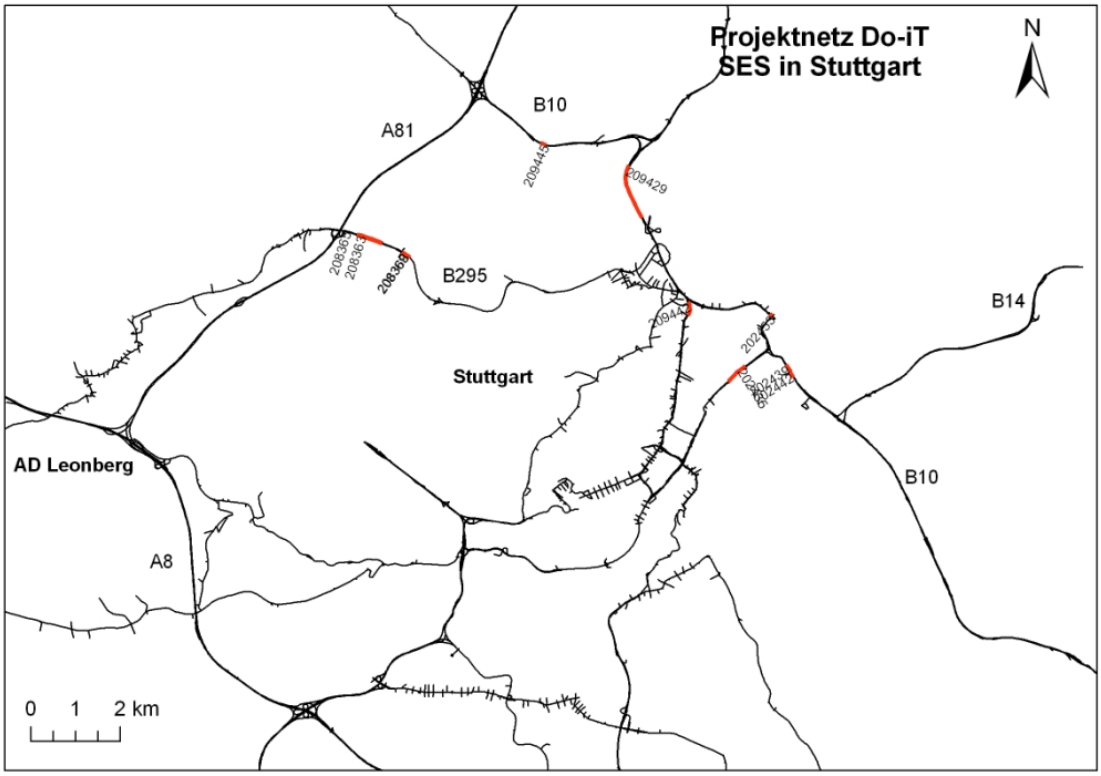


Abbildung 4-13: Übersicht aller verwendeter RMQ in Stuttgart im Projektnetz

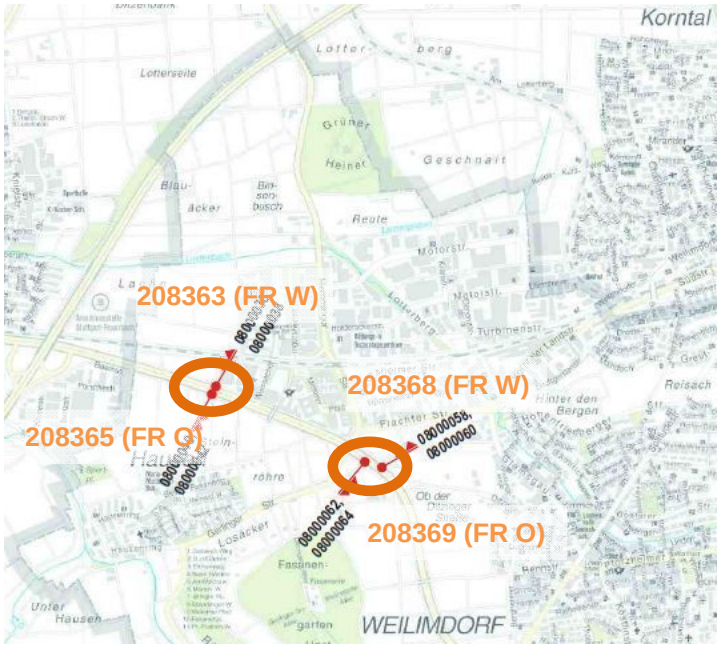


Abbildung 4-14: Lage der ausgewählten Sensoren der LHS – Zubringer B295 zu A81 (Kartenquelle: Stadt Stuttgart)

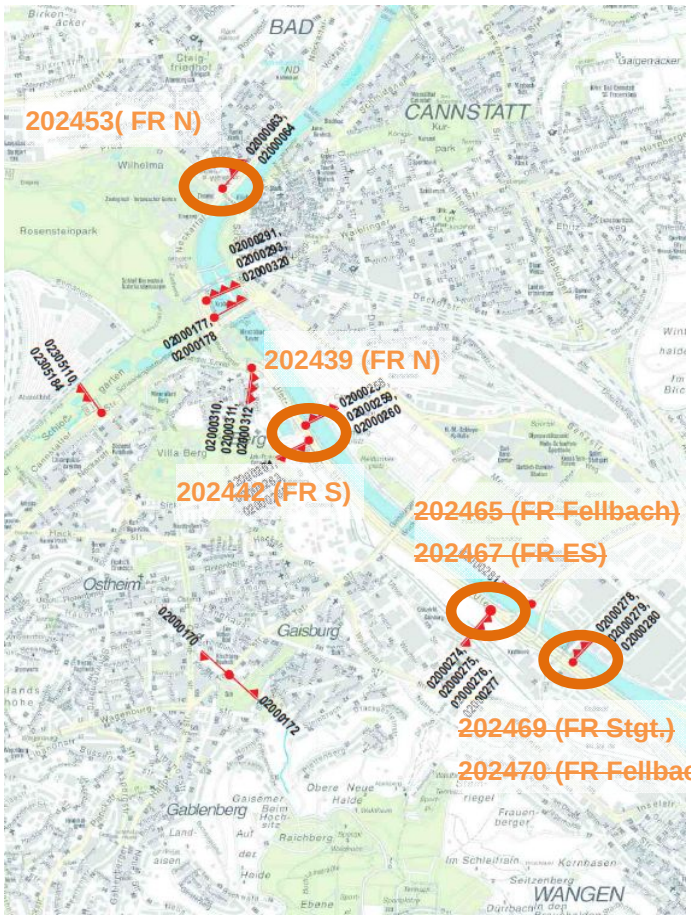


Abbildung 4-15: Lage der ausgewählten Sensoren der LHS – B10, Uferstraße (Kartenquelle: Stadt Stuttgart)

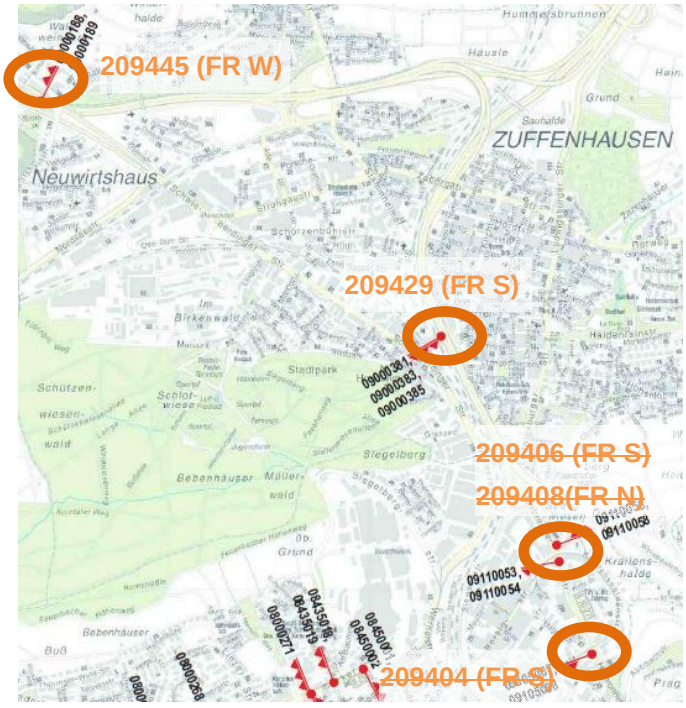


Abbildung 4-16: Lage der ausgewählten Sensoren der LHS – Bereich Pragsattel (Kartenquelle: Stadt Stuttgart)

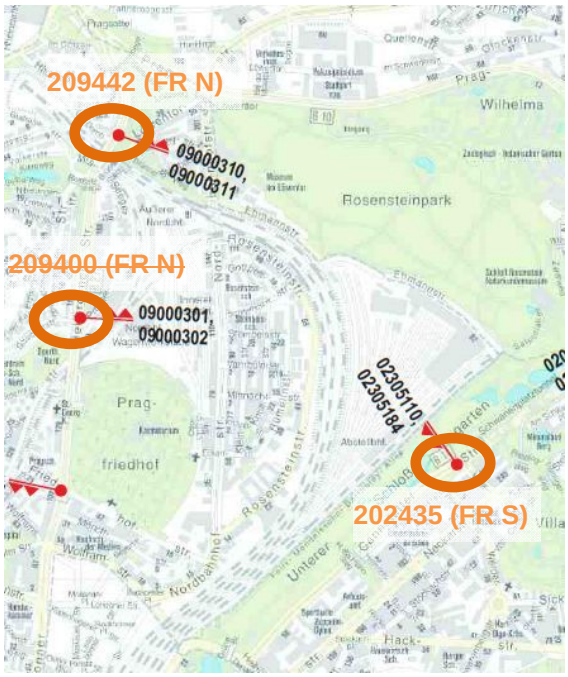


Abbildung 4-17: Lage der ausgewählten Sensoren der LHS – Bereich Nordbahnhof (Kartenquelle: Stadt Stuttgart)

Es wurden Sensoren auf den innerstädtischen Haupttransversalen wie B10/B14 sowie einige Stuttgarter Ein- und Ausfallstraßen wie die B295 gewählt, da auf diese überregionale Bedeutung haben. Daher wurden zwei Sensoren auf dem Autobahnzubringer B295 bei Korntal (208363/65 und 208368/69), wie in Abbildung 4-14 dargestellt, sowie die 202439/42 und 202453 auf der B10 entlang des Neckarufers (vgl. Abbildung 4-15) ausgewählt. Die Sensoren an dem Autobahnzubringer B10/Zuffenhausen wurden ebenfalls genutzt (209445), um die Verteilung des Verkehrs auf die drei alternativen Zufahrten von Stuttgart-Zentrum zur Autobahn A81 nach Norden (über AS Feuerbach, AS Zuffenhausen oder AS Ludwigsburg) erfassen zu können.

Die Erfassung der Verkehrslage auf dem Pragsattel (209429, Abbildung 4-16) sowie auf dessen unmittelbarer Zufahrt (209442, Abbildung 4-17) ist auf Grund der enormen täglichen Verkehrslasten und der damit verbundenen Staugefahr aus verkehrstechnischer Sicht von enormer Bedeutung. Daher wurden hier die gezeigten Sensorstandpunkte ausgewählt. Aus demselben Grund wurde schließlich auch noch der Messquerschnitt 202435 auf der B14/Cannstatter Straße stadteinwärts in die Untersuchungen mit einbezogen (vgl. Abbildung 4-17).

Zur Evaluierung der Grenzen der Mobilfunkortung wurden auch 6 aufeinanderfolgende Messquerschnitte (MQ14-MQ19) auf der B14 im Heslacher Tunnel ausgewählt. Die ungefähre Lage im einspurigen Tunnel ist in der Abbildung 4-18 ersichtlich. Leider waren Ende März keine Daten der RMQ im Heslacher Tunnel verfügbar, daher sind alle RMQ-IDs gestrichen dargestellt.

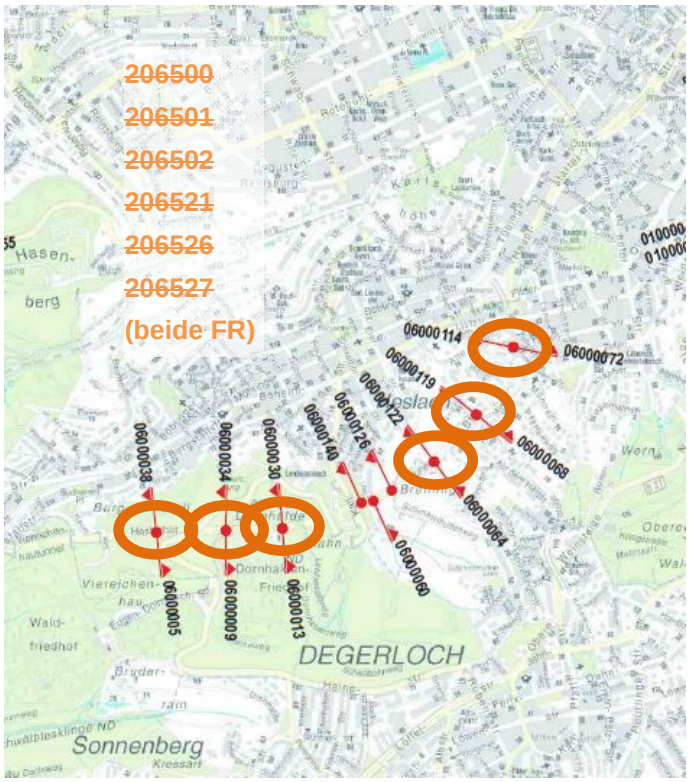


Abbildung 4-18: Lage der ausgewählten Sensoren der LHS – B14/Heslacher Tunnel (Kartenquelle: Stadt Stuttgart)

Die Messdaten der ausgewählten Stationären Erfassungssysteme der Projektpartner standen weitgehend vollständig ab Februar 2009 täglich zur Verfügung. Als Evaluierungszeitraum wurde der März 2009 festgelegt, da in diesem Zeitraum der endgültige Probelauf der A- und Abis-Datenverarbeitung stattgefunden hat und die letzte Softwareversion zur A-Daten Generierung in Fellbach installiert war.

Da der Raum Stuttgart direkt von dem Ausfall der 20 LACs ab dem 24.03. betroffen war, konnten hier für die restliche Zeit des Untersuchungszeitraumes keine FPD-Trajektorien mehr berechnet werden. Der Vergleich mit den Sensoren der Stadt Stuttgart beschränkt sich daher auf Montag, den 23.03.09.

Im folgenden Kapitel werden die vorgestellten Referenzsysteme zur Beurteilung der Qualität der FPD-Trajektorien herangezogen. Die wesentlichen Ergebnisse der Untersuchungen sind dort dargestellt. Für weitergehende Details wird auf den Bericht [AP10-2 2009] verwiesen.

5 Qualitätsbeurteilung

Die Qualitätsbeurteilung der im Rahmen der Projektlaufzeit erzielten Ergebnisse wird durch Vergleiche mit den im Abschnitt 4 vorgestellten Referenzdaten erzielt. Dabei wird unterschieden in die Untersuchung von exemplarisch durchgeführten Testfahrten, die den Vergleich mit GPS als höherwertiges Ortungsverfahren zulassen, jedoch nur einen minimalen Stichprobenumfang darstellen und dem Vergleich mit Daten aus stationär installierten Verkehrsdatenerfassungssystemen. In letztem Fall handelt es sich um eine punktuelle Vollerhebung der Verkehrskenngrößen Verkehrsstärke und –durchschnittsgeschwindigkeit, die nahezu lückenlos während der Evaluierungsphase als Referenz zur Verfügung stand. Es war daher punktuell eine Evaluierung der gesamten aus Mobilfunkdaten berechenbaren Trajektorien und den daraus ermittelten Verkehrsstärken möglich.

5.1 Vergleich mit GPS-Trajektorien

Zur objektiven Beurteilung der aus realen Mobilfunkdaten bestimmten Trajektorien wurden für Abis-Daten im Stadtgebiet Karlsruhe und mit A-Daten im gesamten Projektgebiet Testfahrten mit GPS durchgeführt. Die GPS-Messungen liefern dabei eine Referenz für Zeit und Position der Mobilstation, die sich aufgrund ihrer höheren Genauigkeit dazu eignet, die Qualität der FPD-Trajektorien absolut zu bewerten. Dazu wurden im Qualitätsmodell zur Beurteilung der FPD-Routen die Parameter Zuordnungskorrektheit Typ A und Zuordnungskorrektheit Typ B und zur Beurteilung des Weg-Zeit-Diagramms die Standardabweichung der Geschwindigkeit und die vorzeichenfreie Differenz der Durchschnittsgeschwindigkeit definiert (vgl. Kapitel 3).

5.1.1 Abis-Ebene

Zusätzlich zu diesen aus einer Referenz bestimmbaren Qualitätsparametern kann bei Abis-Daten die Qualität der Route auch ohne GPS-Referenz im Online-Betrieb beurteilt werden anhand der aus den gemessenen Positionen und den Lotfußpunkten berechneten „Mittleren Querabweichung QA_{FPD} [m] (P6.1)“und der „Trajektorienvollständigkeit VO_{FPD} [%] (P3.1)“, der den Anteil der Positionen, die für die Routengenerierung verwendet wurde, angibt.

Abbildung 5-1 stellt diese beiden rein aus den Mobilfunkdaten geschätzten Qualitätsparameter dem Anteil der richtigen Zuordnung auf die GPS-Referenz (Korrektheit Typ B) gegenüber. Der Typ B wurde verwendet, da es sich hierbei um den im Vergleich zum Typ A deutlich kritischeren Qualitätsparameter handelt. Eine eindeutige Korrelation zwischen Qualitätsparametern mit und ohne Referenzdaten ist nicht möglich. Als erste Abschätzung kann gelten, dass die Verwendung von mehr als 97 % der Positionen zu überwiegend (70 %) korrekten Routen führen. Unter Verwendung dieses empirisch

bestimmten Grenzwertes würde sich nur für eine Testfahrt ein Wert von 69 % ergeben. Zieht man als zweiten Parameter die mittlere Querabweichung heran und legt hierfür einen Grenzwert von < 250 m fest, so kann auch diese Trajektorie zu recht ausgeschlossen werden. Allerdings wird auf diesem Wege eine korrekte Zuordnung vom Datenmaterial eliminiert.

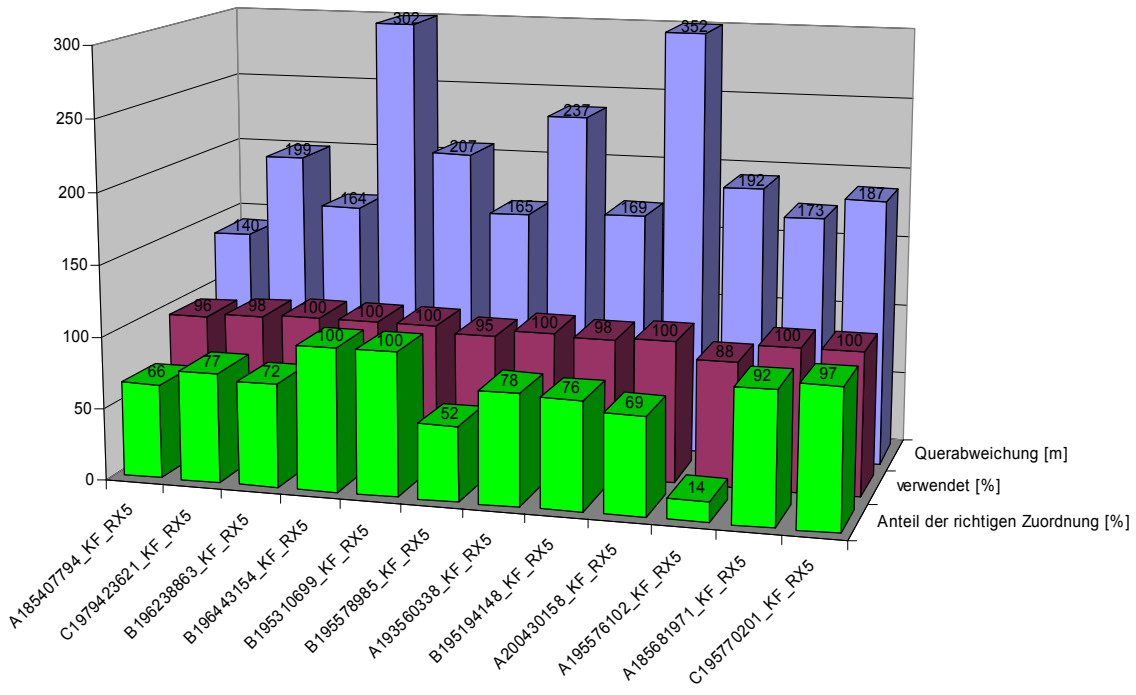


Abbildung 5-1: Qualitätsparameter mit und ohne Referenzdaten, grün = Korrektheit Typ B

Zur Beurteilung des Weg-Zeit-Diagramms werden die Standardabweichung der Geschwindigkeit (vgl. P6.2 im Qualitätsmodell) und die vorzeichenfreie Differenz der Durchschnittsgeschwindigkeit, die ebenfalls im Kapitel 3 definiert ist, verwendet.

Anhand dieser beiden Beurteilungskriterien für die Geschwindigkeit wurde aus den Testfahrten mit GPS-Referenz festgelegt, dass die Diskretisierung des Weg-Zeit-Diagramms am günstigsten mit einem Mittelungsintervall von 30 s erfolgt. Für dieses Diskretisierungsintervall liegen die Absolutwerte der Differenzen der Durchschnittsgeschwindigkeiten im Mittel über alle Testfahrten bei 12.5 km/h (≅3.5 m/s). Diese hohe Genauigkeit ist aber in den relativ langen Trajektorien der Testfahrten begründet und kann nicht auf kürzere Telefonate verallgemeinert werden.

Um den Einfluss der Gesprächsdauer zu untersuchen, wurden die langen Gespräche der Testfahrten in konstante Zeitintervalle von 60 s bis 240 s unterteilt. Indem die Anfangszeiten der Gespräche gleichmäßig über die Testfahrten verteilt wurden, konnten so jeweils 120 verschiedene Datensätze erzeugt werden. Aufgrund von Fahrzeugstillständen bzw. Langsamfahrten konnten bei kurzen Gesprächen rund ein Drittel der Positionsfolgen aufgrund zu geringer Fahrstrecken keine Trajektorien gefunden wer-

den. Ab einer Gesprächsdauer von 2 Minuten konnten bei mehr als 90 % der Positionsfolgen Trajektorien bestimmt werden.

Tabelle 5-1: Gesprächsdauer und Anzahl bestimmter Trajektorien für Beispieltag

Gesprächsdauer [s]	**60**	**80**	**100**	**120**	**140**	**160**	**180**	**210**	**240**	**Σ**
Anzahl Telefonate	120	120	120	120	120	120	120	120	120	1080
Anzahl Trajektorien	78	93	111	108	113	114	113	114	107	951

Korrektheit der Routen

Die richtige Zuordnung der aus Mobilfunkdaten bestimmten Positionsfolgen auf die Karte kann durch die Referenzmessungen mit GPS anhand der Zuordnungskorrektheiten Typ A und B beurteilt werden. In der Abbildung 5-2 zeigt die Korrektheit der Zuordnung keine eindeutige Abhängigkeit von der Gesprächsdauer. Dies liegt an den unterschiedlichen Geschwindigkeiten der einzelnen Testfahrten. Die Korrektheit hängt vom zurückgelegten Weg also der Trajektorienlänge ab. In einigen nachfolgenden Abbildungen wird zur besseren Anschaulichkeit an Stelle der Korrektheit Typ B die entsprechende Fehlerrate bzw. Unkorrektheit (als 1 – Typ B bezeichnet) dargestellt.

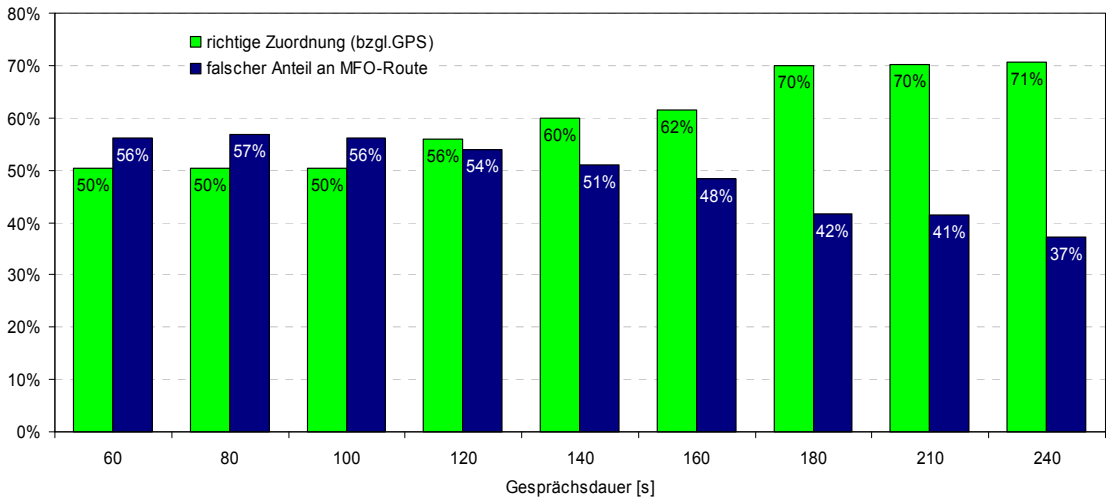


Abbildung 5-2: Korrektheit in Abhängigkeit von der Gesprächsdauer, grün = Typ A, blau = 1 – Typ B (Unkorrektheit)

Die nachfolgende Abbildung zeigt die Korrektheit der Zuordnung bezüglich der mit GPS bestimmten Route (Fehlertyp A): Ein Drittel der 2 km langen Positionsfolgen stimmen zu mindestens 80% mit der GPS-Route überein. Für 3.5 km Länge stimmen mehr als die Hälfte zu mindestens dem gleichen Prozentsatz überein. Eine 50 % Über-

einstimmung mit der GPS-Route wird bei 28 % der Positionsfolgen kürzer 1 km erreicht. Dieser Anteil steigt bei 2 km langen Wegen auf 60% und bei 3 km langen Wegen auf 72 %. Die unzureichende Korrektheit bzgl. der GPS-Route führt zwar zu Informationsverlusten, aber nicht zu Fehlinformationen (Fehlertyp A).

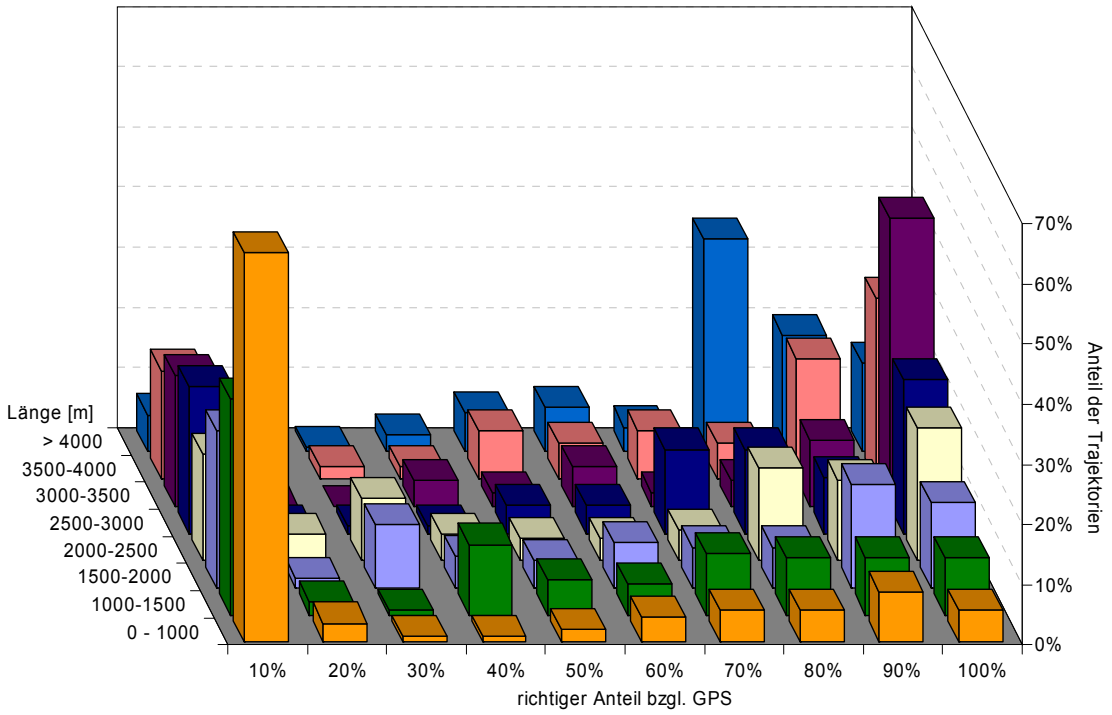


Abbildung 5-3: Anteil der richtigen Zuordnung im Vergleich zu GPS

Zur Beurteilung der Routenteile, die zu Fehlzuordnungen führen (Fehlertyp B), dient die in Abbildung 5-4 dargestellte Unkorrektheit der Zuordnung bezogen auf die mit Mobilfunkortung bestimmte Route.

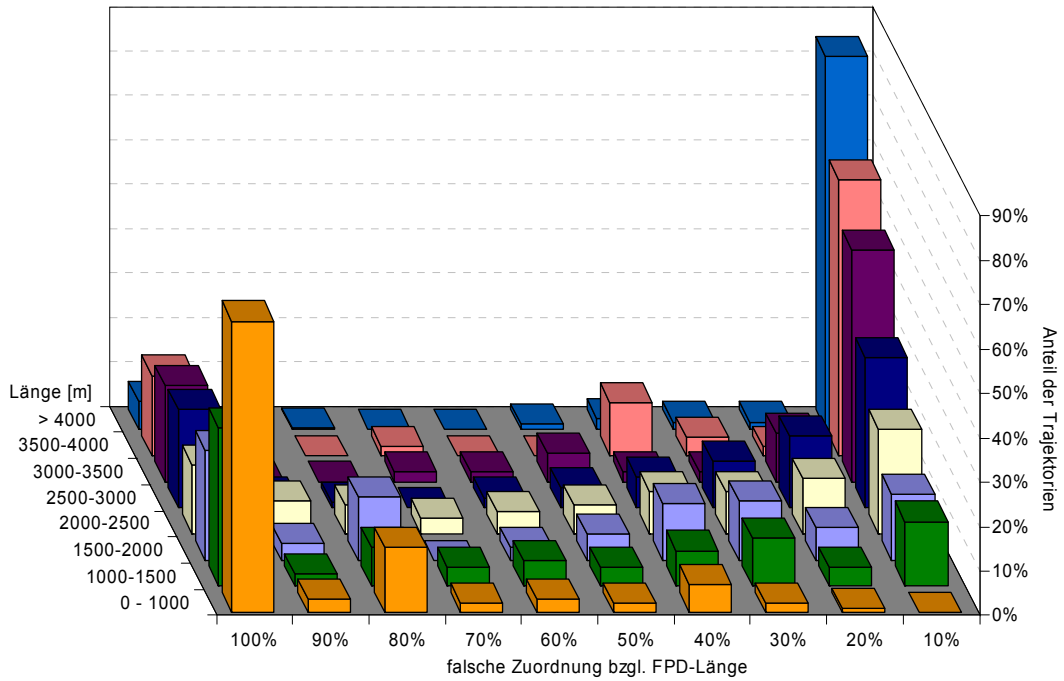


Abbildung 5-4: Anteil der falschen Zuordnung bezogen auf die mit MFO bestimmte Route in Abhängigkeit vom zurückgelegten Weg (1 - Typ B)

Bei den 6 % der über 4 km langen Positionsfolgen, die komplett falsch zugeordnet werden, handelt es sich um Autobahnfahrten, deren Anfangs- und Endpositionen eine so schlechte Genauigkeit aufweisen, dass der Korridor des Suchgebiets keine Autobahnauffahrten enthält und somit bei der Berechnung möglicher Wege die richtige Route nicht gefunden wird. 84 % der 4 km langen Routen können mit Fehlzuordnungen unter 10 % zugeordnet werden. Bei 3 km Länge verringert sich der Anteil der 10%-igen Fehlzuordnungen auf 52 %, bei 2 km auf 24 %.

Wenn 30 % Fehlzuordnungen zugelassen werden sollen, kann dies bereits bei 1-1,5 km langen Routen für 36 % der Trajektorien erreicht werden. Ab 2,5 km Länge erreichen schon über 60% der Routen Fehlzuordnungen von weniger als 30 %. Bei 3 km und 4 km langen Routen ergeben sich Werte von 65 % bzw. 87 %. Das heißt beispielsweise für 3 km lange Routen können 65 % der Trajektorien ausgewertet werden. Die auswertbaren Trajektorien lassen sich durch die Grenzwerte 97 % Trajektorienvollständigkeit und mittlere Querabweichungen kleiner 250 m ohne Referenzdaten identifizieren.

Genauigkeit der Geschwindigkeitsabschätzung

Die Genauigkeit der aus dem Weg-Zeit-Diagramm ermittelten Geschwindigkeit wird durch Vergleich mit der GPS-Referenz anhand der in Kapitel 3 beschriebenen Standardabweichung und der vorzeichenfreien Differenz der Durchschnittsgeschwindigkeit beurteilt. Für die aus der Variation der Gesprächsdauer ermittelten 951 Trajektorien zeigt Abbildung 5-5 unabhängig von Gesprächsdauer und Routenlänge wie diese bei-

den Genauigkeitsgrößen auf Genauigkeitsklassen von jeweils 10 km/h Breite verteilt sind.

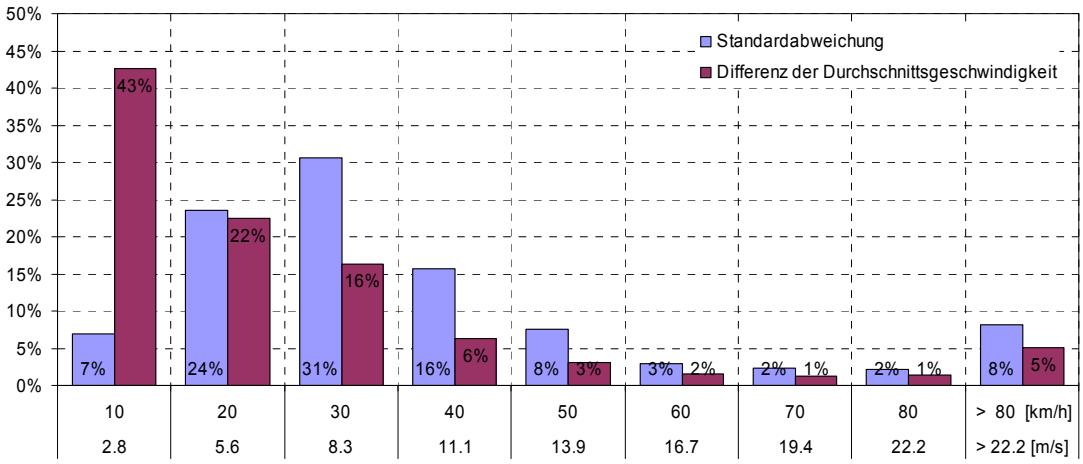


Abbildung 5-5: Genauigkeitsklassen der Geschwindigkeit

Bei 65 % der mit Mobilfunkortung bestimmten Trajektorien erreicht die Differenz der Durchschnittsgeschwindigkeit Werte von unter 20 km/h (≅5.6 m/s). Die Standardabweichung, die den Geschwindigkeitsverlauf innerhalb der Trajektorie beurteilt und somit feiner auflöst, bleibt für über 60 % der Trajektorien unter 30 km/h (≅8.3 m/s). Verglichen mit den Positionsgenauigkeiten, die quer zur Route 200-300 m betragen und in Längsrichtung keine wesentlich besseren Werte erwarten lassen, erscheint dieses Geschwindigkeitsgenauigkeitsniveau realistisch. Die Abhängigkeit der Geschwindigkeitsgenauigkeit vom zurückgelegten Weg ist in den nachfolgenden Abbildungen dargestellt.

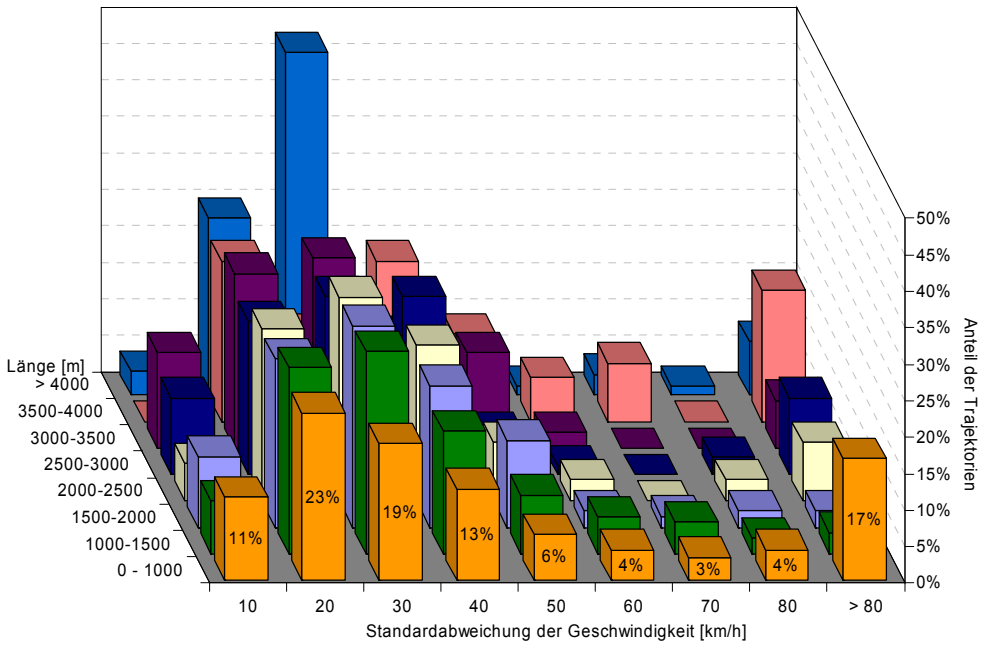


Abbildung 5-6: Standardabweichung der Geschwindigkeit in Abhängigkeit von der Trajektorienlänge

Eine Standardabweichung der Geschwindigkeit von bis zu 30 km/h (≅8.3 m/s) wird nur von 53 % der bis zu 1 km langen Positionsfolgen erreicht. Alle längeren Trajektorien erreichen diese Genauigkeitsklasse in ca. 60 % der Fälle, wobei fast keine Abhängigkeit von der Länge zu erkennen ist. Die Durchschnittsgeschwindigkeit wird in 54 %-70 % der Fälle mit einem Fehler kleiner 20 km/h (≅5.6 m/s) bestimmt; gleichfalls unabhängig von der Trajektorienlänge.

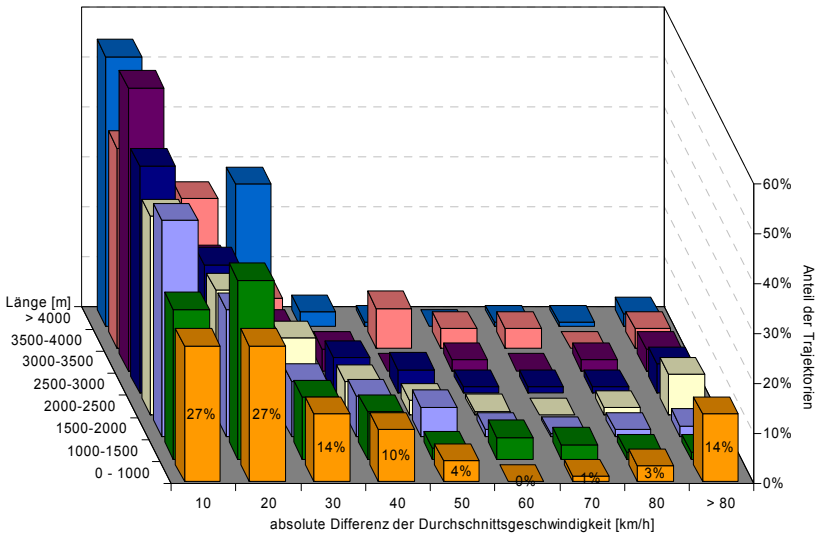


Abbildung 5-7: Beträge der Differenz der Durchschnittsgeschwindigkeit in Abhängigkeit von der Trajektorienlänge

Zusammenfassung der Qualitätsbewertung der Abis-Trajektorien

Zur Beschreibung der Qualität der Abis-Trajektorien dienen die im Anhang A aufgeführten Qualitätsmerkmale Verfügbarkeit, Aktualität, Vollständigkeit, Konsistenz, Korrektheit und Genauigkeit mit ihren dort definierten Parametern.

Die Aktualität der Trajektorien wird durch die Qualitätsparameter „Alter der Trajektorie“ Δt_{FPD} [min] und „Berechnungszeit“ Δt_{ber} [s] beurteilt und aus den Messzeitpunkten und der aktuellen Zeit exakt berechnet. Die Berechnungszeit pro Trajektorie liegt bei Abis-Daten bei etwa 70 Sekunden. Dagegen ist der Parameter Δt_{FPD} [min] mit 27 bis 35 Minuten deutlich länger, da zum einen die Rohdaten nur in 15 Minuten-Blöcken übermittelt und zum anderen umfangreiche Sortieralgorithmen vorgeschaltet werden müssen. Des Weiteren ist zu berücksichtigen, dass der Parameter zwischen dem Beginn der Trajektorie und dem Abschluss der Berechnung bestimmt wird. Das spielt insbesondere bei sehr ausgedehnten Gesprächen (15-Minuten-Block-übergreifend) eine Rolle, da hier zwar nach spätestens 15 Minuten eine Berechnung angestoßen wird, die aber dann nicht mehr auf den ersten Trajektorienzeitpunkt bezogen ist.

Die Parameter Korrektheit der Routen und Genauigkeit der aus den Trajektorien bestimmten Geschwindigkeit können mit den GPS-Testfahrten als Referenz bestimmt werden und dienen zur objektiven Qualitätsbeurteilung der Trajektorien.

Wie die Untersuchungen zeigen, hängt die Korrektheit der Routen (sowohl Typ A als auch Typ B) stark von der Länge des zurückgelegten Wegs ab. Für 70 % richtige Zuordnungen sind die entsprechenden Werte in Abschnitt 5.2 nachzulesen. Tabelle 5-2 gibt dieses auch für andere Prozentzahlen an. Dabei bezieht sich die Spalte „GPS“ auf die mit GPS bestimmte Länge der Route und stellt somit den Korrektheitsratentyp A dar. Die Spalte „MFO“ bezieht sich dagegen auf die mittels Mobilfunkortung bestimmte Länge der Trajektorie. Es handelt sich folglich um den Korrektheitsratentyp B, der als deutlich kritischer einzustufen ist.

Tabelle 5-2: Korrektheit Typ A und Typ B für verschiedene Routenlängen

Länge	**Zuordnung zu mindestens *% richtig**					
	90%		**70%**		**50%**	
	GPS Typ A	**MFO Typ B**	GPS Typ A	**MFO Typ B**	GPS Typ A	**MFO Typ B**
1-2 km	12%	**15%**	33%	**33%**	48%	**48%**
2-3 km	23%	**27%**	50%	**51%**	63%	**66%**
über 3 km	23%	**81%**	71%	**87%**	83%	**92%**

Bei weniger als einem Viertel der über 3 km langen Positionsfolgen können Routen bestimmt werden, die mit der GPS-Route zu über 90 % übereinstimmen. Dies bedeutet, dass nicht über die komplette Gesprächszeit zuverlässige Informationen bestimmt werden können. Die ermittelten Routen enthalten jedoch in über 80 % der Fälle weniger als 10 % falsche Informationen. Will man, wie in Abschnitt 5.2 erläutert, 70 % Korrektheit Typ B erreichen, so gelingt das für 1-2 km lange Trajektorien nur für 33 % der Trajektorien, bei 2-3 km für 51 % und für mehr als 3 km lange Trajektorien für immerhin 87 % der Trajektorien. Es sollten folglich nur Trajektorien, die länger als 3 km sind, einer Auswertung zugeführt werden. Als Kompromiss zwischen Performance, Genauigkeit und Abdeckungsrate wurden im Online-Betrieb aber Trajektorien, die länger als 2 km sind und gleichzeitig den Bedingungen „mehr als 15 Positionen“, „mindestens 60 s Gesprächsdauer“ und „berechnete Geschwindigkeit kleiner als 200 km/h“ genügen, einer Auswertung zugeführt.

Erfreulich ist jedoch in jedem Fall, dass die nur unter Nutzung von Referenzdaten ableitbare Korrektheitsrate Typ B mit dem im Online-Betrieb verfügbaren Parametern „Trajektorienvollständigkeit“ und „durchschnittliche Querabweichung“ in einen empirisch ermittelten Zusammenhang gebracht werden kann. Wenn man 70 % richtig im Sinne der Korrektheit Typ B zugeordnet haben will, so sind nur Trajektorien mit einer Vollständigkeit > 97 % und einer durchschnittlichen Querabweichung < 250 m auszuwerten.

Die aus den FPD-Trajektorien abgeleiteten Standardabweichungen weisen für 30 s Diskretisierungsintervalle Werte von rund 20 km/h auf. Die Durchschnittsgeschwindigkeit kann insbesondere bei längeren Fahrten mit ca. 8 km/h deutlich genauer bestimmt werden. Abweichungen von der tatsächlichen Durchschnittsgeschwindigkeit lassen sich mit den im Online-Betrieb verfügbaren Parametern („Trajektorienvollständigkeit > 97 %“ und „durchschnittliche Querabweichung < 250 m“) so gering wie möglich halten.

Die Untersuchungen haben deutlich gezeigt, dass die FPD-Routen am Anfang und Ende deutlich unzuverlässiger sind als in der Mitte. Um die gefundenen Routen auf den mittleren Teil zu beschränken, wird auf den ersten und letzten 30 % der Strecke eine Geschwindigkeitskontrolle durchgeführt. In diesem Bereich werden Teile abgeschnitten, wenn sie nur mit sehr hohen Geschwindigkeiten zu erreichen sind (größer 170 km/h bzw. größer als die zweifache Kantengeschwindigkeit). Hierbei handelt es sich im Gegensatz zu dem Abschneiden in Kapitel 2.1.4 um eine abschließende Plausibilitätskontrolle.

5.1.2 A-Ebene

Auch die aus A-Daten generierten Trajektorien sollen anhand von Testfahrten mit GPS-Referenz hinsichtlich ihrer absoluten Qualität beurteilt werden.

Korrektheit

Im Gegensatz zu den aus Abis-Daten bestimmten Trajektorien liefert die Routenbestimmung mit A-Daten nicht immer im mittleren Bereich die richtige Route. Um zu mit Abis-Daten vergleichbaren Parametern zur Beurteilung der absoluten Qualität zu gelangen, werden die Längen der richtigen und falschen Teilrouten durch Summation der entsprechenden Kantenlängen bestimmt, so dass die Trajektorie wieder anhand der beiden Qualitätsparameter „Zuordnungskorrektheit Typ A“, der den Anteil der korrekten Route an der mit GPS bestimmten Route angibt, und der „Zuordnungskorrektheit Typ B“, der den prozentualen Anteil der korrekten Route an der aus A-Daten generierten Route angibt.

Bei den Testfahrten wurden jeweils drei Mobilstationen mitgeführt, wobei das erste Telefon M1 ständig aktiv war und damit sehr viele Handover erzeugte, das zweite Telefon M2 viele Telefonate mit Unterbrechungen führte und das dritte Telefon zwar eingeschaltet aber nicht aktiv war und damit nur A-Daten bei den Location Area Updates lieferte. Die nachfolgenden Abbildungen zeigen die beiden Korrektheiten Typ A und Typ B für alle Testfahrten, wobei neben der Bezeichnung der Fahrt die Länge der zugeordneten Route in Kilometer angegeben ist.

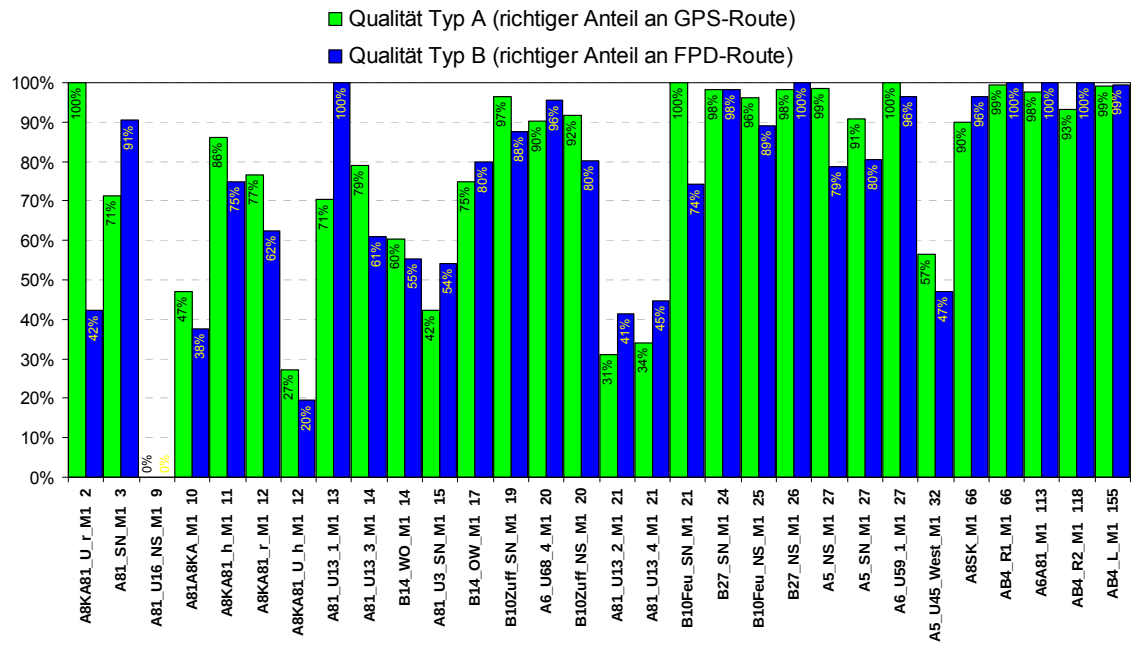


Abbildung 5-8 : A-Daten Qualität (Dauergespräche mit M1)

Die Fahrt auf dem Autobahnviereck im Uhrzeigersinn konnte aufgrund einer Gesprächspause beim Autobahnkreuz Walldorf nicht als durchgehende Trajektorie ausgewertet werden. Um die Testfahrt trotzdem auswerten zu können, wurde die Fahrt aufgeteilt.

Mit Ausnahme der sehr kurzen Umleitungsstrecke U16 konnten alle Fahrten mit der ständig aktiven Mobilstation M1 zumindest in Teilen richtig zugeordnet werden. Im Mittel über die 30 Fahrten erreichte die Zuordnungskorrektheit Typ A (richtiger Anteil der GPS-Route) 76,6 % und Typ B (richtiger Anteil der MFO-Route) 72,9 %. Da lange Trajektorien deutlich besser zugeordnet werden können, ergeben sich für die auf die Länge bezogenen Durchschnitte bessere Werte von 87,0 % (Typ A) und 85,9 % (Typ B).

Die Qualität der aus A-Daten ermittelten Routen nimmt mit geringerer Informationsdichte ab. Für die nicht ständig aktive Mobilstation M2, für die weniger Handover registriert wurden, ergaben sich im Mittel der 30 Fahrten 71,7 % für Typ A und 67,0 % für Typ B und für den auf die Länge bezogenen Durchschnitt 81,1 % (Typ A) und 80,5 % (Typ B).

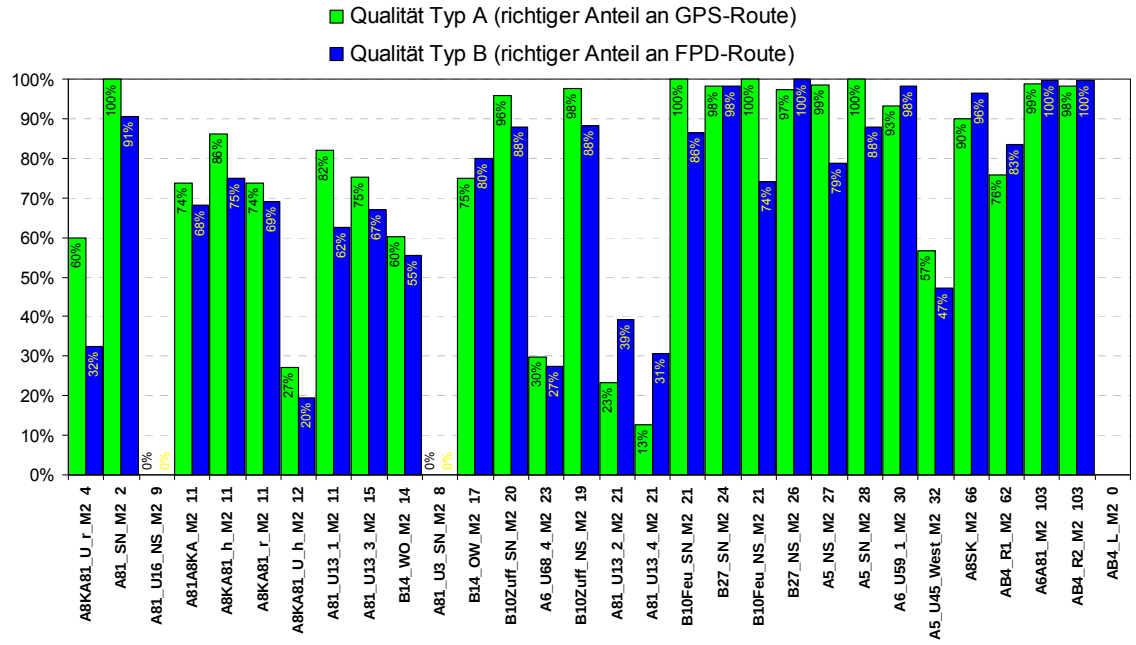


Abbildung 5-9: A-Daten Qualität (mehrere Gespräche mit M2)

Mit der zwar eingeschalteten aber nicht aktiven Mobilstation M3 konnten bei den kurzen Fahrten nur sehr wenige Location Area Updates registriert werden. Abbildung 5-10 zeigt deutlich, wie unzuverlässig die Routengenerierung bei zu wenigen Informationen wird. Bezogen auf die 30 Fahrten erreicht die Zuordnungskorrektheit Typ A nur noch 15,6 % und Typ B 14,2 % und für den auf die Länge bezogenen Durchschnitt 41,4 % (Typ A) und 41,9 % (Typ B).

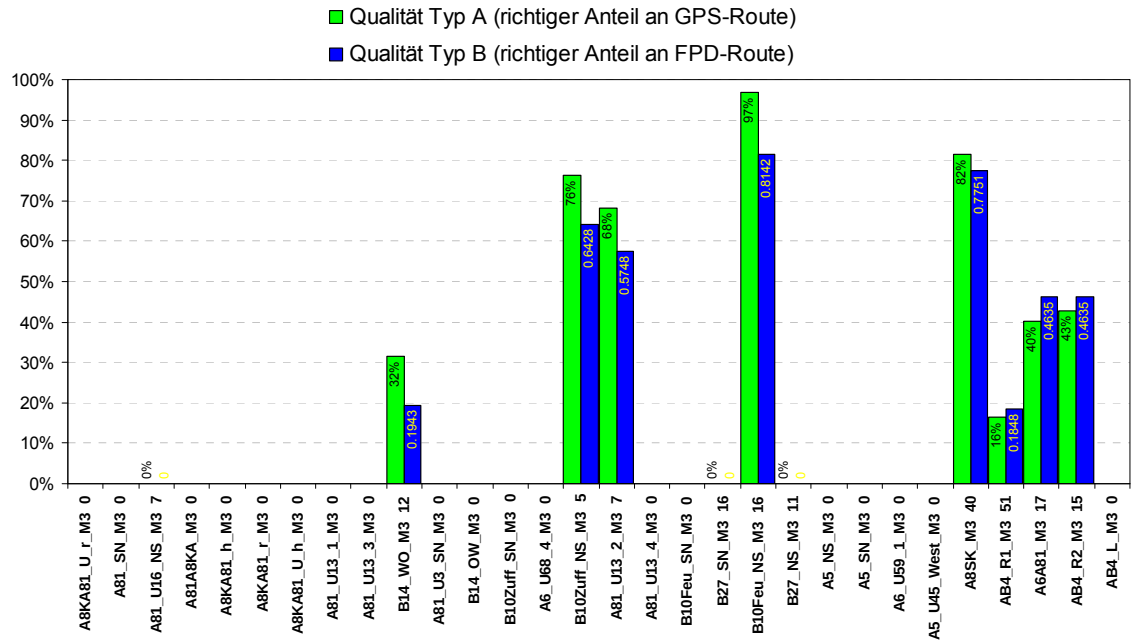


Abbildung 5-10: A-Daten Qualität (nur LUP mit M3)

Die enormen Datenmengen, die in Echtzeit im gesamten Projektgebiet für die A-Daten Prozessierung anfallen, können aufgrund der Rechenkapazität nicht vollständig ausgewertet werden. Deshalb sollen bevorzugt die Datensätze einer Auswertung zugeführt werden, die die hinsichtlich der Zuordnungskorrektheit besten Routen erwarten lassen. Während der kurzen Online-Phase im März 2009 wurden zur Trajektoriengenerierung nur Datensätze mit mehr als 7 verschiedenen Zellen verwendet.

Genauigkeit der Geschwindigkeit

Die aus A-Daten ermittelten Weg-Zeit-Diagramme der Testfahrten mit GPS-Referenz werden mit der in Abschnitt 5.1.1 für Abis-Daten beschriebenen vorzeichenfreien Differenz der Durchschnittsgeschwindigkeit und der Standardabweichung der Geschwindigkeit beurteilt, die die Genauigkeit des Geschwindigkeitsverlaufs innerhalb der Trajektorie angibt. Der zweite Wert ist in den nachfolgenden Abbildungen nur angegeben, wenn genügend Informationen zu seiner Berechnung vorliegen.

Für die Mobilstation M1 mit Dauergesprächen ist die Standardabweichung der Geschwindigkeit für die Fahrt auf der A6/A81 mit 51 km/h am schlechtesten, zeigt für die anderen Fahrten aber deutlich bessere Werte von durchschnittlich 20 km/h. Die vorzeichenfreie Differenz der Durchschnittsgeschwindigkeit liegt bei maximal 41 km/h bei einem Durchschnitt von 10 km/h.

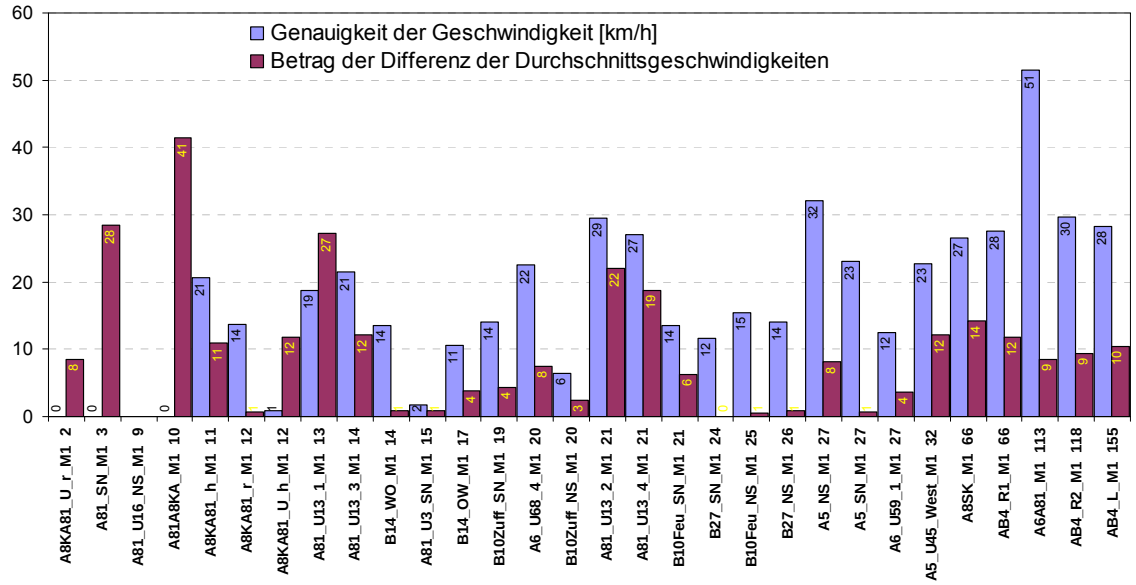


Abbildung 5-11: Genauigkeit der Geschwindigkeit (Dauergespräche mit M1)

Für die Mobilstation M2 mit Standardabweichungen von maximal 51 km/h und durchschnittlich 21 km/h und Differenzen der Durchschnittsgeschwindigkeit von maximal 48 km/h und durchschnittlich 12 km/h verschlechtern sich diese Werte nur unwesentlich, da zur Bestimmung dieser Genauigkeitsangaben nur die Fahrten verwendet werden können, für die eine Zuordnung gefunden wurde.

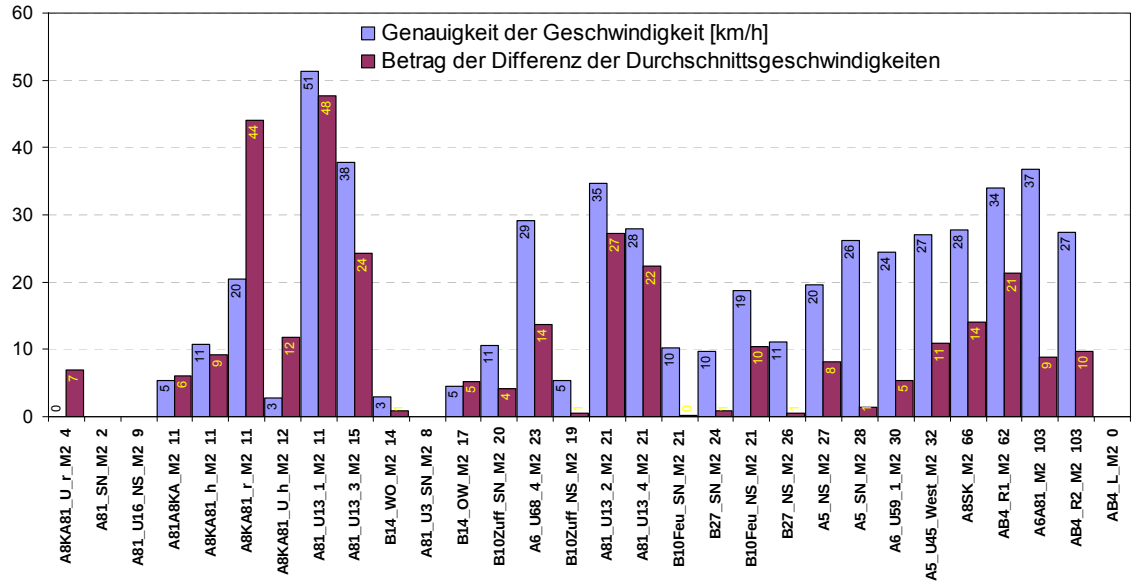


Abbildung 5-12: Genauigkeit der Geschwindigkeit (mehrere Gespräche)

Für die wenigen Fahrten, die mit der Mobilstation M3 ausgewertet werden konnten, wurden Standardabweichungen von 17 km/h erreicht. Die Differenz der Durchschnittsgeschwindigkeit beträgt maximal 42 km/h bei einem Mittel von 16 km/h.

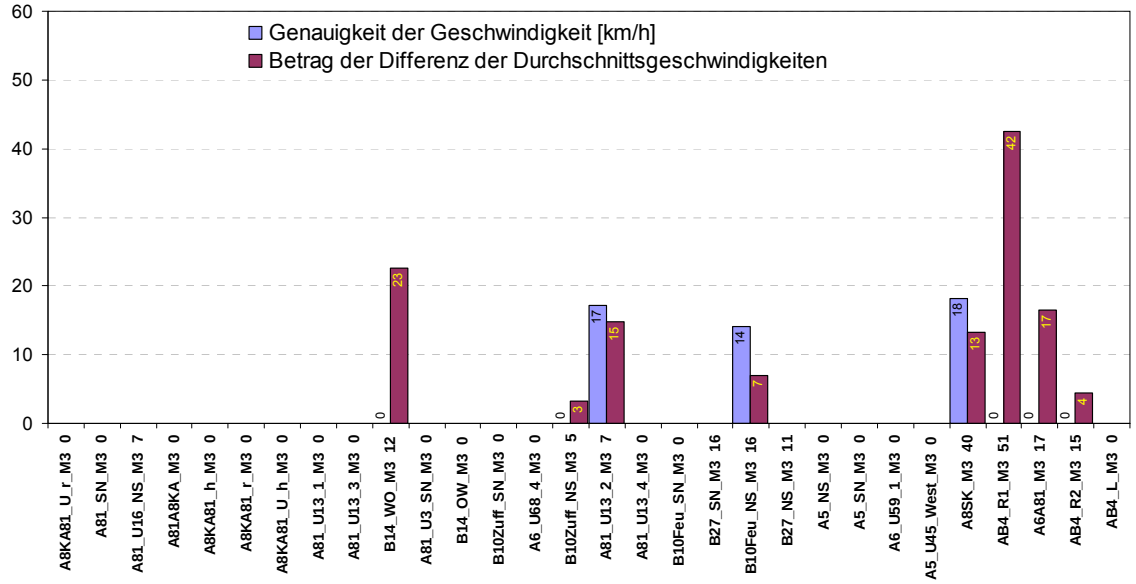


Abbildung 5-13: Genauigkeit der Geschwindigkeit (LUP)

In der Online-Auswertung wurden nach der Berechnung von Trajektorien Plausibilitätstests hinsichtlich der berechneten Geschwindigkeiten durchgeführt, um unkorrekte Trajektorienteile identifizieren und verbessern bzw. eliminieren zu können.

- Rückwärtsfahrten innerhalb einer Trajektorie werden als Fehler gedeutet und durch Elimination der entsprechenden Zeitpunkte verbessert.

- Wenn die aus A-Daten ermittelte Geschwindigkeit das 1,5-fache der auf dem Straßensegment maximal zulässigen Geschwindigkeit übersteigt, wird die Trajektorie verworfen, da vermutlich die falsche Route gefunden wurde.

Zusammenfassung der Qualitätsbewertung der A-Trajektorien

Zur Beschreibung der Qualität der A-Trajektorien dienen die im Anhang A aufgeführten Qualitätsmerkmale. Hier wird sich auf die Korrektheit der Trajektorien und die Genauigkeit der Geschwindigkeitsbestimmung beschränkt, da sie zur unmittelbaren Bewertung der Trajektorien dienen.

Die Aktualität der Trajektorien wird auch hier wieder durch die Qualitätsparameter „Alter der Trajektorie“ Δt_{FPD} [min] und „Berechnungszeit“ Δt_{ber} [s] beurteilt. Δt_{ber} liegt bei A-Daten bei etwa 8 Sekunden; also deutlich geringer als bei A-Daten, da weniger Berechnungsschritte notwendig sind. Demgegenüber ist der Parameter Δt_{FPD} [min] mit 26 bis 37 Minuten wie auf Abis-Ebene deutlich länger. Eine Verbesserung in Relation zum Abis-Ebene gelingt hier nicht, da derselbe Bereitstellungsalgorithmus in 15 Minuten-

Blöcken umgesetzt wird. Außerdem gibt es sicher mehr Trajektorien, die 15-Minuten-Block-übergreifend berechnet werden.

Die Korrektheit der Trajektorienzuordnung (Typ B) ergibt im Idealfall (durchgängiges Telefonat) einen prozentualen Wert von etwa 73 % in Bezug auf die Trajektorienanzahl. Bezieht man den prozentualen Wert auf die gefahrenen Strecken, so steigert er sich sogar auf etwa 86 %. Bei mehrfach unterbrochenen Telefonaten fällt der Wert auf 67 % beziehungsweise 80,5 %. Bleibt das Mobiltelefon passiv, so fällt die Verfügbarkeit von 100 % auf 33 % und die Korrektheit auf 14,2 bzw. 41,9 %. Für diesen Fall ist eine zuverlässige FPD-Generierung nicht möglich.

Die Standardabweichung der Genauigkeit und die mittlere Abweichung derselben liefern ein anders Bild. Bei durchgängigen Telefonaten werden durchschnittlich 20 km/h für die Standardabweichung und 10 km/h für die mittlere Abweichung erreicht. Das ändert sich für unterbrochene Telefonate und passive Telefone nur in geringem Umfang; nur die Verfügbarkeit wird natürlich herabgesetzt.

Mit anderen Worten kann die Korrektheit der Trajektorienzuordnung bei aktiven Mobilfunkteilnehmer zu mehr als 80 % gewährleistet werden; bei passiven Teilnehmern werden deutlich geringere Werte erreicht. Die Geschwindigkeit wird aber bei deutlich herabgesetzter Verfügbarkeit bei aktiven und passiven Telefonen fast gleichgenau geschätzt.

© Do-iT 2009

5.1.3 Abis- versus A-Ergebnisse

Da Abis-Daten nur in Karlsruhe abgegriffen werden konnten, beschränkt sich der direkte Vergleich der beiden Verfahren zur Trajektoriengenerierung auf die Testfahrten in Karlsruhe, die nach der Installation in Fellbach zum Abgriff der A-Daten durchgeführt wurden. Bei den Testfahrten innerhalb der beiden Location Areas mit Abis-Daten-Registrierung handelt es sich um Routen auf dichtem Straßennetz und in relativ kleinen Mobilfunkzellen.

Auf der 4 Mal im Hin- und Rückweg befahrenen Südtangente, die der B10 im Stadtbereich entspricht und in den Abbildungen als Szenario S2 bezeichnet ist, konnten aus Abis-Daten Trajektorien mit einer durchschnittlichen Zuordnungskorrektheit von 85.6 % (Typ A) bzw. 80.2 % (Typ B) bestimmt werden. Die aus A-Daten bestimmten Trajektorien erreichen dagegen nur Zuordnungskorrektheiten von 80.9 % für Typ A und 50.1 % für Typ B.

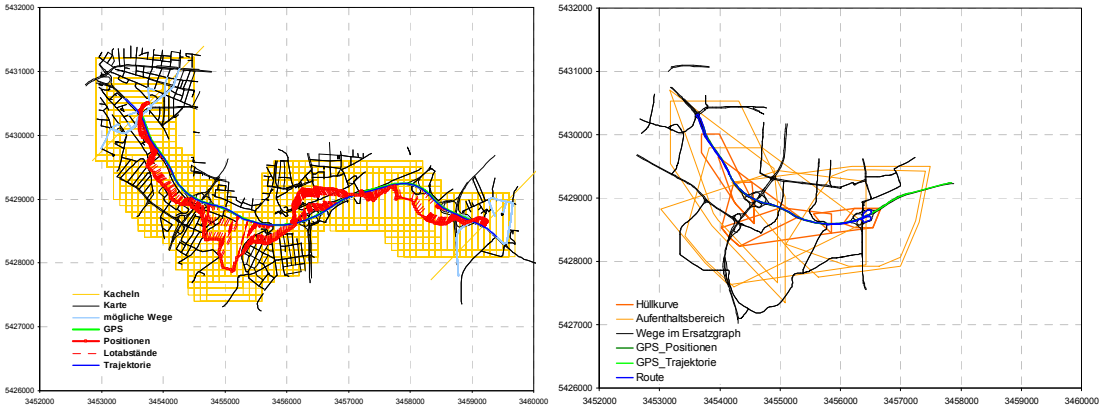


Abbildung 5-14 : Vergleich Abis-/A-Daten (Südtangente)

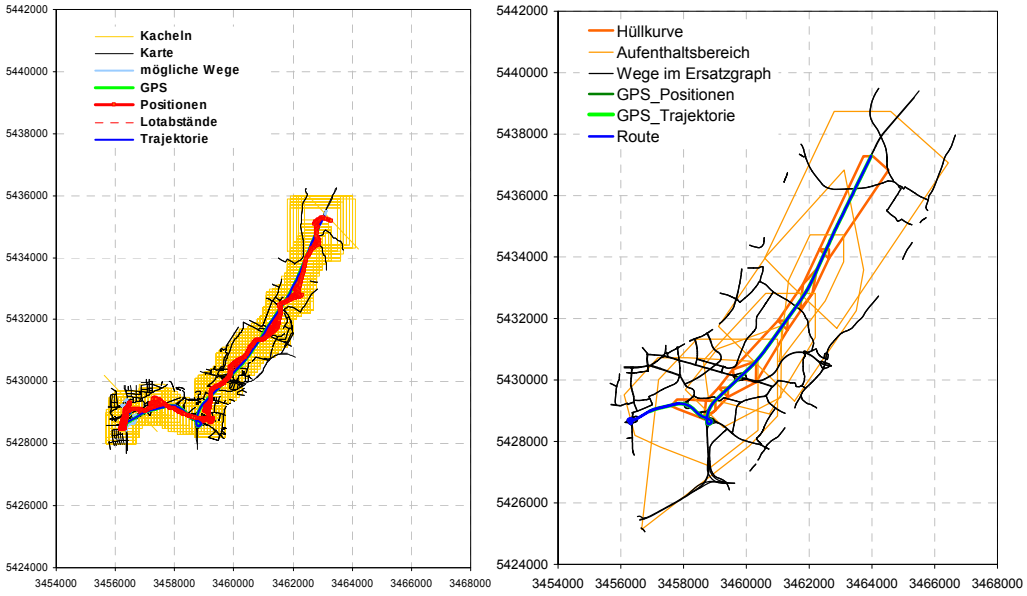


Abbildung 5-15: Vergleich A/Abis-Daten (Autobahn A5)

Für die Autobahnfahrten auf der A5 (in den Abbildungen als Szenario S3 bezeichnet) ergaben sich unterschiedliche lange Aufzeichnungszeiten, da die Abis-Daten im Norden der Testfahrt nur bis zur Grenze der Karlsruher Location Area aufgezeichnet werden konnten. Bezogen auf die jeweilige Aufzeichnungsdauer ergab sich aus den Abis-Daten im Vergleich mit der GPS-Referenz eine durchschnittliche Zuordnungskorrektheit von 97.4 % für Typ A und 94.9 % für Typ B. Dies kann für A-Daten mit 88.3 % bzw. 78.0 % nicht ganz erreicht werden.

Die Testfahrten auf der L561 (S4) und von der B10 zur A8 (S5) konnten mit Abis-Daten problemlos ausgewertet werden und erreichten Zuordnungsgenauigkeiten Typ B von über 75 %. Mit A-Daten wurde in keinem Fall die richtige Route gefunden. Für A-Daten wird das Suchgebiet aus der Hüllkurve um aufeinander folgende Zellwechsel bestimmt, die um einen Sicherheitsabstand verbreitert wird, der sich aus der Breite der ellipsenförmigen Aufenthaltsbereiche bestimmt. Um den maximal in einem Zeitintervall zurückgelegten Weg zu bestimmen, wird die höchste Geschwindigkeit der innerhalb der Hüllkurve gespeicherten Straßensegmente verwendet (vgl. Abschnitt 2.2.4). Da die Routen der Szenarien 4 und 5 nahe der Autobahn liegen, werden aufgrund zu hoher Geschwindigkeiten die Suchgebiete zu breit geschätzt. Dies führt dazu, dass auf dem Graphen innerhalb des Suchgebiets der kürzeste Wege-Algorithmus nicht die richtige Route findet. Leider war es innerhalb der kurzen Zeit zwischen der Installation in Fellbach und Projektende nicht mehr möglich, für diese Problematik eine Lösung zu finden.

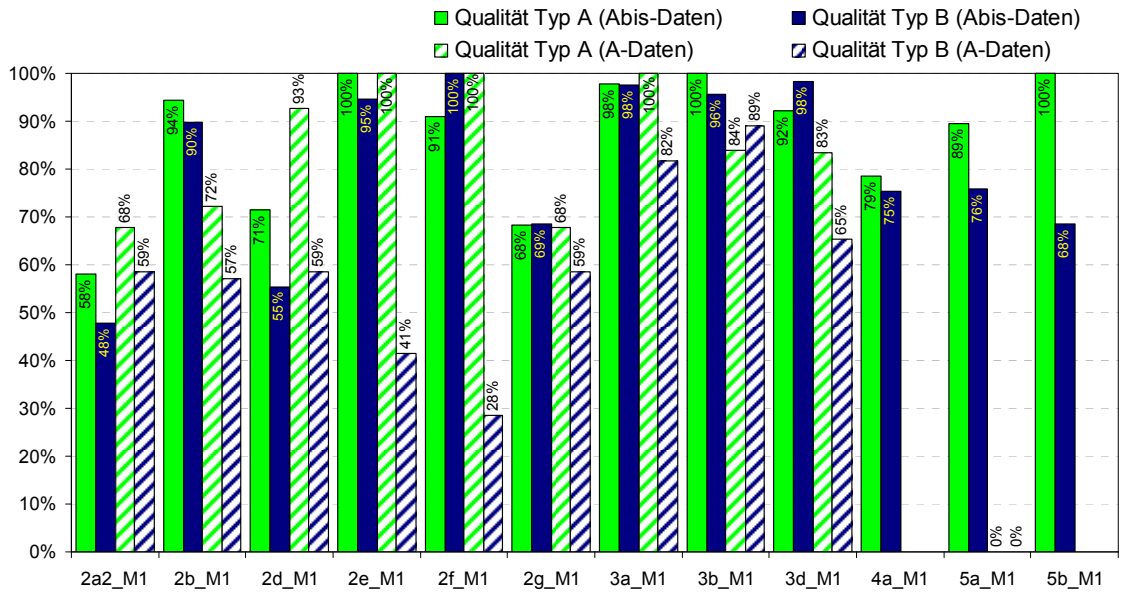


Abbildung 5-16: Vergleich Abis/A-Daten Zuordnungskorrektheit

Abbildung 5-16 zeigt, dass mit einer Ausnahme die Abis-Daten für alle Fahrten deutlich bessere Zuordnungskorrektheiten liefern. In bebauten Gebieten ist mit A-Daten aufgrund des dichten Straßennetzes, der kleinen Mobilfunkzellen und nicht geometriebedingter Umschaltungen keine zuverlässige Zuordnung dieser kurzen Routen mit 7 bis

11 Ereignissen möglich. Im Innenstadtbereich benötigt die Auswertung mit A-Daten aufgrund nicht geometriebedingter Zellwechsel mehr als die 6 für den Online-Betrieb festgelegte minimale Anzahl von Ereignissen, um brauchbare Ergebnisse zu erzielen.

Der Vorteil der Abis-Daten zeigt sich auch in der Genauigkeit der Geschwindigkeit. Aufgrund der wesentlich höheren Informationsdichte bei Abis-Daten, die mit 2Hz abgetastet werden, lassen sich die Weg-Zeitdiagramme genauer bestimmen. Das Weg-Zeit- und das Weg-Geschwindigkeits-Diagramm, die für eine Fahrt auf der Südtangente in Abbildung 5-17 dargestellt sind, zeigen für Abis-Daten eine hohe Übereinstimmung mit der GPS-Referenz. Bei A-Daten liegen nur an wenigen Zeitpunkten Informationen geringer Genauigkeiten vor, die in diesem Beispiel am Anfang zu negativen und im weiteren Verlauf zu extrem hohen Geschwindigkeiten führen. Bei einer Online-Auswertung würden aufgrund der Plausibilitätsprüfung das Weg-Zeit-Diagramm verworfen und keine Trajektorie weitergeben werden.

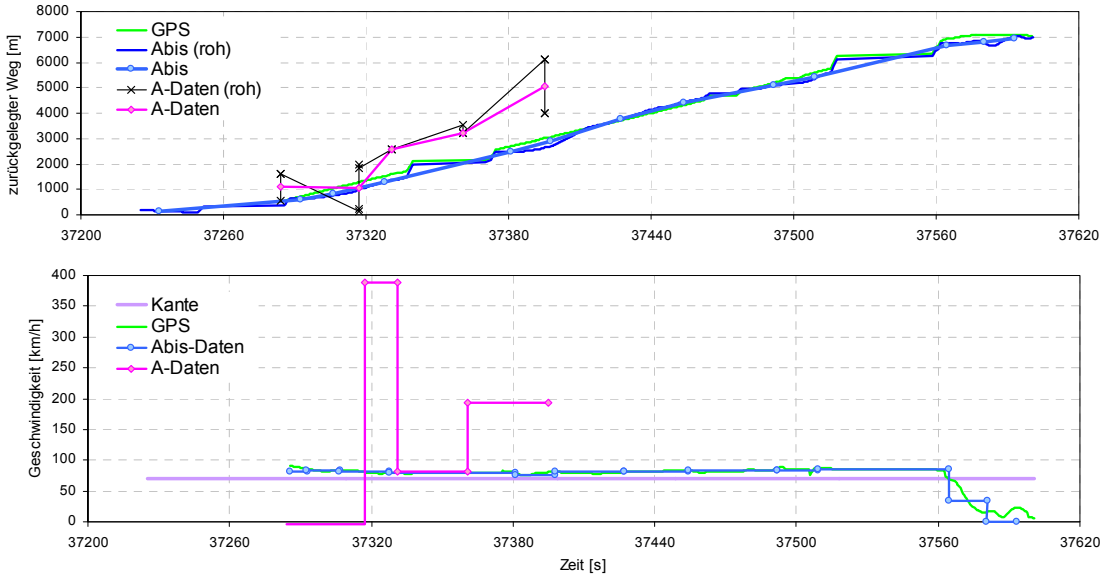


Abbildung 5-17: Vergleich Abis/A Geschwindigkeit (Südtangente)

Von allen Testfahrten in Karlsruhe verbleiben nach den Plausibilitätstests der Online-Auswertung nur die drei Fahrten auf der Autobahn. Im Weg-Zeit-Diagramm der Abbildung 5-18 zeigt sich für beide Verfahren eine gute Übereinstimmung mit der GPS-Referenz. Lediglich am Anfang der A-Daten beinhaltet ein Zellwechsel mehrere weit auseinander liegende Schnittpunkte mit der gefundenen Route, deren Mittelung nicht die richtige Position trifft, so dass in diesem Bereich die Geschwindigkeit mit 140 km/h zu hoch geschätzt wird. Die Plausibilitätsprüfung der Online-Auswertung müsste diesen Teil aus der Trajektorie eliminieren, da die Geschwindigkeit größer als das 1.5-fache der gespeicherten Kantengeschwindigkeit ist.

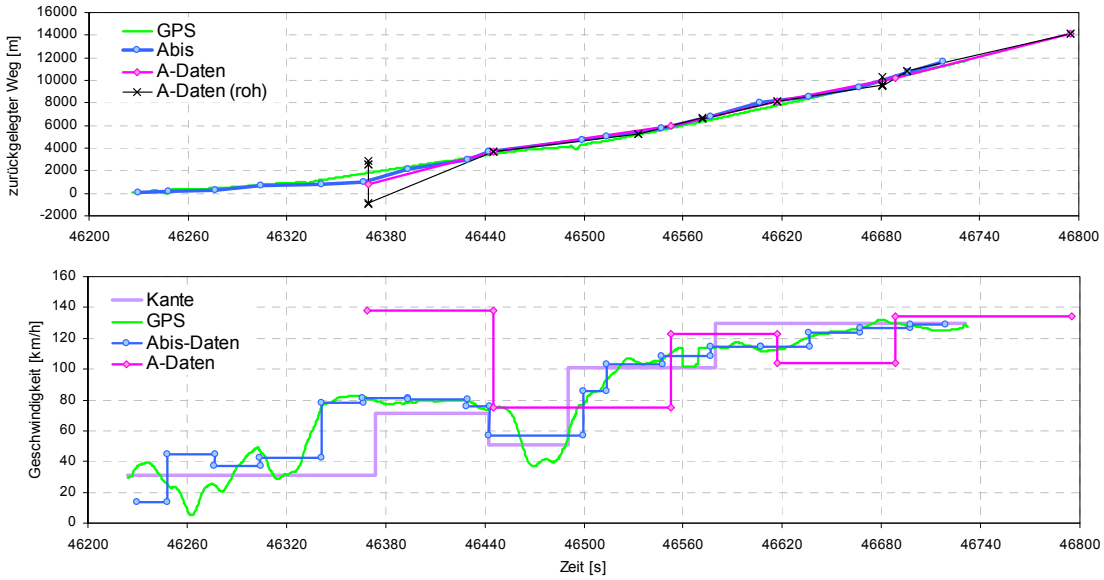


Abbildung 5-18 : Vergleich Abis/A Geschwindigkeit (A5)

Die aus den drei Autobahnfahrten ermittelten Genauigkeiten der Geschwindigkeit sind im Folgenden tabellarisch dargestellt.

Tabelle 5-3: Genauigkeit der Geschwindigkeit

	Abis	A
Standardabweichung der Geschwindigkeit [km/h]	30.1	---
Betrag der Differenz der Durchschnittsgeschwindigkeit [km/h]	5.2	10.2

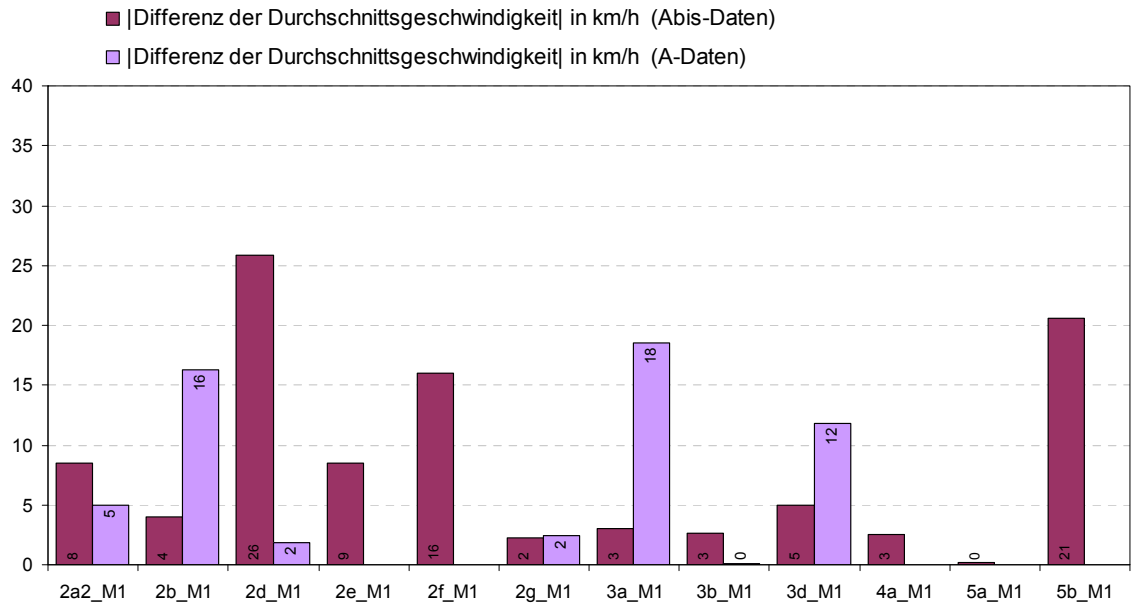


Abbildung 5-19: Vergleich Abis/A Geschwindigkeit (Autobahnfahrt)

Zusammenfassung des Vergleichs A- versus Abis-Trajektorien

Sowohl die Korrektheitsparameter als auch die Genauigkeitsparameter zeigen die deutliche Überlegenheit der Abis-Daten. Hinzu kommt, dass für die Online-Prozedur die Verfügbarkeit für die A-Daten deutlich herabgesetzt ist. Diese Aussagen sich allerdings nicht überzubewerten, da sie zwar stichhaltig sind, sich aber auf eine geringe Stichprobe beziehen.

5.2 Vergleich mit SES

Der Vergleich mit stationären Erfassungssystemen erlaubt die Beurteilung der Durchdringung (P 3.3) als Parameter der Vollständigkeit sowie der Korrelation (P 5.7) als Parameter der Korrektheit der aus FPD-Daten ermittelbaren Verkehrsstärken. Die Eignung der Mobilfunkdaten für den weiteren Haupteinsatzbereich, der Verkehrslageschätzung und -steuerung sowie der Generierung von Quelle-Ziel-Matrizen und damit zur Verkehrsprognose, wurde von den Projektpartnern im Rahmen der Demonstratoren untersucht. Die Ergebnisse dieser Evaluierungen finden sich im Bericht zum AP10 [AP10-2 2009].

Im Folgenden wird die Untersuchung aufgeteilt in den Vergleich mit Trajektorien, die aus Abis-Daten generiert wurden und in Trajektorien aus A-Daten. Da im Projekt nur reale Abis-Daten im Stadtgebiet Karlsruhe, genauer gesagt in den beiden Location Areas Karlsruhe und Ettlingen zur Verfügung standen, war auch nur dort der Vergleich mit SES sowie eine Gegenüberstellung von A und Abis möglich. Es standen daher zur Evaluierung der Verkehrsstärken aus Abis-Daten nur die Sensoren der Stadt Karlsruhe und zusätzlich noch 6 Messquerschnitte der DDG sowie 5 des IM-BW auf der Autobahn A5 als Referenz zur Evaluierung der Abis-Trajektorien zur Verfügung.

A-Daten standen zunächst im gesamten Projektnetz zur Verfügung, daher konnten hier auch alle ausgewählten SES zum Vergleich herangezogen werden. Am 2. Tag der Evaluierungsphase, die vom 23.-31. März 2009 stattfand, fiel jedoch ein großer Teil der LACs und damit auch etwa ein Drittel aller SES aus. Zusätzlich standen aus ungeklärter Ursache am Samstag, den 28. März keine A-Daten zur Verfügung, folglich konnten auch keine A-Trajektorien berechnet werden.

In diesem Bericht beschränkt sich die Darstellung auf die wesentlichen Erkenntnisse der Evaluierung mit SES. Die detaillierten Untersuchungsergebnisse sind im Abschlussbericht zum AP10 zu finden (vgl. [AP10-2 2009]).

5.2.1 Abis-Ebene

Die Qualität der Trajektorien aus Abis-Mobilfunkdaten konnte nur im Raum Karlsruhe untersucht werden. Durch die Verwendung von SES als Referenz, war in erster Linie

der Vergleich mit der Verkehrsstärke und deren Änderung möglich. Dabei handelt es sich um eine der wesentlichen Informationen, die als Grundlage für diverse verkehrstechnische Anwendungen erforderlich ist.

Zur Evaluierung standen in Karlsruhe zum einen die stadteigenen SES zur Verfügung, wie sie in Abbildung 4-11 und Abbildung 4-12 dargestellt sind. Die Sensoren sind insbesondere auf der verkehrstechnisch wichtigen Südtangente als Ost-West-Transversale, sowie auf wichtigen Ein- und Ausfallsstraßen installiert. Zum Anderen fielen 5 Sensoren des IM sowie 6 weitere Sensoren der DDG in das mit Abis-Daten abgedeckte Gebiet. Diese Sensoren lagen hauptsächlich auf der Autobahn A5 zwischen der Anschlussstelle Ka-Durlach im Süden und Ka-Nord im Zentrum des Abis-Testgebiets. Hinzu kamen 3 Sensoren am Rand des Abis-Gebiets, einer auf der A5 kurz vor Bruchsal im Norden von Ka und zwei auf der A8 im Südosten des Testgebiets. Die folgende Abbildung 5-20 zeigt alle verwendeten SES im Gebiet Karlsruhe.

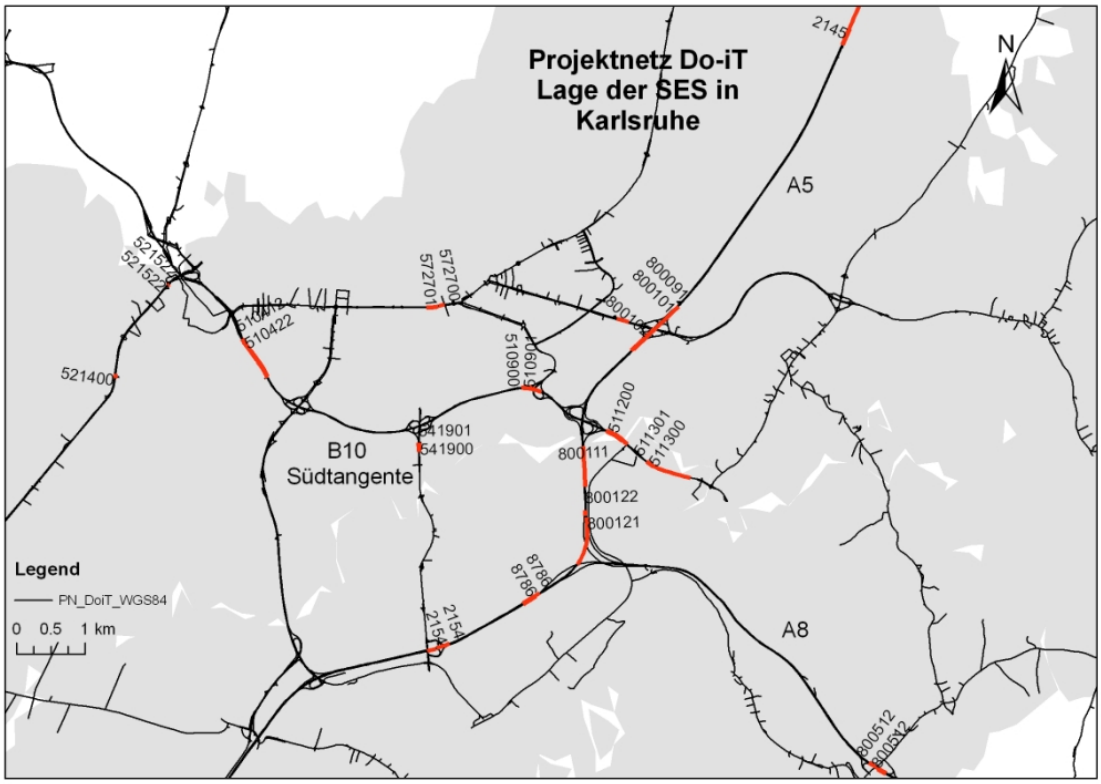


Abbildung 5-20: Alle SES im Abis-Testgebiet Karlsruhe

Die Evaluierung fand Ende März statt und es wurden insgesamt 9 Tage vom Montag 23.03.09 bis zum Projektende am Dienstag 31.03.09 ausgewertet. In diesem Zeitraum lag die durchschnittliche Durchdringung innerhalb 24 h bei 0,4 bis 3,3 ‰ bei einer relativ starken Streuung von 0,2 bis 2,6 ‰. Die Standardabweichung innerhalb dieser 9 Tage lag bei maximal 0,8 ‰. Am Rand des Abis-Gebiets in Fahrtrichtung Karlsruhe nimmt die Durchdringung erwartungsgemäß stark ab, da in das Gebiet einfahrende Fahrzeuge erst nach einer gewissen Fahrtstrecke auswertbar sind.

Die Durchdringungen auf der Autobahn sowie auf der weitgehend kreuzungsfreien Südtangente erreichen mit knapp 2 bis 3 ‰ höhere Durchdringungen als auf dem untergeordneten Straßennetz. In der Innenstadt ist die Berechnung von brauchbaren Trajektorien sehr schwierig, es können hier lediglich 0,4 ‰ bei einer Streuung von ca. 0,6 ‰ detektiert werden.

Zur Beurteilung, in welchem Maße die Abis-Daten zur Detektion von Veränderungen der Verkehrsstärke geeignet sind, wurde die Rangkorrelation zwischen den Tagesganglinien aus Abis und aus der jeweiligen Referenzquelle berechnet. Die Korrelationen blieben neben einzelner tageweiser Ausnahmen im Mittel unter 0,5 bei einer Streuung von über 0,5, bei 6 von 25 SES lag die Streuung sogar über 1,0. Die zugehörigen Standardabweichungen bewegen sich zwischen 0,1 und 0,6.

Die Betrachtung der absoluten Anzahl berechneter Trajektorien in 24 h erklärt diese relativ schlechten Werte. Nur an 3 SES konnten im Mittel mehr als 96 Trajektorien am Tag bzw. 4 Trajektorien pro Stunde berechnet werden. Meist lag die Anzahl weit darunter. Daher führen einzelne berechnete Trajektorien zu sehr großen Streuungen und stark schwankenden Durchdringungen. Eine befriedigende Abbildung der realen Verkehrsstärkeänderungen eines Tages scheint mit dieser geringen Trajektoriendichte nur schwer möglich.

An den SES auf der Autobahn wurden trotz teilweise relativ hoher Trajektorienanzahl nur mäßige Korrelationen ermittelt. Eine mögliche Erklärung für diesen Umstand könnte die fehlende Berücksichtigung des tageszeitabhängigen Telefonierverhaltens der Verkehrteilnehmer sein. Die Auswertung der Teilnehmer hängt sehr stark von der Anzahl und Länge der Telefonate ab, die in den Fahrzeugen geführt werden. Bislang wird von einem konstanten Telefonierverhalten der Teilnehmer ausgegangen. Diese Annahme ist jedoch falsch. Dazu sind jedoch umfassende Untersuchungen zum Telefonierverhalten erforderlich. Nur so ist es möglich diesen Effekt aus den Tagesganglinien zu eliminieren.

Des Weiteren reicht die Auflösung der Abis-Daten im Innenstadtbereich derzeit nicht aus, um in größerem Umfang korrekte Fahrzeugtrajektorien zu generieren. Es ist nicht zu erwarten, dass eine weitere Optimierung der Auswertesoftware alleine diesen technisch bedingten Mangel in den originären Mobilfunkdaten auf Abis-Basis lösen kann.

5.2.2 A-Ebene

Die Beurteilung der Trajektorie aus A-Daten erfolgte auf Grund der sehr großen Anzahl von 173 Referenzsensoren im gesamten Testgebiet nach Gebieten unterteilt. Zunächst wurden die in Abbildung 5-20 dargestellten Sensoren zum Vergleich im Raum Karlsruhe herangezogen. Es folgte dann die Evaluierung in Stuttgart (vgl. Abbildung 4-13 ff.) bevor der Rest des Testgebietes und damit vorwiegend das Autobahnviereck untersucht wurde (vgl. Abbildung 4-9 f).

Der Untersuchungszeitraum erstreckt sich vom Montag, 23.03. bis zum Dienstag, den 31.03. Aus ungeklärter Ursache standen am Samstag, den 28.03. keine A-Daten zur Verfügung, daher beschränkt sich die Evaluierung auf insgesamt 8 Tage.

Die Durchdringung mit A-Daten erreicht bis zu 15 ‰, sinkt in der Innenstadt jedoch auf bis zu 0,4 ‰ ab. Die zugehörigen Streuungen liegen außerorts bei bis zu 9 ‰ und innerorts etwa auf dem Niveau der Durchdringung. Die Standardabweichungen erreichen außerorts bis zu 3 ‰ und Innerorts 1 bis 2 ‰. Auf der Kriegstraße in unmittelbarer Nähe des Stadtzentrums liegt die Durchdringung und die Streuung bei 0,4 bis 0,6 ‰. Hier ist aufgrund der begrenzten Auflösung der A-Daten, die noch großmaschiger ist als bei den Abis-Daten, eine ausreichende Abbildung des Verkehrs derzeit nicht möglich. Auf der Südtangente sowie auf der Autobahn in Stadtnähe hingegen können durchschnittlich 6 – 20 Trajektorien pro Stunde berechnet werden.

Die Korrelationen der A-Trajektorien mit den Verkehrsstärken aus SES liegen auf der Südtangente und auf der Autobahn durchweg über 0,5 und erreichen bis zu 0,8. Die zugehörigen Streuungen liegen meist unter dem halben Wert für die Durchdringung. Die Standardabweichungen von unter 0,2 deuten auf ein homogenes Verhalten hin.

Innerorts scheint eine ausreichende Detektion von Veränderungen der Verkehrsstärke nur schwer möglich zu sein. Das zeigen die Korrelationen von unter 0,35 bei Streuungen von 0,7 und mehr. Bei Betrachtung der absoluten Anzahl berechenbarer Trajektorien erklären sich diese schlechten Werte sofort. Im Mittel konnten nur 2 oder weniger Trajektorien je Stunde berechnet werden. Die Hochrechnung der Verkehrsstärke sowie die Feststellung von signifikanten Änderungen der Verkehrsstärke wird damit nahezu unmöglich.

In Stuttgart konnten aufgrund des Ausfalls der A-Daten-Aufzeichnung am zweiten Tag der Evaluierung bis zum Rest der Projektlaufzeit keine Mobilfunktrajektorien mehr berechnet werden. Die Evaluierung in Stuttgart beschränkt sich daher auf einen bzw. 2 Tage (die Durchdringung konnte für den 24.03. nach dem Ausfall um ca. 11 Uhr noch extrapoliert werden).

Die Durchdringungen, die in Stuttgart ermittelt werden konnten, bestätigen mit weniger als 2 ‰ die Werte in Karlsruhe. Lediglich auf den Ausfallstraßen B295 und B10 nördlich vom Stadtzentrum werden höhere Durchdringungen bis 6,2 ‰ erreicht. An diesen Sensorstandorten ist jedoch die Infrastrukturdichte wesentlich kleiner als bei den anderen Sensoren, die in Stuttgart zur Verfügung standen.

Die Korrelation mit der tatsächlichen Verkehrsstärke erreichen bei den Sensorstandorten außerhalb des Stadtzentrums bis zu 0,83, innerstädtisch liegen sie mit einer Ausnahme bei teilweise weit unter 0,6. Wie bereits in Karlsruhe lässt sich dieser Umstand teilweise durch die geringe Anzahl auswertbarer Teilnehmer auf Grund der geringen Auflösung der A-Daten erklären. Einige SES-Standorte z. B. in der Nähe des Pragsattels zeigen jedoch trotz hoher Straßendichte mit 2 bis 3 Trajektorien pro Stunde und

Korrelationen von 0,6 bessere Werte. Vermutlich ist auf diesen Strecken ein sehr hoher Transitanteil vorhanden, der vor der Einfahrt bzw. nach der Ausfahrt in den Innenstadtbereich berechnet werden kann. Dadurch gelingt es gelegentlich, die mangelhafte Auflösung im Mittelteil durch Verbindung der beiden Anfangs- und Endstücke zu überwinden und die Gesamttrajektorie zu berechnen.

Für Teile des Autobahnnetzes standen ab dem zweiten Tag aufgrund des LAC-Ausfalls ebenfalls keine Mobilfunkdaten mehr zur Auswertung zur Verfügung. Der Ausfall betraf 44 von 136 SES und damit ca. 1/3 der Sensorstandorte. Des Weiteren wurden 11 SES des Innenministeriums identifiziert, über die an keinem der 9 Evaluierungstage FPD-Trajektorien verliefen. Da benachbarte SES hohe Durchdringungen mit A-Trajektorien aufwiesen, liegt die Ursache darin, dass die entsprechenden REID der digitalen Straßenkarte nicht im Lokalisierungsnetz vorhanden waren. Es handelt sich damit nicht um einen Verfahrensfehler. Die entsprechenden Sensoren wurden daher aus der Auswertung ausgeschlossen.

An 82 % aller evaluierbaren SES Standorte auf dem Autobahnnetz wurde eine durchschnittliche Durchdringung von mehr als 5 ‰ ermittelt. Bei der Hälfte der Sensoren liegt die Korrelation über 0,75 bei einer relativ kleinen Streuung von <= 0,4 sowie einer Standardabweichung von < 0,16. Die absolute Anzahl berechneter Fahrzeugtrajektorien erreicht im Mittel 8 bis 25 pro Stunde. Wie die in [AP10-2 2009] dargestellten Beispiele von gegenübergestellten Tagesganglinien bestätigen, ist mit diesen Werten eine Abbildung der Verkehrslage möglich. Ungewöhnliche Schwankungen in der Verkehrsstärke, die auf Stauungen hindeuten, sind innerhalb weniger Minuten detektierbar.

Die Kriterien Durchdringung, Rangkorrelation und absolute Trajektorienanzahl sind zur Beurteilung einzelner Straßenabschnitte hinsichtlich der Überwachung mit A-Daten geeignet. Eine exakte Festlegung der Grenzwerte erscheint jedoch nur in Verbindung mit der Evaluierung weiterer Anwendungen sinnvoll.

5.2.3 A-Ebene versus Abis

Die Beurteilung der unterschiedlichen Datenquellen zur Generierung von Trajektorien auf dem Straßennetz konnte nur in Karlsruhe erfolgen. In Ballungszentren sind auch die wesentlichen Vorteile der höher aufgelösten Abis-Daten zu erwarten. Im Bereich der Autobahn in Karlsruhe sowie der Südtangente, die ein ähnliches Verhalten aufweist, kann mit Abis-Daten lediglich weniger als 1/3 der Durchdringung erreicht werden, wie es bei A-Daten der Fall ist. Dies ist klar, da die Abis-Daten-Auswertung, anders wie bei der A-Daten-Auswertung nur bei durchgängigen Telefonaten durchgeführt werden kann. Derzeit ist aufgrund der mangelnden Berücksichtigung des Telefonierverhaltens der Straßenverkehrsteilnehmer keine höhere Qualität der Abis-Trajektorien feststellbar. Die Korrelationen liegen derzeit weit unter den mit A-Daten erreichbaren. Dies kann sich jedoch durch Anbringen geeigneter Korrekturfaktoren zukünftig ändern.

Im Innenstadtbereich von Karlsruhe liegen sowohl die Durchdringungen, als auch die Korrelationen auf vergleichbar niedrigem Niveau. Hier kann eine Berücksichtigung des Telefonierverhaltens eine leichte Verbesserung der Ergebnisse aus Abis-Daten bringen, allerdings kann die niedrige Durchdringung dadurch nicht erhöht werden. Eine weitere Erhöhung der Rechenleistung kann Abhilfe schaffen.

Der Vorteil der Abis-Daten liegt in der höheren Auflösung der Daten, der Vorteil der A-Daten liegt in der viel höheren Datendichte. Aufgrund der genannten Probleme, kann die Prozessierung der Abis-Daten derzeit seinen Vorteil noch nicht ausspielen. Wie der Demonstrator der Stadt Karlsruhe jedoch gezeigt hat, genügt die derzeitige Qualität der Abis-Trajektorien für eine Detektion der Verkehrslage bereits aus, sofern eine ausreichende Anzahl Trajektorien generiert werden konnte (vgl. [AP10-2 2009]).

Eine Kombination der beiden Verfahren bzw. Datenquellen scheint die beste Lösung zu sein. Außerhalb der Ballungsräume in Gebieten mit weitmaschiger Infrastruktur können A-Daten ausreichend sein, um die Verkehrslage zu ermitteln. In Stadtnähe kann nur eine Ergänzung mit Abis-Daten eine Erhöhung der Auflösung bis hin zu Umgehungsstraßen sowie wichtigen Ein- und Ausfallstraßen ermöglichen.

© Do-iT 2009

6 Zusammenfassende Bewertung

Hier wird die Qualität der erzielten Ergebnisse in konzentrierter, vereinfachter aber trotzdem korrekter Form zusammengefasst. Tabelle 6-1 Beinhaltet die numerischen Werte, die den Evaluierungen der vorstehenden Kapitel entnommen sind.

Tabelle 6-1: Vergleichende Zusammenstellung der Qualitätsparameter für A- und Abis-Daten; zum Teil untergliedert in Straßentyp (A+D =Autobahn und Durchgangsstraßen, I = innerorts) und Mobiltelefonstatus (A = aktiv, P = passiv)

	Abis-Daten	**A-Daten**	**Bemerkung**
Aktualität Berechnungsdauer / Alter der Trajektorie	70s / 27-35min	8s / 26-37min	
Verfügbarkeit	97 %	A: 100 % P: 33 %	Werte nicht direkt vergleichbar
Korrektheit	87 %	A: > 80 % P: < 42 %	Werte gelten für Rahmenbedingungen, die in den jeweiligen Kapiteln genannt sind
Querabweichung	< 250 m	--	für A-Daten nicht berechenbar
Geschwindigkeit RMS / Durchschnitt	20 km/h / 8 km/h	20 km/h / 10 km/h	
Durchdringung	A+D: 2 - 3 ‰ I: 0,4 ‰	A+D: > 5‰ I: 0,4 – 2,0 ‰	
Korrelation	< 0,5	A+D: 0,5 - 0,8 I: < 0,6	

Die Aktualität der Daten ist nicht ausreichend. Dieses ist aber mit dem aktuellen Bereitstellungsalgorithmus zu begründen, der aufgrund der Verzögerung in der Datenbereitstellung durch T-Mobile vereinfacht entwickelt wurde. Dieses ist zukünftig anzupassen.

Die Verfügbarkeit der Trajektorien ist für beide Verfahren bei aktivem Telefon sehr hoch. Das Gleiche gilt für die Korrektheit. Hier sind die Abis-Daten im Vorteil.

Der Genauigkeitsparameter Querabweichung zeigt die aktuelle Grenze des Verfahrens für die Abis-Ebene auf. Straßen mit geringeren Abständen kleiner als 250 m können nur in seltenen Fällen korrekt detektiert werden. Im Gegensatz hierzu sind die A-Daten deutlich ungenauer; eine numerische Festlegung auf km-Werte ist nicht sinnvoll. Hier kann auch der Hauptvorteil der Abis-Daten gesehen werden: die höhere räumliche Auflösung.

Die Genauigkeitsparameter für die Geschwindigkeit zeigen auf, dass insbesondere die mittleren Abweichungswerte durchaus in Konkurrenz zu den Ergebnissen der Stationären Erfassungssysteme treten können. Die Abis-Daten sind wiederum in geringem Umfang überlegen.

Bezüglich Durchdringung und Korrelation zeigt sich dagegen die Überlegenheit der A-Daten; zumindest auf Autobahnen und Durchgangsstraßen. Eine Hochrechnung (Durchdringung) und eine Abbildung des Verkehrsflusses (Korrelation) gelingt hier. Für den innerstädtischen Bereich liefert der Durchdringungsparameter mit A-Daten an ausgewählten Stellen interpretierbare Ergebnisse. Mit Abis-Daten können Autobahnen und Durchgangsstraßen hochgerechnet werden.

A-Daten können den Verkehrsfluss und die mittlere Geschwindigkeit auf Autobahnen und Durchgangsstraßen überwiegend korrekt und mit geringer räumlicher Auflösung abbilden. Die höhere Genauigkeit und damit die verbesserte räumliche Auflösung sowie die erhöhte Korrektheit lassen die Abis-Ebene als Ergänzung im städtischen Bereich als sehr sinnvoll erscheinen. Für sinnvolle Hochrechnungen ist jedoch noch das charakteristische Telefonierverhalten von Straßenverkehrsteilnehmern zu berücksichtigen.

Die Ergänzung der A-Daten um die genaueren, detaillierten und korrekteren Abis-Positionen und Geschwindigkeiten ist insbesondere unter dem Gesichtspunkt positiv zu bewerten, dass zum einen die Rechnergeschwindigkeiten weiter steigen werden und zum anderen beide Datenquellen in der GSM-Nachfolgetechologie UMTS an einer Stelle des Mobilfunknetzes abgegriffen werden und somit immer zusammen zur Verfügung stehen werden. Eine integrierte Auswertung drängt sich – bei zunehmender am Markt verfügbarer Rechnerleistung – geradezu auf.

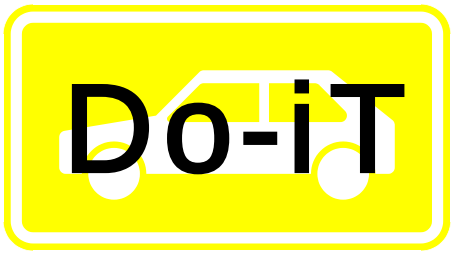
7 Literatur

AP10-1 (2008)	Do-iT: Evaluierungskonzept. Interner Do-iT Bericht zum AP10-1, Bestandteil des ZMS3, 2008.
AP10-2 (2009)	Do-iT: Evaluierungsbericht. Interner Do-iT Bericht zum AP10-2, Bestandteil des ZMS4, 2009.
AP3-5 (2008)	Schollmeyer, R.: Bericht zur Identifizierung aktiver Verkehrsteilnehmer. Interner Do-iT-Bericht AP3, Bestandteil des ZMS3, 2008.
AP3-6 (2009)	Schollmeyer, R.: Bericht zur Qualitätsbewertung der Identifizierung aktiver Verkehrsteilnehmer. Interner Do-iT-Bericht AP3, Bestandteil des ZMS4, 2009.
AP4-5 (2008)	Czommer, R.: Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren. Interner Do-iT-Bericht AP4, Bestandteil des ZMS3, 2008.
AP8-2 (2006)	Laufer, R., Wiltschko, T. (2006): Qualitätsmodell für Do-iT (V1.0). Interner Do-iT Bericht AP8-2, Bestandteil des ZMS1, 2006.
AP8-7 (2009)	Laufer, R., Schwartz, S. (2009):Nachweis der Wirksamkeit der Qualitätssicherungsmethoden. Interner Do-iT Bericht AP8-7, Bestandteil des ZMS4, 2009.
Karrer (2008)	Karrer, K.: Untersuchung der Qualität von stationären Erfassungssystemen am Beispiel der Autobahn A81. Studienarbeit am IAGB, Universität Stuttgart, nicht veröffentlicht, 2008.
Ramm & Schwieger (2007)	Ramm, K., Schwieger, V.: Mobile Positioning for Traffic State Acquisition., Proceedings on 5th Symposium on Mobile Mapping, Padua, 2007.
Sachse (2002)	Sachse, T.: Die Messwertqualität autarker Detektionssysteme am Beispiel der Traffic Eyes Universal. Straßenverkehrstechnik, Nr. 9, Seiten 463-470, Kirschbaum-Verlag GmbH : Bonn, 2002.
Schollmeyer (2007)	Schollmeyer, R.: Spezifikation Datenaufbereitung für den Datenimport ins Lokalisierungsnetz. Interner Do-iT-Bericht Version 017 vom 03.07.2007.
Schwieger et al. (2007)	Schwieger, V., Ramm, K., Czommer, R., Möhlenbrink, W.: Mobile Phone Positioning for Traffic State Acquisition, Journal of Applied Geodesy, Walter de Gruyter, Berlin – New York, Vol. 1, Iss. 3 11/2007.
TLS (1997)	Bundesanstalt für Straßenwesen BaSt: Technische Lieferbedingungen für Streckenstationen TLS, 1997.

Arbeitspaket 3 Bericht

Datenoptimierung für integrierte Telematik

Do-iT



Förderkennzeichen: 19 B 5007

AP3 Bericht zur

Identifizierung aktiver Verkehrsteilnehmern

2008

Ralf Schollmeyer

Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen

Dokument-Informationen

Version	Datum	Bearbeiter	Bemerkung
0001	**21.06.07**	**Schollmeyer**	• Erstellung und Zusammenführung der Vorarbeiten
0059	**10.08.08**	**Schollmeyer**	• Überarbeitung für Endbericht
0069	**18.08.08**	**Schwieger**	• Review

Inhaltsverzeichnis

1 Glossar 5

2 Einleitung 6

3 Aufgabenstellung 7

3.1 Anforderungen an die FPD-Identifizierung 7

3.2 Vorgehensweise der Identifizierung aktiver Verkehrsteilnehmer 7

3.2.1 Auswahl geeigneter Fahrzeugklassen 9

3.2.2 Übersicht der in Do-iT entwickelten Identifizierungsverfahren 11

4 Methoden der FPD-Identifizierung und -Klassifizierung 13

4.1 FPD-Identifizierung auf GSM-Daten Ebene 13

4.1.1 IMEI-(TAC-) Kennung 13

4.1.2 Räumliche Sektor-Maske, MS in interessierender Region ? 15

4.1.3 Priorisierung von Mobilfunkdaten für die Online-Auswertung 20

4.2 Ortungsinformationen berechnet aus GSM-Mobilfunkdaten 23

4.3 Diskretisierung der Geodaten 24

4.3.1 Definition des Kachelkoordinatensystems 25

4.3.2 Auswahl der Kachelgröße 26

4.4 Aufenthaltswahrscheinlichkeit eines FPD in einer Kachel 27

4.5 Bewegungsanalyse 30

4.5.1 Bewegungsanalyse mittels transformierter GSM-Daten 30

4.5.2 Clusterverfahren zur Berechnung von Positionsschwerpunkten für die Identifikation von stehenden und sich bewegenden FPD 36

4.5.3 Klassifizierung der FPD-Bewegung mit FPD-Schwerpunkten 39

4.6 FCM – Koordinatenschwerpunktbildung bewegter FPD 41

4.7 Aufenthaltsanalyse, POI 43

4.7.1 Geodatenintegration 43

4.7.2 Vorprozessierung für die FPD-Echtzeitanalyse 45

4.8 Klassifizierung von FPD in Öffentlichen Verkehren 46

4.8.1 Vorgehensweise 46

4.8.2 Planungsdaten der Öffentlichen Verkehre 47

4.8.3 Fuzzyfizierung der Fahrplandaten 53

4.8.4 Vorprozessierung für die FPD-Echtzeitanalyse 58

4.8.5 Datenmodellierung für Ganzjahresfahrpläne 59
4.8.6 PlanÖV-Klassifizierung der FPD 61
4.9 Verkehrsnetzgestützte Verkehrsteilnehmerklassifizierung 65
4.9.1 Definition typischer Verkehrsnetze von Verkehrsteilnehmern 65
4.9.2 Datenvorprozessierung 68
4.9.3 Bestimmung der FPD-Verkehrsklassengruppenwahrscheinlichkeit 69
4.10 Topologische Überprüfung der Verkehrsklassenwahrscheinlichkeit 72
4.10.1 Ein Verfahren zur logischen Überprüfung der Verkehrsklassenwahrscheinlichkeit: Routensuche mit Restriktionen aus den FPD-Schwerpunkten 72
4.10.2 Räumliche Mindestauflösung 80
4.10.3 Verkehrsklassenbewertung anhand der gefundenen Routen 81
5 Literatur 83

1 Glossar

aktive Verkehrsteilnehmer		Verkehrsteilnehmer welche für die Erfassung der IV Verkehrslage nutzbringende Informationen liefern
VT	**V**erkehrs**t**eilnehmer	
IV	**I**ndividual**v**erkehr	
MIV	**M**otorisierter **I**ndividual**v**erkehr	
FPD	**F**loating **P**hone **D**ata	Aus Mobilfunkdaten gewonnene Daten zur Erfassung der Verkehrslage, bestehend aus Trajektorien
Trajektorie		Positionsfolge mit zeitlichem Bezug und Referenzierung auf die digitale Karte
MFT	**M**obil**f**unk**t**eilnehmer	
Do-iT		
BSC	**B**ase **S**tation **C**ontroller	
BTS	**B**ase **T**ransceiver **S**tation	Funkantenne im Mobilfunknetz
CGI	Cell Global Identity	Zellortung / globale Zell-Identifikation
CI	Cell Identity	Zell-ID
LA	Location Area	
LAC	Location Area Code	Location Area-ID
MS	Mobile Station	Mobiles Endgerät (Handy)
MSC	Mobile Switching Center	
RXLEV		Signalstärke
TA	Timing Advance	
TMC	Traffic Message Channel	
HLR	Home Location Register	
VLR	Visitor Location Register	
LMU	Location Measurement Unit	
Verkehrsklassengruppe		Zusammenfassung mehrerer Verkehrsklassen

2 Einleitung

Mobilfunktelefone können durch Auswertung der Kommunikationssignale auf der Seite der Infrastruktur geortet werden. Die daraus ableitbaren und als **FPD** (Floating Phone Data) bezeichneten Informationen, stützen die Erfassung der Verkehrslage und Ermittlung des Routenwahlverhaltens von Verkehrsteilnehmern. Die datenschutzrechtlich bedingte Anonymisierung der Kommunikationssignale, die Nutzung verschiedener Verkehrsmittel durch einen Mobilfunkteilnehmer oder verkehrlich inaktive MFT, erfordern die Identifikation der verkehrstechnisch relevanten FPD. Können in einem Gebiet bspw. mehrere Verkehrsmittel genutzt werden, wie in Abbildung 2-1 exemplarisch dargestellt, ist für die Verkehrslageerfassung eine verkehrstechnische Identifikation der FPD notwendig. Solang Mobilfunkgeräte zum Beispiel in öffentlichen Verkehren, Gebäuden oder in Freizeiteinrichtungen nicht von denen des Individualverkehrs getrennt wurden, könnten sie die Zuverlässigkeit der Verkehrslageerfassung negativ beeinträchtigen.

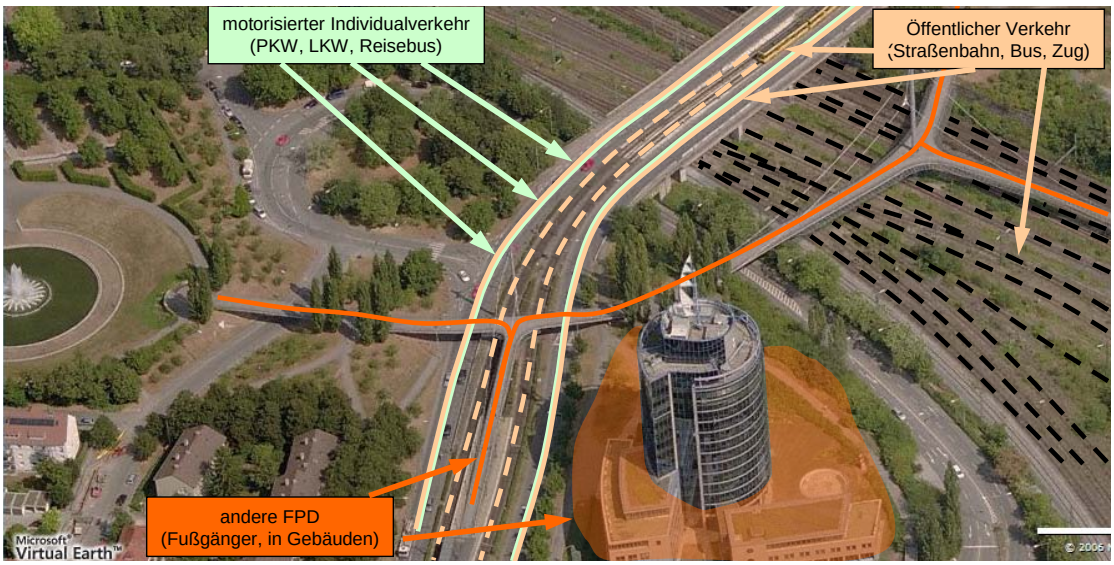


Abbildung 2-1: Problemstellung der verkehrstechnischen Identifikation von FPD

Im vorliegenden Bericht werden die im Projekt Do-iT entwickelten Verfahren zur verkehrstechnischen Identifizierung der FPD vorgestellt. Mit den Identifikationsverfahren wird eine Verbesserung durch Elimination von Verkehrsteilnehmern welche nicht am motorisierten Individualverkehr teilnehmen angezielt. Zudem sollen verkehrstechnisch interessante, sogenannte „aktive Verkehrsteilnehmer“ selektiert werden.

3 Aufgabenstellung

3.1 Anforderungen an die FPD-Identifizierung

In der FPD-Identifizierung sind aus den Mobilfunkdaten diejenigen herauszufiltern, die keine verkehrstechnische Relevanz für die Verkehrslageerfassung haben. Da FPD in der zeitlichen Folge verschiedene Verkehrsmodi nutzen bzw. keine Partizipierung mit dem Verkehrgeschehen aufweisen können, sind auch zeitliche Änderungen des Status zu berücksichtigen.

Die Zielsetzungen machen die zeitnahe Ausführung der Rechenoperationen erforderlich. Ebenso sollen keine menschlichen Eingriffe in die Prozessierung erforderlich sein.

Im Projekt Do-iT sind vorrangig nur Verkehrsteilnehmer des individuellen motorisierten Verkehrs von Interesse. FPD die diesen Verkehrsklassen nicht angehören, sollen, besonders für die Echtzeitanalysen, rechtzeitig herausgefiltert werden. Ebenso können FPD die keine Partizipation am Verkehrsgeschehen aufweisen, frühzeitig selektiert werden, wodurch zu Gunsten der verkehrstechnisch relevanten FPD, Rechenressourcen eingespart werden.

3.2 Vorgehensweise der Identifizierung aktiver Verkehrsteilnehmer

Zu Beginn der Identifizierung startet jedes FPD als passiv bis der Status aktiv mittels der Verfahren eindeutig identifiziert werden konnte.

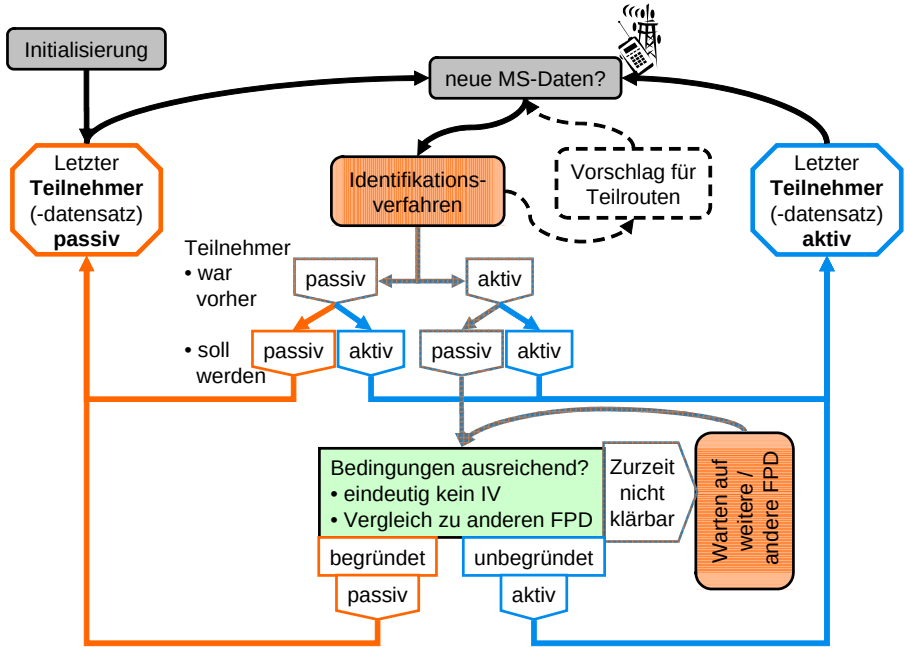


Abbildung 3-1 Ablaufdiagramm für die Online-Identifizierung

Für die weiteren Identifizierungen werden die GSM-Daten transformiert in kartesische Koordinaten bzw. Koordinatenfolgen. Dieser Vorgang wird Ortung oder Positionsbestimmung genannt. Die Ortung erfolgt mit den in AP4 entwickelten Algorithmen, wobei zwischen der Datenquellen der GSM-Abis- und GSM-A-Schnittstelle unterschieden wird. Bei der Datenquelle aus der GSM-Abis-Schnittstelle werden Positionsaktualisierungen von ca. 0,48 sec erreicht und es sind Genauigkeiten mehreren 100m zu erwarten. Für Ortungen der GSM-A-Schnittstelle ist mit einer deutlich geringeren Positionsaktualisierung, da GSM-Informationen nur bei einem Hand-Over erzeugt werden, und es ist mit einer geringeren Koordinatengenauigkeiten von bis zu mehreren Kilometern auszugehen.

Die folgenden Verfahren stützen sich im Wesentlichen auf die Koordinaten aus der GSM-Abis-Schnittstelle. Da die hohe Positionsaktualisierungsrate im Vergleich zur geringen Koordinatengenauigkeit ineffizient für die Analysen ist, wird im ersten Schritt eine Reduzierung der Koordinaten auf signifikante Koordinatenschwerpunkte durchgeführt, die im Abschnitt „Berechnung signifikanter FPD-Koordinatenschwerpunkte“ erläutert wird.

Die weiteren Analysen werden in kartenunabhängige und kartengestützte Analyse gegliedert. Beginnend mit der kartenunabhängigen Geschwindigkeitsanalyse wird eine Trennung von stehenden und sich bewegenden Mobilfunkteilnehmern durchgeführt, um eine verkehrliche Unterscheidung zu gewährleisten und um nicht erforderliche Analyseschritte einzusparen. Wird ein Teilnehmer als sich bewegend eingestuft, wird anschließend mit der kartenunabhängigen Geschwindigkeitsanalyse die Bewegungsgeschwindigkeit bzw. der Verlauf der Bewegungsgeschwindigkeit sowie die daraus ableitbare Verkehrsteilnehmerklasse ermittelt.

In den kartengestützten Analysen werden weitere Informationsquellen, wie digitale Straßennetze, ÖV-Fahrpläne und POI, zur Verbesserung der Verkehrsklassenidentifikation verwendet. In der „kartengestützten Geschwindigkeitsanalyse“ werden die vom Mobilfunkteilnehmer vermeintlich gewählten Routen mittels map-matching-Technologien zu einer Gesamtroute zusammengeführt, wodurch die Bestimmung des zurückgelegten Weges und damit die Berechnung der Geschwindigkeit verbessert werden kann. Unter der Annahme, dass die Fortbewegungen auf Verkehrswegen von Verkehrsnetzen stattfinden, werden die FPD-Positionen mit dem Vorhandensein und der Nutzungsmöglichkeit dieser Verkehrsnetze in der „verkehrsnetzgestützten Verkehrsteilnehmerklassifizierung“ untersucht. Mittels der Prognose von Fahrzeugpositionen des ÖV wird anschließend geprüft, ob Mobilfunkteilnehmer einen ÖV-Verkehrsmodus gewählt haben. In der „Aufenthaltsanalyse“ wird angenommen, dass an bestimmten POI nur ausgewählte Verkehrsteilnehmer anzutreffen sind. Wird ein FPD in der Nähe eines solchen POI geortet, so kann auf dessen Verkehrsteilnehmerklasse zurück geschlossen werden.

Die bisherigen Analysen liefern sich redundante, sich ergänzende, aber auch konkurrierende Ergebnisse. Mit einer Plausibilitätsanalyse werden diese Einzelergebnisse zusammengeführt und, soweit möglich, eine abschließende Verkehrskehrsteilnehmerklasse ermittelt. Da im Laufe der Bewegung vom Mobilfunkteilnehmer unterschiedliche Verkehrsmodi gewählt werden können, ist die Betrachtung von Teilrouten erforderlich. Teilrouten werden in der FPD-Identifizierung als Trajektorienabschnitte eine FPD mit verschiedener Verkehrsmoduswahl betrachtet.

Nutzen mehrer FPD dieselben Fahrzeuge, ist verkehrstechnisch gesehen eine Zusammenfassung dieser FPD, das so genannte Clustern von FPD, anzustreben. In der abschließenden „Teilnehmer übergreifenden Klassifizierung“ werden hierfür Clusterverfahren vorgestellt, die nach ähnlichen Identifikationsmerkmalen der FPD suchen.

3.2.1 Auswahl geeigneter Fahrzeugklassen

Das Kraftfahrt-Bundesamt (KBA, http://www.kba.de/) nutzt den Begriff Fahrzeugklasse für die Unterscheidung von Krafträdern (KRAD), Personenkraftwagen (PKW), Kraftomnibussen (Reise-, ÖV- Bus), Lastkraftwagen (LKW), Zugmaschinen und sonstige Kfz. Bis auf Busse des Öffentlichen Verkehrs gehören alle dem motorisierten Individualverkehr an. Auf Basis dieser Definitionen sowie der Hinzunahme werden für das Projekt Do-iT die Einteilungen erweitert um die Unterteilungen des Öffentlichen Verkehrs und einer Gruppe, die ohne Verkehrsbeteiligung ist. Aus Sicht der Verkehrslageerfassung lassen sich darauf aufbauend, vgl. Abbildung 3-2, FPD in die drei übergeordneten Gruppen von Mobilfunkteilnehmern „ohne Verkehrsbeteiligung“, im „Individualverkehr“ sowie im „Öffentlichen Verkehr“ einteilen.

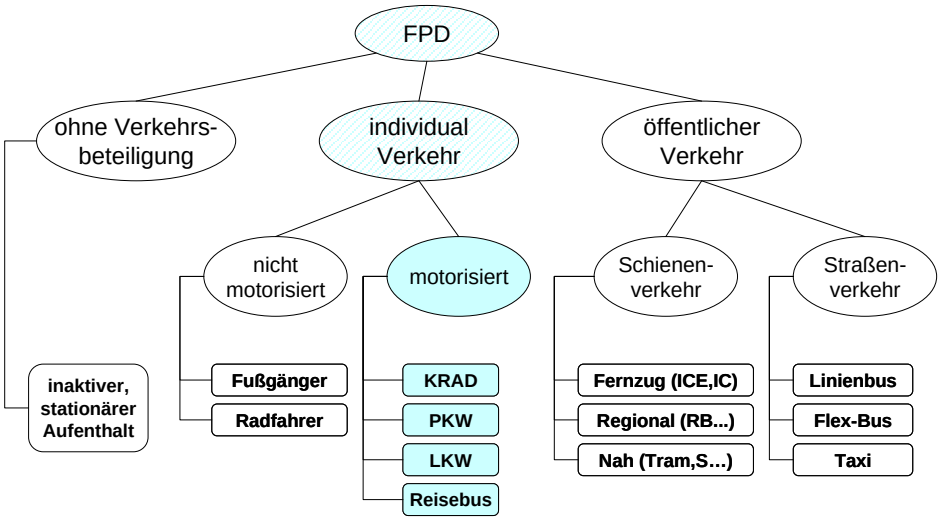


Abbildung 3-2: Unterscheidung von Teilnehmerklassen im Projekt Do-it

Für die Aufgaben im Projekt Do-iT sind ausschließlich die FPD relevant, die im Individualverkehr der Gruppe des motorisierten Verkehrs angehören. Diese FPD werden folgend als **aktive FPD** und FPD in einer anderen Gruppe als **inaktive FPD** bezeichnet. Verfahren der Identifizierung können somit entweder FPD:

1. die der Gruppe aktive FPD angehören selektieren oder
2. die der Gruppe inaktive FPD zugeordnet werden können, ausschließen.

Je nach Funktionsweise der Verfahren werden ausgewählte Verkehrsklassen selektiert oder ausgeschlossen. Die Ergebnisse der einzelnen Verfahren sind anschließend auf Plausibilität zu untersuchen und für eine abschließende Bewertung zusammenzuführen.

Mobilfunkteilnehmer können verschiedene Transportmittel wählen, Transportmittel wechseln oder sich zwischenzeitlich stationär aufhalten. Identifizierungen sind daher für jede Ortungsinformation erforderlich. Aufgrund der hohen Positionsaktualisierung von bis zu ca. 2 Hz ist die Bildung von so bezeichneten **Teilrouten**, also Folgen von Positionen sinnvoll (vergl. Abb. 3.1). Teilrouten beinhalten FPD-Positionsfolge mit einer Identifizierungsinformation, ob sie zu den aktiven bzw. inaktiven FPD gehören und zu welchen Verkehrsklassen sie einzuordnen sind.

Wie einleitend genannt, haben die FPD-Ortungsinformationen eine gewisse Qualität, die in Form von Genauigkeitsmaßen angegeben wird. Die Qualität der darauf aufbauenden Methoden ist demzufolge abhängig von der Qualität der FPD-Ortungsinformationen. Neben der Zuordnung der FPD zu einer Verkehrsklasse, soll weiterhin die Qualität der Identifikation durch geeignete Maße, wie z.B. Wahrscheinlichkeiten oder Zugehörigkeiten, ermittelt werden.

Kann mit genügender Sicherheit der Ausschluss von Mobilfunkdaten für die Verkehrserfassung festgestellt werden, ist ein Abbruch der weiteren Analysen und damit eine Einsparung von Rechenressourcen möglich. In Abbildung 3-3 ist der Ablauf der Identifizierungsverfahren dargestellt. Die einzelnen Verfahren werden in den folgenden Abschnitten erläutert.

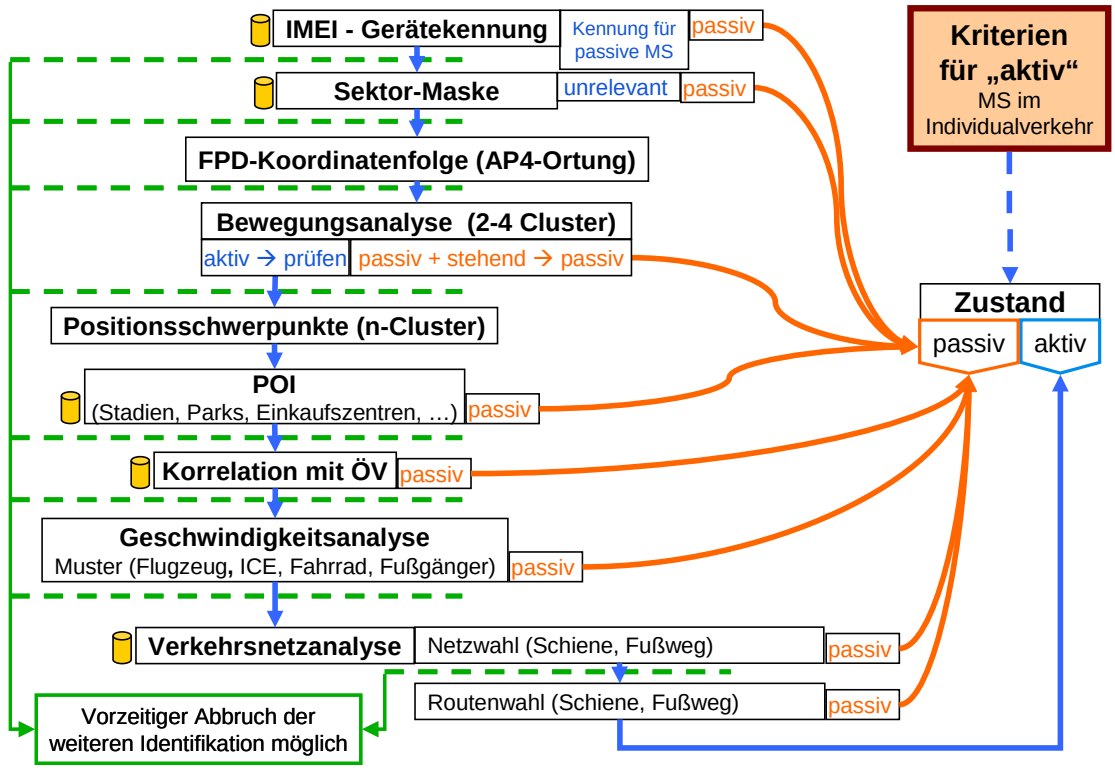


Abbildung 3-3: Ablaufreihenfolge der Identifikationsverfahren im Online-Betrieb

3.2.2 Übersicht der in Do-iT entwickelten Identifizierungsverfahren

Tabelle 1: Stand der im Rahmen von Do-iT implementierten Verfahren

Identifizierungsverfahren (Abschnitt)	ge-plant	In Arbeit	imple-mentiert	In Betrieb
IMEI – Gerätekennung (4.1.1)	X	X		
Sektor Maske (4.1.2)	X	X	X	X
xy – Bewegungsanalyse (4.5)	X	X	X	X
xy – Positionsschwerpunkte (4.5.2)	X	X	X	X
xy - POI – Analyse (4.7)	X	X		
xy - Korrelation mit ÖV (4.8)	X	X	X	X
xy - Geschwindigkeitsanalyse ()	X	X		
xy - Verkehrsnetzanalyse (4.9)	X	X	X	X
xy - Unscharfes Routing (4.10)	X	X	X	X
A-Daten - Unscharfes Routing ()	X	X	X	

In vorstehender Tabelle 1 wird für die verschiedenen, in diesem Dokument beschriebenen Identifikationsverfahren der aktuelle Zustand dokumentiert. „In Arbeit“ bedeutet dabei, dass das Verfahren entweder neu entwickelt wird oder zur

Zeit optimiert wir, obwohl es bereist implementiert ist. Die Spalte „in Betrieb“ dokumentiert, dass das entsprechende Verfahren bereits auf Abis-Ebene auf dem FPD-Server im Online-Betrieb läuft.

4 Methoden der FPD-Identifizierung und -Klassifizierung

Die Mobilfunkteilnehmer werden anonymisiert und ohne Information auf eine Verkehrsklassenzugehörigkeit erfasst. Zur Ableitung und Interpretation der Verkehrssituation werden die FPD auf Partizipierung mit dem Individualverkehr, Öffentlichen Verkehr oder keiner Beteiligung am Verkehrsgeschehen verkehrstechnisch untersucht. Im Folgenden werden Methoden zur Prozessoptimierung für den Online-Einsatz von Geodaten und die Reduzierung der Mobilfunkdaten sowie den daraus ermittelten FPD-Positionen vorgestellt. Darauf aufbauend werden Verfahren erläutert, die eine verkehrsrelevante Identifizierung des Mobilfunkteilnehmers vornehmen, wobei entweder FPD in relevanten Verkehrsklassen hervorgehoben oder in nicht relevanten ausgeschlossen werden. Das Grundprinzip für die Online-Identifizierung basiert auf dem Ausschlussprinzip von FPD passiver Verkehrsteilnehmer.

4.1 FPD-Identifizierung auf GSM-Daten Ebene

Beim Abgriff der Mobilfunkdaten aus dem Telekommunikationsdatennetz entstehen erhebliche Datenmengen in kurzer Zeit. Eine frühzeitige Reduktion uninteressanten Daten bereits auf Mobilfunkdaten-Telegramm-Ebene (GSM-Daten-Ebene) wird daher zur effizienten Nutzung der Rechenressourcen angestrebt. Für die online-Auswertung werden drei Verfahren auf GSM-Daten-Ebene eingesetzt. Im ersten Verfahren wird die Kennung der Endgerätenummer zur Selektion verwendet. Im zweiten Verfahren wird aus einem kleinen Teil der GSM-Daten sowie mit geringer räumlicher Auflösung Aufenthaltsregionen abgeleitet und GSM-Daten in nicht interessierenden Regionen herausgefiltert. Für die übrig bleiben Mobilfunkdaten sind nun weitergehende Identifizierungen erforderlich. Um die Rechenressourcen effizient zu nutzen, wird eine Priorisierung der GSM-Daten anhand einer Aktiv-Inaktiv-Liste von Mobilfunkteilnehmern und weiterer Mobilfunkdaten-Kriterien vorgenommen.

4.1.1 IMEI-(TAC-) Kennung

In der Spezifikation der GSM-Netze ist eine Gerätetypenunterscheidung anhand der Gerätetyp-Kennung IMEI vorgesehen. Jedes Endgerät besitzt somit eine im GSM-Netz eindeutige Kennziffer. Sie setzt sich im Wesentlichen aus einer Hersteller-Kennung und einer Gerätemodell-Kennung (TAC) sowie der eindeutigen Endgeräte-Kennung (SNR) zusammen. Die folgende Abbildung 4-1 zeigt den Aufbau der IMEI, wie er seit 01.04.2004 verwendet wird, im Detail.

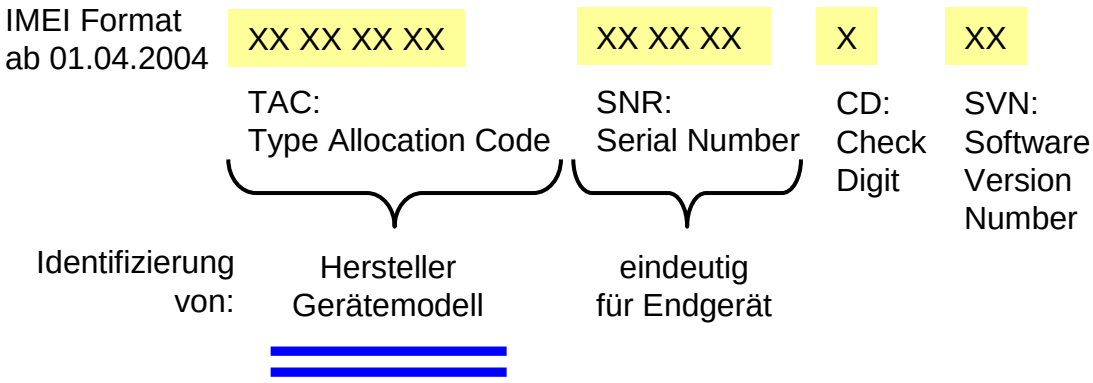


Abbildung 4-1: Datenformat der IMEI-Endgerätekennung (seit 01.04.2004)

Aus datenschutzrechtlichen Gründen ist im Projekt Do-iT die Nutzung und Speicherung der Endgeräte-Kennung SNR nicht gestattet. Betrachtet man nun die Hersteller- und Gerätemodell-Kennung TAC, so kann man bei ausreichender Anonymität eine Korrelation zwischen TAC und MIV mittels Historiendaten bestimmten. Wenn für spezielle TAC eine häufige MIV Nutzung nachgewiesen werden kann, können Mobilfunkdaten mit diesen TAC gezielt priorisiert oder selektiert werden. Ebenso können TAC die nie vom MIV genutzt werden, von der Online-Analyse ausgeschlossen werden.

Die Grundidee der TAC Nutzung für die Identifizierung basiert somit auf der Kenntnis der MIV-Aktivität spezifischer TAC. Diese Information kann bspw. aus der Offline-Analyse oder aus Expertenwissen in Form einer Datenbank vorgehalten werden. Im Online-Betrieb kann schließlich zeiteffizient aus der Datenbank die prognostizierte MIV Nutzung abgefragt und die weitere Online-Analyse damit gesteuert werden. Eine Prinzipdarstellung dieser Vorgehensweise zeigt Abbildung 4-2.

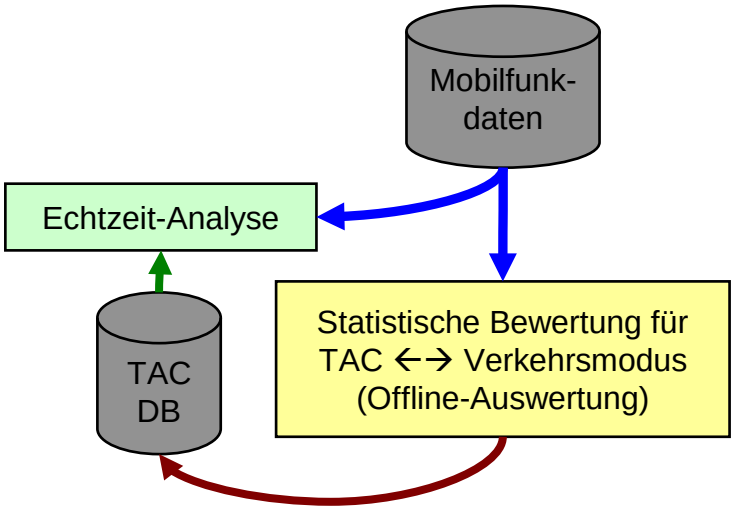


Abbildung 4-2: Prinzipdarstellung der Befüllung und Nutzung der TAC-MIV Datenbank zur Identifizierung auf GSM-Daten-Ebene

Die klassifizierten TAC der Offline-Analyse können in Form einer black - white – Liste gesammelt werden. Aus den Historiedaten kann man für jede TAC mit der Grundgesamtheit und der Anzahl der MIV Nutzungen die Sicherheitswahrscheinlichkeit mit (3-1) für die TAC-Identifizierung berechnen.

$$P_{TAC}^{MIV} = \frac{i_{TAC}^{MIV}}{n_{TAC}} \tag{3-1}$$

Unterschreitet ein FPD eine vorzugebende Mindestwahrscheinlichkeit $P_{TAC,\min}^{MIV}$, kann es von der weiteren Online-Analyse ausgeschlossen werden. Wird mindestens diese Grenze erreicht, ist zu prüfen, ob es den Kriterien eines aktiven Verkehrsteilnehmers genügt.

4.1.2 Räumliche Sektor-Maske, MS in interessierender Region ?

Die Grundidee der Sektor-Maske Analyse basiert auf der Einteilung der Kommunikationsregionen des GSM-Netzes in flächenhafte Sektoren. Ist bekannt welche Nutzungsart oder Nutzungsarten in diesen Sektoren stattfinden, kann auf das Verhalten des FPD rückgeschlossen werden. Liegen geeignete Geodaten vor, kann für die Sektoren des GSM-Netzes die Nutzungsart mittels mapmatching Verfahren ermittelt und vorprozessiert in einer Datenbank abgelegt werden. Eine Identifizierung ist somit auf GSM-Daten-Ebene möglich. Ebenfalls ist eine räumliche Filterung möglich, so dass Mobilfunkdaten interessierender Regionen auf GSM-Datenebene effizient ausgewählt werden können.

Die Mobilfunk-Sektoren sind durch den Öffnungswinkel der BTS-Antenne definiert. Ein Sektorring wird durch die zusätzliche Nutzung des TA-Wertes, also der Distanz zwischen BTS und MS definiert. Der Öffnungswinkel ist je Antennenbauart verschieden und kann bis zu 360° für Rundstrahler betragen. Der TA-Wert ist in ganzzahligen Schritten von 0 bis 63 in der GSM-Spezifikation (3gpp GSM 05.10) festgelegt. Mit der Funkwellenausbreitungsgeschwindigkeit *c* und der Dauer eines maximal ausgesendeten Kommunikationspaketes (Burst) von $3{,}69 \cdot 10^{-6}\,\mathrm{sec}$ im GSM [Walke] ergibt sich eine räumliche Distanz von 554 m je TA-Schritt. Da die Auflösung volle TA-Wert Schritten (n*554 m) beträgt, ist mindestens von einem Bereich von je einer halben TA-Schrittweite weniger und einer halben TA-Schrittweite zusätzlich zum TA-Wert radial entfernt vom Antennenstandort auszugehen. Die Abbildung 4-3 veranschaulicht den sich ergebenen Sektorring. In diesem Bericht wird im Weiteren der Begriff Sektor aus Vereinfachungsgründen als synonym zu Sektorring betrachtet. Verwechselungen sind nicht zu befürchten. Auch der Begriff der Sektor-Maske ist im Sinne einer Sektorring-Maske zu interpretieren.

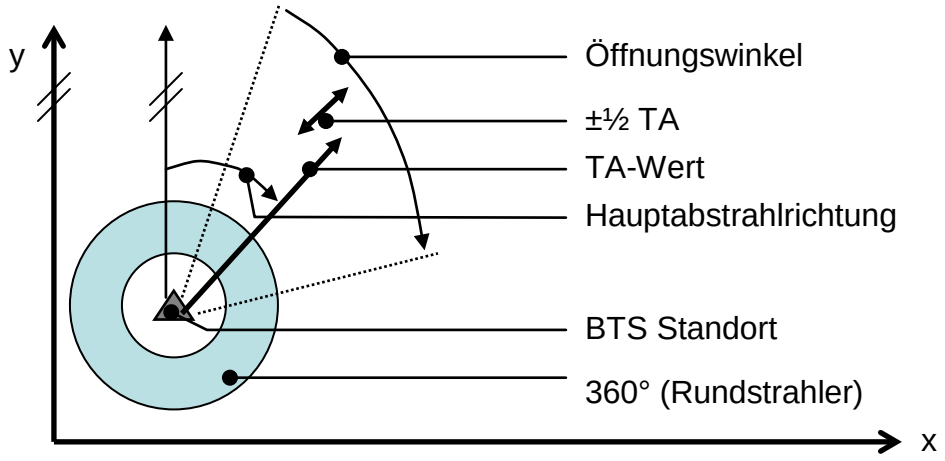


Abbildung 4-3: Sektordefinition mittels Abstrahlrichtung und TA-Wert an BTS

Befindet sich ein Sektor vollständig innerhalb einer verkehrlich nicht interessanten Region, so können FPD dieses Sektors herausgefiltert und als verkehrlich nicht interessant gekennzeichnet werden. Dieses Vorgehen entspricht einem „black-list“ -Verfahren, wobei die „black-list“ aus verkehrlich nicht relevanten Sektoren besteht, siehe Abbildung 4-4. Ebenso kann mit Hilfe von verkehrlich relevanten Regionen, wie z.B. Straßen, eine „white-list“ aufgestellt werden.

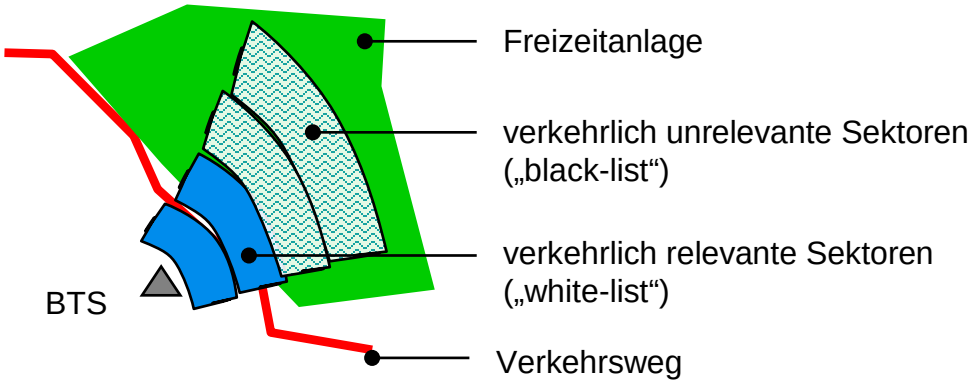


Abbildung 4-4: Für Verkehrslageerfassung interessante (blau) und nicht interessante (grün) GSM-Sektoren

Im Projekt Do-iT ist die Identifizierung verkehrlich relevanter FPD über das Verfahren der räumlichen Sektor-Maske mittels der „white-list“ implementiert. Hiermit werden sich ausschließlich in den interessierenden Regionen und auch nur zu Teilen in diesen Regionen aufhaltende FPD aus den Mobilfunkdaten selektiert.

4.1.2.1 Genauigkeit der Planungsdaten

Mit dem Standort der BTS xy_{BTS}, der Hauptabstrahlrichtung der Antenne ϕ_{BTS} und dem Öffnungswinkel $\Delta\phi_{BTS}$ aus den Planungsdaten des Mobilfunknetzes

(Daten der TMD) lassen sich die Sektorgrenzen „scharf“ berechnen. Aufgrund von Abweichungen bei der Montage, der Antennenkonstruktion, etc. gegenüber den Planungsdaten sind Fehlereinflüsse mit einzubeziehen. Im Projekt werden die folgenden Fehlereinflüsse eingeführt:

- Öffnungswinkel: 360° Strahler ±0°, sonst ±25°
- Hauptabstrahlrichtung: ±25°
- BTS-Standort: als fehlerfrei angenommen
- TA-Wert: ±½ TA-Schrittweite

Die aufgeführten Maße der Abweichungen sind Näherungen. Die Abweichungen der Planungsdaten werden als zusätzlicher „unscharfer“ Bereich außerhalb des „scharfen Sektors“ berücksichtigt und gemeinsam als „erweiterter Sektor“, siehe Abbildung 4-5, bezeichnet. Eine für jeden Sektor gültige Verteilungsfunktion ist nicht bekannt und die empirische Ermittlung muss als nicht praktikabel verworfen werden. Allein für das abis-Testgebiet Karlsruhe wären die Verteileilungen von bis zu 12 000 Sektoren zu bestimmen. In Do-iT werden daher die erweiterten Sektoren wie folgt definiert: Ist zum Zeitpunkt t ein FPD im Sektor x, dann ist die Aufenthaltswahrscheinlichkeit:

- innerhalb des „scharfen Sektors“ x P=100%
- innerhalb des „unscharfen Bereiches“ des „unscharfen Sektors“ x P=$\left(\frac{a}{b}\cdot 100\right)$%, mit a und b siehe Abbildung 4-5
- außerhalb des „erweiterten Sektors“ x P=0%

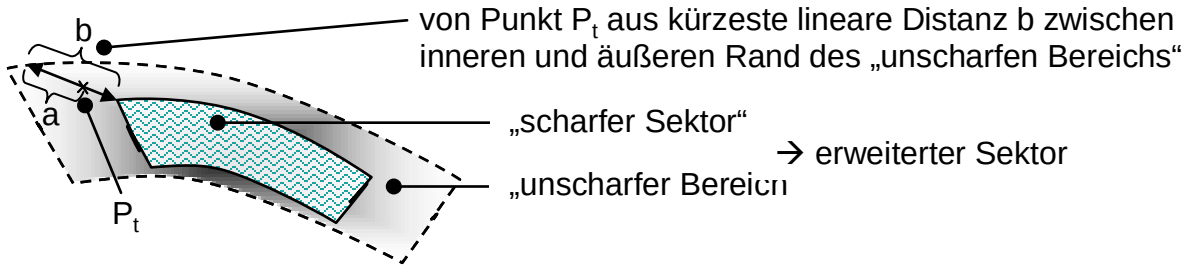


Abbildung 4-5: Sektor mit scharfen und unscharfen Bereichen

Die erweiterten Sektoren können nun mit nutzerspezifischen Regionen überlappt werden. Für erweiterte Sektoren, die vollständig innerhalb einer nutzerspezifischen Region sind, kann eine eindeutige Zuordnung erfolgen. In Abbildung 4-6 sind diese erweiterten Sektoren schraffiert und als „verkehrlich nicht relevante“ Sektoren gekennzeichnet. Für erweiterte Sektoren, die nicht vollständig innerhalb von nutzerspezifischen Regionen liegen, muss eine mögliche

verkehrliche Nutzung geprüft werden. In Abbildung 4-6 sind diese erweiterten Sektoren gepunktet dargestellt.

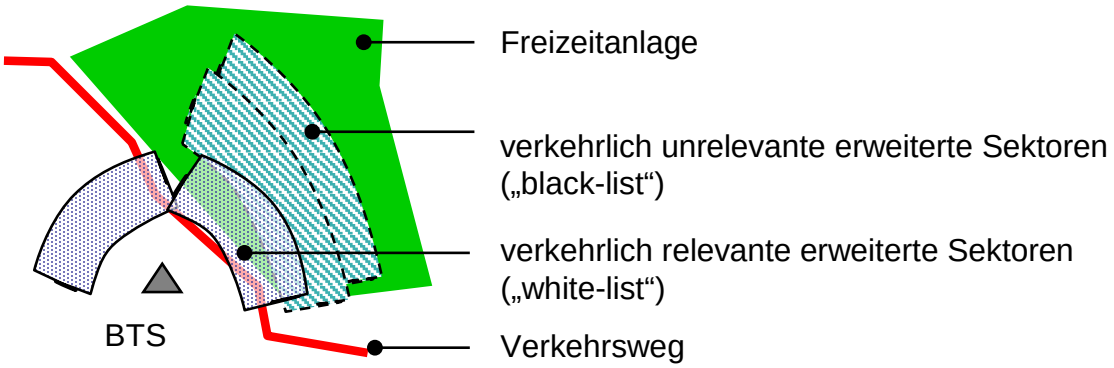


Abbildung 4-6: GSM-Daten basierte Sektormaske zur Filterung von GSM-Daten in unrelevante Regionen

Wie noch in Abschnitt 4.3 anhand von georeferenzierten Daten näher gezeigt wird, können die erweiterten Sektoren in ein Kachelkoordinatensystem transformiert werden. Neben einer Überprüfung, ob sich in den betroffenen Kacheln verkehrlich relevante bzw. nicht relevante Objekte befinden, sind ebenfalls die Wahrscheinlichkeiten für die verkehrliche Nutzung berechenbar. Grundlage für die Wahrscheinlichkeitsberechnung ist die oben definierte Verteilungsfunktion.

Die Transformation der erweiterten Sektoren in das Kachelkoordinatensystem sowie Bestimmung der Wahrscheinlichkeiten für die verkehrliche Nutzung können vorprozessiert und in einer lookup-Tabelle bereitgestellt werden. Ergebnis sind Werte für die maximale Wahrscheinlichkeit der verkehrlichen Nutzung für die erweiterten Sektoren. Die prinzipielle Datenprozessierung hierzu zeigt Abbildung 4-7.

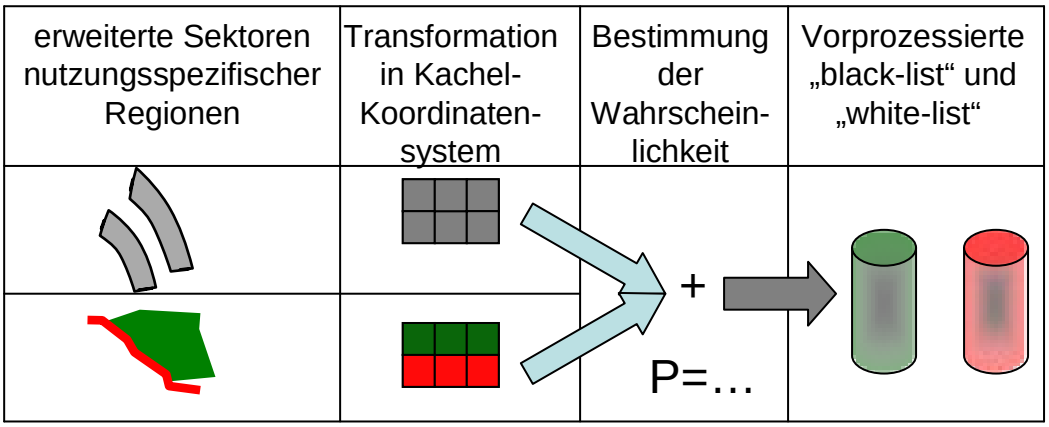


Abbildung 4-7: Datenprozessierung zur Bestimmung von verkehrlich relevanten („white-list“) und nicht relevanten („black-list“) erweiterte Sektoren

Die lookup-Tabellen wurden zur effizienten Datenabfrage in Do-iT mit der Cell-ID und TA-Wert indiziert erstellt. Basis sind das Lokalisierungsnetz und die Planungsdaten des Mobilfunknetzes (BUF). Abbildung 4-8 zeigt einen Auszug einer vorprozessierten lookup-Tabelle. Für das Abis-Testgebiet im Raum Karlsruhe mit 200 BTS, die im T-Mobile GSM-Netz 200 Funkzellen entsprechen, und den TA-Werten von 0 bis 63, ergeben sich maximal 12.800 Einträge. In der „white-list“ würde diese Anzahl jedoch nur erreicht werden, wenn in jedem Sektor eine verkehrliche Nutzung möglich wäre.

Cell-ID	TA-Wert	Wahrscheinlichkeit „white list“
24562	0	0,23
24562	1	1
…	…	…
28673	0	0
28673	1	0
…	…	

Abbildung 4-8: Ausschnitt einer vorprozessierten „white list“ mit Wahrscheinlichkeiten je erweitertem Sektor

4.1.2.2 Geodaten der räumlichen Sektor-Maske

Zur Prozessierung der „black list“ oder „white list“ sind, wie bereits oben genannt, die verkehrlich relevanten bzw. nicht relevanten Objekte im Kachelkoordinatensystem erforderlich. Die Datenprozessierung erfolgt für beide Listen nach demselben Prinzip aus Abschnitt 4.1.2.1. Die berechneten Wahrscheinlichkeitswert der „black list“ und der „white list“ müssen in der Summe jedoch nicht 100% ergeben. In einer Kachel können sich verkehrlich relevante und nicht relevante Objekte gleichrangig befinden. Daraus folgt, dass in einem erweiterten Sektor die „black list“ und „white list“ Wahrscheinlichkeiten gleichrangig nebeneinander, siehe Abbildung 4-6, stehen. Folglich ist es mit dem hier gezeigten Verfahren nur möglich Mobilfunkdaten zu selektieren, deren erweiterte Sektoren unter anderem verkehrliche Objekte enthalten. Verwendet ein FPD jedoch vorrangig erweiterte Sektoren der „black list“ und selten welche der „white list“, kann dies zwar die räumlichen Nähe zu verkehrlich relevanten Objekten aufzeigen, aber zur Identifizierung eines passiven Verhaltens des FPD führen.

Zur Bestimmung der „black list“ und „white list“ können unter anderem die folgenden Objekte der Tabelle 2 verwendet werden. In Do-iT ist die Grundlage der Verkehrsnetze das Lokalisierungsnetz. In Abbildung 4-9 sind entlang der Auto-

bahnen die verkehrlich relevanten erweiterten Sektoren dargestellt. Das dargestellte Gebiet umfasst das Demonstratorgebiet der A-Daten des Projektes Do-iT.

Tabelle 2 Objekte der "black list" und der "white list" zur Bildung der Sektor-Maske

„black list“	*„white liste“*
Gebäude	Autobahnen
Parkplätze	Bundesstraßen
Schienen	Landes-, Kreis- und Kommunalwege
Parkanlagen, Erholungsgebiete	
Stadien, Freizeitanlagen	

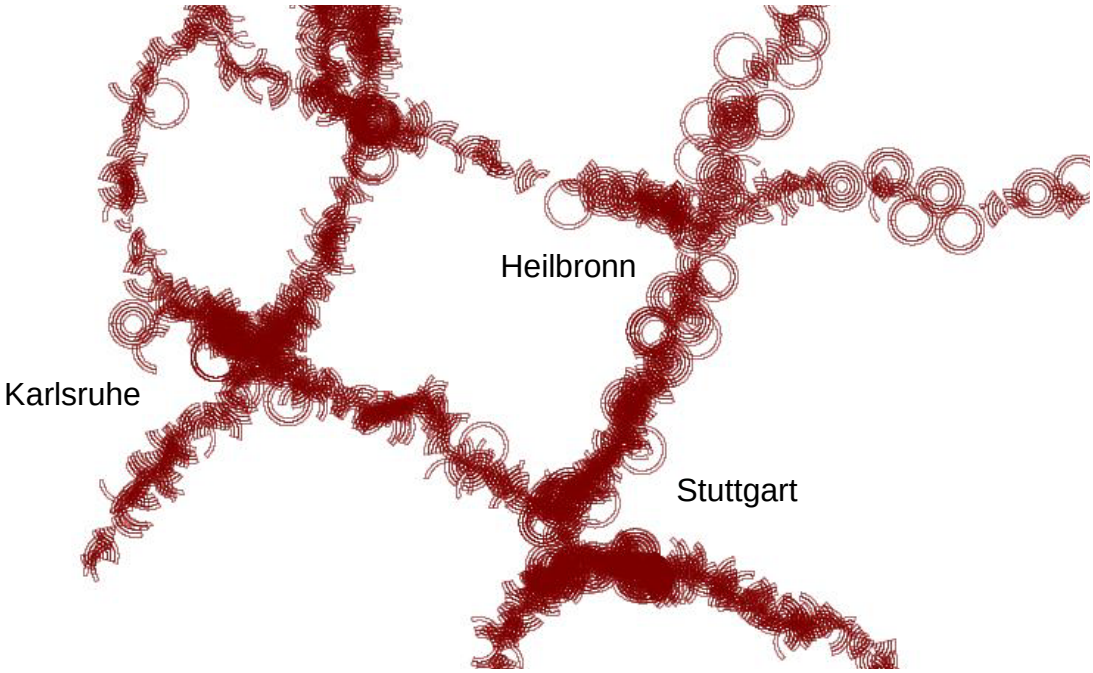


Abbildung 4-9: Verkehrsrelevante Sektoren (für TA-Werte von 0 bis 6) für die Auswahl Autobahnen des Demonstratorgebietes des Projektes Do-iT

4.1.3 Priorisierung von Mobilfunkdaten für die Online-Auswertung

Wie bereits eingangs erläutert, ist die Menge der im Online-Betrieb anfallenden Mobilfunkdaten erheblich, so dass im Sinne einer effizienten Verkehrslageerfassung die Rechenressourcen möglichst auf die relevanten Mobilfunkdaten zu richten sind. Hierfür ist eine Priorisierung der Mobilfunkdaten erforderlich, die z.B. aus Erfahrungen zurückliegender Analysen (Historie) oder einer Priorisie-

rung aus den Mobilfunkdaten selbst zu berechnen ist. Für das Projekt Do-iT sollen externe Eingriffe in den Auswerteprozess nicht stattfinden.

Die Einordnung der Priorisierung in den Ablauf der Online-Auswertung zeigt die folgende Abbildung 4-10.

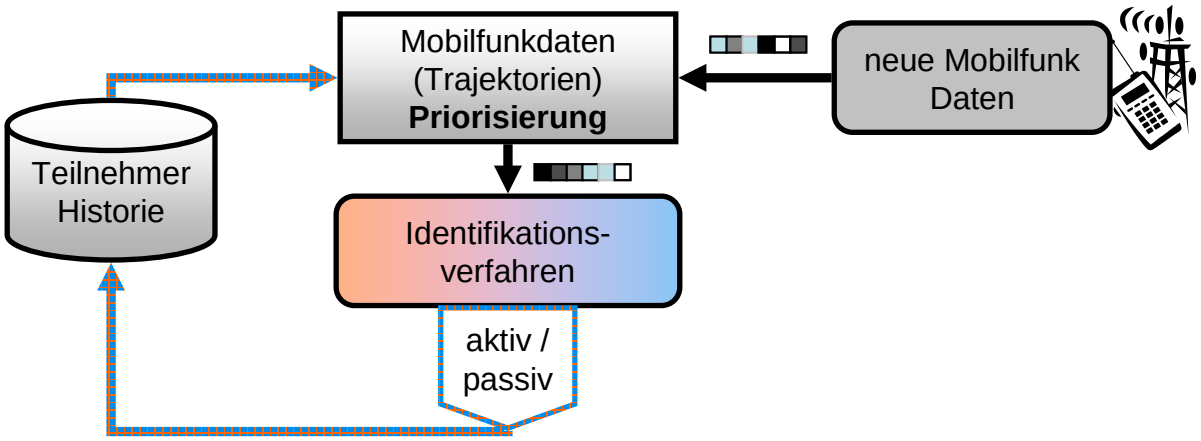


Abbildung 4-10: Einordnung der Mobilfunkdaten-Priorisierung in den Ablauf der Online-Auswertung

Eine Historieninformation ist bspw., ob ein Teilnehmer in einer der letzten Berechnungen bereits als „aktiv“ eingestuft worden ist. Kommen weitere Mobilfunkdaten dieses Teilnehmers, könnten diese eher von Interesse sein als von einem bislang „passiven“ Teilnehmer. Für eine Rangliste aus den Mobilfunkdaten selbst, wurden bereits in den letzten Abschnitten Funktionen erläutert. Diese und weitere Kriterien für die Priorisierung sind folgend aufgelistet.

1. H_{ap}: Grad der IV-Aktivität (**H**istorie): Teilnehmer war **a**ktiv = 1, **p**assiv = 0
2. Grad der Verkehrsrelevanz:
   - TAC_{Vr}: **TAC** Identifizierung für **V**erkehrs**r**elevanz (0..1)
   - P_{Sektor}: arithmetisches Mittel der Wahrscheinlichkeit (**P**) der Verkehrsrelevanz bestimmt mittels der erweiterten **Sektor**en (0..1)
3. Grad der Bewegung:
   - Anzahl unterschiedlicher Sektoren (0,1,2 ... $n_{Sektoren}$)
   - Anzahl Messelemente (0,1,2 … $n_{Messelemente}$)

Zur effizienten Abarbeitung der Rangliste wird aus den Kriterien eine Priorisierungszahl (PZ) mit der Formel (3-2) abgeleitet. Je höher diese Zahl ist, desto höher ist die Priorität des Teilnehmers für die Auswertung.

$$PZ = \mathrm{int}\left(\begin{array}{lll} & H_{ap} & \cdot 100\,000 \\ + & TAC_{Vr} & \cdot 1000\,000 \\ + & P_{Sektor} & \cdot 10\,000 \\ + & \frac{n_{Sektoren}}{n^{max}_{Sektoren}} & \cdot 1000 \\ + & \frac{n_{Messelemente}}{n^{max}_{Messelemente}} & \cdot 100 \end{array} \right) \qquad (3\text{-}2)$$

mit : $n^{max}_{Sektoren}$ bislang maximale Anzahl Sektoren je Trajektorie

$n^{max}_{Messelemente}$ bislang maximale Anzahl Messwerte je Trajektorie

In den folgenden Abschnitten werden der Ablauf der Identifizierung und die darin enthaltenen Verfahrensschritten näher erläutert. Diese Verfahren arbeiten vorzugsweise mit den georeferenzierten Mobilfunkdaten. Der Prozessierungsschritt der Mobilfunkortung wird im Projekt Do-iT im Arbeitspaket 4 bearbeitet. Weiterführende Informationen zu diesem Thema sind in den dort angefertigten Dokumenten nachzulesen.

4.2 Ortungsinformationen berechnet aus GSM-Mobilfunkdaten

Die Abbildung 4-11 zeigt das Datenmodell der Ortungsinformationen (siehe Do-iT AP4). Zusätzlich sind statistische Maße für die Genauigkeiten der Koordinaten mit ausgewählten fehlerzeigenden Figuren einschließlich ihrer Parameter enthalten. In Do-iT wurde als Bezugssystem das WGS84 gewählt, wobei die Koordinaten mittels Gauß-Krüger-Abbildung mit dem 9. östlichen Längengrad als Hauptmeridian in ein kartesisches Koordinatensystem überführt wurden. In Do-iT wird eine derartige Position eines Teilnehmers als „**FPD-Position**“ bezeichnet.

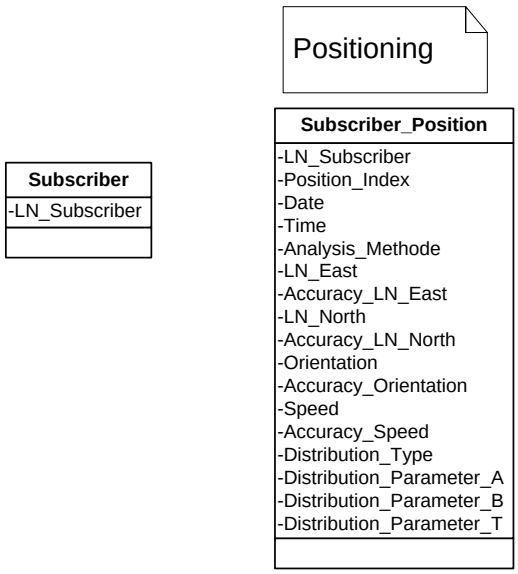


Abbildung 4-11: Datenmodell der FPD-Position

In der folgenden Tabelle 3 sind die Erläuterungen zu den übergebenen Werten zusammengefasst. Neben der Angabe einer eindeutigen, aber anonymen Bezeichnung des Mobilfunkteilnehmers, des Datum und der Uhrzeit sowie der aufsteigenden Nummerierung der Positionen, sind auch Qualitätsparameter mit der Bezeichnung „Accuracy“ für die Genauigkeit und „Distribution“ für die statistische Verteilungsfunktion zu finden. Im Feld „Analysis_Method“ werden die bei der Ortung eingesetzten funktionalen Modelle sowie verwendete Filterfunktionen vermerkt.

Tabelle 3: Erläuterung des Datenmodells der FPD-Position

Kurzname	Type	Beschreibung
Tabelle: Subscriber_Position		
Element	Datentyp	Beschreibung
LN_Subscriber	Integer	Bezeichnung des Mobilfunkteilnehmers
Position_Index	Integer	Innerhalb eines Teilnehmers eindeutiger, aufsteigender Index
Date	Integer	Datum des Eintrages (YYYYMMDD, Bsp. 20060331)
Time	Double	Sekunden ab Mitternacht (Bsp.: 14h:51m:15,123s=53475,123)
Analysis_Method	Integer	-1 (default) keine Methode 1 Bogenschnitt 2 Signalmatching 3 Ringsektorortung (Schwerpunkt) 4 „Pseudo-Bogenschnitt“ (gewichtetes Mittel 2er BTS) 5 „Pseudo-Bogenschnitt“ (gewichtetes Mittel mehrerer BTS) x1 Ausgleichung x2 geglättete Ausgleichung x3 Kalman-Filter (mit x für die Ziffer des verwendeten funktionalen Modells)
LN_East	Double	Kartesische Koordinate East in [m], Gauß-Krüger Abb. auf WGS84-Ellipsoid
Accuracy_LN_East	Double	Standardabweichung der Koordinate East in [m] (-1=keine Standardabweichung berechenbar)
LN_North	Double	Kartesische Koordinate North in [m], Gauß-Krüger Abb. auf WGS84-Ellipsoid
Accuracy_LN_North	Double	Standardabweichung der Koordinate North in [m] (-1=keine Standardabweichung berechenbar)
Orientation	Double	Orientierung des Teilnehmers (Bewegungsrichtung) in Grad [°], von Nord (=0°) rechtsläufig, (-1=keine Orientierung berechenbar)
Accuracy_Orientation	Double	Standardabweichung der Orientation in [°], (-1=keine Orientierung berechenbar)
Speed	Double	Geschwindigkeit des Teilnehmers in [m/s], (-1=keine Geschwindigkeit berechenbar)
Accuracy_Speed	Double	Standardabweichung der Geschwindigkeit Speed in [m/s], (-1=keine Geschwindigkeit berechenbar)
Distribution_Type	Integer	Verwendete Verteilungsfunktion der statistischen Parameter -1 (default) keine 1 Normalverteilung (Form Ellipse) 2 Gleichverteilung (Form Rechteck)
Distribution_Parameter_A	Double	Maß der großen Halbachse [m] (default=-1)
Distribution_Parameter_B	Double	Maß der kleinen Halbachse [m] (default=-1)
Distribution_Parameter_T	Double	Maß der rechtsläufigen Orientierung der großen Halbachse [°]

4.3 Diskretisierung der Geodaten

Der Einsatz von raumbezogenen Daten unterstützt einige der hier vorgestellten Identifizierungsverfahren oder macht sie zum Teil erst möglich. Meist werden jedoch nur kleine, räumlich begrenzte Ausschnitte benötigt. Die hohen Datenmengen der digitalen Karten verursachen allerdings bei der Suche und Selektion von benachbarten Kartenobjekten ineffiziente Rechenoperationen.

Für den effizienten Datenzugriff wurden daher die Geodaten räumlich gekachelt. Ein Beispiel für die entstehenden Kacheln und die dadurch repräsentierten Geodaten zeigt Abbildung 4-12. Ist eine Kachel Menge oder Teilmenge eines Geodatenobjektes, so repräsentiert die Kachel das Kartenobjekt. Für eine

räumliche Suche in der Nähe der Kachel würden nun nur die Kartenobjekte untersucht werden, die in der gleichen Kachel oder in einer Kachel der unmittelbaren Nachbarschaft sind. Die Anzahl der zu untersuchenden Kartenobjekte wird dadurch wesentlich verringert. Die Effizienz bei der räumlichen Suche und der Kartenobjektselektion steigt dementsprechend.

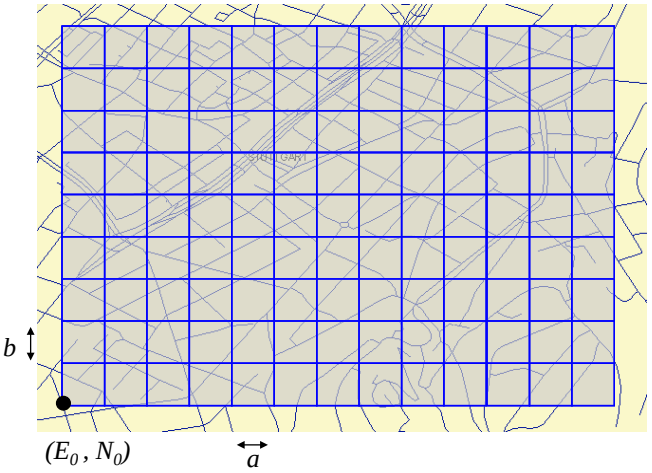


Abbildung 4-12: Beispiel einer Verkachelung von Geodaten

4.3.1 Definition des Kachelkoordinatensystems

Mit der Definition der kartesischen Ursprungskoordinaten für die Ostrichtung (East) E_0 und die Nordrichtung (North) N_0, sowie der Seitenlängen a und b für eine rechteckige Kachel, kann für eine beliebige geografische Koordinaten (E_j, N_j) die Kachelkoordinate (n_j^E, n_j^N) mit den Formeln (3-3) berechnet werden.

$$n_j^E = \text{int}\left(\frac{E_j - E_0}{a} + 1\right) \quad \text{und} \quad n_j^N = \text{int}\left(\frac{N_j - N_0}{b} + 1\right) \tag{3-3}$$

Zur Vereinfachung der Datenverwaltung können weiterführend die Kachelkoordinaten (n_j^E, n_j^N) zu einer impliziten Koordinatenzahl EN (**E**ast**N**orth) zusammengesetzt werden. Für das Lokalisierungsnetz im Projekt Do-iT errechnet sich die Koordinatenzahl EN mit Formel (3-4).

$$EN = E \cdot 100\,000 + N \tag{3-4}$$

Die zuvor zweidimensional mit *E* und *N* gespeicherten Kacheln sind mittels der Koordinatenzahl EN eindimensional verwaltbar. Die Koordinatenzahlen werden anschließend sortiert und können nun mittels Suchalgorithmen wie Quick-Search effizient gefunden sowie die angefragten Kartenobjekte selektiert werden. Die Abbildung 4-13 zeigt einen Ausschnitt der Kacheldarstellung für die Verkehrsnetze des Lokalisierungsnetzes.

LN_Kachel_EN	LN_Kachel_E	LN_Kachel_N	LN_FromNode_ID	LN_Edge_ID	LN_ToNode_ID	LN_Length	LN_Form_Of_Way
7900968	79	968	722962	856158	722961	644.8	2
7900972	79	972	722970	856167	722968	373.731	30
7900972	79	972	722968	856168	722970	373.731	30
7900973	79	973	722970	856167	722968	373.731	30
7900973	79	973	722968	856168	722970	373.731	30
7900974	79	974	722968	856165	722969	35.858	30
7900974	79	974	722969	856166	722968	35.858	30
7900995	79	995	729865	864800	729866	807.209	30
7900995	79	995	729866	864801	729865	807.209	30
7901010	79	1010	3632	6832	3633	1027.674	32

Abbildung 4-13: Werte der Kacheldarstellung des Lokalisierungsnetzes

4.3.2 Auswahl der Kachelgröße

Grundsätzlich gilt, dass je kleiner die Kachelfläche ist, desto mehr Kacheln sind für vollständige Abdeckung erforderlich. In Do-iT werden jedoch nur die Kacheln mit einer Referenz auf ein Geoobjekt gespeichert. Daher erhöht sich die Anzahl nicht quadratisch mit Halbierung der Seitenlängen, wie es bspw. beim Quadtree Verfahren der Fall ist. Vielmehr fließt die Kartenobjektdichte als Funktion mit ein. In Abbildung 4-14 ist die Anzahl der Kacheln in Abhängigkeit zu verschiedenen Seitenlängen der Kacheln für das Lokalisierungsnetz dargestellt. Die Seitenlängen a und b wurden jeweils gleichlang gewählt.

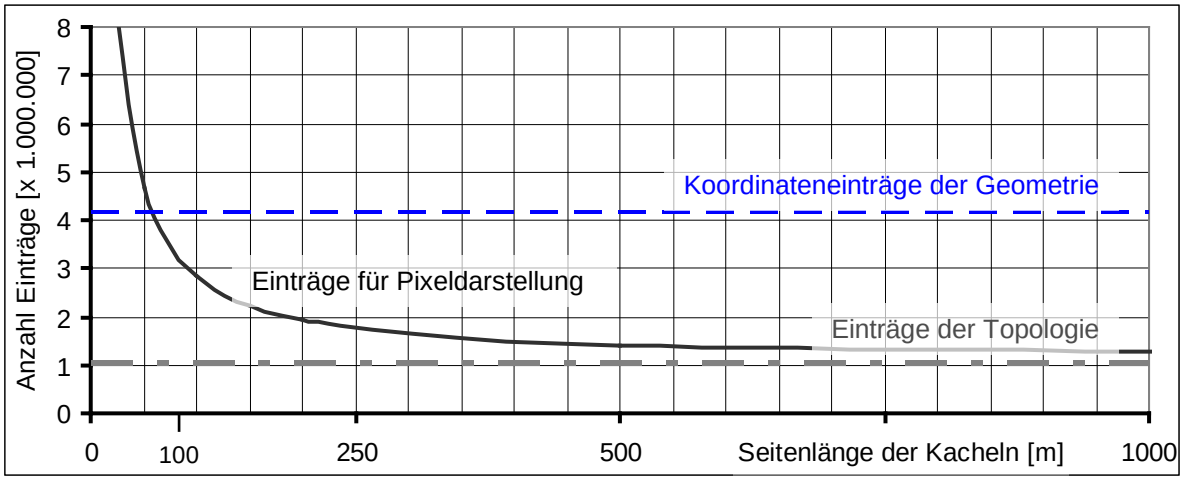


Abbildung 4-14: Abhängigkeit der Anzahl der Tabelleneinträge von der Seitenlänge der Kacheln (Lokalisierungsnetz Version 017 Projekt Do-iT)

Das Datenmodell des Lokalisierungsnetzes sieht die separate Speicherung der Kantengeometrien je Fahrtbeziehung vor. Im Gegensatz zum GDF ist die bauliche Trennung nicht von Bedeutung. Die minimale Anzahl möglicher Kacheleinträge ist somit begrenzt durch die Anzahl der Fahrtbeziehungen (Topologie). In Abbildung 4-14 ist die Anzahl der Topologieeinträge gestrichelt dargestellt. Ab

einer Kachelgröße von weniger als 200 m je Seitenlänge (4000 m^2) steigt die Anzahl der Kacheleinträge deutlich an und übertrifft ab ca. 50 m (2500 m^2) die Anzahl der ursprünglich benötigten Koordinateneinträge für die Geometriedarstellung. Als praxistauglichen Kompromiss zwischen der Anzahl der Kartenobjekte je Kachel und der Anzahl der Kacheleinträge wird für das Lokalisierungsnetz eine Kachelgröße von 2500 m^2 (50 m) bis 4000 m^2 (200 m) gesehen. Je nach Auflösung kann ca. ein Drittel bis die Hälfte an Einträgen eingespart werden, wodurch die verringerte Datenmenge die Prozesseffizienz steigern kann.

Der größere Effizienzvorteil liegt nunmehr darin, dass mittels der Koordinatenzahl EN ohne räumliche Suchoperationen Kartenobjekte direkt selektiert werden können. Die Koordinatenzahl EN wiederum kann durch die vergleichweise einfachen Rechenoperationen mit Formel (3-3) und (3-4) ermittelt werden.

4.4 Aufenthaltswahrscheinlichkeit eines FPD in einer Kachel

Die qualitative Beschreibung der FPD-Positionen erfolgt, wie oben bereits genannt, durch die Angabe von fehlerzeigenden Figuren (Ellipse, Rechteck), der Verteilungsfunktion (Normalverteilung, Gleichverteilung) sowie Maßen für die Ausdehnung der Figuren. Innerhalb einer Kachel wird überall die gleiche Aufenthaltswahrscheinlichkeit angenommen. Abbildung 4-15 zeigt diesen Vorgang am Beispiel einer FPD-Position mit Normalverteilung und der fehlerzeigenden Figur einer Ellipse. Je dunkler die Farbe desto höher ist die Aufenthaltswahrscheinlichkeit. Transformiert man die FPD-Positionen in die Kacheldarstellung, ist bei Überschreitung der Kachelgrenzen die **Aufenthaltswahrscheinlichkeit** des FPD in der benachbarten Kachel zu berechnen.

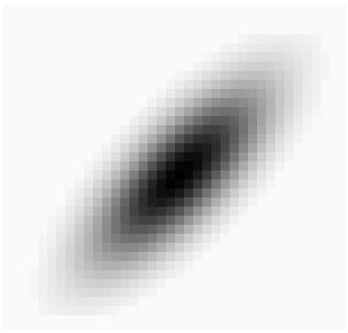


Abbildung 4-15: Aufenthaltswahrscheinlichkeiten einer FPD-Position für die fehlerzeigende Figur einer Ellipse und der Normalverteilung

Für die Ermittlung der Aufenthaltswahrscheinlichkeiten der Kacheln der FPD-Position dienen die folgenden Figuren sowie Verteilungen aus AP4:

Figur 1 Fehlerellipse mit Normalverteilung, Angabe der großen und kleinen Halbachse sowie der Orientierung der großen Halbachse bei einer Sicherheitswahrscheinlichkeit von 95%

Figur 2 Rechteck mit Gleichverteilung, Angabe der Ausdehnung für große und kleine Seitenlänge sowie der Orientierung der großen Seitenlänge

Für die Figur 1 (Fehlerellipse mit Normalverteilung) werden mittels der Schnittgeradenmethode [Rauber93] die Kacheln des Randbereiches berechnet. Einfließende Parameter sind die großen und kleinen Halbachse sowie die Orientierung der Fehlerellipse. Die eingeschlossen Kacheln ergänzen die Kachelmenge. Für den Wert der Aufenthaltswahrscheinlichkeit der gesamten Kachel wird der Kachelschwerpunkt angenommen.

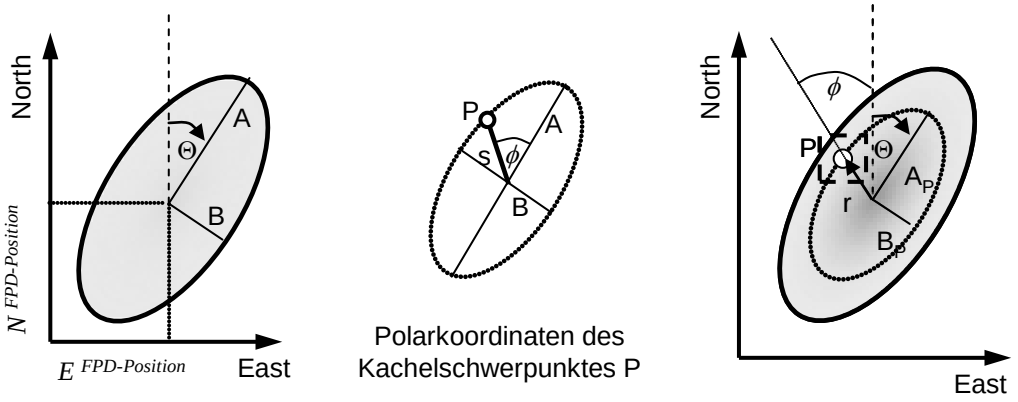


Abbildung 4-16: Ermittlung der Aufenthaltswahrscheinlichkeit für beliebige Punkte einer Fehlerellipse
a) Fehlerellipse der FPD-Position
b) Polarkoordinaten eines Punktes P ausgehend von der FPD-Position
c) Ellipse gleicher Wahrscheinlichkeiten wie für den Punkt P

In der Abbildung 4-16 ist der Ablauf der Bestimmung für die Aufenthaltswahrscheinlichkeit in einer Kachel einer FPD-Positionfehlerellipse skizziert. Bild a) zeigt die Daten der FPD-Position aus AP4. Für die umgebenen Kacheln der FPD-Position werden nun nacheinander die Aufenthaltswahrscheinlichkeiten der Kachelschwerpunkte berechnet. Dazu werden die Parameter einer durch den Kachelschwerpunkt gehenden Ellipse mittels der Polarkoordinaten ausgehend von den Koordinaten der FPD-Position ermittelt. Das Bild b) der Abbildung 4-16 zeigt diesen geometrischen Zusammenhang. Anschließend wird basierend auf der Funktion der Normalverteilung die Aufenthaltswahrscheinlichkeit auf dieser Ellipse abgeleitet, siehe Bild c). Der Wert Aufenthaltswahrscheinlichkeit der betreffenden Kachel wird folgend mit $P^{FPD-Kachel}$ bezeichnet

Für die Figur 2 (Rechteck mit Gleichverteilung) werden wie bei der Figur 1 die Randkacheln berechnet und mit den eingeschlossenen Kacheln des angenommenen Aufenthalts ergänzt, siehe Abbildung 4-17. Mit Annahme der Gleichverteilung erhalten alle Kacheln die gleiche Aufenthaltswahrscheinlichkeit. Mit der Anzahl der Kacheln n wird $P^{FPD-Kachel}$ mit Formel (3-5) ermittelt.

$$P^{FPD-Kachel} = \frac{100\%}{n} \tag{3-5}$$

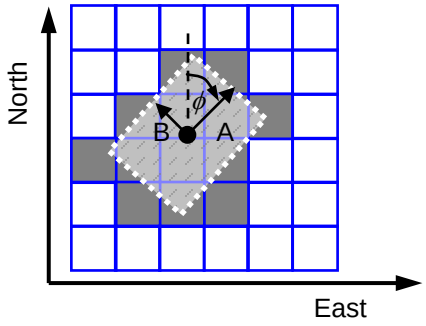


Abbildung 4-17: Aufenthaltswahrscheinlichkeiten einer FPD-Position für die fehlerzeigende Figur eines Rechtecks und der Gleichverteilung

4.5 Bewegungsanalyse

Eine Grundvoraussetzung für die Nutzung von Mobilfunkdaten zur Verkehrslageerfassung ist die Unterscheidungsmöglichkeit von stehenden und sich bewegenden FPD. In den folgenden Abschnitten werden die Möglichkeiten und Einschränkungen aufgrund der Genauigkeiten abgeschätzt und Verfahren zur Trennung von sich bewegenden und stehenden FPD vorgeschlagen. Hierfür werden die kartesischen Koordinaten verwendet, die in AP4 des Projektes Do-iT aus den Mobilfunkdaten prozessiert wurden.

4.5.1 Bewegungsanalyse mittels transformierter GSM-Daten

4.5.1.1 Grenzdistanz

Unter der Annahme, dass sich ein MFT im xy-Raum (geometrisch) bewegt, ermittelt sich die Distanz durch Addition der zurückgelegten Wege zwischen den FPD-Positionen. Ist die Distanz (vgl. Formel 4.2-1) größer null, ist die Annahme einer Bewegung begründet. Aufgrund der bereits oben beschriebenen FPD-Positionsgenauigkeiten ist fehlertheoretisch zu prüfen, ob ein signifikanter Distanzwert überschritten wurde. Dazu benötigt werden:

- die Standardabweichungen der zurückgelegten Wege $s_{d,\,i}$, mit Formel (3-7), abgeleitet aus den FPD-Positionen mit Formel (3-6)
- die Genauigkeit der Gesamtdistanz $s_{d,\,gesamt}$, mit Formel (3-8)
- der sich ergebene Grenzwert $G_{d,\,P}$ für eine angenommene Sicherheitswahrscheinlichkeit P.

$$d_i = \sqrt{(x_i - x_{i-1}) + (y_i - y_{i-1})} \qquad d_{gesamt} = \sum_{i=1}^{n} d_i \tag{3-6}$$

$$s_{d,i} = \sqrt{\left(\frac{\partial d_i}{\partial x_{i-1}}\right)^2 \cdot \left(s_{x_{i-1}}\right)^2 + \left(\frac{\partial d_i}{\partial x_i}\right)^2 \cdot \left(s_{x_i}\right)^2 + \left(\frac{\partial d_i}{\partial y_{i-1}}\right)^2 \cdot \left(s_{y_{i-1}}\right)^2 + \left(\frac{\partial d_i}{\partial y_i}\right)^2 \cdot \left(s_{y_i}\right)^2} \tag{3-7}$$

$$s_{d,gesamt} = \sqrt{\sum_{i=1}^{n} s_{d,i}^2} \tag{3-8}$$

Nun kann die erforderliche Distanzgrenze $d_{gesamt}^{signifikant}$ mit Formel (3-9) berechnet werden, ab der eine Bewegung bei der vorliegenden Standardabweichung $s_{dgesamt}$ mit der Sicherheitswahrscheinlichkeit P signifikant ist. Unter der Annahme, dass die FPD-Positionen normalverteilt sind, gelten die Quantile der Normalverteilung y_P. Beispielwerte können der Tabelle 4 entnommen werden.

$$d_{gesamt}^{signifikant} = s_{d,gesamt} \cdot y_P \tag{3-9}$$

In Umkehrung kann mit Formel (3-10) bei vorliegender Gesamtdistanz und deren Standardabweichung die Sicherheitswahrscheinlichkeit P für die Annahme berechnet werden, dass sich der MFT bewegt.

$$P = \int_{-\infty}^{d_{gesamt}} \frac{1}{s_{d,gesamt} \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{x^2}{2 \cdot s_{d,gesamt}^2}} dx \tag{3-10}$$

Abschätzung der zu erwartenden Werte:

Mit der Annahme, dass die Genauigkeit der jeweiligen FPD-Positionen sowohl für die x als auch für y Komponente gleich sind, gilt Formel (3-11). Die Standardabweichung der Distanz zwischen zwei Positionen kann nun mit Formel (3-12) und schließlich die Standardabweichung der Gesamtdistanz mit Formel (3-13) berechnet werden.

$$s_x = s_{x_{i-1}} = s_{x_i} = s_{y_{i-1}} = s_{y_i} \tag{3-11}$$

$$s_d = s_x \cdot \sqrt{\left(\frac{x_1 - x_2}{d}\right)^2 + \left(\frac{x_2 - x_1}{d}\right)^2 + \left(\frac{y_1 - y_2}{d}\right)^2 + \left(\frac{y_2 - y_1}{d}\right)^2} \tag{3-12}$$

$$s_{d,gesamt} = \sqrt{\sum_{i=1}^{n} s_{d,i}^2} = \sqrt{\sum_{i=1}^{n} 2 \cdot s_x^2} = s_x \cdot \sqrt{2 \cdot n} \tag{3-13}$$

Zur Abschätzung, ab welcher Gesamtdistanz eine Bewegung signifikant ist, kann mit Formel (3-14) je nach gewählter Sicherheitswahrscheinlichkeit *P* mit dem Quantil y_P der Standardnormalverteilung die Grenzdistanz d_{min} berechnet werden.

$$d_{\min} = s_d \cdot y_P = s_x \cdot \sqrt{2} \cdot y_P \tag{3-14}$$

In der folgenden Tabelle 4 sind für ausgewählte Sicherheitswahrscheinlichkeiten die Werte der Grenzdistanzen zwischen zwei Positionen aufgeführt. In der Abbildung 4-18 sind für verschiedene Sicherheitswahrscheinlichkeiten die Grenzdistanzen bis zur Bewegungssignifikanz für ausgewählte FPD-Positionsgenauigkeiten dargestellt.

Tabelle 4: Ausgewählte Sicherheitswahrscheinlichkeiten und deren Grenzdistanzen

Sicherheitswahrscheinlichkeit *P*	Quantil y_P	Grenzdistanz d_{min} zwischen 2 Positionen für s_{FPD}=150m
65%	0,39	82,7 m
80%	0,85	180,3 m
90%	1,28	271,5 m
95%	1,65	350,0 m
97,5%	1,96	415,8 m
99%	2,33	494,3 m
99,5%	2,58	547,3 m
99,9%	3,10	657,6 m

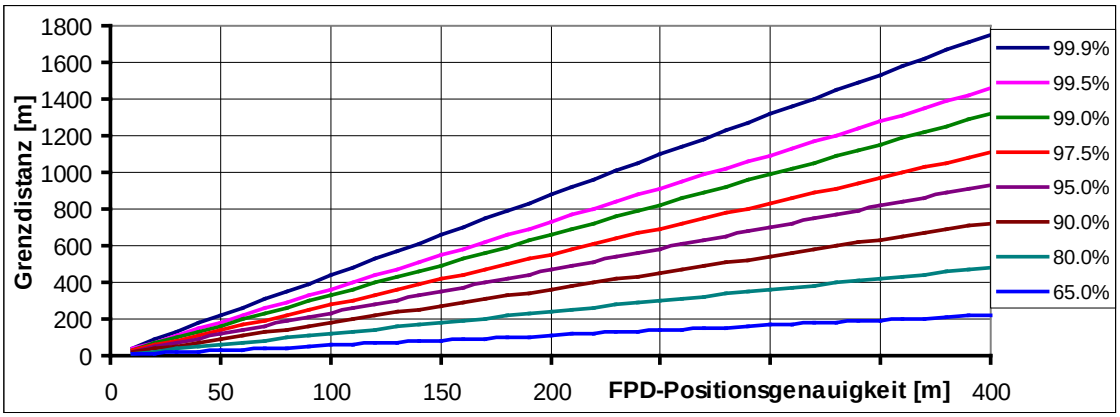


Abbildung 4-18: Abhängigkeit der Grenzdistanz zur Aufdeckung von Bewegungen bei verschiedenen FPD-Positionsgenauigkeiten

4.5.1.2 Grenzgeschwindigkeit

Als eine weitere Variante der Bewegungserkennung ist unter der Annahme einer Bewegung eine signifikante Geschwindigkeit zu prüfen. Neben der zurückgelegten Distanz ist zusätzlich die verstrichene Zeit zu berücksichtigen, so dass bei Überschreitung einer signifikanten Grenzgeschwindigkeit eine Bewegung angenommen werden kann.

Die Abis-Datenschnittstelle liefert während eines Telefonates für einen Teilnehmer alle 0,48 Sekunden einen neuen Measurement-Report, woraus jeweils eine neue Position des Teilnehmers berechnet werden kann. Je Position ist mit einer Koordinatengenauigkeit von wenigen 10 m bis zu mehreren 100 m zu rechnen. Die Genauigkeit des Measurement-Report Zeitstempels ist aufgrund der GSM Systemspezifikationen wenige Millisekunden. Für den Verwendungszweck der Geschwindigkeitsbestimmung ist festzustellen, dass die zeitliche und die geometrisch Genauigkeit sehr unterschiedlichen Einfluss auf das Ergebnis

haben. Bei einer angenommenen maximalen Geschwindigkeit von 250 km/h legt ein Fahrzeug ca. 35 m in 0,48 Sekunden zurück. Bei einem angenommenen Fehler von 10 ms ist der geometrische Einfluss durch die Zeitmessung selbst bei dieser hohen Geschwindigkeit mit ca. 0,35 m deutlich geringer als die zu erwartenden Fehler durch die Einzelpositionsberechnung von bis zu einigen 100 m.

Aufgrund der vergleichsweise hohen Positionsbestimmungsrate und der geringen geometrischen Genauigkeit ist der effektive Informationsgehalt je Einzelposition gering. Es treten sowohl Positionsfolgen mit sehr hohen und nahezu identischen systematischen Einflüssen, so genanntes farbiges Rauschen (hohe Korrelation), sowie Positionsfolgen mit fast ausschließlich zufälligen Einflüssen, weißes Rauschen (keine Korrelation), auf.

Abschätzung der zu erwartenden Genauigkeiten und Signifikanzgrenzen

Für die Abschätzungen wird angenommen, dass die FPD-Positionen statistisch normalverteilt sind. Ebenso wird aus den oben genannten Gründen der Einfluss der Zeitmessgenauigkeit vernachlässigt.

In Abbildung 4-19 sind für verschiedene Distanzgenauigkeiten, abgeleitet aus FPD-Positionsgenauigkeiten, die Abhängigkeiten zwischen Genauigkeit der Geschwindigkeit und dem erforderliche Zeitintervall aufgetragen. Wenn beispielsweise eine Genauigkeit der Geschwindigkeit von ca. ±10 m/s erzielt werden soll, ist bei einer typischen Distanzgenauigkeit von ±150 m zwischen zwei FPD-Positionen ein Zeitintervall von ca. 15 Sekunden erforderlich. Je genauer die Geschwindigkeit bestimmt werden soll, desto länger muss das entsprechende Zeitintervall sein. Da die FPD-Positionsgenauigkeiten sehr variieren können, sind diese Beispielrechnungen natürlich nur als Abschätzung der erzielbaren Genauigkeiten zu verstehen.

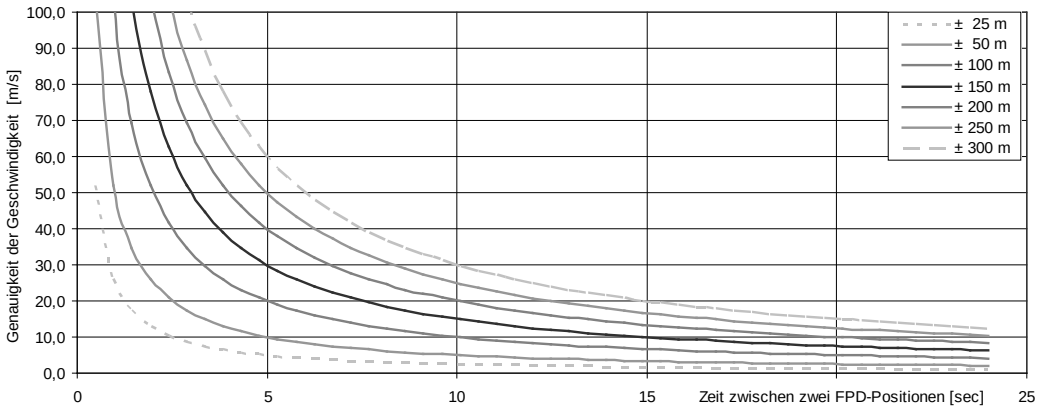


Abbildung 4-19: Erwartete Genauigkeiten für die Geschwindigkeitsbestimmung aus zwei FPD-Positionen bei unterschiedlichen Zeitabständen (unterste Kurve für ±25 m, oberste Kurve für ±300 m Distanzgenauigkeit)

Aus der Abbildung 4-19 können weiterführend folgende Aspekte interpretiert werden:

1. Die Geschwindigkeitsbestimmungen im gleichen Zeitintervall wie die Abis Positionsbestimmung mit ca. 2 Hz sind nicht sinnvoll. Vielmehr ist je nach FPD-Positionsgenauigkeit und geforderter Geschwindigkeitsgenauigkeit, ein minimal abzuwartendes Zeitintervall festzusetzen.
2. Durch Definition einer Grenzgeschwindigkeit ist prinzipiell die Erkennung von stehenden MFT möglich. Geschwindigkeiten über dieser Grenze, können ein sich bewegender MFT identifizieren.

Bestimmung von Grenzgeschwindigkeiten

Für die folgende Betrachtung wird von normal verteilten FPD-Position und einer FPD-Positionsgenauigkeit von ±150 m ausgegangen. Mit Vorgabe einer Sicherheitswahrscheinlichkeit *P* kann nun die erforderliche Zeit t bis zur Signifikanz einer Bewegung ermittelt werden. Der Schwellwert der signifikanten Bewegung wird folgend als Grenzgeschwindigkeit bezeichnet. In Abbildung 4-20 sind für ausgewählte Sicherheitswahrscheinlichkeiten *P* zwischen 65% und 99,9% die Kurven der Grenzgeschwindigkeit aufgetragen. Beispielsweise muss ein FPD mit einer Messdauer von 50 Sekunden, eine tatsächliche Bewegungsgeschwindigkeit von mindestens 25 km/h haben, um die Bewegung mit *P* = 95% identifizieren zu können. Im Gegenzug ist ableitbar, dass man die Bewegung eines FPD mit durchschnittlich 7 km/h nach 50 Sekunden mit *P* = 65% ermitteln kann.

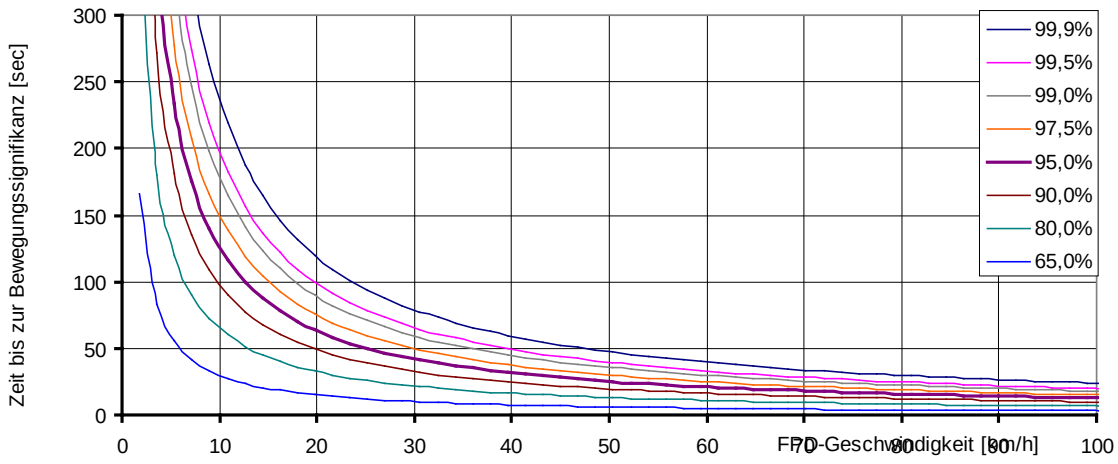


Abbildung 4-20: Zeitintervalle für die Signifikanz von FPD-Bewegungsgeschwindigkeiten für verschiedene Sicherheitswahrscheinlichkeiten bei einer FPD-Positionsgenauigkeit von 150 m

Für die Ermittlung der Zeit bis eine signifikanten Aussage der Bewegung eines FPD bei definierter Sicherheitswahrscheinlichkeit *P* getroffen werden kann, zeigt Abbildung 4-21 für ausgewählte FPD-Positionsgenauigkeiten und der Sicherheitswahrscheinlichkeit *P* = 95% die Kurven der Mindestgeschwindigkeiten. Sollen beispielsweise FPD mit 10 km/h als bewegend detektiert werden können, so sind bei FPD-Positionsgenauigkeiten von ±150 m mehr als 120 s FPD erforderlich. Sind FPD mit ±50 m verfügbar, so kann diese Aussage bereits nach 40 s getroffen werden.

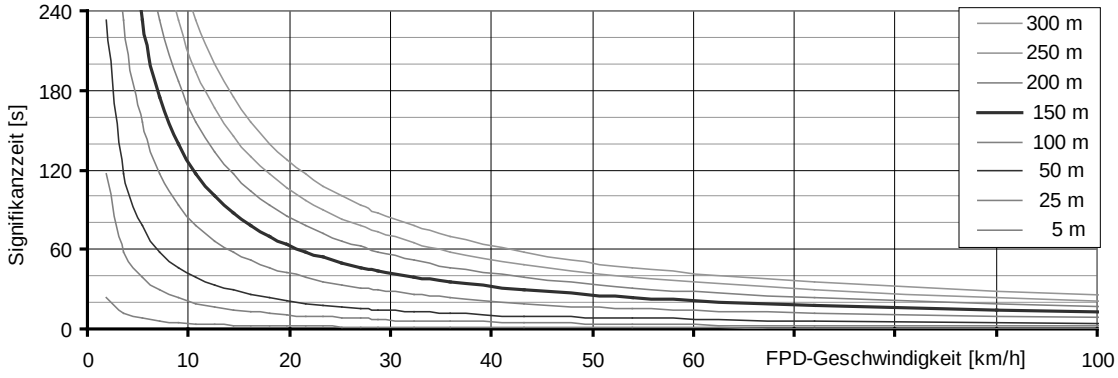


Abbildung 4-21: Zeit bis zur signifikanten Ermittlung einer Geschwindigkeit (*P* = 95% Sicherheitswahrscheinlichkeit) bei verschieden FPD-Positionsgenauigkeiten (von ±5 m bis ±300 m)

Die erforderliche Messzeit, ab der bewegte FPD signifikant sind, ist somit abhängig von der FPD-Positionsgenauigkeit, der gewünschten Sicherheitswahrscheinlichkeit *P* und der durchschnittlichen Mindestgeschwindigkeit v_{min}. Zur Optimierung der Detektionsrate könnten die Parameter (*P* und v_{min}) beispiels-

weise auch an regionale Gegebenheiten angepasst werden. Auf Autobahnen könnten höhere Sicherheitswahrscheinlichkeiten und höhere Mindestgeschwindigkeiten als innerorts mit geringeren Sicherheitswahrscheinlichkeiten und geringeren Mindestgeschwindigkeiten zur effizienteren Verkehrslagebestimmung und vor allem zur besseren Nutzung von kurzen FPD-Messreihen beitragen.

4.5.2 Clusterverfahren zur Berechnung von Positionsschwerpunkten für die Identifikation von stehenden und sich bewegenden FPD

Wie bereits oben erwähnt, ist der geometrische Informationsgehalt aufgrund der hohen Aktualisierungsrate im Vergleich zur Positionsgenauigkeit gering. Der Grundgedanke des hier behandelten Clusterverfahrens ist daher die zahlreichen FPD-Positionen mit geringer Information auf wenige Positionsschwerpunkte zu konzentrieren.

Für die Identifikation, ob der MFT steht oder sich bewegt, wird wie im Abschnitt zuvor vorerst eine Bewegung des FPD angenommen und im Anschluss überprüft, ob die Annahme zutrifft. Fand eine Bewegung des FPD statt, so soll das Clusterverfahren dieses durch räumlich genügend weit auseinander liegende Positionsschwerpunkte aufzeigen können.

4.5.2.1 Szenarien zur Bewegungserkennung mittels Positionsschwerpunkten

Die Eigenschaft einer FPD-Bewegung zeigt sich in der Signifikanz der Gesamtdistanz für den zurückgelegten Weg. Die Gesamtdistanz kann mittels Aufaddierung der Koordinatendifferenzen der Positionsschwerpunkte berechnet werden. Die Gesamtdistanz ist somit abhängig von der Generalisierung der Geometrie der FPD-Positionsfolge und somit von der Anzahl und der räumlichen Anordnung der berechneten Positionsschwerpunkte. Der Einfluss der Genauigkeiten der Einzelortungen steigt letztlich mit der Anzahl der zu berechnenden Positionsschwerpunkte, ebenso steigt der erforderliche Rechenbedarf.

Das Ziel ist daher die Repräsentation der generalisierten Positionsfolge im Sinne der Gesamtdistanzbestimmung möglichst realistisch und die Anzahl der Positionsschwerpunkte dabei möglichst gering zu halten.

Aus geometrischen Gesichtspunkten genügen für geradlinige Bewegungen zwei Positionsschwerpunkte, siehe Abbildung 4-22 Szenario „geradlinig“, zur Repräsentation. Finden Fahrten mit Hin- und Rückweg statt, muss mindestens ein weiterer Positionsschwerpunkt berechnet werden. In Abbildung 4-22 wird diese Szenario mit „gebogen 1. Art“ bezeichnet. Würden nur zwei Positionsschwerpunkte berechnet werden, könnten sie räumlich in unmittelbarer Nähe liegen, wodurch die Distanz falsch geschätzt werden würde. Bei einer ange-

nommen Bewegung mit zweifacher Kreuzung des Weges, wie es in Abbildung 4-22 im Szenario „gebogen 2. Art“ dargestellt ist, sind bereits mindestens vier Positionsschwerpunkte erforderlich. Je komplizierter die Bewegungsformen werden, bei denen in die Nähe früherer Positionen zurückgekehrt wird oder Wege gekreuzt werden, desto mehr Positionsschwerpunkte sind zur Aufdeckung der Gesamtdistanz notwendig. Kompliziertere Wegformen als „gebogen 2. Art“ werden in Do-iT nicht betrachtet.

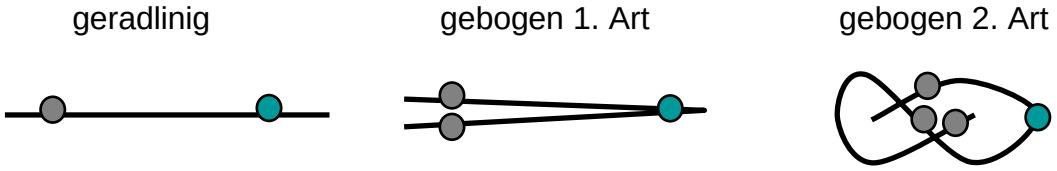


Abbildung 4-22: Szenarienformen zur Erkennung von Bewegungen

Je nach zeitlicher Länge der FPD-Positionsfolge werden für die Szenarienformen untersucht, ob die Sicherheitswahrscheinlichkeit für die Bewegung mindestens den vorgegeben Schwellwert erreicht. Die Wahl der FPD-Schwerpunktanzahl wird in Abschnitt 4.6 näher erläutert. Die Szenarienform „geradlinig“ wird für FPD-Positionsfolgen bis 60 Sekunden, zusätzlich „gebogen 1. Art“ bis 90 Sekunden und für mehr als 90 Sekunden ebenfalls „gebogen 2. Art“ untersucht. Im Auswerteprozess ist implementiert, dass sobald auch die Szenarienform „gebogen 2. Art“ berechnet wird, mindestens zwei der drei Szenarienformen den Schwellwert erreichen müssen. Eine funktionale Erklärung konnte bislang nicht gefunden werden. Jedoch zeigt sich empirisch, dass stehende FPD besser erkannt werden können. Ursache könnten stärkere systematische Fehlereinflüsse sein. Für bewegte FPD wurden dadurch keine Einschränkungen bei der Erkennung beobachtet.

4.5.2.2 FCM-Clusterverfahren zur FPD-Schwerpunktberechnung

Allen Clusterverfahren gemein ist die Eigenschaft, dass die Cluster zueinander sich maximal unterscheiden sollen und die Objekte in einem Cluster möglichst ähnliche Kriterien aufweisen sollen. Übertragen auf die hier gestellte Problemstellung, sind räumlich und zeitlich nahe beieinander liegende FPD-Positionen zu FPD-Schwerpunkten zusammen zu fassen. Neben den kartesischen Koordinaten X und Y, wie sie in Abbildung 4-23 für einen sich nicht bewegenden Teilnehmer gezeigt werden, ist die Zeit t als dritte Achse zu verwenden. Eine Form der Ermittlung von Nachbarschaftsbeziehungen ist die kartesische Distanzbestimmung, die auch hier eingesetzt werden soll. Die Distanzwerte sind die Vektorlängen der FPD-Positionen zu den Clusterschwerpunktkoordinaten.

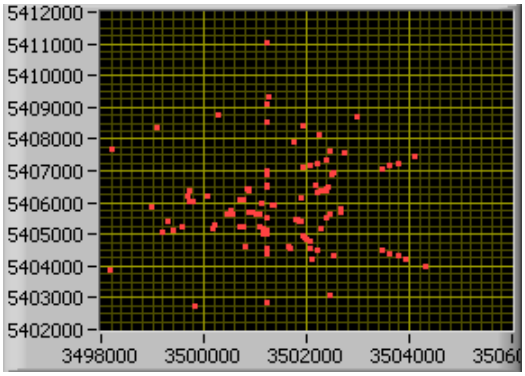


Abbildung 4-23: FPD-Positionen eines stehenden Teilnehmers

Zu beachten ist, dass die Dimensionseinheiten metrische Länge und Zeit diese Berechnung nicht ohne weiteres gestatten. Um den Zugehörigkeitsgrad einer FPD-Position zu einem Clusterschwerpunkt zu bestimmen, genügt es jedoch ein „Maß“ der Zugehörigkeit zu ermitteln. Die Ermittlung dieses Maßes sollte wiederum repräsentativ für die Beziehung zwischen den Eingangsdaten und den gewünschten Ergebnissen sein. In Do-iT wurden empirisch mehrere Anpassungen zwischen den Koordinaten und der Zeit untersucht. Folgendes Vorgehen zeigte bislang die günstigsten Eigenschaften.

Die Achsen X und Y werden unverändert metrisch belassen. Da die Distanzberechnung im FCM alle drei Koordinatenachsen gleich gewichtet, wird die Zeitachse entsprechend der maximalen Ausdehnung von X oder Y skaliert. Zusätzlich zeigte sich, dass durch eine weitere Skalierung der Zeitachse um einen Wert von 2,5 bis 3,5 bessere Clusterschwerpunktstreuungen erreicht wurden. Innerhalb dieses Bereiches werden häufig sehr ähnliche Ergebnisse für die Clusterschwerpunktkoordinaten berechnet. Durch die zusätzliche Skalierung üben Positionen, die zeitlich weit entfernt vom Clusterschwerpunkt liegen, einen deutlich geringeren, und somit weniger verzerrenden Einfluss aus. Begründet ist dieses Verhalten in der geringeren Koordinatengenauigkeit gegenüber der hohen Genauigkeit der Zeit. Systematische Störeinflüsse (hohe Anzahl von FPD-Positionen mit ähnlichen Koordinaten, trotz Teilnehmerbewegung) wirken sich geringer aus, da sie gegebenenfalls durch einen eigenen, meist weit entfernten Clusterschwerpunkt repräsentiert werden und mit einem Ausreißertest eliminiert werden können.

In Abbildung 4-24 sind die FPD-Schwerpunkte für das oben gezeigt Bespiel eines stehenden FPD dargestellt. Die linke Abbildung zeigt je FPD-Schwerpunkte farbig hervorgehoben die jeweilig zugehörenden FPD-Positionen. In der rechten Abbildung sind die berechneten FPD-Schwerpunkte weiß hervorgehoben.

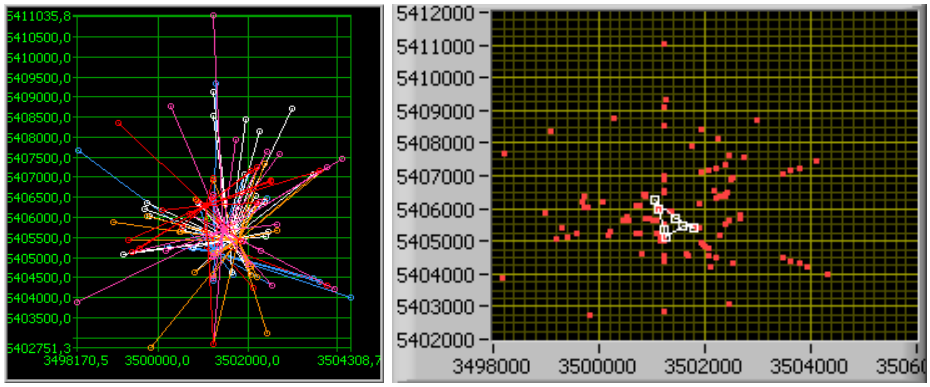


Abbildung 4-24: Bestimmung der FPD-Schwerpunkte mittels FCM-Clusterbildungsverfahren für ein stehendes FPD

4.5.3 Klassifizierung der FPD-Bewegung mit FPD-Schwerpunkten

Unter der Annahme, dass die FPD-Schwerpunkte statistisch normal verteilt sind, kann für die Bewegungserkennung bspw. die bereits oben erläuterte Methode mittels Grenzdistanz und einer gewünschten Sicherheitswahrscheinlichkeit durchgeführt werden. Bei Versuchsfahrten zeigte sich jedoch, dass die Bewegungserkennung meist auch bei kleineren Grenzdistanzen, wie unten näher erläutert wird, möglich ist. Die Annahme der normal verteilten FPD-Positionen und der FPD-Schwerpunkte scheint daher nicht uneingeschränkt zu zutreffen. Infolge dessen wurde mit den Ergebnissen der Versuchsfahrten eine Fuzzy-Logik erstellt, die einen Zugehörigkeitswert anstatt einer Wahrscheinlichkeit für die Bewegung errechnet.

Für den Entwurf der Fuzzy-Logik wurden die Daten der Versuchsfahrten nach stehend und bewegt zugeordnet sowie die Gesamtdistanzen zwischen der FPD-Schwerpunkten berechnet. Abbildung 4-25 zeigt die Häufigkeitspolygone für stehende und bewegte Versuchsfahrten, wobei die aus den FPD-Schwerpunkten berechneten Gesamtdistanzen in Klassen mit einer Klassenbreite von 25 m gruppiert wurden. Aus der Abbildung 4-25 sind bezogen auf die Klassen der Gesamtdistanzen folgende Aspekte erkennbar:

a) kein bewegter Versuch hatte eine Gesamtdistanz von weniger als 100 m
b) kein stehender Versuch hatte eine Gesamtdistanz von mehr als 350 m
c) für Gesamtdistanzen zwischen 100 m und 350 m existieren sowohl stehende als auch bewegte Versuche

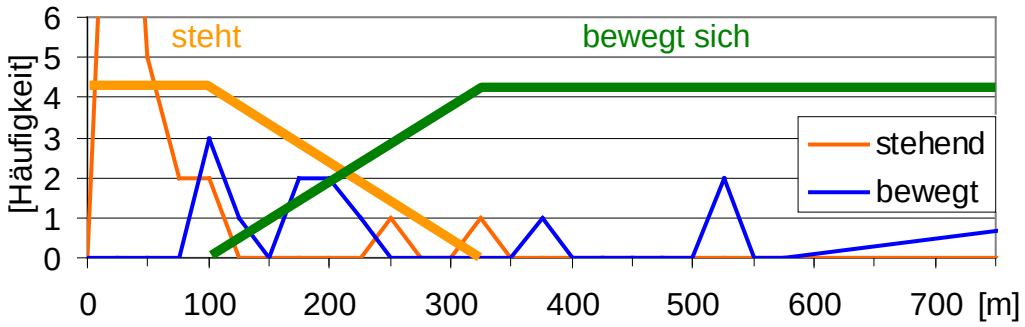


Abbildung 4-25: Häufigkeitspolygone für stehende und bewegte Versuchsfahrten sowie die Zugehörigkeitsfunktionen zur Fuzzyfizierung der Gesamtdistanz aus den FPD-Schwerpunkten

Die genannten Aspekte können zur Fuzzyfizierung der Eingangsvariable „Gesamtdistanz“ angewendet werden. Die Zugehörigkeiten können funktional bspw. durch Trapez- oder Dachfunktionen beschrieben werden. Die Zusammenfassung der Zugehörigkeitsfunktionen ergibt schließlich die Fuzzyfunktion für die Gesamtdistanz. Abbildung 4-25 zeigt die modellierten Zugehörigkeitsfunktionen farbig gekennzeichnet mit orange für „steht“ und grün für die Zugehörigkeit zu „bewegt“. Im Überlappungsbereich wird ein linearer Verlauf angenommen, wobei größere Gesamtdistanzen eher eine Bewegung kennzeichnen als kleinere. Da nur eine Eingangsgröße verwendet und ein linearer Verlauf im unsicheren Bereich angenommen wird, kann auf eine Inferenz und die Modellierung einer gesonderten Defuzzifierung verzichtet werden. Der Zugehörigkeitswert für die „Bewegung“ des FPD kann mit Formel (3-15) berechnet werden, wobei Z_{steht} der Wert der Zugehörigkeitsfunktion für „steht“ und Z_{bewegt} der Wert der Zugehörigkeitsfunktion für „bewegt“ ist. Aufgrund der linearen Abhängigkeit gilt $Z_{bewegt} = 1\text{-}Z_{steht}$ für Z_{steht} und $Z_{bewegt} = 0..1 \in R$. Die eingesetzten Werte gelten natürlich ausschließlich für die oben geschlossenen Erkenntnisse aus den Versuchsfahrten.

$$Z_{Bewegung} = \begin{cases} 0 & \text{für} \quad d_{gesamt} \leq 100m \\ Z_{bewegt} \text{ bzw. } 1 - Z_{steht} & \text{für} \quad 100\text{m} < d_{gesamt} > 350\text{m} \\ 1 & \text{für} \quad d_{gesamt} \geq 100m \end{cases} \qquad (3\text{-}15)$$

Für $Z_{Bewegung}$ kann analog der oben beschriebenen Sicherheitswahrscheinlichkeit ein Schwellwert gesetzt werden. Angemerkt sei: die Gesamtdistanzen und der darauf abzustimmende Schwellwert ist nur für die FPD-Schwerpunkte mit dem oben beschriebenen Clusterverfahren sowie den aus den Versuchsfahrten abgeleiteten Fuzzyfizierungsfunktionen gültig.

4.6 FCM – Koordinatenschwerpunktbildung bewegter FPD

Folgt aus der Bewegungsanalyse, dass sich das FPD bewegt, ist die Annahme mehrerer FPD-Schwerpunkte für die Positionsfolge des FPD gerechtfertigt. In der Bewegungsanalyse wurden basierend auf den drei behandelten Szenarienformen jedoch nur maximal vier Positionen berechnet. Um die geometrische Form und den Verlauf der Geschwindigkeit repräsentativer zu bestimmen, sind besonders längere Positionsfolgen zeitlich und räumlich höher aufzulösen. Bei der Wahl der FPD-Schwerpunktanzahl sind unter anderem die folgenden Aspekte abzuwägen: je höher die FPD-Schwerpunktanzahl desto höher ist die räumliche und zeitliche Auflösung, desto mehr können sich Ungenauigkeiten auswirken, desto geringer ist die Informationsdichte und desto mehr Rechenressourcen benötigt das Clusterverfahren. Berücksichtigt man die zu erwartenden Genauigkeiten der FPD-Positionen und stellt die erforderliche Messzeit bis zur Aufdeckung einer möglichen Geschwindigkeit gegenüber, siehe Abschnitt 4.5.1.2, ergeben sich bspw. im städtischen Gebiet für die Identifizierung von Fußgängern und PKW mit einer Sicherheitswahrscheinlichkeit von 80% Messzeiten von ca. 20 Sekunden bis ca. 80 Sekunden. Als Kompromiss zwischen beiden Teilnehmerklassen, wird in Do-iT ein zeitlicher Abstand von 30 Sekunden gewählt.

Für die Berechnung der FPD-Schwerpunkte wird das gleiche Clusterverfahren FCM eingesetzt, wie es bei der Bewegungsanalyse verwendet wurde. Für die FCM Berechnung ist die Anzahl der zu ermittelnden Cluster vorzugeben. In Do-iT wird diese Anzahl anhand der zeitlichen Länge der FPD-Positionsfolge berechnet, sodass alle 30 Sekunden ein FPD-Schwerpunkt generiert wird. Zu beachten ist, dass die zeitlichen Abstände der berechneten FPD-Schwerpunkte des FCM ebenfalls abhängig von der räumlichen Lage sind und daher variieren können. Die Abbildung 4-26 zeigt reale Beispieldaten für die berechneten FPD-Schwerpunkte, rote Punkte mit roten Verbindungslinien. Die zu den jeweiligen FPD-Schwerpunkten am meisten zugehörigen FPD-Positionen sind durch grüne Punkt mit grünen Verbindungslinien dargestellt.

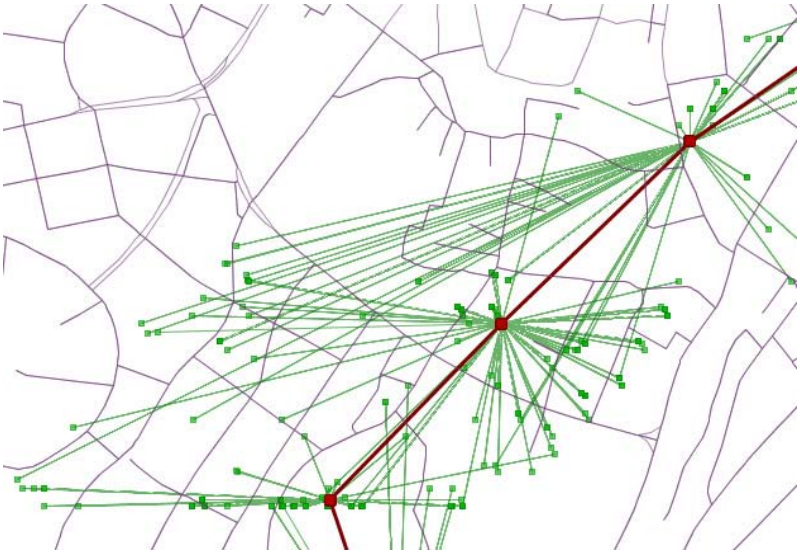


Abbildung 4-26: FPD-Positionsschwerpunkte (rot) für sich bewegende FPD mit Verbindungslinien zu den zugeordneten FPD-Positionen (grün)

4.7 Aufenthaltsanalyse, POI

Es kann davon ausgegangen werden, dass an bestimmten Orten das Vorhandensein einzelner Verkehrsklassen ausgeschlossen oder nur spezielle Verkehrsklassen angetroffen werden. So werden z.B. in Erholungsparks oder Fußgängerzonen keine PKW, jedoch Fußgänger erwartet. Das Ziel der POI-Aufenthaltsanalyse ist es nach Korrelationen der FPD-Position zu besonderen Geoobjekten (POI - Point Of Interest) zu suchen und zu prüfen, ob Verkehrsklassen hervorgehoben oder ausgeschlossen werden können.

4.7.1 Geodatenintegration

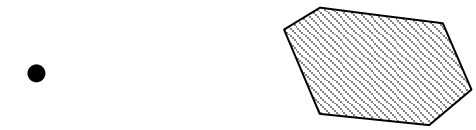
Basis der Aufenthaltsanalyse sind POI, die einen Ausschluss oder das Vorhandensein von Verkehrsklassen oder Verkehrsklassengruppen gestatten. Grundlage für die Verkehrsklassen sind die Abbildung 3-2 gezeigt Teilnehmerklassen. Die Datenquelle für die POI sind die für Do-iT definierten Daten von Navteq 2005/Q3. Im Projekt Do-iT wurden daraus die in Tabelle 5 zusammengefassten Geoobjektarten herangezogen. Auf der rechten Tabellenhälfte sind entsprechend der Nutzungsart die Verkehrsklassen genannt, die aufgrund der Geoobjektart ausgeschlossen oder bestätigt also hervorgehoben werden können.

Die räumliche Beschreibung, die so genannte Georeferenzierung, der POI kann in Form von postalischen Adressen (z.B. Geschwister-Scholl-Str. 24D) oder in Form von geografischen Koordinaten erfolgen. In den Navteq Daten werden gemäß der GDF Spezifikation neben der geografischen Koordinate auch eine dazu gehörende Straßenkante angeführt. Für die POI-Aufenthaltsanalyse werden ausschließlich die geografischen Koordinaten eingesetzt.

Obwohl die meisten POI eine räumliche Ausdehnung aufweisen, wird meist jedoch nur eine Koordinate (punktueller Raumbezug) angegeben, wie z.B. der Eingang eines Parks oder der Mittelpunkt von Gebäuden. Die räumliche Korrektheit ist somit nur eingeschränkt gewährleistet. Zur Einbindung von POI in die Aufenthaltsanalyse können prinzipiell Punkte und Flächen verwendet werden. Flächen werden hierbei als geschlossene Flächenpolygone verstanden, siehe Abbildung 4-27. Für punktuelle POI ist allerdings die Berechnung der Aufenthaltswahrscheinlichkeit nur bedingt möglich, siehe in Folgenden.

Tabelle 5: Geoobjektarten für den Ausschluss bzw. das Hervorheben von Verkehrsklassen

Geoobjektart des POI (Navteq NAVSTREETS Street Data, 2005/Q3)		**Verkehrsklasse**	
		Ausschluss	Hervorheben
Waterway Polygons, Waterway Segments (WaterPoly / WaterSeg)		Alle	
Building / Landmark Features (Landmark)			Ohne Verkehrsbeteiligung, Fußgänger
Hospitals			
Parks and Recreation			
Shopping			
Restaurant			
Entertainment			
Financial Instituts			
Business Facilities			
Community Service Centres			
Education Institutions			
Post Office			
Auto Maintenance, Service, and Petrol		Motorisierter Individualverkehr	
Parking			
Transportation Hubs - Airport - Bus Station - Train Station - Ferry Terminal - Commuter Rail Station			Öffentlicher Verkehr, Ohne Verkehrsbeteiligung



Punkt

geschlossenes Flächenpolygon

Abbildung 4-27: Eingangsdaten für die räumliche Beschreibung von POI

4.7.2 Vorprozessierung für die FPD-Echtzeitanalyse

In der Vorprozessierung wird die räumliche Suche für eine effiziente Analyse im Online-Betrieb vorgearbeitet. Grundgedanke ist, wie auch in den vorherigen Abschnitten, die Verwendung der Kacheldarstellung der Geodaten. Für die entsprechenden POI Kacheln werden die Wahrscheinlichkeiten entsprechend der räumlichen Verteilung für die jeweiligen Verkehrsklassen berechnet. Da für punktuelle POI diese Art der Wahrscheinlichkeitsangabe nicht repräsentativ ist, wurde dieses Verfahren in Do-iT vorerst zurückgestellt. Die Erfassung der Flächendarstellung für ausgewählte POI wird derzeit für ausgewählte POI vorgenommen. Die POI-Analyse ist derzeit in Do-iT daher noch nicht im Einsatz.

4.8 Klassifizierung von FPD in Öffentlichen Verkehren

Das Grundprinzip des Verfahrens stützt sich darauf, dass der Öffentliche Verkehr nach festen vordefinierten Fahrplänen und auf definierten Fahrwegen Personen transportiert. Flexibler Öffentlicher Verkehr wie Taxi, Anrufsammeltaxis und routenflexible Angebote, die den Start- und Zielpunkt der Fahrten sowie den Fahrweg erst unmittelbar bei Fahrtbeginn festlegen oder frei wählen, seien vorerst nicht berücksichtigt. Wie noch gezeigt wird, ist eine dahingehende Erweiterung flexibler Daten durchaus gewährleistet. Zur Abgrenzung von den flexiblen Öffentlichen Verkehren ohne festen Fahrplan (flexÖV), wird im Weiteren von Öffentlichen Verkehre mit vordefiniertem Fahrplan (PlanÖV) gesprochen.

Ausgangsinformationen der Mobilfunkteilnehmer für die ÖV-Analyse sind die 2-dimensionalen geografischen Koordinaten (FPD-Position) und die Erfassungszeitpunkte (Zeit). Diese räumlichen und zeitlichen Informationen werden daraufhin analysiert, ob ein oder mehrere Fahrzeuge des PlanÖV an derselben Position zur selben Zeit sind. Da sowohl die FPD-Positionen Ungenauigkeiten aufwesen und auch die PlanÖV Fahrzeuge Verfrühungen bzw. Verspätungen unterliegen können, ist die Berücksichtigung von Unschärfen unumgänglich.

4.8.1 Vorgehensweise

Um die oben genannte zeitliche und räumliche Analyse mit PlanÖV Fahrzeugen durchzuführen, werden die PlanÖV Fahrzeugpositionen auf einer digitalen Karte prognostiziert. Aus den Fahrplänen können hierfür die An- und Abfahrzeiten der Fahrzeuge an den Haltestellen und somit auch die Fahrzeiten zwischen den Haltestellen entnommen werden.

Zur Prognose der Fahrzeugpositionen zwischen den Haltestellen werden die Fahrwege oder Linienwege auf der digitalen Karte geroutet. Als digitale Karte wird das Lokalisierungsnetz eingesetzt. Da das Lokalisierungsnetz vorrangig Verkehrsnetze des Individualverkehrs beinhaltet, sind Verkehrsnetze, besonders Schienennetze, durch eine Georeferenzierung neu einzufügen.

Mit den Fahrplaninformationen sowie Annahmen für typische Verfrühungen und Verspätungen werden zwischen den Haltstellen die Fahrzeugpositionen prognostiziert. Da die PlanÖV-Klassifizierung in Echtzeit durchzuführen ist, werden die Prognosen in Vorprozessierungen berechnet und für den Online-Betrieb effizient gespeichert.

4.8.2 Planungsdaten der Öffentlichen Verkehre

4.8.2.1 Georeferenzierung der Fahrpläne

Mit den Fahrplänen der Verkehrsverbünde und –betriebe werden die Zu- bzw. Aussteigeorte sowie die Zeiten, an denen in die Fahrzeuge ein bzw. aus den Fahrzeugen ausgestiegen werden kann, angegeben. Die Ketten der Ein- und Aussteigepunkte ergeben die Routen der Fahrzeuge. Aus Sicht der Fahrzeugprognose geben die Fahrpläne an, zu welchen Zeitpunkten sich welche Fahrzeuge an welchen Positionen aufhalten. Wie auch in Abbildung 4-28 gezeigt, kann durch Interpolation die Fahrzeugposition zwischen den Haltepunkten auf den Fahrwegen bestimmt werden.

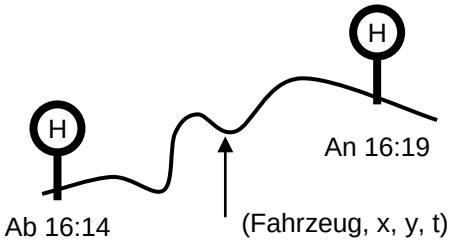


Abbildung 4-28: Schemadarstellung der Interpolation von prognostizierten PlanÖV Fahrzeugpositionen

Wesentliche Komponenten sind die Angabe der Ab- und Ankunftszeiten aus Fahrplänen, der Aufenthaltsorte als Haltepunkte und soweit vorhanden auch Zwischenpunkte zur genaueren Beschreibung des geplanten Fahrweges. Im Projekt Do-iT dient das Lokalisierungsnetzes zur Beschreibung und Haltung von Verkehrsnetzen. Da die Haltepunkte in diesen Netzen bislang nicht vorhanden sind, werden sie im Fall der Linienbusse auf die bestehenden Straßennetze referenziert. Im Fall der Schienenwege werden zumeist vollständig neue Netze einschließlich der Haltepunkte neu eingefügt. Diese Daten wurden vom Projektpartner VuV vorbereitet und bereitgestellt.

4.8.2.2 Referenzierung der Haltepunkte

Strukturbedingt existiert bis zum heutigen Tag keine einheitliche, flächendeckende und vor allem vollständige Datenbasis für Haltepunktdaten der Verkehrsverbünde und Verkehrsbetriebe. Viele Unternehmen halten jedoch seit Einführung des NAVSTAR GPS neben der Ortsbeschreibung durch die postalische Adresse auch die GPS-Koordinaten vor. Für die Referenzierung im Lokalisierungsnetz werden die GPS-Koordinaten sowie eine eindeutige Bezeichnung der Haltepunkte vorausgesetzt. Ist die GPS-Koordinate nicht vorhanden, so ist die nachträgliche Erfassung mittels GPS-Empfänger oder die Georeferenzierung bspw. mittels eines GIS durchzuführen.

Das Lokalisierungsnetz beinhaltet das Straßennetz mit Verkehrsknoten und Straßenkanten, die sowohl topologisch als auch geometrisch beschrieben sind. Die Haltepunkte sind zusätzliche Elemente und ergänzen daher das Lokalisierungsnetz. Für die Aufrechterhaltung der topologischen und geometrischen Beziehungen im Lokalisierungsnetz, werden die Haltepunkte als Verkehrsknoten betrachtet und in die Straßenkanten eingeschoben. Dazu wird die nächstliegende Kante zum Haltepunkt gesucht und der Lotfußpunkt darauf berechnet. Wie in Abbildung 4-29 a) und b) gezeigt, wird die ursprüngliche Koordinate des Haltepunktes verworfen. Sehr häufig existiert zur durchbrochenen Straßenkante auch eine Kante in entgegensetzter Richtung (bidirektionale Befahrbarkeit der Straße). Ist dies der Fall und für den Haltepunkt existiert keine besonderen Fahrtrichtungsangaben, dann wird auch diese Kante durchbrochen und der Haltepunkt als zusätzlicher Verkehrsknoten eingefügt. Abbildung 4-29 c) zeigt das Einfügen eines Haltepunktes aus topologischer Sicht.

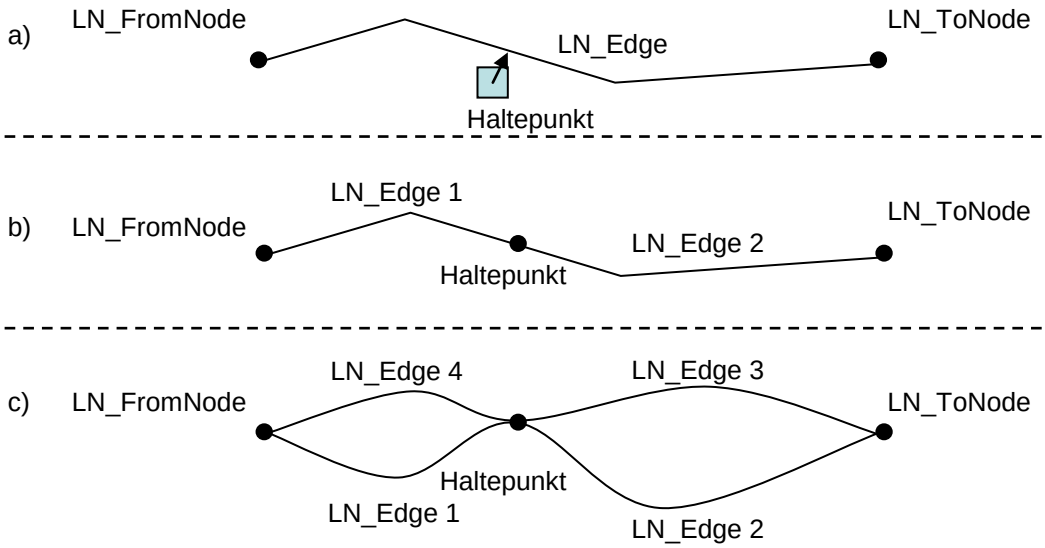


Abbildung 4-29: Referenzierung eines Hlatepunkts im Lokalisierungsnetz

Topologisch sind die Haltepunkte nun gleichzusetzen mit allen anderen Verkehrsknoten. Sie besitzen jedoch die besondere Eigenschaft, dass sie jeweils Start- und Zielpunkt von PlanÖV-Routen bzw. PlanÖV-Teilrouten sind.

4.8.2.3 Georeferenzierung von ÖV-Verkehrsnetzen

Verkehrsverbünde oder Verkehrsbetriebe verwenden zunehmend Verkehrsnetze für die Erstellung von Fahrplänen oder zur Fahrgastinformation. Häufig sind die Netze auch routingfähig, sodass sie topologisch ähnlich dem Lokalisierungsnetz aufgebaut sind. Diese Daten sind besonders wertvoll, da die Wege zwischen den Haltepunkten eindeutig definiert sind. Zur Übernahme und Refe-

renzierung dieser Daten in Lokalisierungsnetz sind in der Praxis zwei wesentliche Datenbestände zu unterscheiden.

Zum einen wurde ein spezielles Straßennetz oder ein Straßennetz eines anderen Verkehrsnetzanbieters als Basisdatenbestand verwendet. Die Koordinaten der Knotenpunkte sowie der Straßenkanten unterscheiden sich somit von denen des Lokalisierungsnetzes. Gegebenenfalls existieren zusätzliche Knoten und Kanten, die im Lokalisierungsnetz nicht vorhanden sind.

Zum anderen können Verkehrsnetzdaten übernommen werden, die vom gleichen Verkehrsnetzanbieter stammen, jedoch worin zusätzliche Elemente ergänzt oder der Datenbestand leichte Abweichungen als der des Lokalisierungsnetzes aufweist. Sofern die Objekt Identifier (Objekt_ID) und die Geometrie der Objekte übereinstimmen, müssen die entsprechenden Objekte nicht im Lokalisierungsnetz ergänzt werden. Stimmt eines der beiden Attribute nicht überein, ist eine räumliche und topologische Prüfung erforderlich, um neue Objekte zu ergänzen oder Objekte zu ersetzen.

Es ergeben sich in der Praxis auch Mischformen, sodass ein einheitliches Vorgehen für die Übernahme der Netze meist nicht vorausgesetzt werden kann. In Do-iT wurde daher eine Methode entwickelt, die auf folgenden Grundprinzipen basisiert:

1. Bestehende Netze werden nur verdichtet, d.h. keine Objekte gelöscht oder geometrisch verändert.

2. Neue Verkehrsnetze werden als komplett neue Netze eingefügt. Diese Netze besitzen in der Regel, wie in Abbildung 4-30 schematisch dargestellt, keine topologischen Übergänge zu bestehenden Netzen.

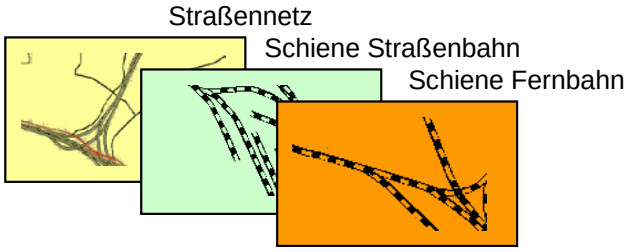


Abbildung 4-30: Ergänzte PlanÖV-Verkehrsnetze im Lokalisierungsnetz

4.8.2.4 Routing der Fahrwege

Sobald die Haltepunkte und ggf. auch die zusätzlichen Straßenelemente im Lokalisierungsnetz eingearbeitet sind, können die Routen der PlanÖV-Fahrzeuge berechnet werden. Den Fahrplänen werden dazu die Haltestellenketten entnommen. Mittels einer Routenbildung von Start- bis Endhaltestelle über alle dazwischen liegenden Haltestellen, lässt sich der Fahrweg auf dem Lokalisie-

rungsnetz berechnen. Die Haltstellenketten werden im Do-iT-Datenbestand als *Servicetrips* bezeichnet.

Bei Übernahme von kompletten ÖV-Netzen kann die korrekte Berechnung der Fahrwege angenommen werden, da nur die zulässigen topologischen Beziehungen enthalten sind. Fahrwege über untergeordnete Verkehrsnetze, z.B. auf Feldwegen oder für Linienbusse zu schmalen Gassen, können damit ausgeschlossen werden.

Liegen jedoch nur die Haltepunkte und keine zusätzlichen Routenwahlkriterien vor, wird zu Gunsten der automatisierten Prozessierung angenommen, dass das Fahrzeug jeweils zwischen den Haltepunkten die kürzeste Route wählt. Diese Annahme ist in einigen Fällen sicherlich nicht zutreffend, da besonders große Fahrzeuge aufgrund der eingeschränkten Manövrierfähigkeit Sonderwege nutzen oder schmale Wege sowie Abzweigungen meiden müssen. Aufgrund der Ortungsungenauigkeiten, Fahrplanunschärfen sowie dem geringen Einfluss kurzer Fehlroutungen, wie unten gezeigt wird, wird ein vergleichsweise geringer Einfluss auf die Güte der Klassifizierung von FPD in PlanÖV erwartet.

Im Laufe der Vorprozessierung werden in Do-iT die vollständigen Fahrwege der PlanÖV mit den jeweiligen Ab- und Ankunftszeiten aus den Fahrplandaten generiert. Ist zwischen zwei Haltestellen kein konsistenter Fahrweg vorhanden, wird ein Teilweg im Lokalisierungsnetz geroutet. Da Haltepunkte wie Verkehrsknoten im Lokalisierungsnetz bekannt sind, genügt zur vollständigen Beschreibung des Fahrweges einer Fahrt (LN_OeV_ServiceTripID) eine sortierte Liste der befahrenen Kanten (LN_Edge_ID). Zur effizienteren Bearbeitung der Interpolation und Verunschärfung der Fahrplanzeiten, siehe folgende Abschnitte, werden temporär auch die Geometriepunkte der Fahrwege mit abgelegt. Das für die Fahrwegsbeschreibung verwendete Datenformat ist in Abbildung 4-31 dargestellt.

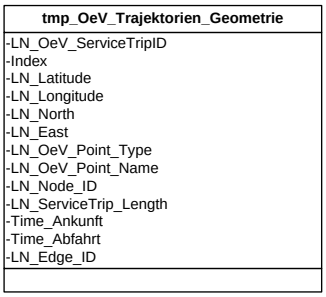


Abbildung 4-31: Datenmodell der Fahrwege mit geometrischem Verlauf während der PlanÖV Datenvorprozessierung

4.8.2.5 Berechnung der Antreffzeiten der PlanÖV-Fahrzeuge an Knoten

Basisinformationen der Fahrzeiten sind die An- und Abfahrzeiten der Fahrzeuge an den Haltestelen. Unter Verwendung der im Lokalisierungsnetz gespeicherten Durchschnittsgeschwindigkeiten der Straßenelemente ist die Fahrzeit zwischen den Haltepunkten ermittelbar. Diese Fahrzeit entspricht jedoch nicht der tatsächlichen Fahrzeit des ÖV-Fahrzeugs, da sich die Durchschnittsgeschwindigkeiten zumeist auf den Transport im Individualverkehr beziehen. Die Fahrzeiten aus den Fahrplänen zwischen den Haltepunkten wurden durch die Verkehrsbetriebe ermittelt und sind häufig über längere Zeit austestet sowie an die örtlichen Gegebenheit angepasst worden. D.h. die im Fahrplan hinterlegten Fahrzeiten zwischen Ab- und Ankunftszeit entsprechen mit einer guten Approximation den real auftretenden Fahrzeiten. Für die Prognose, wann sich das Fahrzeug an welchem Punkt des Fahrweges aufhält, kann eine Skalierung der Fahrzeit aus dem Lokalisierungsnetz t_{LN}^{Fz} mit der Fahrzeit aus den Fahrplandaten t_{Plan}^{Fz} vorgenommen werden. Der mit Formel (3-17) berechenbare Faktor m_{LN}^{Plan} kann somit als eine gute Approximation für das Zeitverhalten eines PlanÖV auf dem Lokalisierungsnetz gegenüber dem Fahrplan angenommen werden.

Mittels Interpolation der skalierten Durchschnittsgeschwindigkeiten des Lokalisierungsnetzes kann für jeden Punkt des Fahrweges die Antreffzeit berechnet werden. Im folgenden Abschnitt wird gezeigt, dass die Berechnung des Antreffzeitpunktes und –ortes vereinfacht werden kann. Aufgrund der unterschiedlichen Durchschnittsgeschwindigkeiten der überfahrenen Kanten, werden daher vorerst nur die Antreffzeitpunkte an den überfahrenen Knoten berechnet, die sich zwischen zwei Haltestellen des Fahrplans befinden. Die Abbildung 4-32 veranschaulicht die Berechnung der Antreffzeit.

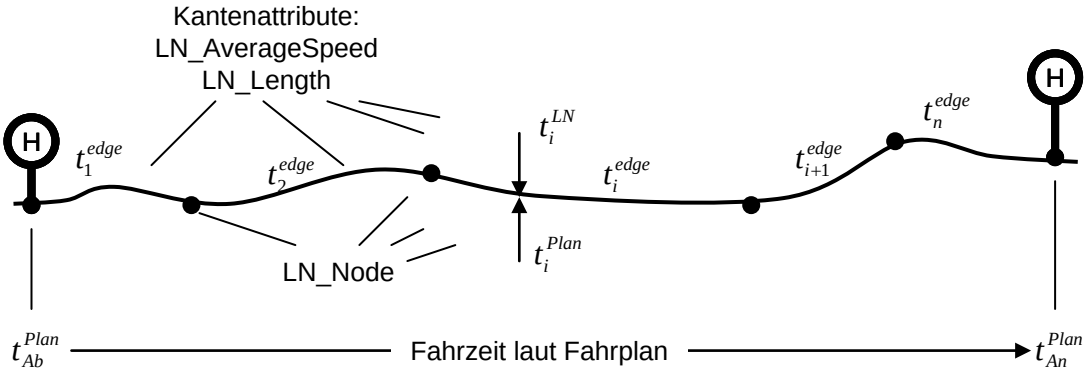


Abbildung 4-32: Interpolation der Fahrplanzeiten für Fahrzeugpositionen auf dem Fahrweg zwischen Haltestellen

Der Antreffzeitpunkt t_i^{Plan} des Fahrzeugs an einem Knoten i kann durch Interpolation zwischen der Abfahrtzeit t_{Ab}^{Plan} an der ersten Haltestelle und der Ankunftszeit t_{An}^{Plan} an der zweiten Haltestelle sowie der überfahrenen Kantenlängen (LN_Length) und der Durchschnittsgeschwindigkeit je Kante (LN_AverageSpeed) berechnet werden. Wie oben gezeigt, ist vorerst der Maßstab zwischen der PlanÖV-Fahrzeit und der Fahrzeit laut Verkehrsnetz mit Formel (3-17)zu bestimmen. Dieser Maßstab ist jedoch nur zwischen den beiden fahrplanzeitlich bekannten Punkten als gültig anzusehen. Die PlanÖV-Fahrzeit t_{Plan}^{Fz} wird aus der Differenz zwischen Abfahrtzeit t_{Ab}^{Plan} und der Ankunftszeit t_{An}^{Plan} berechnet. Die Fahrzeit aus dem Verkehrsnetz kann mittels der Kantenlängen (LN_Length) und der Durchschnittsgeschwindigkeit (LN_AverageSpeed) mit Formel (3-16) berechnet werden:

$$t_{LN}^{Fz} = \sum_{i=1}^{n} \mathrm{t}_{LN_i}^{Fz} = \sum_{i=1}^{n} \left(\frac{\mathrm{LN_AverageSpeed}}{\mathrm{LN_Length}} \right)_i \tag{3-16}$$

Der Maßstab zwischen der PlanÖV-Fahrzeit und der Verkehrsnetzfahrzeit wird berechnet mit:

$$m_{LN}^{Plan} = \frac{t_{LN}^{Fz}}{t_{Plan}^{Fz}} \tag{3-17}$$

Für jede Kante wird nun proportional ihrer ursprünglich angegebenen Fahrzeit die Fahrzeit t_i^{Plan} berechnet mit:

$$t_{Plan_i}^{Fz} = m_{LN}^{Plan} \cdot t_{LN_i}^{Fz} \tag{3-18}$$

Für jeden zwischen den Haltestellen liegenden Knoten k kann anschließend die prognostizierte Antreffzeit des PlanÖV-Fahrzeugs mit Formel (3-19) berechnet werden:

$$t_{LN_Node_k}^{Plan} = \sum_{i=1}^{k} t_{Plan_i}^{Fz} \tag{3-19}$$

Da der Maßstab der Fahrzeit für alle Kanten zwischen zwei Haltestellen gleich ist, wird vorausgesetzt, dass die ursprünglichen Geschwindigkeiten gleichwertig skaliert werden dürfen. Die Geschwindigkeitsunterschiede zwischen den Kanten werden dadurch mit skaliert. An einzelnen Kanten ist es somit möglich, dass rechnerisch Maximalgeschwindigkeiten überschritten werden, bzw. geringere Geschwindigkeiten auf Kanten angegeben werden, bei denen erfahrungsgemäß höhere Geschwindigkeiten gefahren werden.

4.8.3 Fuzzyfizierung der Fahrplandaten

Wie bereits angesprochen wurde, sind die Antreffzeiten aus den Fahrplandaten nur Näherungen, da die Ab- und Ankunftszeiten aus empirischen Optimierungen von Seiten der Verkehrsbetriebe und –verbünden ermittelt wurden. Zudem unterliegen Fahrzeiten verkehrlichen Störungen. Die vermeintlich rechnerisch scharfen Antreffzeiten sind daher zu verunschärfen (unscharf engl. fuzzy → dengl. Fuzzyfizierung).

Für den Einsatz der PlanÖV-Klassifizierung im Online-Betrieb ist der oben beschriebene Rechenaufwand sowie Datenzugriffsaufwand zur Interpolation der Antreffzeit ungeeignet. Zudem ist zu erwarten, dass häufig wiederkehrenden Berechnungen durchzuführen sind. Zur Effizienzsteigerung werden daher die PlanÖV-Fahrzeugpositionen sowohl räumlich als auch zeitlich vorprozessiert.

4.8.3.1 Diskretisierung der Fahrwege

Zur Optimierung für den Online-Betrieb wurde ein Verfahren entwickelt, das die Fahrwege in Kacheln unterteilt. Wie in Abschnitt 4.3 erläutert, erhalten alle Kartenelemente innerhalb des Kachel K_i dieselben Koordinaten n und m mit K_i (n,m). Die Größe der Kachel definiert die Auflösung der Karte und damit die erforderliche Anzahl der Kacheln. Die geometrische Genauigkeit wird zwar auf diese Weise reduziert, jedoch wird die Zugriffsgeschwindigkeit erhöht.

Ermittlung der Ein- und Ausfahrpunkte

Die zuvor durch Roadelementketten definierten Linienwege werden nun durch Ketten von Kacheln repräsentiert. Erforderlich dafür ist die Ermittlung der Kacheln, bzw. Kachelkoordinaten. Zusätzlich sind die Koordinaten der jeweiligen Ein- und Ausfahrpunkte zu berechnen, um für die Prognose der Fahrzeugposition den Ein- und Ausfahrzeitpunkt zu ermitteln.

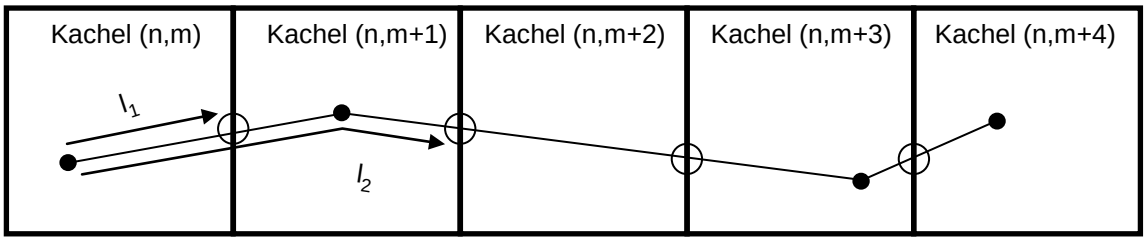


Abbildung 4-33: Diskretisierung der Linienwege

In Abbildung 4-33 ist ein Beispiel für die Verkachelung eines Linienweges dargestellt. Die schwarzen Punkte entsprechen Verkehrsknoten, Konturpunkten oder Haltepunkten. Die Kreise signalisieren die Schnittpunkte des Linienweges mit den Kachelgrenzen. Wie bei der sortierten Liste der befahrenen Kanten, siehe Abschnitt 4.8.2.4, werden die kompletten geometrischen Verläufe der

Fahrwege, nun jedoch als Kachelketten gespeichert. Das Datenmodell wird erweitert um die Elemente in Tabelle 6.

Tabelle 6: Erläuterungen zum Datenmodell für die Zwischenspeicherung der „scharfen“ Linienwege in der Kacheldarstellung

Kurzname	**Type**	**Beschreibung**
Tabelle: tmp_Kachel_OeV_Trajektorien_Geometrie		
Element	Datentyp	Beschreibung
bisherige Elemente siehe Abbildung 4-31)	…	…
LN_OeV_Point_Type	Integer	Typ des gespeicherten Punktes (bei Kombinationen werden Werte addiert) 1 – Verkehrsnetzknoten 10 – Geometriepunkt 100 – Haltepunkt 1000 – Schnittpunkt Route mit Kachel
LN_OeV_Point_Name	Text	Nur für Haltepunkte: der Name des Haltepunktes, sonst NULL
Siehe oben …	…	…
LN_Kachel_N	Integer	Kachelindex in nordweisender Richtung
LN_Kachel_E	Integer	Kachelindex in ostweisender Richtung

Ermittlung des Start- und Endzeitpunktes der Kachelgrenzen

Unter Verwendung der in Abschnitt 4.8.2.5 beschriebenen Prognose der Antreffzeiten der PlanÖV-Fahrzeuge zwischen den Haltepunkten, wird der Zeitpunkt für das Einfahren in eine Kachel und das Ausfahren aus einer Kachel berechnet. Diese Zeitpunkte können durch Interpolation zwischen den Knoten ermittelt werden. Da zumeist kein Geschwindigkeitsprofil auf den Kanten verfügbar ist, wird eine lineare Interpolation durchgeführt. Beschleunigungsvorgänge an Verkehrknoten und bei der An- sowie Abfahrt an Haltepunkten werden nicht berücksichtigt.

Die Interpolation der Fahrzeiten je Kachel ergibt einen Zeitpunkt für die Einfahrt in eine Kachel „LN_Ankunft“ und Ausfahrt aus der Kachel „LN_Abfahrt“. Die beiden Felder werden dem oben genannten Datenmodell hinzugefügt.

Tabelle 7: Erläuterungen zur Erweiterung des Datenmodells von Tabelle 6, um die Zeiten für die Kachelein- und Kachelausfahrt

Kurzname	**Type**	**Beschreibung**
Tabelle: tmp_Kachel_OeV_Trajektorien_Geometrie		
bisherige Elemente siehe Tabelle 6	…	…
LN_Ankunft	Double	Einfahrt in die Kachel
LN_Abfahrt	Double	Ausfahrt aus der kachel

Für die weitere Verarbeitung im Kachelkoordinatensystem sind ausschließlich die Kachelkoordinaten sowie die Zeitpunkte für die Ein- und Ausfahrt erforderlich. Der neue temporäre Datensatz basiert auf dem, in Abbildung 4-34 dargestellten, Datenmodell:

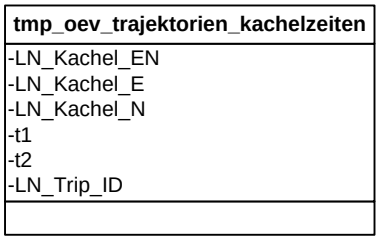


Abbildung 4-34: Datenmodell für die Zwischenspeicherung der "scharfen" Kachelein- und Kachelausfahrten

Die Einfahrzeit wird mit t1 und die Ausfahrzeit mit t2 bezeichnet. Zusätzlich ist die Bezeichnung des ÖV-Fahrzeugs mit LN_Trip_ID (ServiceTrip) als Fahrzeugkennung, LN_Trip_ID angegeben.

4.8.3.2 Fuzzifizierung der Fahrplanzeiten

Die eingangs genannten Störeinflüsse aufgrund von Fahrplanungenauigkeiten oder verkehrlichen Einflüssen des PlanÖV-Fahrzeugs verursachen eine zeitliche Unschärfe, da die Fahrgeschwindigkeiten und somit auch die Fahrzeiten beeinflusst werden. Voraussetzung für die folgende Betrachtung ist, dass keine Umleitungsrouten sondern stets auf dem geplanten Fahrweg des PlanÖV gefahren wird. Räumliche Unschärfen werden daher vernachlässigt.

Zur funktionalen Beschreibung der Auswirkungen von Störeinflüssen und Ungenauigkeiten müssen die funktionalen Zusammenhänge der einzelnen Einflussgrößen auf die Fahrzeit bekannt sein. In der Regel ist dieses jedoch nicht der Fall. Auf Basis von empirischen Boabachtungen und Expertenwissen können typische Bewegungsverhalten beschrieben und näherungsweise gemessen werden. In Do-iT wurden aus Versuchsfahrten in verschiedenen Verkehrsmit-

teln Näherungswerte für Verfrühungen und Verspätungen bestimmt und mittels Expertenwissen Verallgemeinerungen erarbeitet.

Empirische Beobachtungen der Fahrplanabweichungen

Aus Versuchsfahrten in Linienbussen im Stuttgarter Raum (Luo, 2007) zeigt sich an Haltestellen folgende Verteilung der Fahrzeitabweichungen:

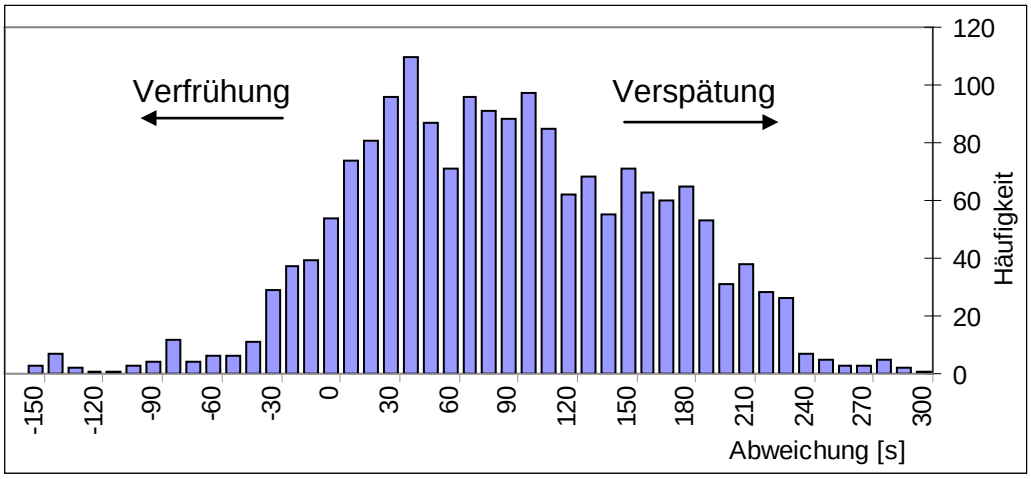


Abbildung 4-35: Histogramm der Fahrplanabweichungen aus Versuchfahrten in Linienbussen im Raum Stuttgart (Studienarbeit Dan Luo, 2007)

Signifikante tageszeitliche Abhängigkeiten konnten nicht gefunden werden, was zeigt, dass die Fahrpläne bereits gute Näherungen des Verkehrsverhaltens von typischen Schwachlastzeiten und Zeiten mit hohem Verkehrsaufkommen enthalten. Die durchschnittliche Fahrplanabweichung lag bei 87 Sekunden Verspätung.

Expertenwissen und funktionale Modellierung von Fahrplanabweichungen

Auch wenn bereits der Fahrgastwechsel an der Haltestelle abgeschlossen ist, sind die Fahrzeugführer zumeist angewiesen, die Haltestelle nicht früher als es die Abfahrtzeit im Fahrplan angibt, zu verlassen. Zu sehen ist dieses Verhalten auch an dem oben gezeigten Histogramm in Abbildung 4-35. Die Verteilung der Fahrplanabweichungen ist deutlich stärker in Richtung Verspätung (Abweichungen mit positivem Vorzeichen) als in Richtung Verfrühung (Abweichungen mit negativem Vorzeichen). Betrachtet man den Verlauf der Fahrplanabweichungen während der Fahrten, zeigt sich, dass häufig an einer Haltestelle eine Verspätung vorliegt, jedoch bereits an den folgenden Haltestellen diese wieder durch den Fahrzeugführer korrigiert werden kann.

Für eine Verallgemeinerung der Fahrplanabweichungen wird ein vereinfachtes Modell verwendet, vgl. Abbildung 4-36, das aus je einem linearen Bestandteil

für die Beschreibung möglicher Verfrühungen zwischen den Zeitpunkten t_0 und t_{Plan} sowie möglicher Verspätungen zwischen den Zeitpunkten t_{Plan} und t_3 besteht.

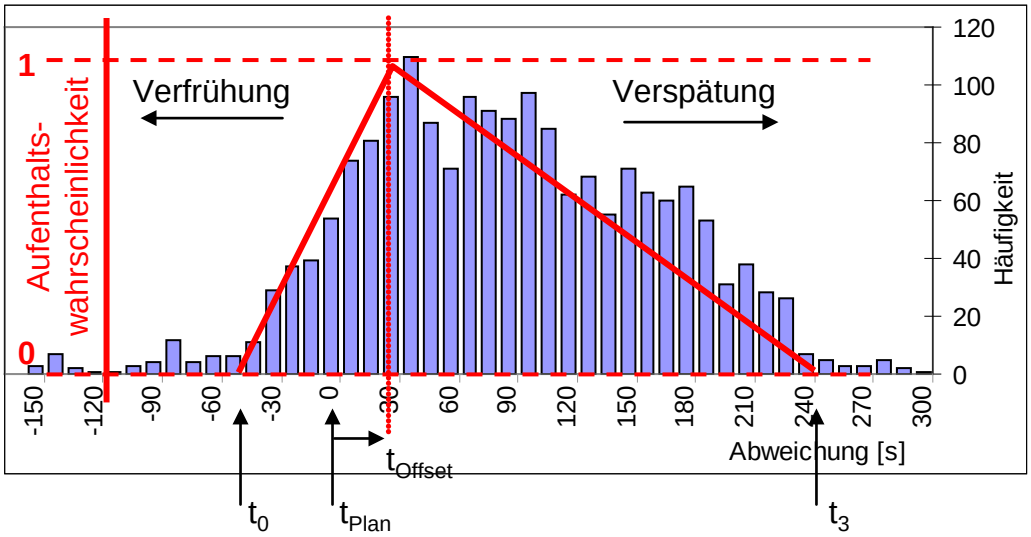


Abbildung 4-36: Verallgemeinerung der Fahrplanabweichungen mit linearen Bestandteilen

Dieses Modell der Fuzzyfierung der Fahrplanzeit wird verwendet, um die Antreffwahrscheinlichkeit eines PlanÖV-Fahrzeugs an einem beliebigen Punkt auf dem Fahrweg abzuleiten. Zum Zeitpunkt der angegebenen Fahrplanzeit t_{Plan} + t_{Offset} wird dazu interpretiert, dass sich das PlanÖV-Fahrzeug mit maximaler Antreffwahrscheinlichkeit an der entsprechenden Haltestelle aufhält. Je früher der zu untersuchende Zeitpunkt ist, also je weiter der Zeitpunkt in Richtung maximaler Verfrühung t_0 ist, desto unwahrscheinlicher ist die Antreffwahrscheinlichkeit des PlanÖV-Fahrzeugs an dieser Haltestelle. Ebenso wird interpretiert, dass die Antreffwahrscheinlichkeit sinkt, je weiter der zu untersuchende Zeitpunkt in Richtung maximaler Verspätung t_3 liegt. Die maximale Antreffwahrscheinlichkeit wird definiert mit 1 und die minimale Antreffwahrscheinlichkeit, außerhalb der Zeitpunkte t_0 und $t_{3,}$ wird definiert mit 0.

Anpassung der unscharfen Fahrplanzeiten an die Diskretisierung der Fahrwege

In Abschnitt 4.8.3.1 wurde die kachelweise Darstellung der Fahrwege erläutert. Verwendet man das zeitliche Antreffen der PlanÖV-Fahrzeuge auf diese Darstellung an, so sind die Ein- und Ausfahrzeitpunkte in der jeweiligen Kachel von Interesse. In der Zeit zwischen Ein- und Ausfahrt kann angenommen werden, dass sich das PlanÖV-Fahrzeug vollständig innerhalb der Kachel befindet. Die Antreffwahrscheinlichkeit des PlanÖV-Fahrzeug in dieser Kachel ist zu diesem Zeitpunkt somit maximal.

Die Ein- und Ausfahrzeitpunkte t_1 und t_2 je Kachel wurden, wie in Abschnitt 4.8.3.1 erläutert, bereits aus Fahrplanzeiten berechnet. Zur Berücksichtigung von Verfrühungen kann nun der Zeitpunkt t_1 als Fahrplanzeit t_{Plan} + t_{Offset} betrachtet und entsprechend t_0 ausgehend von t_1 ermittelt werden. Ebenso können Verspätungen ausgehend von der Fahrplanzeit t_2 + t_{Offset} mit t_3 berücksichtig werden. Das gesamte Zugehörigkeitsdiagramm zeigt Abbildung 4-37 mit der maximalen Antreffwahrscheinlichkeit des PlanÖV-Fahrzeugs in der Kachel als schraffierte Fläche.

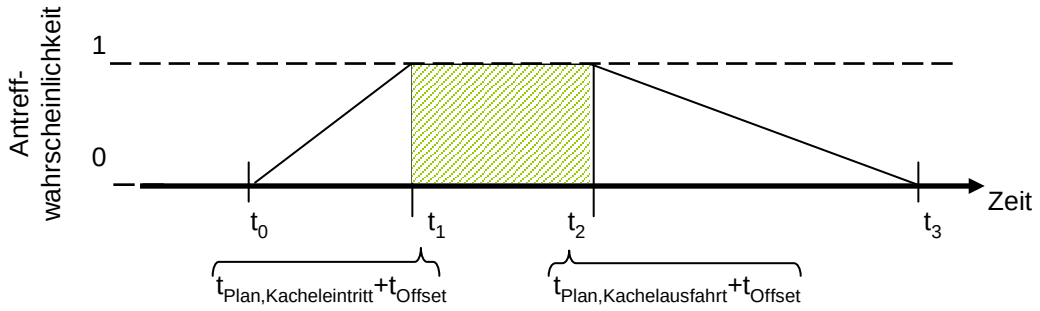


Abbildung 4-37: Zugehörigkeitsdiagramm für die Antreffwahrscheinlichkeit eines PlanÖV-Fahrzeugs in einer Kachel mit Berücksichtigung von Verfrühungen und Verspätungen

4.8.4 Vorprozessierung für die FPD-Echtzeitanalyse

Transportangebote in Öffentlichen Verkehren sind häufig in Linien organisiert, die meist durch mehrere Fahrten am Tag realisiert werden. Im Allgemeinen finden diese Fahrten auch auf gleichen Fahrwegen in Hin- und Rückrichtung statt. Kacheln die durch Öffentliche Verkehre befahren werden, werden somit häufig, zum Teil mehrfach pro Tag von PlanÖV-Fahrzeugen aufgesucht. Zur Effizienzsteigerung der Berechnung, mit welcher Antreffwahrscheinlichkeit zu einem spezifischen Zeitpunkt ein PlanÖV-Fahrzeug in einer zu analysierenden Kachel ist, können die Antreffwahrscheinlichkeiten aller PlanÖV-Fahrzeuge einer Kachel im Zeitbereich aneinandergekettet werden. Abbildung 4-38 zeigt schematisch den Grundgedanken.

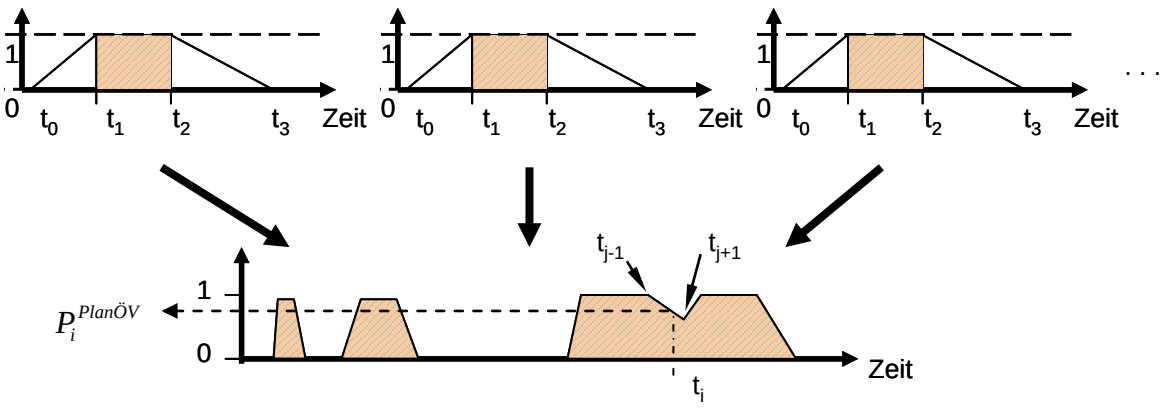


Abbildung 4-38: Zusammenführung der Antreffwahrscheinlichkeiten aller PlanÖV-Fahrzeuge einer Kachel

Mit der Angabe der Zeit t_i einer Kachel für eine FPD-Position kann nun durch Suche des nächst größeren Zeitpunktes t_{j+1} und des nächst kleinen Zeitpunktes t_{j-1} die Antreffwahrscheinlichkeit eines PlanÖV-Fahrzeuges interpoliert werden.

Für die Anforderungen im Projekt Do-iT ist es nicht erforderlich zu ermitteln, um welches PlanÖV-Fahrzeug es sich handelt. Vielmehr ist die Tatsache und der Wahrscheinlichkeitswert von Interesse, dass ein oder mehrere PlanÖV-Fahrzeuge zur selben Zeit in derselben Kachel wie der FPD-Teilnehmer vorhanden sind. Trifft dieser Sachverhalt zu, kann das FPD als passiv eingestuft und von der weiteren Analyse ausgeschlossen werden.

4.8.5 Datenmodellierung für Ganzjahresfahrpläne

Vorwiegend werden die Fahrplandaten von den Verkehrsbetrieben für einen Gültigkeitszeitraum von ca. einem Jahr (Deutsche Bahn AG ½ jährlich) bereitgestellt. Um den Datenumfang während der PlanÖV-Klassifizierung zu begrenzen, werden die vorprozessierten Daten in Blöcke unterteilt. Die Blockgröße ist auch abhängig von der Menge der befahrenen Kacheln und der Häufigkeit der Menge der PlanÖV-Fahrten. Für eine übersichtliche Verwaltung werden in Do-iT je nach ÖV-Angebot Blockgrößen von einer Woche bis zu einem Monat gewählt. Die Datenblöcke werden in Do-iT nach Verkehrsmitteln und Regionen, wie in Tabelle 8 gezeigt, gegliedert.

Tabelle 8: Gliederung der PlanÖV-Datenblöcke nach Zeitraum, Verkehrsmittel und Region im Projekt Do-iT

Zeitraum	**Verkehrsmittel**		**Region**
start_ / end_validity_time	*classification_type*	*transport_vehicle*	*region*
1 Woche bis 1 Monat	1	Linienbus	Stuttgart
	2	Straßenbahn	Karlsruhe
	3	Regionalbahn (RB,RE)	
	4	Fernbahn (IR,IC,ICE)	

Für jede Kombination aus Zeitraum, Verkehrsmittel und Region wird eine eigene Datentabelle *classification_xxxx* erzeugt und ein Eintrag in die Metadaten-Tabelle *oev_classification_meta* vorgenommen. Die Bezeichnung xxxx steht für einen fortlaufenden Index, der während der Datenblockprozessierung vergeben wird. In der Abbildung 4-39 ist das Datenmodell für die Modellierung der Metadaten-Tabelle und der für die effiziente Suche der zu analysierenden Kachel sowie der Zeiten der Antreffwahrscheinlichkeiten vorprozessierten PlanÖV-Datenblöcke dargestellt. In der Tabelle 9 und der Tabelle 10 sind die entsprechenden Erläuterungen zum Datenmodell zusammengefasst. Das Kürzel pt steht für public transport (Öffentlicher Verkehr).

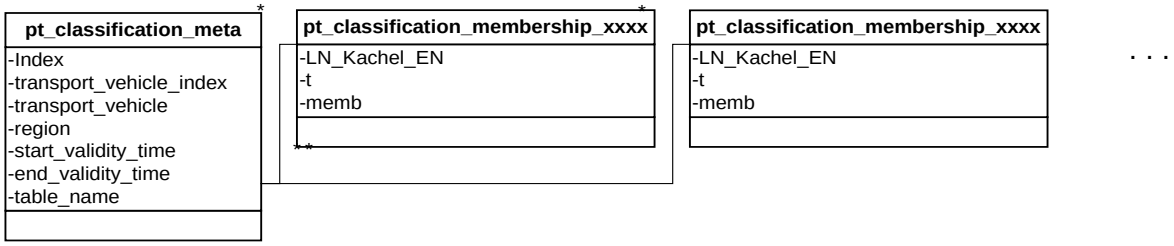


Abbildung 4-39: Datenmodell der vorprozessierten Datenblöcke für die PlanÖV-Klassifizierung

Tabelle 9: Erläuterungen zum Datenmodell der Meta-Daten für die Datenblöcke der PlanÖV Klassifikation

Kurzname	**Type**	**Beschreibung**
Tabelle: pt_classification_meta		
Element	Datentyp	Beschreibung
index	Integer	Fortlaufender Zähler
transport_vehicle_index	Integer	Numerische Kennzeichnung der Klassifizierungsdaten, mit: 1 - Linienbus 2 - Straßenbahn 3 – Regionalbahn (RB, RE) 4 – Fernbahn (IRE, IR, IC, ICE)
transport_vehicle	Varchar	Langname des *transport_vehicle_index* (z.B. Linienbus Stuttgart)
start_validity_time	Char	Anfangsdatum und -uhrzeit des Zeitraumes für die Gültigkeit der Daten (dd.mm.yyyy hh:mm:ss)
end_validity_time	Char	Enddatum und –uhrzeit des Zeitraumes für die Gültigkeit der Daten (dd.mm.yyyy hh:mm:ss)
table_name	Varchar	Name der Datentabelle

Tabelle 10: Erläuterungen zum Datenmodell der Datenblöcke der PlanÖV Klassifikation

Kurzname	**Type**	**Beschreibung**
Tabelle: pt_classification_xxxx		
Element	Datentyp	Beschreibung
LN_Kacheln_EN	Integer	Kachelnummer E * 100 000 + N = EN, mit E und N als Kachelkoordinaten
t	Integer(32)	Zeit in zehntel Sekunden ab start_validity_time für den fuzzyfizierten Zugehörigkeitswert memb (t=t/10+start_validity_time)
memb	Integer(8)	Zugehörigkeitswert der Kachel LN_Kacheln_EN zum Zeit t als Zahl zwischen 0 – 255, wobei 0 = 0% und 255 = 100% entsprechen (memb=memb/255)

4.8.6 PlanÖV-Klassifizierung der FPD

Ziel der PlanÖV-Klassifizierung ist die Bestimmung eines PlanÖV-Wahrscheinlichkeitswertes für eine FPD-Positionsfolge. Dazu sind die Antreffwahrscheinlichkeiten der PlanÖV-Fahrzeugen in den jeweiligen Kacheln der FPD-Positionen zu berechnen. Für eine FPD-Positionsfolge können somit Wahrscheinlichkeitswertverläufe (schematische Darstellungen zeigt Abbildung 4-40) für die untersuchten PlanÖV-Angebote ermittelt werden.

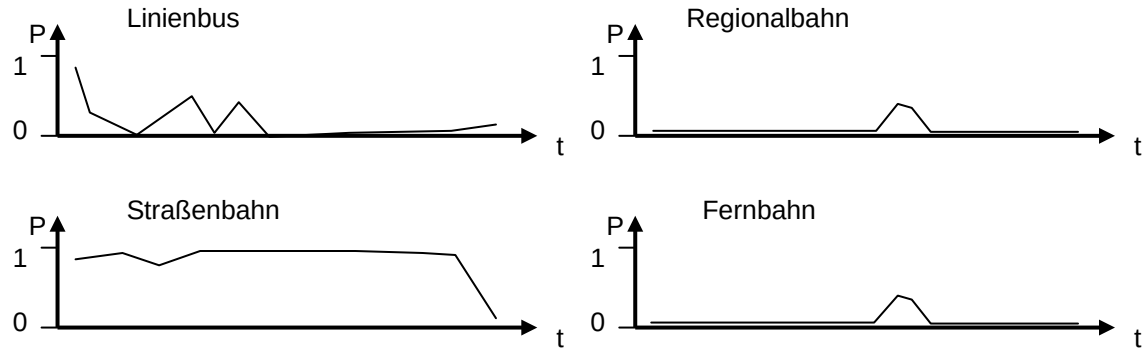


Abbildung 4-40: Wahrscheinlichkeitsverläufe für die Partizipierung von FPD-Positionen mit PlanÖV-Fahrzeugen (schematisierte Darstellung)

4.8.6.1 Bestimmung der PlanÖV-Wahrscheinlichkeit je Kachel

Wie bereits erläutert, werden die FPD-Positionen einschließlich ihrer räumlichen, genauigkeitsbedingten Verteilung in die Kachel-Darstellung umgewandelt. Je FPD-Position existiert somit eine Menge von Kacheln mit Angabe der Aufenthaltswahrscheinlichkeit. Zu jeder Kachel ist zu ermitteln, wie hoch die Antreffwahrscheinlichkeit eines PlanÖV-Fahrzeuges ist. Da weiterhin angenommen werden kann, dass sich FPD auf Verkehrsnetzen bewegen, werden für die fol-

genden Analysen nur die Kacheln verwendet, die mindestens einem Verkehrsnetz angehören.

Aus den vorprozessierten Datenblöcken, siehe Abschnitt 4.8.5, werden mit Angabe der Kachelkoordinate und der Zeit die zeitlich nahe liegenden Antreffwahrscheinlichkeitswerte gelesen. Existiert die Kachelkoordinate nicht oder der gesuchte Zeitpunkt ist außerhalb der abgelegten Zeiten, liegt eine Wahrscheinlichkeit gleich 0 vor. Konnten sowohl die Kachelkoordinate als auch die nahe liegenden Antreffwahrscheinlichkeitswerte, in Abbildung 4-41 mit den Punkten E und F bezeichnet, gefunden werden, kann für die FPD-Positionszeit der Antreffwahrscheinlichkeitswert interpoliert werden.

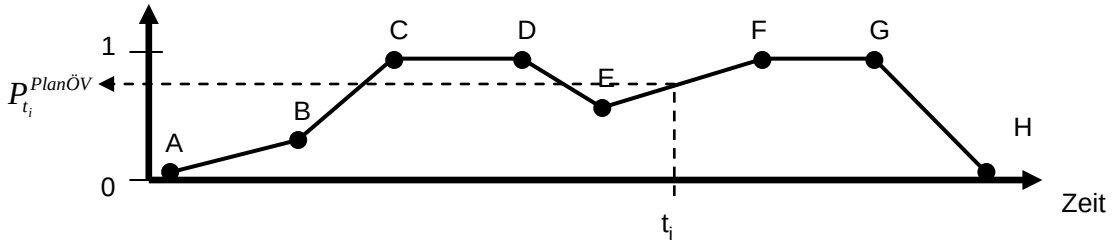


Abbildung 4-41: Interpolation des ÖV-Zugehörigkeitswertes

Die Wahrscheinlichkeit $P_{Kachel,t}^{PlanÖV}$ für Nutzung des PlanÖV in der untersuchten Kachel zum Zeitpunkt t ergibt sich mittels Formel (3-20) aus der Multiplikation interpolierte PlanÖV-Antreffwahrscheinlichkeit $P_{Kachel,t_i}^{PlanÖV}$ mit der Aufenthaltswahrscheinlichkeiten P_{Kachel_t,t_i}^{FPD} des FPD in der Kachel.

$$P_{Kachel,t}^{PlanÖV} = P_{Kachel,t_i}^{PlanÖV} \cdot P_{Kachel_t,t_i}^{FPD} \tag{3-20}$$

4.8.6.2 Bestimmung der ÖV-Zugehörigkeit je Position

Für die Bestimmung der Wahrscheinlichkeit eines PlanÖV je FPD-Position $P_{PlanÖV}^{FPD,Position}$ ist zuerst die Summe aller Wahrscheinlichkeiten $P_{PlanÖV}^{FPD,Kachel}$, getrennt nach PlanÖV-Angebot zu ermitteln. Dieser Wert wird in [Sachs] als die „günstigen Fälle“ bezeichnet und wird mit Formel (3-21) berechnet.

$$\begin{aligned} &\text{"günstige Fälle"}_{PlanÖV}^{FPD,Position} = \sum_{j}^{n} P_{PlanÖV}^{FPD,Kachel_j} \\ &\text{mit} \quad j = 1,2...n \quad \text{Kachelindex der Position} \\ &\qquad n \qquad\quad \text{Anzahl der Kacheln je Position} \end{aligned} \tag{3-21}$$

Der maximale Wert existiert, wenn jede Kachel einen Wahrscheinlichkeitswert von 100% für die Aufenthaltswahrscheinlichkeit sowie 100% für die PlanÖV-Wahrscheinlichkeit hätte. Dieser Wert wird in (Sachs, 2004) als die „möglichen

Fällen“ bezeichnet, und berechnet sich mit Formel (3-22) aus der Anzahl n der Kacheln der FPD-Position.

$$\begin{aligned} &\text{"mögliche Fälle"}_{PlanÖV}^{FPD,Position} = n \cdot P\left(\text{maximale } P_{Kachel}^{FPD}\right) \cdot P\left(\text{maximale } P_{PlanÖV}\right) \\ &\text{mit} \quad n \qquad\qquad \text{Anzahl der Kacheln der Position} \\ &\qquad P\left(\text{maximale } P_{Kachel}^{FPD}\right) \;=\; 100\% \\ &\qquad P\left(\text{maximale } P_{PlanÖV}\right) \;=\; 100\% \end{aligned} \tag{3-22}$$

Die Wahrscheinlichkeit dass ein PlanÖV-Angebot in der Position genutzt wurde, errechnet sich mit Formel (3-23) aus dem Verhältnis der „möglichen“ PlanÖV zu den „günstigen“ PlanÖV Fällen, getrennt nach PlanÖV-Angeboten:

$$P_{PlanÖV}^{FPD,Position} = \frac{\text{"günstige Fälle"}_{PlanÖV}^{FPD,Position}}{\text{"mögliche Fälle"}_{PlanÖV}^{FPD,Position}} \tag{3-23}$$

4.8.6.3 Bestimmung der ÖV-Zugehörigkeit je FPD-Positionsfolge

Die Wahrscheinlichkeit der Nutzung eines PlanÖV während der gesamten FPD-Positionsfolge, ergibt sich analog zum vorherigen Abschnitt aus dem Verhältnis der „möglichen Fälle“, zu den ermittelten, also den „günstigen Fällen“. Der maximal mögliche Wert eines PlanÖV errechnet sich aus der Anzahl der FPD-Positionen. Hierbei wird wieder angenommen, dass maximal 100% PlanÖV Wahrscheinlichkeit für die Positionen vorliegen würden.

$$\begin{aligned} &\text{"mögliche Fälle"}_{PlanÖV}^{FPD,Positionsfolge} = n \cdot P\left(\text{maximale } P_{PlanÖV}\right) \\ &\text{mit} \quad n \qquad\qquad \text{Anzahl der Positionen} \\ &\qquad P\left(\text{maximale } P_{PlanÖV}\right) \;=\; 100\% \end{aligned} \tag{3-24}$$

Der Wert für die „günstigen Fälle“ ergibt sich aus der Summe der PlanÖV-Wahrscheinlichkeiten aller Positionen.

$$\begin{aligned} &\text{"günstige Fälle"}_{PlanÖV}^{FPD,Positionsfolge} = \sum_{i}^{m} P_{PlanÖV}^{FPD,Position_i} \\ &\text{mit} \quad i = 1,2...m \quad \text{Positionsindex der Positionsfolge} \\ &\qquad m \qquad\qquad \text{Anzahl der Positionen in der Positionsfolge} \end{aligned} \tag{3-25}$$

Die Wahrscheinlichkeit dass ein PlanÖV-Angebot in der FPD-Positionsfolge gewählt wurde, wird hier bezeichnet als $P_{PlanÖV}^{FPD,Positionsfolge}$ und wird mit Formel (3-27) berechnet.

$$P_{PlanÖV,absolut}^{FPD,Positionsfolge} = \frac{\text{"günstige Fälle"}_{PlanÖV}^{FPD,Positionsfolge}}{\text{"mögliche Fälle"}_{PlanÖV}^{FPD,Positionsfolge}} \tag{3-26}$$

Für die Identifizierung der FPD mit PlanÖV Nutzung kann ein Schwellwert $P_{Schwellwert}$ für die gewünschte Sicherheitswahrscheinlichkeit definiert werden.

Wird für eine FPD-Positionsfolge mindestens dieser Schwellwert erreicht kann der MFT als Nutzer eines PlanÖV identifiziert werden. Im Online-Betrieb von Do-iT wird ein Wert von 80% eingesetzt. Wird mindestens dieser Wert erreicht, gilt das FPD als passiv und wird von der weiteren Identifizierung ausgeschlossen.

4.9 Verkehrsnetzgestützte Verkehrsteilnehmerklassifizierung

Verkehrsteilnehmer bewegen sich zumeist auf Verkehrswegen wie Straßen, Schienenwegen oder Fußwegen. Durch eine Analyse, welche Verkehrsnetze vom Mobilfunkteilnehmer genutzt wurden, kann wiederum auf die zugehörige Verkehrsteilnehmerklasse geschlossen werden. Da gleiche Verkehrsnetze von unterschiedlichen Verkehrsteilnehmerklassen verwendet werden können (z.B. Straßen von PKW, LKW, Linienbussen), sind aus Sicht der Verkehrsnetznutzung nur Gruppen von Verkehrsteilnehmerklassen unterscheidbar. Eine Gruppenbildung für Verkehrsteilnehmerklassen ist z.B. bei PKW und LKW erforderlich, da Straßen prinzipiell von beiden gleichwertig genutzt werden können. Zusammengefasste Verkehrsklassen werden folgend als Verkehrsklassengruppe bezeichnet.

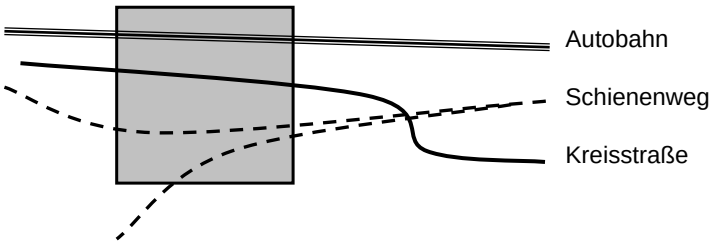


Abbildung 4-42: Verkehrsnetze in einer Kachel

Ziel der Verkehrsnetzgestützten Analyse ist es Wahrscheinlichkeitswerte für Verkehrsklassengruppen auf Basis der genutzten Verkehrsnetze zu ermitteln, um unwahrscheinliche Verkehrsklassen von den weiteren Analysen ausschließen zu können. Die Wahrscheinlichkeitswerte der Verkehrsklassengruppen je FPD-Position werden abschließend bewertet und je Verkehrsklasse ein Wahrscheinlichkeitswert berechnet.

4.9.1 Definition typischer Verkehrsnetze von Verkehrsteilnehmern

Datengrundlage der Verkehrsnetzanalyse sind die Verkehrsnetzelemente des Lokalisierungsnetzes. Darin enthalten sind Straßen, Schienenwege und Fußwege. Das Verkehrsnetzelement ist die topologische Kante zwischen zwei topologischen Knoten. Um die Analyse im Echtzeitbetrieb rechenzeitoptimiert zu ermöglichen, sind diese Daten für die späteren Datenabfragen aufzubereiten und zugriffsoptimiert zu speichern. Die Verkehrsnetz werden dazu nach dem in Abschnitt 4.3 beschriebene Verfahren verkachelt, siehe Abbildung 4-42. Für jede Kachel werden die Wahrscheinlichkeiten der jeweiligen Verkehrsklassegruppe schließlich vorprozessiert in einer zugriffsoptimierten Datenbank abgelegt. Wie einleitend erläutert, sind vorerst jedoch die zu untersuchenden Verkehrsklassen zu definieren.

4.9.1.1 Definition von Verkehrsklassengruppen und Verkehrsnetze

Die Definition der Verkehrsklassen orientiert sich an den Verkehrsteilnehmerklassen, die durch die Nutzung der verschiedene Verkehrsnetze bzw. Verkehrsnetzelemente unterschieden werden können. Für die Definition ist zu berücksichtigen, dass aufgrund der tatsächlichen Nutzung und der hier vorliegenden Verkehrsnetze folgende Abhängigkeiten bestehen können:

1. ein Verkehrsnetzelement kann mehreren Verkehrsnetzen angehören (z.B. eine Straße kann auch als Fußweg gewidmet sein),
2. ein Verkehrsnetz kann mehreren Verkehrsklassen angehören (z.B. eine Straße kann von PKW und LKW genutzt werden).

Da Verkehrsnetzelemente Bestandteil mehrerer Verkehrklassen einer Verkehrsnetzgruppe sein können (Fall 1), wird definiert:

Ein Verkehrnetzelement kann maximal einmal
einer Verkehrsklassengruppe zugeordnet werden.

Aus räumlicher Sicht ist dies so zu begründen: in einer Kachel sollen einzelne Verkehrsklassengruppen keine höhere Wahrscheinlichkeit erhalten, nur weil sie mehrere Verkehrsnetze nutzen können, wenn die Verkehrsklassen jedoch alle durch ein einziges Verkehrsnetzelement repräsentiert werden.

Die Zuordnung der Verkehrsnetze zu den unterscheidbaren Verkehrsklassen bzw. Verkehrsklassengruppen ist mittels Expertenwissen festzulegen. Auf Basis der im Lokalisierungsnetz vorhandenen Verkehrsnetze wurden folgende Festlegungen getroffen (Tabelle 11).

Tabelle 11: Definition der Verkehrsklassen/-gruppen für die Verkehrsnetzklassifizierung

Nr.	Name der Verkehrs-netzgruppe	Beschreibung	
		Verkehrsnetz	Verkehrsklasse(n)/-gruppe
0	other	Sonstige Verkehrsnetze	keine zugeordnete Verkehrsklasse
1	pedestrian	Verkehrsnetze für Fußgänger	Fußgänger; Straßenkategorien der Navteq-Verkehrsnetze
2	car_truck	Fahrzeuge des Individualverkehrs	PKW, LKW, Bus; Straßenkategorien der Navteq-Verkehrsnetze
3	tram	Verkehrsnetz der ÖV Straßenbahnen	Schienennetze der Straßenbahnen Stuttgart und Karlsruhe
4	long_dist_rail	Schiene der Fern- und Regionalbahnen	Regional und Fernzüge (DB)

Zur Bestimmung der zu den Verkehrsklassengruppen zugehörigen Verkehrsnetze ist festzulegen, mit welcher Logik die Verkehrsnetze die entsprechenden

Verkehrsklassen identifizieren. Als Logik werden die Operatoren GLEICH sowie ODER verwendet.

- Der Operator GLEICH wird verwendet, wenn die Verkehrsklasse ausschließlich in der genannten Kombination durch das bzw. die Verkehrsnetze definiert ist.
- Der Operator ODER wird verwendet, wenn ein Verkehrsnetzelement zumindest Bestandteil des genannten Verkehrsnetzes sein muss. Das Verkehrsnetzelement kann aber auch Bestandteil eines anderen Verkehrsnetzes sein.

Unter Verwendung der Vergleichoperatoren werden anschließend den Verkehrsklassengruppen die Verkehrsnetze zugeordnet. In Tabelle 12 sind die im Lokalisierungsnetz vorhanden Verkehrnetze aufgeführt.

Tabelle 12: Verkehrsklassen des Lokalisierungsnetzes

Werte: LN_Form_Of_Way	
Wert (2^x) x=	**Beschreibung (Nutzungsart)**
0	Nicht definiert
1	Fußgänger
2	PKW
3	LKW
4	Bus
5	Schienen-Fernverkehr (GDF-Daten)
6	Schienen-Fernverkehr (DB – VUV-Daten)
7	Schienen-Straßenbahnen (VUV-Daten)

In der Tabelle 12 sind die entsprechenden Verknüpfungen der Verkehrsklassengruppen, wie sie im Projekt Do-iT eingesetzt werden, aufgeführt. Die Verkehrsklasse „other“ wird vergeben, wenn ein Verkehrsnetzelement ohne eine Verknüpfung zu einer Verkehrsklassengruppe gefunden wird.

Tabelle 13: Verknüpfungen der Verkehrsnetze zu den Verkehrsklassengruppen

Verkehrsklasse	**Verknüpfung-/ Vergleichsoperator**	**Erläuterung**
Other (nicht definiert)		Zuordnung wenn kein Bestandteil einer anderen Verkehrsklasse
Pedestrian (Fußgänger)	GLEICH 1	Ausschließlich Kartenelemente dieses Verkehrsnetzes
car_truck (PKW oder LKW)	2 ODER 3 ODER 4	Mindestens Bestandteil eines oder mehrer dieser Verkehrsnetze
Tram (Straßenbahn)	ODER 7	Mindestens Bestandteil dieses Verkehrsnetzes
Long_dist_rail (Fernzug)	5 ODER 6	Mindestens Bestandteil eines oder mehrerer dieser Verkehrsnetze

4.9.2 Datenvorprozessierung

4.9.2.1 Berechnungen der Verkehrsklassenwahrscheinlichkeit

Ausgangspunkt für die Berechnung der Verkehrsklassenwahrscheinlichkeit ist die Kacheldarstellung des Lokalisierungsnetzes (siehe Definition des Lokalisierungsnetzes). Die Wahrscheinlichkeit der Verkehrsklassengruppe in der Kachel lässt sich berechnen, durch die Länge der innerhalb der Kachel verlaufenden Verkehrsnetze. Die höchste Wahrscheinlichkeit liegt vor, wenn die gesamten Verkehrsnetzelemente einer einzelnen Verkehrsklassengruppe zuzuordnen sind. Für die Berechnung werden je Verkehrsklassengruppe die Längen der zugeordneten Verkehrsnetzelemente mit Formel (3-27) summiert.

$$L^{Verkehrsklassengruppe} = \sum_{i=0}^{n} L_i^{Verkehrsklassengruppe}$$

$$\text{mit} \quad i = 1,2,...n \qquad (3\text{-}27)$$

$$n \quad \text{Anzahl der zur Verkehrsklassengruppe gehörenden Verkehrsnetzelemente}$$

Die Summe aller Verkehrnetzelementlängen von allen Verkehrklassengruppen ergibt den maximal möglichen Anteil:

$$L^{LN}_{Verkehrsklassengruppen,\ gesamt} = \sum_{j=1}^{m} L^{LN}_{Verkehrsklassengruppe,\ j}$$

$$\text{mit} \quad j = 1,2,...m \qquad (3\text{-}28)$$

$$m \ \text{Anzahl Verkehrsklassengruppen}$$

Die Wahrscheinlichkeit für eine Verkehrsklassengruppe kann nun nach Formel (3-29) berechnet werden aus dem Verhältnis der Verkehrsnetzelementlänge je Verkehrsklassengruppe zu Gesamt Länge aller Verkehrsnetzelemente aller Verkehrsklassengruppen in der Kachel.

$$P^{LN}_{Verkehrsklassengruppe} = \frac{L^{LN}_{Verkehrsklassengruppe}}{L^{LN}_{Verkehrsklassengruppen,\ gesamt}} \qquad (3\text{-}29)$$

4.9.2.2 Vorprozessierung der Verkehrsklassengruppenwahrscheinlichkeit

Die Verkehrsnetzelemente und somit auch die Längen je Verkehrsklassengruppen können, solange keine Änderung der Verkehrsnetzelemente vorgenommen wird, als statisch angenommen werden. Die Berechnung der Verkehrklassengruppenwahrscheinlichkeit in der Kachel ist somit im Vorfeld möglich. Das Ergebnis ist eine Tabelle mit den Angaben der Verkehrklassengruppenwahrscheinlichkeit für die jeweilige Verkehrsklassengruppe in Prozent und der Verkehrselementlänge je Kachel. Das Datenmodell der Tabelle *trafficnetwork_probability* zeigt Abbildung 4-43.

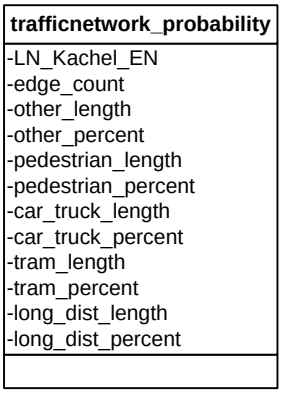


Abbildung 4-43: Datenmodell der vorprozessierten Tabelle zur Verkehrsnetzanalyse

Kurzname	**Type**	**Beschreibung**
Tabelle: trafficnetwork_probability		
Element	**Datentyp**	**Beschreibung**
LN_Kacheln_EN	Integer	Kachelnummer E * 100 000 + N = EN, mit E und N als Kachelkoordinaten
edge_count	Integer	Typ des gespeicherten Punktes (bei Kombinationen werden Werte addiert) 1 – Verkehrsnetzknoten 10 – Geometriepunkt 100 – Haltepunkt 1000 – Schnittpunkt Route mit Kachel
other_length	Double	Verkehrsnetzelementlänge der Verkehrsklassengruppe
other_percent	Double	Prozentualer Anteil der Verkehrsklassengruppe in der Kachel
pedestrian_length	Double	Verkehrsnetzelementlänge der Verkehrsklassengruppe
pedenstrian_percent	Double	Prozentualer Anteil der Verkehrsklassengruppe in der Kachel
car_truck_length	Double	Verkehrsnetzelementlänge der Verkehrsklassengruppe
car_truck _percent	Double	Prozentualer Anteil der Verkehrsklassengruppe in der Kachel
tram_length	Double	Verkehrsnetzelementlänge der Verkehrsklassengruppe
tram _percent	Double	Prozentualer Anteil der Verkehrsklassengruppe in der Kachel
long_dist_length	Double	Verkehrsnetzelementlänge der Verkehrsklassengruppe
long_dist_percent	Double	Prozentualer Anteil der Verkehrsklassengruppe in der Kachel

4.9.3 Bestimmung der FPD-Verkehrsklassengruppenwahrscheinlichkeit

Für die Bestimmung der Verkehrsnetzzugehörigkeit werden die Kacheln der FPD-Positionen einschließlich der Aufenthaltswahrscheinlichkeiten verwendet. Im ersten Schritt werden die Kacheln jeder Positionen einzeln analysiert und bewertet, wonach im zweiten Schritt die Ergebnisse für die gesamte FPD-Positionsfolge zusammengefasst und bewertet werden. Je FPD-Positionsfolgen wird somit für jede der oben definierten Verkehrsklassengruppe ein Wahrscheinlichkeitswert für die Nutzung der entsprechenden Verkehrsnetze berechnet.

Die Häufigkeit der Verkehrsnetzelemente der zu analysierenden Verkehrsnetze ist stark verschieden. Verkehrswege für PKW oder LKW sind bspw. deutlich häufiger als Schienenwege. In Abbildung 4-44 sind die Anteile der Verkehrsklassengruppen je Kachel mit einer räumlichen Verteilung dargestellt.

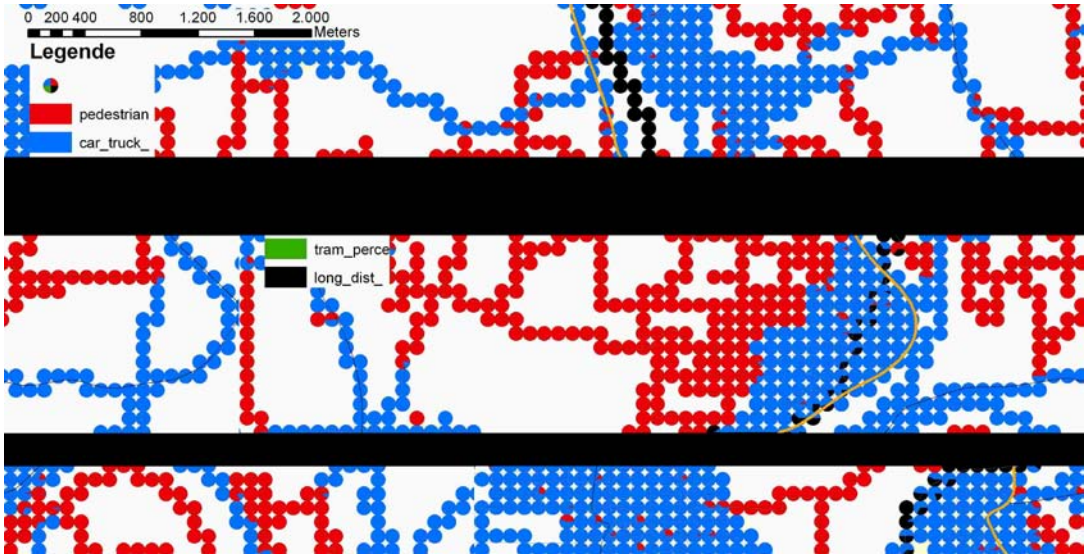


Abbildung 4-44: Verteilung der Verkehrsklassengruppen (Lokalisierungsnetz nahe Ludwigsburg)

Zu erkennen ist, dass Schienenwege (schwarze bzw. grüne Anteile der Kacheln) besonders innerorts von Straßen umgeben sind. Auch wenn FPD Schienenwege nutzen, ist der Wahrscheinlichkeit der Verkehrsklassengruppe PKW-LKW sehr hoch. Nach der Bestimmung der räumlich analysierten Verkehrsnetznutzung wird am Ende dieses Abschnittes ein Verfahren zur logischen Überprüfung der prognostizierten Nutzung vorgestellt.

4.9.3.1 Bestimmung der Verkehrsnetzzugehörigkeit je Kachel

In der Vorprozessierung wurde gezeigt, wie die Verkehrsklassengruppenwahrscheinlichkeit je Kachel berechnet und in der Tabelle *trafficnetwork_probability* gespeichert wird. Je höher die Aufenthaltswahrscheinlichkeit in der Kachel und je höher die Verkehrsklassenwahrscheinlichkeit ist, desto wahrscheinlicher ist die Wahl der entsprechenden Verkehrsklasse. Um die FPD-Verkehrsklassengruppenwahrscheinlichkeit der jeweiligen Kachel zu bestimmen, werden die Aufenthaltswahrscheinlichkeiten der zur FPD-Position gehörenden Kachel multipliziert mit den vorprozessierten Verkehrsklassengruppenwahrscheinlichkeiten:

$$P^{FPD-Kachel}_{\text{Verkehrsklassegruppe}} = P^{LN}_{Verkehrsklassengruppe} \cdot P^{FPD-Kachel}$$

mit $P^{FPD-Kachel}$ Aufenthaltswahrscheinlichkeit des FPD

$P^{LN}_{Verkehrsklassengruppe}$ vorprozessierte Verkehrsklassengruppen-wahrscheinlichkeit

$P^{FPD-Kachel}_{Verkehrsklassegruppe}$ Verkehrsklassengruppenwahrscheinlichkeit in der Kachel (3-30)

Kacheln, die keinen Anteil an einem Verkehrsnetz haben, werden nicht in die Berechnung einbezogen.

4.9.3.2 Bestimmung der Verkehrsklassengruppenwahrscheinlichkeit je FPD-Position

Für die Berechnung der Verkehrsklassengruppenwahrscheinlichkeit je FPD-Position werden die Ergebnisse aus Abschnitt 4.9.3.1 zusammengefasst. Unter der Annahme, dass die maximal mögliche Verkehrsklassengruppenwahrscheinlichkeit dann erreicht wird, wenn die maximal mögliche Verkehrsklassengruppenwahrscheinlichkeit von 100% und die maximal mögliche FPD-Aufenthaltswahrscheinlichkeit von 100% vorliegt, kann die Verkehrsklassengruppenwahrscheinlichkeit je FPD-Position z.B. mit Formel (3-31) durch das arithmetische Mittel der Verkehrsklassengruppenwahrscheinlichkeit je FPD-Kachel berechnet werden.

$$P_{Verkehrsklassengruppe}^{FPD-Position} = \frac{\sum_{i}^{n} P_{\text{Verkehrsklassegruppe}}^{FPD-Kachel_i}}{n} \qquad (3\text{-}31)$$

$$\text{mit} \quad i = 1,2...n \quad \text{Kachelindex der FPD - Position}$$
$$n \quad \text{Anzahl der FPD - Kachel}$$

4.9.3.3 Bestimmung der Verkehrsklassengruppenwahrscheinlichkeit je FPD-Positionsfolge

Analog zu Abschnitt 4.9.3.2 können z.B. durch das arithmetische Mittel die Verkehrsklassengruppenwahrscheinlichkeiten der FPD-Positionen mit Formel (3-32) zusammengefasst werden:

$$P_{Verkehrsklassengruppe}^{FPD-Positionsfolge} = \frac{\sum_{i}^{n} P_{\text{Verkehrsklassegruppe}}^{FPD-Position_i}}{n} = P_{Verkehrsklassengruppe,absolut}^{FPD-Positionsfolge} \qquad (3\text{-}32)$$

$$\text{mit} \quad i = 1,2...n \quad \text{Positionsindex der FPD - Positionsfolge}$$
$$n \quad \text{Anzahl der FPD - Positionen}$$

Die Verkehrsklassengruppenwahrscheinlichkeit $P_{Verkehrsklassengruppe}^{FPD-Positionsfolge}$ wird ausschließlich aus der räumlichen Lage des jeweiligen Verkehrsnetzes und den FPD-Positionen berechnet. Sie wird folgend daher auch als Absolute Verkehrsklassenwahrscheinlichkeit ($P_{Verkehrsklassengruppe,absolut}^{FPD-Positionsfolge}$) bezeichnet. Unter der Annahme, dass mindestens ein Verkehrsnetz genutzt wird, kann die Verkehrsklassengruppe mit der rechnerisch höchsten Wahrscheinlichkeit auf den Wert 100% gesetzt werden. Die Wahrscheinlichkeitswerte der weiteren Verkehrsklassengruppen können nun proportional zu dieser Verkehrsklasse skaliert werden. Die

daraus resultierenden Maße werden folgend als Relative Verkehrsklassenwahrscheinlichkeit ($P^{FPD-Positionsfolge}_{Verkehrsklassengruppe,relativ}$) bezeichnet und mit Formel (3-33) berechnet.

$$P^{FPD-Positionsfolge}_{Verkehrsklassengruppe,relativ,j} = \frac{P^{FPD-Positionsfolge}_{Verkehrsklassegruppe,absolut,j}}{\max\left(P^{FPD-Positionsfolge}_{Verkehrsklassegruppe,absolut,j}\right)}$$

mit $j = 1,2...n$ Index der absoluten Verkehrsklassengruppenwahrscheinlichkeit (3-33)

m Anzahl der Verkehrsklassengruppen

Aufgrund der FPD-Positionsungenauigkeiten und der daraus resultierenden räumlich verteilten Menge von Kacheln je FPD-Kachel, wird der Maximalwert von 100% für eine einzelne Verkehrsklasse in Gebieten mit dichten Verkehrsnetzen, wie es im Do-iT Demostratorgebiet der Fall ist, nur gelegentlich erreicht. Die Verkehrsklassen können daher nur bedingt identifiziert werden. Sehr wohl ist aber die Abgrenzung von einigen Verkehrsklassen und damit Verkehrsnetzen möglich.

4.10 Topologische Überprüfung der Verkehrsklassenwahrscheinlichkeit

Besonders bei Verkehrsklassen die Netze geringer Dichte (z.B. Schienennetze), nutzen und in deren Nähe häufig andere Verkehrsnetze (z.B. Straßennetze) verlaufen, ist die Verkehrsklassentrennung mittels räumlicher Vergleichsverfahren, wie es in Abschnitt 4.9.3 erläutert wurde, nur bedingt möglich. Obwohl sich bspw. ein FPD auf einem dünnen Verkehrsnetz bewegt (vergleiche Abbildung 4-44 der schwarz eingefärbte Verlauf der Schienenwege) werden hohe Wahrscheinlichkeitswerte für die Verkehrsnetze höherer Dichte ermittelt. Eine Ursache kann in der vereinfachten Betrachtung der räumlichen Verteilung der Netze gefunden werden. Die Verkehrsnetze des Lokalisierungsnetzes besitzen zusätzlich auch topologische Informationen, die in dem oben beschriebenen Verfahren nicht genutzt wurden.

4.10.1 Ein Verfahren zur logischen Überprüfung der Verkehrsklassenwahrscheinlichkeit: Routensuche mit Restriktionen aus den FPD-Schwerpunkten

Ziel des Verfahrens ist es zu prüfen, ob es entlang der FPD-Positionsfolge eine Route auf einem Verkehrsnetz gibt, die sowohl den räumlichen, zeitlichen und topologischen Restriktionen entspricht. Das in Abschnitt 4.9.3 erläuterte Verfahren untersucht nur das räumliche Vorhandsein eines Verkehrsnetzes, kann jedoch zur Vorauswahl der Infrage kommenden Verkehrsnetze genutzt werden. Wenn mindestens eine Route gefunden werden kann, kommt das untersuchte

Verkehrsnetze prinzipiell als mögliches Transportnetz in Betracht. Kann keine Route ermittelt werden, sind das Verkehrsnetz und die zugeordneten Verkehrsklassen zu verwerfen.

Die hier gestellte Problemstellung kann auch als Spezialfall der Tourenplanung aufgefasst werden, da die Aufenthaltsorte, die Aufenthaltszeitpunkte und die Reihenfolge der aufgesuchten Orte in die Routenberechnung als Optimierungskriterium mit einfließen. Abweichend von der Tourenplanung sind bereits bekannt:

- die Anzahl der Fahrzeuge, gleich 1,
- die Reihenfolge der aufzusuchenden Orte,
- die Zeitfenster für den Aufenthalt an spezifischen Orten.

Infolge der FPD-Positionsgenauigkeit sind jedoch die aufzusuchenden Orte abweichend von den Standardaufgaben der Tourenplanung nur unscharf bekannt. Daraus ergibt sich eine Abwandlung der Tourenplanungsziele, so dass eine Route zu finden ist, die:

- mindestens über ein Verkehrselement je FPD-Position führt,
- der Reihenfolge der FPD-Positionen und deren Verkehrsnetzelementen entspricht,
- nur Verkehrselemente passiert, die innerhalb der Zeitfenster der FPD-Positionen sind,
- dem Zeitbedarf (Gesamtkosten) der Route innerhalb der gemessenen Zeit zwischen der ersten und der letzten FPD-Position genügt.

1. Restriktion: Räumliche Ausdehnung und Gewichtung der Verkehrsnetzelemente

Mit den FPD-Positionen sind Koordinaten und fehlerzeigende Figuren für die räumliche Ausdehnung der Aufenthaltswahrscheinlichkeit gegeben. Basierend auf diesen Flächen sind die Verkehrsnetzelemente, auf denen der MFT zu den Zeiten der FPD-Positionen gewesen sein könnte, bekannt. Unter der Annahme, dass sich das FPD zwischen zwei FPD-Positionen auf dem Verkehrsnetz und der Bedingung, dass es sich maximal mit der verkehrsnetztypischen Fortbewegungsgeschwindigkeit bewegt hat, kann ein Korridor für den Aufenthalt des FPD ermittelt werden.

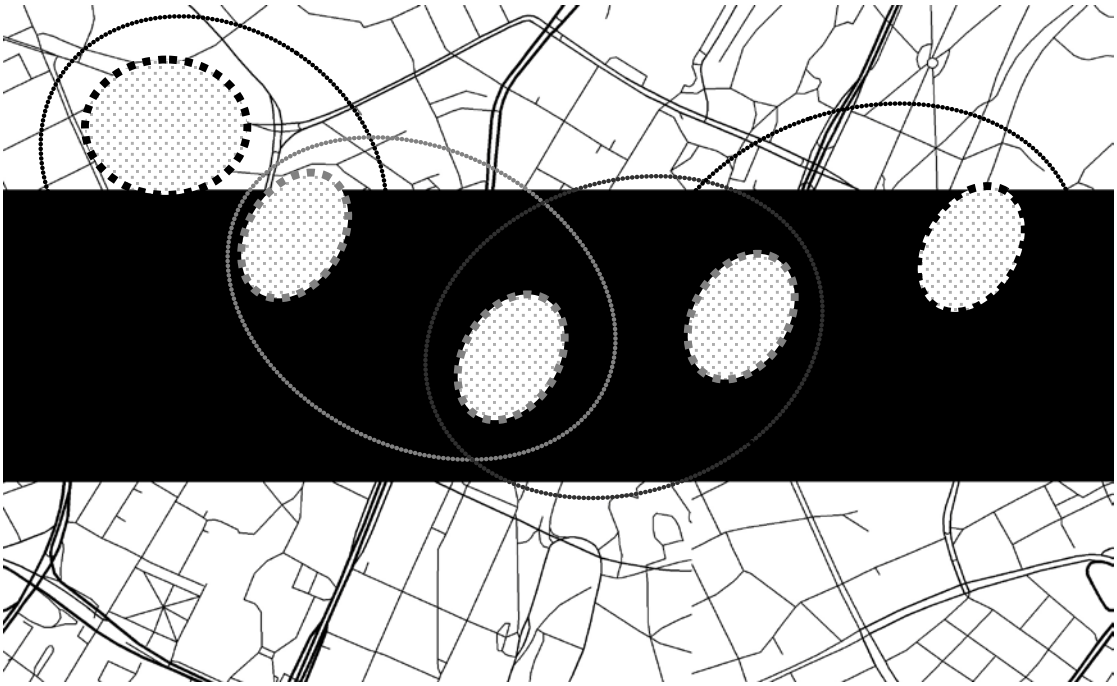


Abbildung 4-45: Mögliche Korridore einer FPD-Positionsfolge

Zur Berechnung dieses Korridors wird, adaptiert von Kohlen (2003) eine ellipsenförmige Korridorfläche berechnet, unter Zuhilfenahme

- der aufeinander folgenden FPD-Positionen,
- der Geschwindigkeit zwischen den beiden FPD-Positionen aus dem euklidischen Abstand und der verstrichenen Zeit,
- der maximalen Fortbewegungsgeschwindigkeit im Verkehrsnetz und
- einem Aufschlag in den Halbachsen, um die FPD-Positionsgenauigkeit zu berücksichtigen

Der zusätzliche Aufschlag im Radius aufgrund der FPD-Positionsgenauigkeit wird in Kohlen (2003) empirisch bestimmt bzw. aus Erfahrungswerten ermittelt. In Do-iT können hierfür die Parameter der fehlerzeigenden Figur verwendet weden. Dazu werden die Brennpunkte der Korridorellipse um den Abstand des Wahrscheinlichkeitsgrenzmaßes (5%, entspricht dem großen Halbmesser der FPD-Konfidenzellipse) verschoben, vgl. Abbildung 4-46. Die Auswahl der Verkehrselemente ist damit genügend groß, um die möglichen Routen zu beinhalten. Gleichfalls ist der Datenumfang ausreichend gering zur effizienten Routensuche.

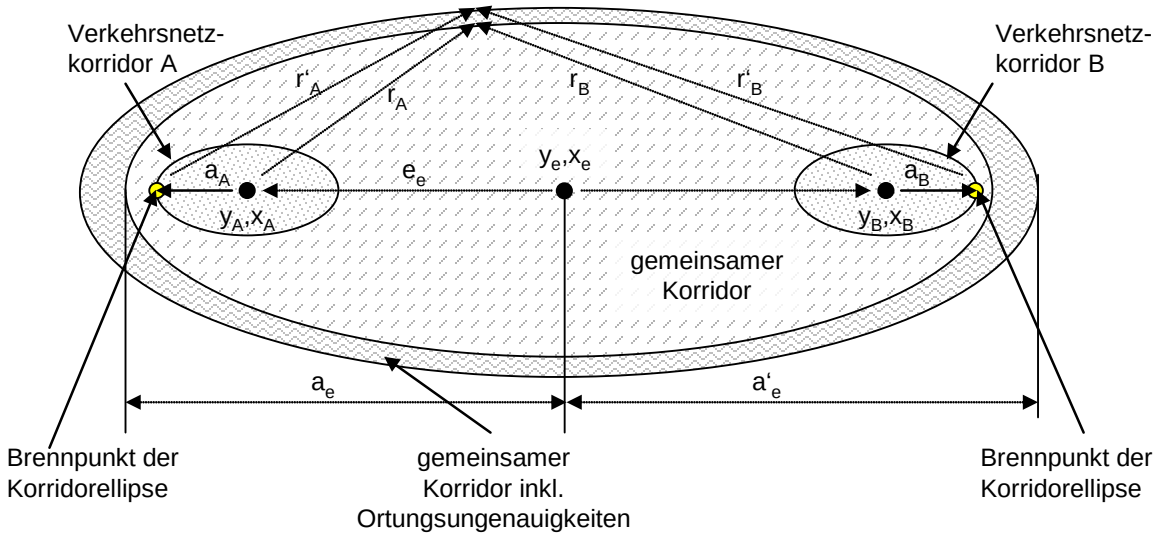


Abbildung 4-46: Ermittlung des verkehrsnetzabhängigen Korridors zwischen zwei FPD-Positionen

Es ist zu unterscheiden zwischen Verkehrselementen die a) ausschließlich Bestandteil der ersten FPD-Position, b) ausschließlich Bestandteil der zweiten FPD-Position, c) ausschließlich Bestandteil des Korridors zwischen den FPD-Positionen und d) gleichfalls Bestandteil der ersten und zweiten FPD-Position sind. In der Abbildung 4-47 werden diese Unterteilungen veranschaulicht, wobei die erste FPD-Position mit A und die folgende FPD-Position mit B gekennzeichnet sind.

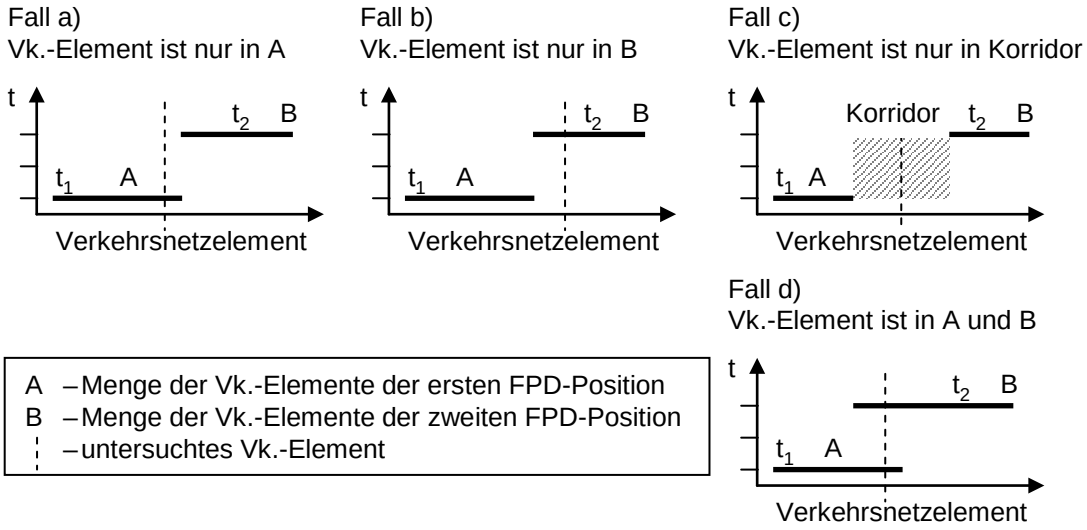


Abbildung 4-47 : Varianten der Verkehrsnetzelementbeteiligung für die Zeitfensterberechnung der topologischen Verkehrsnetzanalyse

Sind Korridorellipsen zwischen allen FPD-Positionen berechnet, werden die Verkehrsnetzelemente aus dem Verkehrsnetz selektiert. Den Verkehrsnetzele-

menten innerhalb der FPD-Positionen können zusätzlich, wie in Abbildung 4-48 gezeigt, die Aufenthaltswahrscheinlichkeiten in den Kacheln zugeordnet werden. Mit dem gewichten Verkehrsnetz kann ein Optimierungsverfahren zur Routesuche die wahrscheinlichsten Routen ermitteln. Im Rahmen der Identifizierung ist dieses nicht Bestandteil. Vielmehr würde das Vorhandensein einer beliebigen Route bereits zeigen, dass das Verkehrsnetz und die entsprechenden Verkehrsklassen nicht auszuschließen sind.

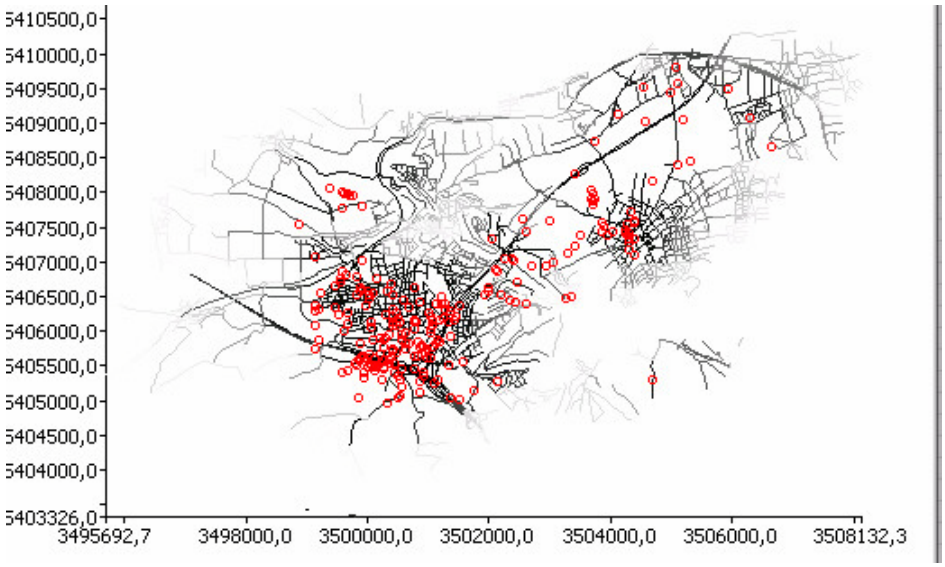


Abbildung 4-48: Beispiel für die Gewichtung von Verkehrsnetzelementen in Abhängigkeit von der FPD-Aufenthaltswahrscheinlichkeiten

2. Restriktion: Reihenfolge / Sequenz

Entgegen dem Optimierungskriterium der Tourenplanung, bei der eine optimale Reihenfolge der anzufahrenden Zwischenziele zu finden ist, ist die Reihenfolge der aufzusuchenden Knoten bereits durch die FPD-Positionsfolge bekannt. Infolge der Ortungsungenauigkeiten ist jedoch häufig nicht nur ein Knoten für eine FPD-Position repräsentativ, sondern eine Menge von Knoten, vgl. auch Abbildung 4-45.

Vorraussetzung der klassischen Routingalgorithmen, wie Dijkstra, Bellman-Ford oder Floyd und Warshall, ist die Einmaligkeit von Knoten in einem Verkehrsnetz. Aufgrund der Ortungsungenauigkeiten und da auch die tatsächliche Route eines FPD das mehrfache Vorhandensein derselben Knoten in der Route bedingen kann, wäre zur Aufrechterhaltung dieser Bedingung die Unterteilung der Routensuche in mehrere Teilrouten denkbar. Da letztlich jede Relationskombination in und zwischen allen Korridoren zu berechnen ist, bedingt dieses einen hohen Rechenaufwand.

In Do-iT wird hingegen eine temporäre Anpassung der Datengrundlage vorgenommen. Die Korridore zwischen jeweils zwei FPD-Positionen werden vorerst als eigenständige Verkehrsnetze betrachtet. Den in den verschiedenen Korridoren mehrfach vorhandenen Knoten wird dazu eine eindeutige temporäre Knotennummer zugewiesen. Diejenigen Knoten die innerhalb der Fehlerellipse einer FPD-Position sind, werden außerdem mit Relationen der Fahrzeiten = 0 ergänzt. Die Relationen verbinden somit die Knoten der ersten FPD-Position des Korridors mit den Knoten der zweiten FPD-Position des Vorgängerkorridors. Unter Anderem wird dadurch auch das Verbleiben des FPD an diesen Knoten bei Stillstand gestattet. Abbildung 4-49 zeigt diesen Zusammenhang an einem Beispiel mit vier FPD-Positionen.

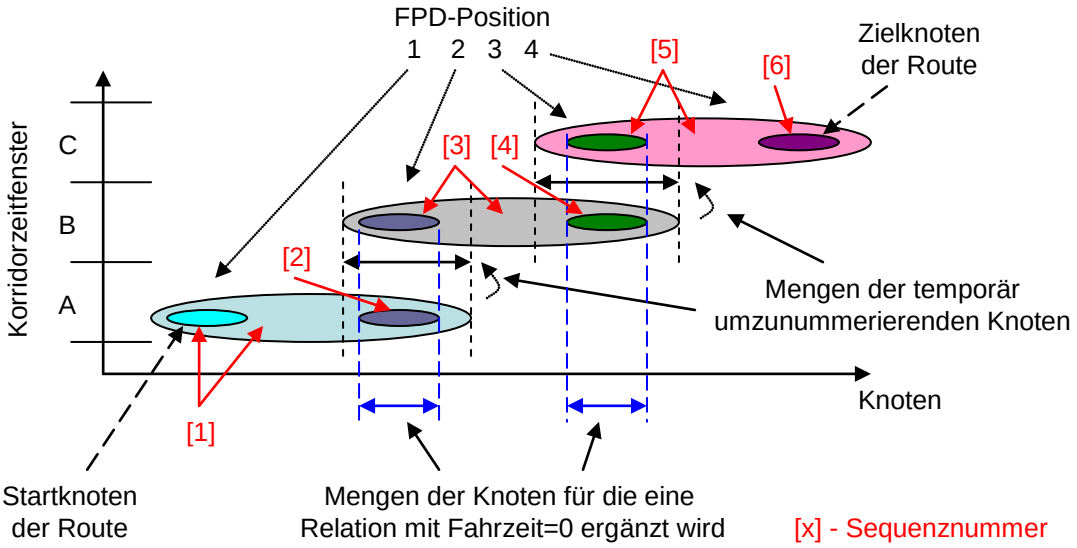


Abbildung 4-49: Temporäre Umnummerierung der sich in verschiedenen Korridoren zeitlich überlappenden Knoten

Jeweils zwischen zwei FPD-Positionen, entspricht einem Korridorzeitfenster, existieren somit eindeutige Verkehrsnetze. Die topologisch getrennten Verkehrnetze werden mit Relationen der Fahrzeit = 0 ergänzt und damit topologisch miteinander verbunden. Da die Knoten eindeutige Knotennummern besitzen, kann man die einzelnen Korridorverkehrsnetze zu einem neuen, routingfähigem Verkehrsnetz verschmelzen.

3. Restriktion: Zeitfenster

Betrachtet man ausschließlich die Verkehrsnetzknoten des Korridors, so kann angenommen werden, dass sich das FPD innerhalb des Zeitfensters $t_a - t_b$ auf oder zwischen diesen Knoten bewegt haben kann. Der Aufenthalt, vgl. auch Abbildung 4-50, eines FPD kann somit außerhalb dieses Zeitfensters ausgeschlossen werden.

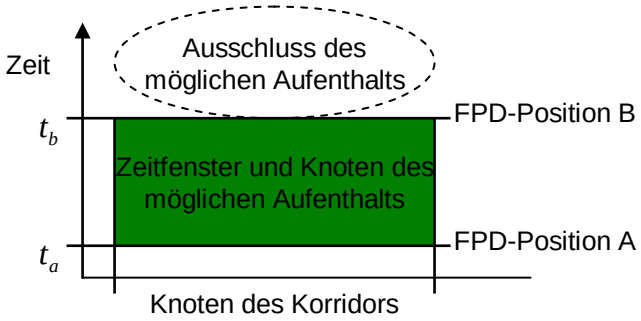


Abbildung 4-50: Durch zwei aufeinander folgende FPD-Positionen räumlich und zeitlich definierter Ausschluss bzw. möglicher Aufenthalt des FPD

Eine Relation zwischen zwei Knoten ist daraufhin nur dann gültig, wenn das FPD den Zielknoten spätestens zum Zeitpunkt t_b unter Berücksichtigung der minimalen Fahrzeit dt erreichen kann. Abbildung 4-51 veranschaulicht diese Bedingung wobei t_1 und t_2 die Zeitgrenzen des Startknotenzeitfensters sowie t_3 und t_4 die Zeitgrenzen des Zielknotenzeitfensters sind. Eine mögliche Bewegung des FPD kann somit nur innerhalb des Zeitfensters stattgefunden haben.

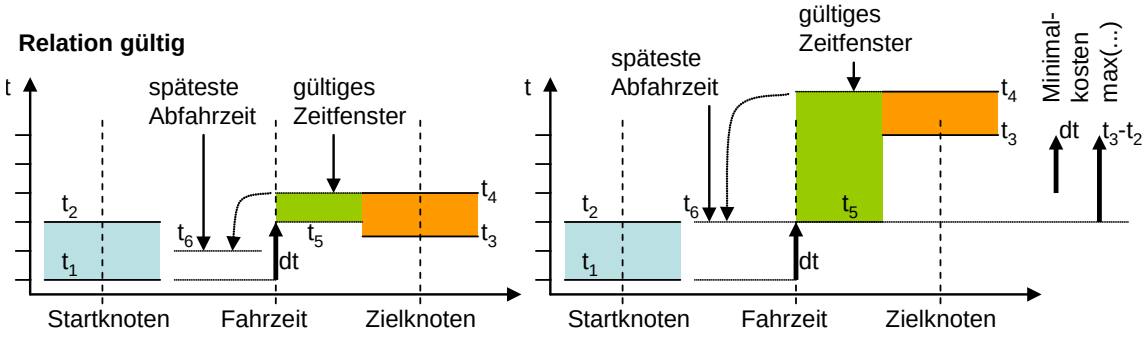


Abbildung 4-51: Zeitrestriktionen für die Gültigkeitsüberprüfung und Minimalkostenbestimmung (Fahrzeit) von Knotenrelationen

Demgegenüber ist eine Relation zwischen zwei Knoten als ungültig zu bewerten, wenn selbst bei Abfahrt zum frühesten Zeitpunkt t_1 vom Startknoten und bei Berücksichtigung der geringsten Fahrzeit dt, der Zielknoten nicht bis spätestens zum Zeitpunkt t_4 erreicht werden kann, wie es beispielsweise in Abbildung 4-52 gezeigt wird.

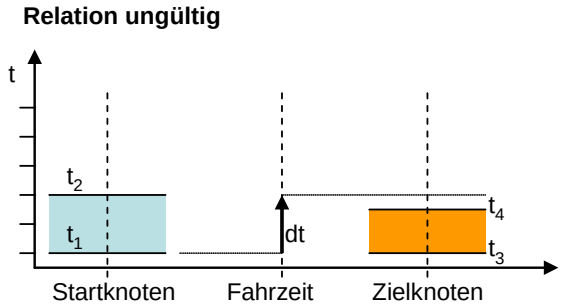


Abbildung 4-52: Ungültige Knotenrelationen aufgrund unzulässiger Zeitfenster

Aus den Fällen a) bis d) der Abbildung 4-47 können die Zeitfenstergrenzen unter Berücksichtigung der Gültigkeitsüberprüfung und der Minimalkostenbestimmung aus Abbildung 4-51 abgeleitet werden. Tabelle 14 zeigt die zusammengefassten Bedingungen und die Berechnung der Zeitfenstergrenzen für die Knotenrelationen.

Tabelle 14: Zeitfenstergrenzen der Verkehrselemente auf Basis des angenommenen Aufenthalts des FPD

Fall	Zeitfenstergrenzen
a) Element ist nur in A	Frühestens ab dem Zeitpunkt t_1 Spätestens bis zum Zeitpunkt t_2 - schnellste Verbindung von A bis B
b) Element ist nur in B	Frühestens ab Zeitpunkt t_1 + schnellste Verbindung von A bis B Spätestens zum Zeitpunkt t_2
c) Element ist nur in Korridor	Frühestens ab t_1 + schnellste Verbindung von A bis Element Spätestens bis t_2 - schnellste Verbindung von Element bis B
d) Element ist in A und B	Frühestens ab Zeitpunkt t_1 Spätestens bis Zeitpunkt t_2

Für jedes Verkehrselement können mit den Bedingungen aus Tabelle 14 die gültigen Zeitfenstergrenzen berechnet werden. Die Zeitpunkte t_1 sowie t_2 werden aus den FPD-Positionen übernommen und werden den Verkehrnetzelementen als Zusatzinformation angefügt. Die als „schnellste Verbindung“ bezeichnete Zeitspanne ist vom Routingalgorithmus während der Routenberechnung zu berechnen. Da die Obergrenze des Zeitfensters der maximal gültigen Fahrdauer der Route entsprechen darf, kann der Router bereits während der Berechnung Routen mit zeitlich längeren Wegen verwerfen. In Tabelle 15 wird ein Datenbeispiel für die Knotennummerierung in den Korridoren sowie für die ergänzten Zeitfenstergrenzen und den Relationen der Fahrzeit = 0 gezeigt. Die verwendete Bezeichnung „Sub_...“ wird im folgenden Abschnitt erläutert.

Tabelle 15: Datenbeispiel für die ergänzten Zeitfenstergrenzen und den ergänzten Relationen der Fahrzeit = 0 zwischen den Korridoren im Datenmodell

	relationale Daten des Lokalisierungsnetzes und ergänzte Relationen mit Fahrzeit = 0			Ergänzte Restriktionen der Zeitfenster	
Korridor	Sub_From_Node	Sub_To_Node	Sub_Journey_Time	Zeitfernstergrenze	
				untere	obere
	[LN_Node_ID]	[LN_Node_ID]	[s]	[s]	[s]
1	1	2	12	0	28
	1	7	22	0	28
…	…	…	…	…	…
2	7	20	0	28	50
	20	21	7	28	50
…	…	…	…	…	…
3	20	30	0	50	83
	21	31	0	50	83
	30	32	8	50	83
…	…	…	…	…	…

4.10.2 Räumliche Mindestauflösung

Werden, wie im hier vorgestellten Verfahren, ausschließlich die Knoten der Verkehrselementrelationen eingesetzt, ist die räumliche Verteilung der Aufenthaltswahrscheinlichkeiten abhängig von der räumlichen Verteilung der Verkehrsnetzknoten. Sind die beiden Knoten einer Kante beispielsweise außerhalb der FPD-Positionsfigur, geometrisch verläuft die Kante jedoch auch innerhalb dieser Figur, vgl. Abbildung 4-53, so wird der mögliche Aufenthalt auf der Kante trotz vorhandener Aufenthaltswahrscheinlichkeit nicht berücksichtigt.

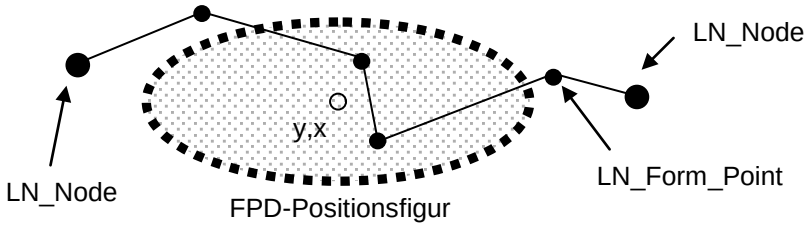


Abbildung 4-53: Knoten (LN_Node) ohne Aufenthaltswahrscheinlichkeit obwohl die Kante Bestandteil der FPD-Positionsfigur ist

Eine Variante dieser Fehlinterpretation entgegen zuwirken, ist die Verwendung der bereits oben genannten Kacheln der FPD-Positionsfigur. Mittels einer Vorprozessierung können die Schnittpunkte der Kanten an den Kachelgrenzen berechnet, in Abbildung 4-54 mit Kreisen gekennzeichnet, und als neue Knoten der Kante ergänzt werden. Da die Kanten unterteilt werden, wird für die neuen

Verkehrsnetzelemente die Bezeichnung *Sub_Node* und *Sub_Edge* mit *Sub_From_Node* als Startknoten und *Sub_To_Node* als Zielknoten einer Kante *Sub_Edge* in der Kachel gewählt.

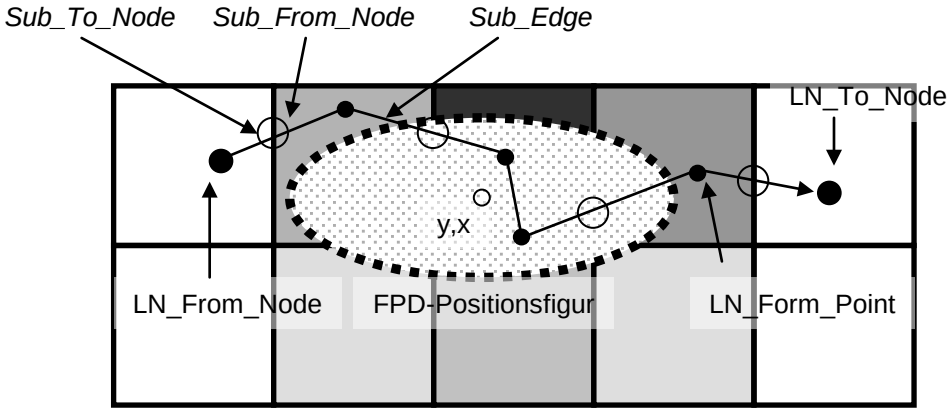


Abbildung 4-54: Unterteilung der Sub_Edges

Für die Sub_Edges kann nun die Aufenthaltswahrscheinlichkeit des FPD in der Kachel übernommen werden. Wie bereits oben erwähnt, fließt in Do-iT dieser Aufenthaltswahrscheinlichkeitswert nicht in die Routenberechnung ein.

4.10.3 Verkehrsklassenbewertung anhand der gefundenen Routen

Konnte auf keinem Verkehrsnetz eine den Restriktionen genügende Route gefunden werden, hat das FPD entweder ein unbekanntes Verkehrsnetz benutzt oder die FPD-Ortungen spiegeln auch unter den berücksichtigten Unschärfen den Verlauf der Bewegung nicht genügend wider. Das hier vorgestellte Verfahren kann dann keine Verkehrsklasseneingrenzung liefern.

Konnte nur auf einem Verkehrsnetz eine gültige Route gefunden werden, spricht dieses für die Eingrenzung auf die entsprechende Verkehrsklasse. Wenn das Verkehrsnetz prinzipiell von mehren Verkehrsklassen genutzt werden kann, dann spricht es natürlich nur für die entsprechende Verkehrsklassengruppe. Gegebenenfalls konnten jedoch bereits mit einem der voran erläuterten Verfahren Verkehrsklassen dieser Gruppe ausgeschlossen werden.

Wenn auf mehreren Verkehrsnetze eine Route gefundenen werden konnte, ist eine Eingrenzung der Verkehrsklassen nur auf die Verkehrsklassen oder Verkehrsklassengruppen der entsprechenden Verkehrsnetze möglich. In Abbildung 4-55 ist ein derartiges Beispiel eines MFT dargestellt. Innerhalb der Zeiten für die FPD-Schwerpunkte sowie der räumlichen Eingrenzung werden auf dem Straßennetz und auch auf dem Schienennetz Routen gefunden. Verkehrklassen beider Verkehrsnetze gelten damit weiterhin als möglich. Die Entscheidung für das FPD als aktiv ist somit nicht eindeutig möglich.

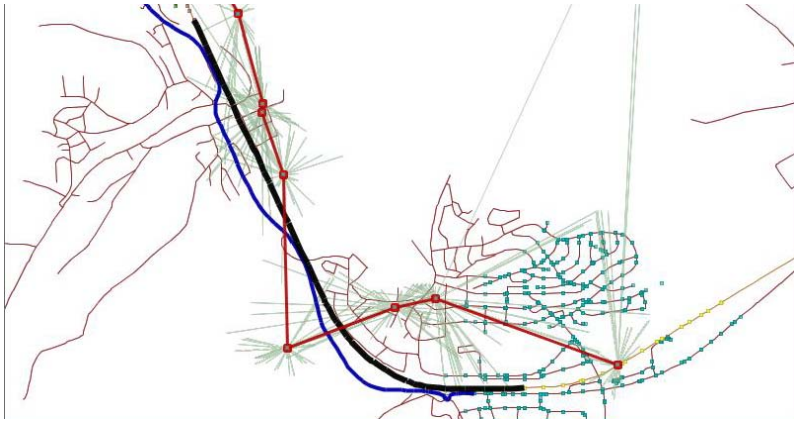


Abbildung 4-55: Beispiel für die Identifizierung von gültigen Routen eine FPD auf den Straßennetz (blau) und dem Schienennetz der DB (schwarz)

5 Literatur

(Kohlen, 2003) Routenidentifikation in Verkehrsnetzen auf Grundlage unscharfer Ortungsinformationen (GSM), Ralf Kohlen, Heft 52, Shaker Verlag, Aachen, 2004

(Luo, 2007) Erfassung und Analyse von Fahrplanabweichungen ausgewählter ÖPNV-Angebote, Studienarbeit IAGB Universität Stuttgart (unveröffentlicht), 2007

(Rauber, 1993) Algorithmen in der Computergrafik, Thomas Rauber, B. G. Teubner Stuttgart, 1993

(Sachs, 2004) Angewandte Statistik, Anwendungen statistischer Methode, Lothar Sachs, Springer-Verlag Berlin Heidelberg New York, 2004

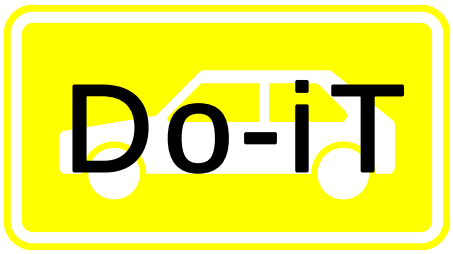
(Walke, 2001) Mobilfunknetze und ihre Protokolle 1, Grundlagen, GSM, UMTS und andere zellulare Mobilfunknetze, 3. Auflage, Teubner, 2001.

(3gpp GSM 05.10) Digital cellular telecommunications system (Phase 2+);Radio subsystem synchronization (GSM 05.10 version 6.2.0 Release 1997), www.3gpp.org, 13.07.2008

Arbeitspaket 3 Bericht

Datenoptimierung für integrierte Telematik

Do-iT



Förderkennzeichen: 19 B 5007

Qualitätsbewertung zur Identifikation aktiver Verkehrsteilnehmer

Ralf Schollmeyer, Volker Schwieger

August 2009

Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen

Dokument-Informationen

Version	Datum	Bearbeiter	Bemerkung
0.1	**05.02.09**	**Schwieger**	• Gliederung und Struktur
0.2	**02.03.09**	**Schwieger**	• Einarbeitung Eingaben R. Schollmeyer und Korrektur
0.3	**23.04.09**	**Schweitzer**	• Einarbeitung Testfahrt Karlsruhe (Kapitel 3)
0.4	**25.05.09**	**Schwieger**	• Einarbeitung Input R. Schollmeyer
0.5	**26.05.09**	**Schweitzer**	• Einfügen der Trajektorien der Testfahrten Sep. 2007 u. Dez. 2007
0.9	**26.06.09**	**Schollmeyer**	• Überarbeitung Kapitel 4, 5 und 6
1.0	**09.07.09**	**Schwieger**	• Überarbeitung und Kommentierung des aktuellen Stands
1.1	**22.07.09**	**Schollmeyer**	• Überarbeitung 4.3 bis 7, Einpflegen von Kommentaren aus 1.0
1.2	**04.08.09**	**Schwieger**	• Überarbeitung und Kommentierung ab 4.3
1.3	**05.08.09**	**Schollmeyer**	• Einarbeitung der Kommentare von 1.2
1.4	**13.08.09**	**Schwieger**	• Abschlusskontrolle mit Frau Reichert

Gliederung

1 Einleitung 4

2 Qualitätsbeurteilung 4

3 Beschreibung der durchgeführten Testfahrten 6
3.1 Basisdaten der Identifikationsverfahren 6
3.2 Beschreibung der Versuchsfahrten 9

4 Analyse der Ergebnisse 23
4.1 Identifizierungsstrategie 23
4.2 Kurzbeschreibung der Identifizierungsschritte 24
4.2.1 Sektormaske 24
4.2.2 Bewegungsanalyse 25
4.2.3 Positionsschwerpunkte 26
4.2.4 Korrelation mit Öffentlichem Verkehr 28
4.2.5 Verkehrsnetzanalyse 30
4.3 Vergleich mit GPS-Referenzdaten und empirische Analyse 34
4.3.1 Bewegungsanalyse 35
4.3.2 Positionsschwerpunkte 38
4.3.3 Korrelation mit Öffentlichem Verkehr 42
4.3.4 Verkehrsnetzanalyse 46
4.3.5 Analyse des Gesamtverfahrens 54
4.4 Analysen der verketteten Identifizierung von MFT-Datensätzen 57
4.5 Statistiken zur 24h Auswertung des 19.12.2007 62
4.5.1 Statistik zur tageszeitlichen Trajektorienverteilung 64
4.5.2 Flächenabdeckung und räumliche Verteilung 65
4.5.3 Rechengeschwindigkeit 69

5 Zusammenfassende Qualitätsbewertung 71

Literatur 74

6 Anlage 75

1 Einleitung

Im Projekt Do-iT sollen für die Verkehrslageerfassung und -prognose auf Basis anonymisierter Mobilfunkdaten Ortungsinformationen in Form so genannter „Floating Phone Daten“ (FPD) erzeugt werden, die den von einem Mobiltelefon auf dem digitalen Straßennetz zurückgelegten Weg mit zeitlichem Bezug beschreiben. Von entscheidender Bedeutung hierbei sind die Algorithmen zur Identifikation aktiver Verkehrsteilnehmer, die es gestatten aus der Vielzahl der Mobiltelefone die zu extrahieren, die aktiv am Individualverkehr teilnehmen. Die Verfahren sind im Bericht zum Arbeitspaket 3.5 (AP3.5, [2]) beschrieben. Sie nehmen innerhalb der Prozessierungskette einen zentralen Platz ein (vergleiche Abbildung 1-1). Zu erwähnen ist in diesem Zusammenhang, dass im Rahmen des Projekts die kompletten Algorithmen nur auf Abis-Ebene implementiert wurden. Aus diesem Grunde erfolgt die Qualitätsbewertung gleichfalls nur für die Abis-Daten. Zurzeit läuft eine deutlich reduzierte Identifizierung auf A-Ebene. Die komplette Implementierung für die A-Daten konnte aufgrund der Verzögerungen des Projekts nicht mehr vorgenommen werden und wird im Rahmen einer Dissertation realisiert werden.

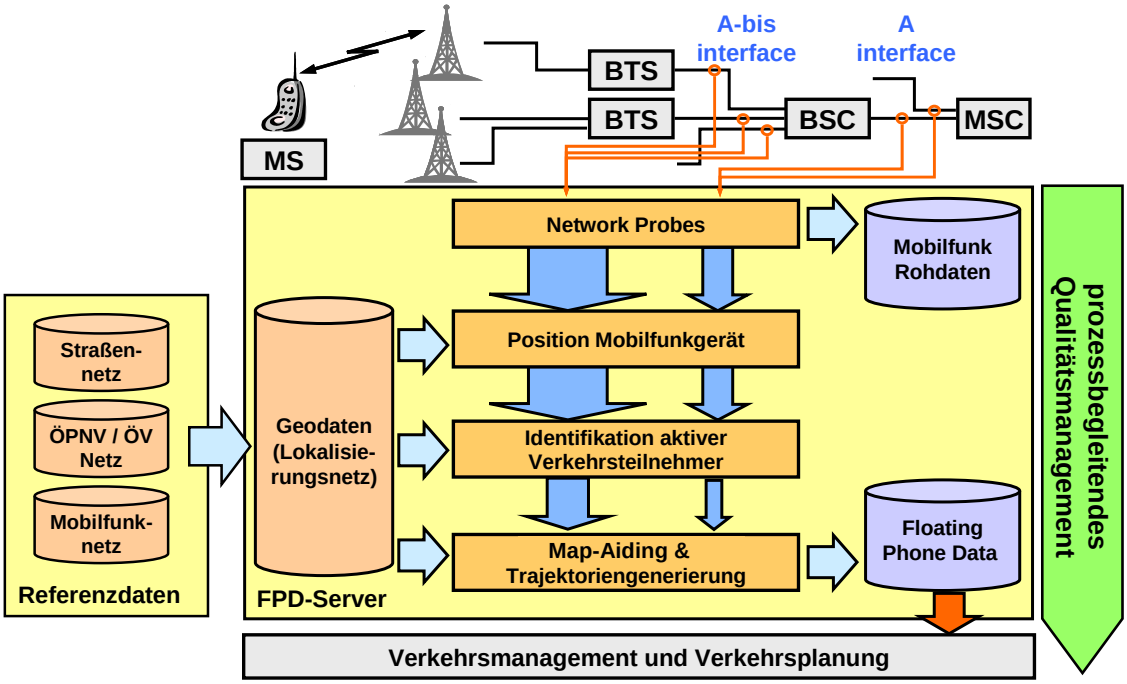


Abbildung 1-1: Systemarchitektur und Ablaufschema der Erzeugung von FPD

2 Qualitätsbeurteilung

Im Rahmen des Projekts Do-iT findet ein durchgehendes Qualitätsmanagement Anwendung. Grundbedingungen hierfür sind definierte Qualitätskriterien und Qualitätsparameter. Dieser Bericht beschäftigt sich mit der Qualität der Identifikation der aktiven Verkehrsteilnehmer. Hierfür sind im Bericht (AP10.1, [1]) zwei Qualitätsparameter vorgesehen

(vergleiche Tabelle 2-1). Beide beschreiben im Prinzip die Zugehörigkeit eines Mobilfunkteilnehmers zu einer Teilnehmerklasse; im einfachsten Fall „aktiv“ oder „passiv“.

Tabelle 2-1: Qualitätsparameter zur Identifikation aktiver Verkehrsteilnehmer

Nr.	**Merkmal**	**Parameter**	**Referenzdaten**
P5.1	Korrektheit	Klassifizierungskorrektheit	mit
P5.2	Korrektheit	Mittlere Klassifizierungskorrektheit	mit

Im Bericht (AP10.1, [1]) ist aufgeführt, dass die Qualitätsparameter P5.1 und P5.2 ohne Referenzdaten bestimmt werden sollen. Es hat sich aber im Laufe des Projekts herausgestellt, dass Aussagen zur Korrektheit nur mit externen Referenzdaten möglich sind. Aus diesem Grunde ist Tabelle 2-1 entsprechend angepasst.

Die Klassifizierung einzelner Identifikationsschritte ist gleichfalls mit Referenzdaten zu vergleichen. Es können, wie in vorhergehenden Berichten des IAGB (z.B. AP4.5, [3]), GPS-Trajektorien herangezogen werden, die im Individualverkehr, im öffentlichen Verkehr als auch von inaktiven unbewegten Teilnehmern erzeugt werden. Die Korrektheit der Zuordnung ist anhand der Referenzdaten dann eindeutig zu bewerten. Für jede Trajektorie mit ja oder nein. Des Weiteren ist für den Fall der fehlerhaften Identifikation für jede einzelne Identifikationsmaßnahme (vergleiche AP3.4, 2008) die Korrektheit zu überprüfen, um zu identifizieren, welche Maßnahme fehlerhafte Ergebnisse geliefert hat. Als Einzelmaßnahmen sollen nach Beicht (AP3.4, 2008) die folgenden Punkte aufgefasst werden:

- Sektormaske
- Bewegungsanalyse
- Positionsschwerpunkte
- Korrelation mit öffentlichen Verkehr
- Verkehrsnetzanalyse

Für die mittlere Klassifizierungskorrektheit kann keine Korrektheitsrate ohne den Vergleich mit Referenzdaten aller an einem Tage entstehenden Teilnehmerdaten angegeben werden. Dieses ist aus praktikablen Gründen nicht möglich, da nicht für alle Teilnehmer eines Tages Referenzdaten erzeugt oder beschafft werden können. Ersatzweise wird die Rate für alle realisierten Referenztrajektorien, der in diesen Bericht verwendeten Versuchfahrten berechnet.

© Do-iT 2009

3 Beschreibung der durchgeführten Testfahrten

3.1 Basisdaten der Identifikationsverfahren

Für die Auswertung der Versuchsfahrten wurden die nachfolgend genannten Softwareversionen verwendet:

- IN-TRADE-DB_GSM-Ident_OUT-Ortung-file_005.1 Stand: Dez. 2008
- Identifizierung_012c2 Stand: Apr. 2009

Auch die Datengrundlage sei im Folgenden angegeben. Sie umfasst das Lokalisierungsnetz und die Daten der Fahrpläne des Öffentlichen Verkehrs (ÖV-Daten). Die folgende Tabelle 3-1 fasst die Datenquellen der Verkehrsnetze im Lokalisierungsnetz zusammen.

Tabelle 3-1: Verkehrsnetze und Datenquellen des Lokalisierungsnetzes in der Version LN_019:

Art	**Quelle / Hersteller**	**Stand**
Straßennetz	NAVTEQ	2005 / Q3
KVV Schienennetz	VuV	März 2007
VVS Schienennetz	VuV	Juni 2007
DB Schienennetz	VuV	Nov. 2007

Für die topologische Verbindung des Baden-Württembergischen und des Rheinlandpfälzischen Teils wurden die „gdf_node_conversation“ Knotenidentifizierungen in Form einer Lookup-Tabelle der DDG verwendet. Diese Daten sind in den NAVTEQ-Daten im Format des ESRI-shape, welche vom IAGB beschafft wurden, nicht enthalten.

ÖV-Daten und Fahrplandatenhochrechnungen

Die Vorarbeiten zur Integrierung der Fahrplandaten des Öffentlichen Verkehrs (ÖV), getätigt durch das VuV Universität Stuttgart, umfassen die Busverkehre und Straßen- sowie Ubahnverkehre der KVV (Karlsruher Verkehrsverbund), des VVS (Verkehrs- und Tarifverbund Stuttgart) sowie die Bahnverkehre der DB (Deutsche Bahn AG). Geometrische und topologische Grundlage bildete das Lokalisierungsnetz des IAGB. Im Anschluss wurden die Fahrplandaten und ggf. erforderliche Ergänzungen der Topologie zusammengestellt und in Form von Servicetrip-Listen dem IAGB zur Weiterverarbeitung bereitgestellt. Ein Servicetrip beinhaltet zeitlich aufsteigend sortierte Knotenpunkte mit Tageszeitangabe für Ankunfts- und Abfahrtszeitpunkt. Diese Form der Modellierung ist von der sonst üblichen linien-orientierten Darstellung unabhängig und ist je Tag fahrzeugscharf. Zusätzlich werden die Betriebstage des Fahrplanjahres angegeben, an denen der Servicetrip durchgeführt wird. Letztlich

können hiermit die Aufenthaltsorte und die entsprechenden Antreffzeiten der ÖV-Fahrzeuge im Lokalisierungsnetz prognostiziert werden.

Die vom VuV bereitgestellten Daten sind in der folgenden Tabelle 3-2 zusammengefasst. Sie umfassen jeweils einen fahrplanbedingten Gültigkeitszeitraum, der der Spalte 4 entnommen werden kann. Da nur einige dieser Zeiträume mit den Zeiten der Versuchsfahrten und keine mit denen der in Do-iT vorgesehenen Demonstrationszeiträumen zusammen fallen, mussten die ÖV-Daten zur Nutzbarmachung in die erforderlichen Zeiträume hochgerechnet werden. Die Zeiträume der vom IAGB durchgeführten Hochrechnungen sind der Spalte 5 zu entnehmen. Angemerkt sei, dass die Betriebstage von ÖV-Linien und ÖV-Fahrzeugen unter anderem auch von Ferienzeiten und Feiertagen beeinflusst sind. Die Merkmale der Betriebstagszuordnung konnten den Angaben in den ÖV-Daten nicht entnommen werden, da sie zumeist direkt bei den Verkehrsverbünden intern verwaltet werden. Das VuV und damit auch das IAGB haben auf diese Datenquelle keinen direkten Zugriff. Eine Anpassung an die jahresspezifischen Betriebstage konnte somit nicht erfolgen. Die Hochrechnungen wurden folglich mit der Annahme durchgeführt, dass die Betriebstage in den Hochrechnungsjahren identisch mit denen des Ursprungszeitraumes sind.

Tabelle 3-2: ÖV-Fahrpläne mit Gültigkeits- und Hochrechnungszeiträumen

Art	**Quelle**	**Stand**	**Fahrplanzeitraum**	**Hochrechnungen für Demonstratorzeiträume:**
KVV Bus	VuV	April 2007	10.12.2006 - 08.12.2007	09.12.2007 – 13.12.2008 14.12.2008 – 12.12.2009
KVV Straßenbahn	VuV	März 2007	10.12.2006 - 08.12.2007	09.12.2007 – 13.12.2008 14.12.2008 – 12.12.2009
VVS Bus	VuV	Juni 2007	10.12.2006 - 08.12.2007	09.12.2007 – 13.12.2008 14.12.2008 – 12.12.2009
VVS Straßen-/ Stadtbahn	VuV	Juni 2007	10.12.2006 - 08.12.2007	09.12.2007 – 13.12.2008 14.12.2008 – 12.12.2009
DB	VuV	Nov. 2007	10.12.2006 - 09.06.2007	10.06.2007 - 08.12.2007 09.12.2007 – 13.12.2008 14.12.2008 – 12.12.2009

Nachbearbeitung der Schienengeometrien

Das von NAVTEQ gelieferte Schienennetz weist, wie in der Abbildung 3-1 unten links erkennbar (schwarze teils unterbrochene oder fehlende Linien gegenüber den ergänzten Linien in rot), sowohl topologische als geometrische Datenfehlstände aus. Die für die ÖV-Fahrzeugprognosen relevanten jedoch fehlenden topologischen Verknüpfungen, wurden vom VUV nachdigitalisiert. In diesem Arbeitsschritt wurden in einigen Bereichen auch geometrisch vorhandene Teile der NAVTEQ Karte wieder ausgeblendet. Bei gleichen geometrischen Informationen sollten die NAVTEQ-Daten in schwarz und die vom VuV nachbearbeiteten Daten in rot übereinander liegen. Fehlen NAVTEQ-Informationen, erkennt man dies an freiliegenden roten Elementen. Die Bereiche mit sehr grob aufgelöster Digitalisierung zeichnen sich durch freiliegende NAVTEQ-Schienenverläufe in schwarz und gene-

ralisiertem VuV-Schienennetz in rot in räumlicher Nähe aus. In dem dargestellten Gebiet treten räumliche Abweichungen von vorhandenen NAVTEQ-Geometrien und nachdigitalisierten Schienennetzelementen des VuV von bis zu 1400m auf.

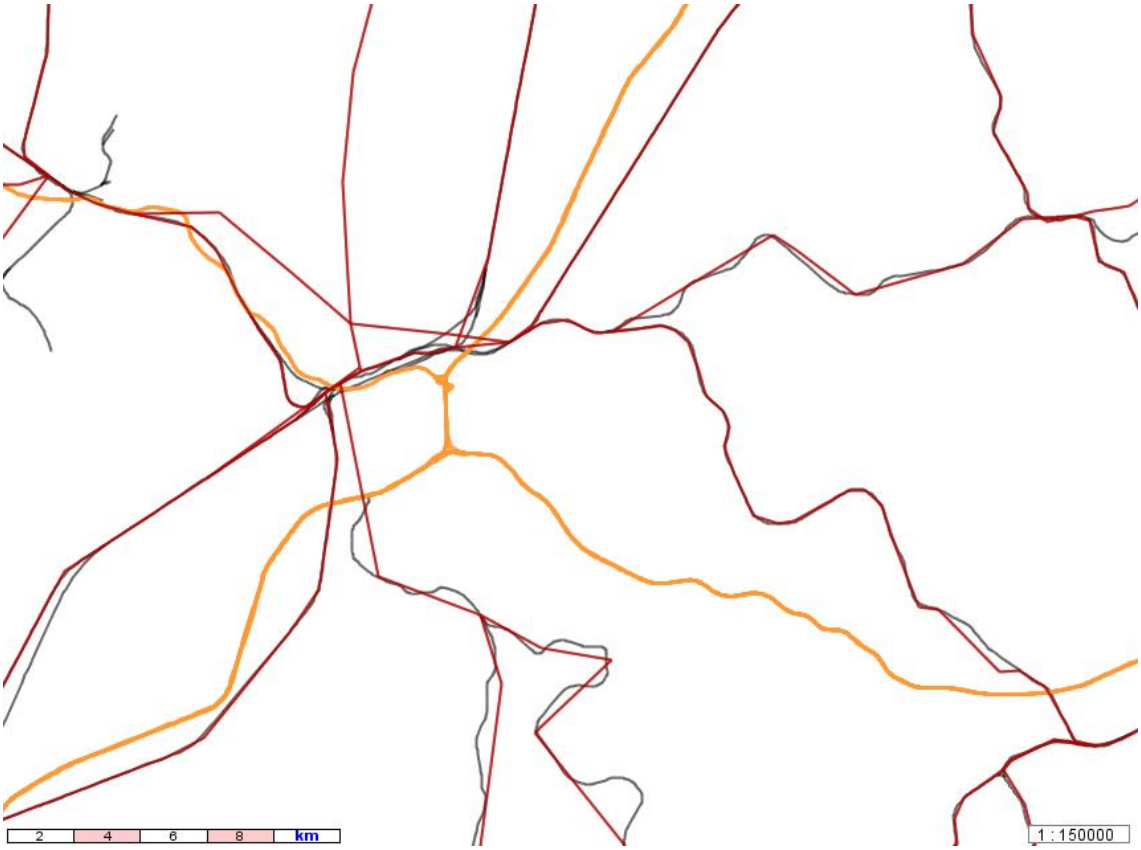


Abbildung 3-1: Kartenmaterial von NAVTEQ (schwarz) und durch VuV topologisch nachbearbeitete Kartenelemente (rot) des Schienennetzes, gelb Autobahnnetz mit Karlsruhe etwa in Kartenmitte

Wie unten noch gezeigt wird, entspricht das nicht dem geometrischen Genauigkeitsniveau der FPD-Daten. Ebenso wären bei unkorrigiertem Schienennetz Fehlentscheidungen bei Vorlage von GPS-Referenztrajektorien zu erwarten. Aus diesen Gründen wurden in einem zusätzlichen Arbeitsschritt die Schienendaten vom VuV auf Geometrieebene wieder mit den Schienendaten von NAVTEQ kombiniert und korrigiert. Die folgende Abbildung 3-2 zeigt den gleichen Kartenausschnitt wie Abbildung 3-1, jedoch in grün die korrigierten Schienennetzelemente. Beim Korrekturverfahren wurden ähnliche Knotenpunkte in beiden Datenbeständen gesucht und die Kantengeometrien entsprechend ergänzt. In den NAVTEQ-Daten nicht vorhandene Geometrien, vgl. rote Linien in Bildmitte, oder nicht vorhandene Topologien, vgl. schwarze Linien am linken Bildrand, wurden nicht ergänzt und als neuer Datenbestand vom VuV unverändert eingepflegt.

© Do-iT 2009

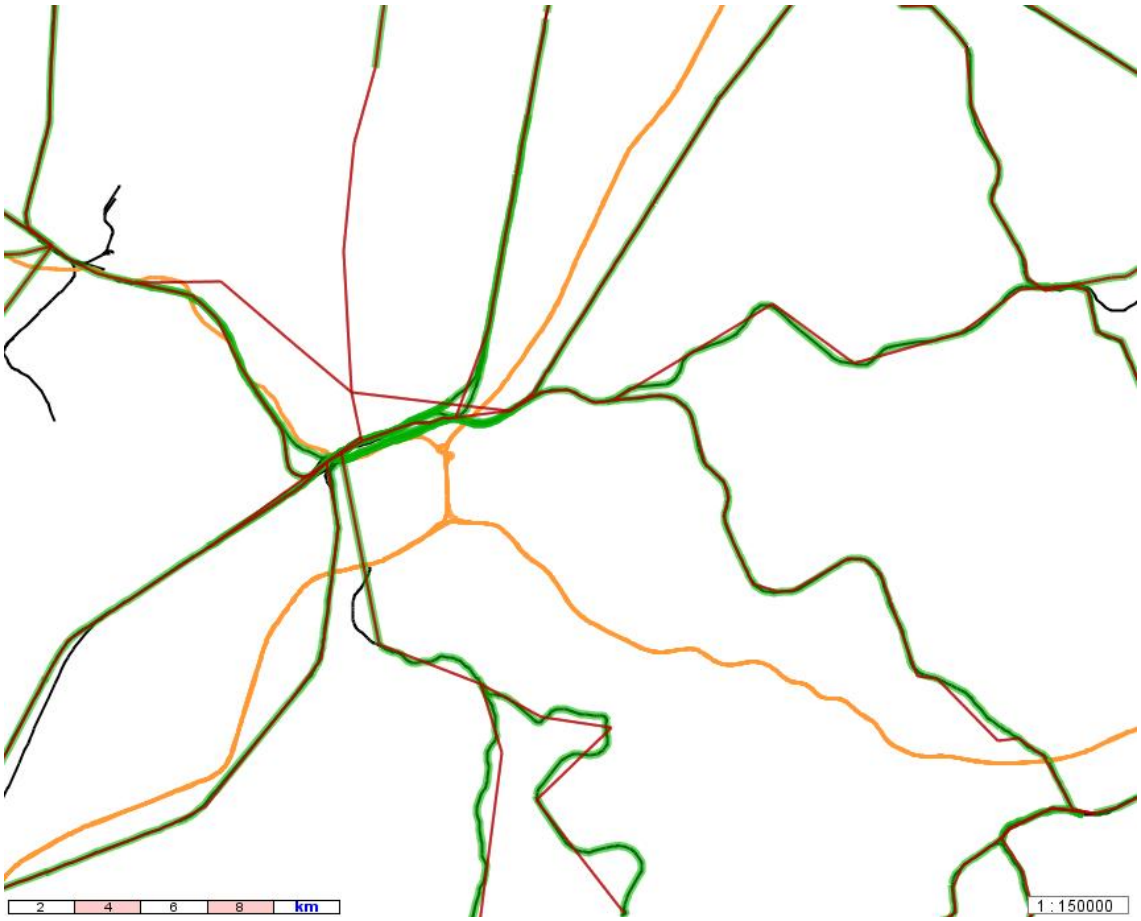


Abbildung 3-2: Korrektur der VuV-Schienennetzdaten mit den Geometrien der NAVTEQ-Schienennetzdaten

3.2 Beschreibung der Versuchsfahrten

Im Folgenden wird auf die einzelnen Versuchsfahrten eingegangen, die jeweils mit GPS als geometrische Referenz durchgeführt wurden. Die Auswertung der Identifizierung beschränkt sich auf Ortungen auf der Basis von Abis-Daten des Mobilfunknetzes. Der Abis-Abgriff erfolgte im Raum Karlsruhe innerhalb der nördlichen LAC 29453 für die Fahrten vom 13.09.2007, 07.11.2007 und 19.12.2007 sowie zusätzlich der südlichen LAC 29454 für die Fahrten am 27.02.2009.

Für die Soll-Identifikation der Versuchsfahrten wird eine Kategorisierung vorgenommen. Diese Einordnung wurde während den Versuchsfahrten handschriftlich protokolliert. Es wird zwischen „aktiv“ (A) und „passiv“ (P) im Sinne eines Verkehrsteilnehmers des MIV (motorisierter Individual-Verkehrsteilnehmer) unterschieden. Die nachfolgende Tabelle 3-3 zeigt auf, dass nur als PKW oder LKW identifizierte Mobilfunkteilnehmer als aktiv gekennzeichnet werden. An dieser Stelle sei angemerkt, dass zwischen aktiv und passiv im Sinne der Teilnahme am motorisierten Individualverkehr und den technischen Begriffen aktives (Gespräch, SMS, etc.) und passives Mobiltelefon zu unterscheiden ist.

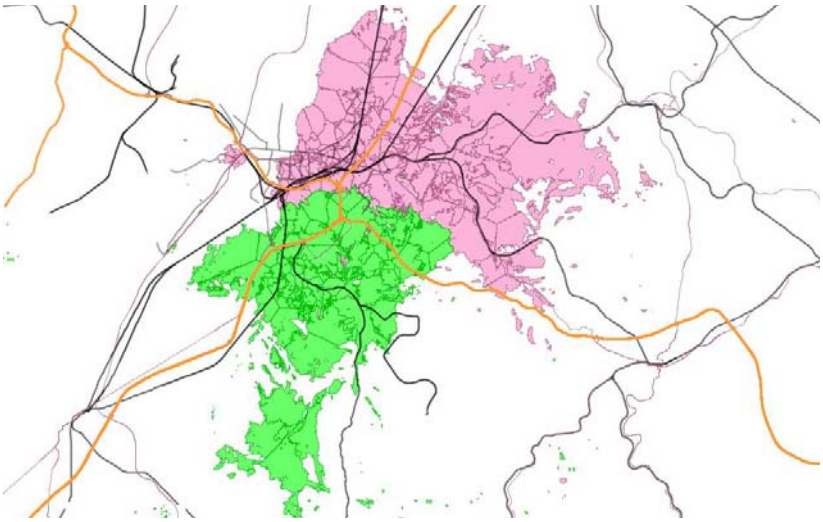


Abbildung 3-3: Abis-Demonstratorgebiet Karlsruhe (Stand Mitte 2008)

Bei einigen Referenzfahrten wurden mehrere Handys eingesetzt. Ist das der Fall, dann findet sich an der Szenarienbezeichnung ein Kürzel mit der entsprechenden Nummer, z.B. für den ersten Mobilfunkteilnehmer (MFT) Handy 1 → H1 und den zweiten MFT Handy 2 → H2. Da die Ortung für jeden MFT separat durchgeführt wird, wird jeder MFT als eigene Referenzfahrt aufgeführt und gezählt.

Nicht alle Mobilfunkdaten der unten aufgeführten Referenzerfassungen konnten erhoben werden. Gründe sind Aufenthalte außerhalb der Reichweite der angeschlossenen Funkstationen sowie Zuordnungsschwierigkeiten wechselnder MFT-Kennungen im Mobilfunknetz. Konnten Referenzdaten vollständig oder zu sehr großen Teilen nicht in den Mobilfunkdaten wiedergefunden werden, so werden sie mit durchgestrichen (~~durchgestrichen~~) gekennzeichnet und nicht ausgewertet.

Die folgende Tabelle 3-3 gibt eine Übersicht zur Anzahl und Art der durchgeführten Versuchsfahrten bei denen die Mobilfunkdaten und in gleicher Weise GPS-Daten als Referenz ausgewertet werden konnten. Bei den Mobilfunkdaten wurde versucht, möglichst viele Versuchsfahrten der Testtage mit einzubeziehen, insofern die handschriftliche Protokollierung der Soll-Identifizierung zugeordnet werden konnte. Wegen Signalstörungen konnten nicht zu allen Mobilfunkdaten auch GPS-Daten als Vergleich herangezogen werden. In den Datenzellen der Tabelle bedeutet der erste Wert die Anzahl der auswertbaren GPS-Datensätze und der zweite die Anzahl der auswertbaren Mobilfunkdatensätze.

Tabelle 3-3: Überblick über alle Versuchsfahrten mit GPS / FPD und deren Kategorisierung im Sinne aktiver Verkehrsteilnehmer

	Stand	Fuß-gänger	Fahr-rad	PKW	LKW	ÖV – Linienbus	ÖV – Straßen-bahn	ÖV – DB
Sollidentifizierung:	P	P	P	A	A	P	P	P
Erfassungstag								
13.09.2007	8 / 8	0 / 0	0 / 0	40 / 40	0 / 0	0 / 0	0 / 0	0 / 0
07.11.2007	3 / 3	3 / 3	0 / 0	0 / 0	0 / 0	2 / 2	2 / 2	0 / 0
19.12.2007	0 / 0	12 / 12	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	9 / 11
27.02.2009	0 / 0	4 / 4	1 / 1	0 / 0	0 / 0	0 / 0	2 / 2	0 / 1
Summe:	**11 / 11**	**19 / 19**	**1 / 1**	**40 / 40**	**0 / 0**	**2 / 2**	**4 / 4**	**9 / 12**

Auf den folgenden Seiten werden die einzelnen Versuchsfahrten aufgelistet und bei Besonderheiten näher erläutert. Die Versuchsfahrten sind tabellarisch zusammengefasst und die Lage sowie der Verlauf der Fahrten werden in Karten dargestellt.

Fahrten am 13.09.2007

Tabelle 3-4: Überblick über die Testfahrten am 13.09.2007

Nr.	Bezeichnung	Sollident.	Szenario	Datum/Uhrzeit	Dauer
1	Autobahn_A5_1_H1	A / PKW	Autobahn	2007.09.13 / 09:29	6 min
2	Autobahn_A5_1_H2	A / PKW	Autobahn	2007.09.13 / 09:29	6 min
11	Autobahn_A5_2_H1	A / PKW	Autobahn	2007.09.13 / 09:33	8 min
12	Autobahn_A5_2_H2	A / PKW	Autobahn	2007.09.13 / 09:33	8 min
21	Autobahn_A5_3_H1	A / PKW	Autobahn	2007.09.13 / 09:57	4 min
22	Autobahn_A5_3_H2	A / PKW	Autobahn	2007.09.13 / 09:57	4 min
31	Autobahn_A5_4_H1	A / PKW	Autobahn	2007.09.13 / 10:01	7 min
32	Autobahn_A5_4_H2	A / PKW	Autobahn	2007.09.13 / 10:01	7 min
41	Bundesstrasse_B3_1_H1	A / PKW	Bundesstrasse	2007.09.13 / 10:29	11 min
42	Bundesstrasse_B3_1_H2	A / PKW	Bundesstrasse	2007.09.13 / 10:29	11 min
51	Bundesstrasse_B3_2_H1	A / PKW	Bundesstrasse	2007.09.13 / 10:42	10 min
52	Bundesstrasse_B3_2_H2	A / PKW	Bundesstrasse	2007.09.13 / 10:42	10 min
61	Bundesstrasse_B3_3_H1	A / PKW	Bundesstrasse	2007.09.13 / 10:51	10 min
62	Bundesstrasse_B3_3_H2	A / PKW	Bundesstrasse	2007.09.13 / 10:51	10 min
71	Bundesstrasse_B3_4_H1	A / PKW	Bundesstrasse	2007.09.13 / 11:04	11 min
72	Bundesstrasse_B3_4_H2	A / PKW	Bundesstrasse	2007.09.13 / 11:04	11 min
81	Stop_a_Go_1_H1	A / PKW	Stau	2007.09.13 / 11:20	12 min
82	Stop_a_Go_1_H2	A / PKW	Stau	2007.09.13 / 11:20	12 min
91	Stop_a_Go_2_H1	A / PKW	Stau	2007.09.13 / 11:33	11 min
92	Stop_a_Go_2_H2	A / PKW	Stau	2007.09.13 / 11:33	11 min
101	Stop_a_Go_3_H1	A / PKW	Stau	2007.09.13 / 11:44	11 min
102	Stop_a_Go_3_H2	A / PKW	Stau	2007.09.13 / 11:44	11 min
111	Stop_a_Go_4_H1	A / PKW	Stau	2007.09.13 / 11:56	12 min
112	Stop_a_Go_4_H2	A / PKW	Stau	2007.09.13 / 11:56	12 min
121	Standphase_1_H1	P / Stand	Stand	2007.09.13 / 12:18	5 min
122	Standphase_1_H2	P / Stand	Stand	2007.09.13 / 12:18	5 min
131	Standphase_2_H1	P / Stand	Stand	2007.09.13 / 12:28	7 min
132	Standphase_2_H2	P / Stand	Stand	2007.09.13 / 12:28	7 min
141	Standphase_3_H1	P / Stand	Stand	2007.09.13 / 12:41	12 min
142	Standphase_3_H2	P / Stand	Stand	2007.09.13 / 12:41	12 min
151	Standphase_4_H1	P / Stand	Stand	2007.09.13 / 13:30	8 min
152	Standphase_4_H2	P / Stand	Stand	2007.09.13 / 13:30	8 min
161	Kurve_kl_1_H1	A / PKW	Parkplatzsuche	2007.09.13 / 13:41	9 min
162	Kurve_kl_1_H2	A / PKW	Parkplatzsuche	2007.09.13 / 13:41	9 min
171	Kurve_kl_2_H1	A / PKW	Parkplatzsuche	2007.09.13 / 13:50	6 min
172	Kurve_kl_2_H2	A / PKW	Parkplatzsuche	2007.09.13 / 13:50	6 min
181	Kurve_kl_3_H1	A / PKW	Parkplatzsuche	2007.09.13 / 13:57	6 min
182	Kurve_kl_3_H2	A / PKW	Parkplatzsuche	2007.09.13 / 13:57	6 min
191	Kurve_kl_4_H1	A / PKW	Parkplatzsuche	2007.09.13 / 14:03	6 min
192	Kurve_kl_4_H2	A / PKW	Parkplatzsuche	2007.09.13 / 14:03	6 min
201*	Kurve_gr_1_H1	A / PKW	Parkplatzsuche	2007.09.13 / 14:27	18 min
202*	Kurve_gr_1_H2	A / PKW	Parkplatzsuche	2007.09.13 / 14:27	18 min
211*	Kurve_gr_2_H1	A / PKW	Parkplatzsuche	2007.09.13 / 14:46	19 min
212*	Kurve_gr_2_H2	A / PKW	Parkplatzsuche	2007.09.13 / 14:46	19 min
221	Kurve_gr_3_H1	A / PKW	Parkplatzsuche	2007.09.13 / 15:05	15 min
222	Kurve_gr_3_H2	A / PKW	Parkplatzsuche	2007.09.13 / 15:05	15 min
231	Kurve_gr_4_H1	A / PKW	Parkplatzsuche	2007.09.13 / 15:20	13 min
232	Kurve_gr_4_H2	A / PKW	Parkplatzsuche	2007.09.13 / 15:20	13 min

© Do-iT 2009

241	Handy_an_Handy_1_H1	A / PKW	PKW	2007.09.13 / 16:04	14 min
242	Handy_an_Handy_1_H2	A / PKW	PKW	2007.09.13 / 16:04	14 min
251	Handy_an_Handy_2_H1	A / PKW	PKW	2007.09.13 / 16:27	23 min
252	Handy_an_Handy_2_H2	A / PKW	PKW	2007.09.13 / 16:27	23 min

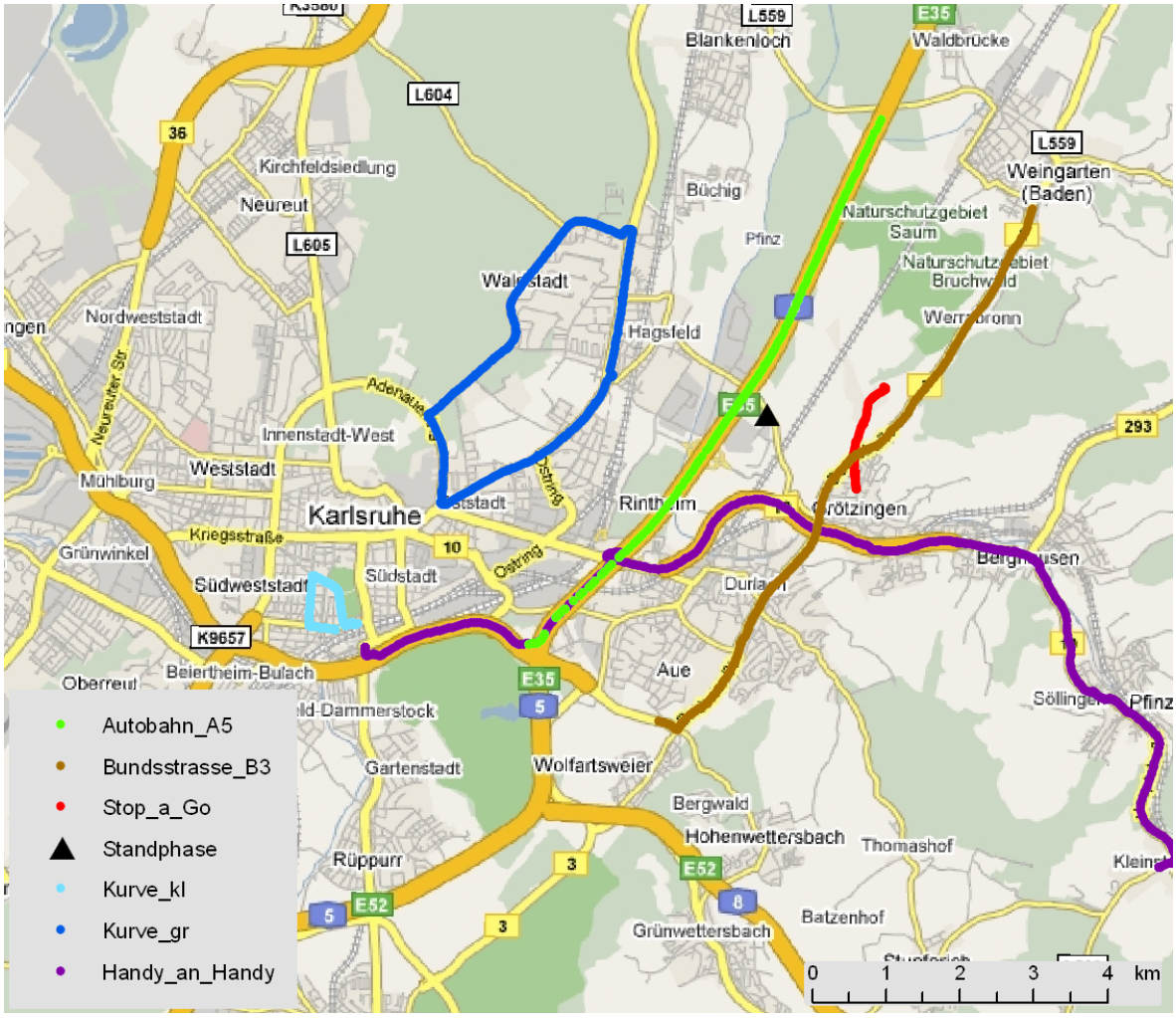


Abbildung 3-4: Trajektorien der Testfahrten vom 13.09.2007

Bemerkung: Die Referenzfahrten 81 bis 112 (Stop_a_Go) wurden, um den fließenden Verkehr nicht zu behindern, auf einer Straße abseits der üblichen Fahrwege durchgeführt. Dies führte jedoch zur Befahrung eines als Fußweg ausgewiesenen Wegabschnittes. Die Abbildung 3-5 zeigt den betreffenden Ausschnitt der Referenzfahrt 81, der repräsentativ für die Fahrten 81-112 ist. Dargestellt sind in blau das Straßennetz, in gelb das Fußwegenetz und in grün die GPS-Positionen sowie in rot die Schwerpunkte der GPS-Positionsfolge.

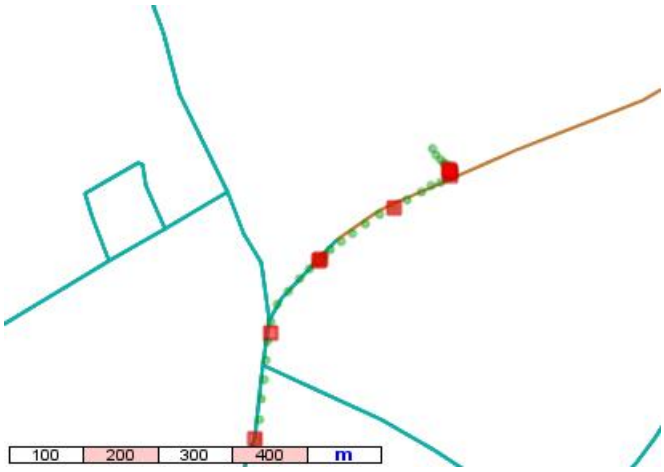


Abbildung 3-5: Ausschnitt der Referenzfahrt 81 vom 13.09.2007

In der folgenden Abbildung 3-6 ist zu den Zeitpunkten der Schwerpunkte 18 bis 22 deutlich die Nutzung der Fußwege anstatt des Straßennetzes zu erkennen. Die Grafik entstammt der Analyse der Verkehrsnetzwahrscheinlichkeit.

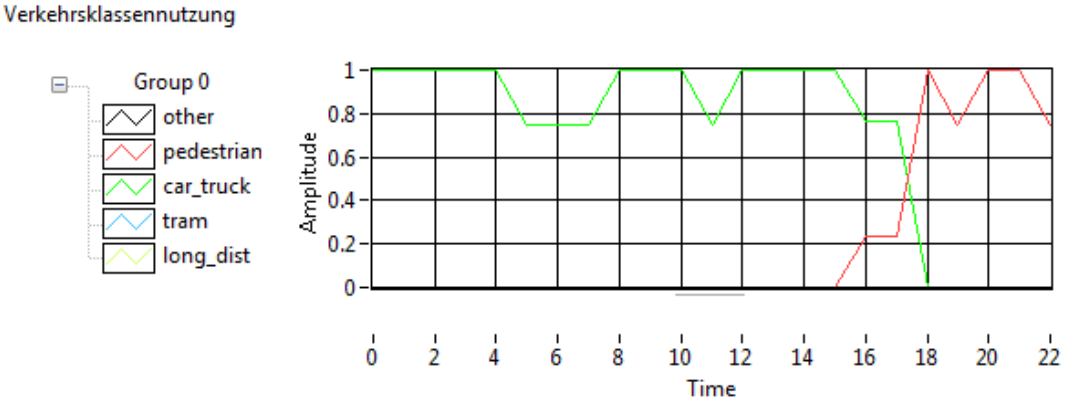


Abbildung 3-6: Verkehrsnetzzugehörigkeit der Verkehrsklassen Fußgänger und PKW-LKW (pedestrian, car_truck) bei den ungekürzten Stop_a_Go Fahrten

Eine Trennung von zusammenhängenden Datensätzen aufgrund wechselnder Verkehrsklassenzugehörigkeiten ist in der derzeitigen Version der Identifizierungsverfahren nicht implementiert. Um die Referenzfahrten dennoch auszuwerten, wurden die Fußweg betreffenden Positionen aus den Daten gelöscht. Das betrifft 81 und 82 ab 11:28:20Uhr, 91 und 92 bis 11:35:20Uhr; 101 und 102 ab 11:51:30Uhr, 111 und 112 bis 11:59:50Uhr. Die Änderung der Verkehrsnetzzugehörigkeit ist in der folgenden Abbildung 3-7 und auf der Karte in Abbildung 3-8 dargestellt. Man erkennt deutlich, dass der Datensatz nun durchgängig der Straße zugeordnet werden kann (grüne Linie). Für die Referenzfahrten 82 bis 112 gilt dies analog.

© Do-iT 2009

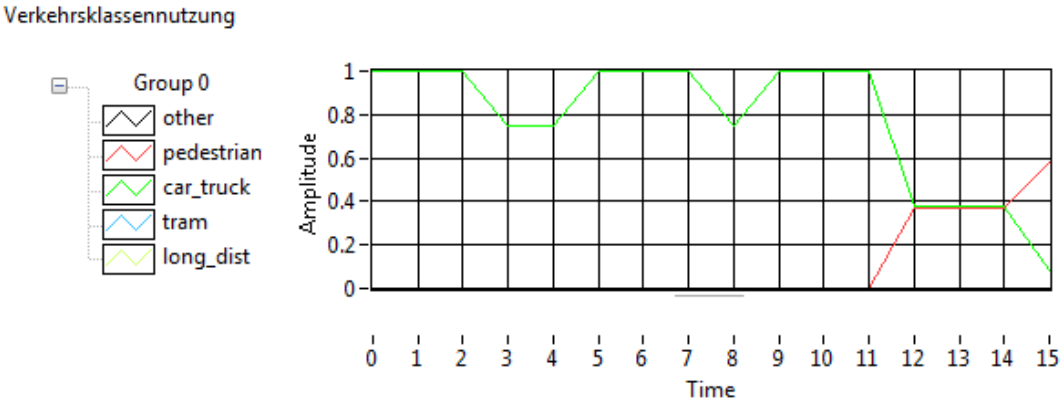


Abbildung 3-7: Verkehrsnetzzugehörigkeit für die Referenzfahrt 81 vom 13.09.2007 nach der Änderung des Datensatzes

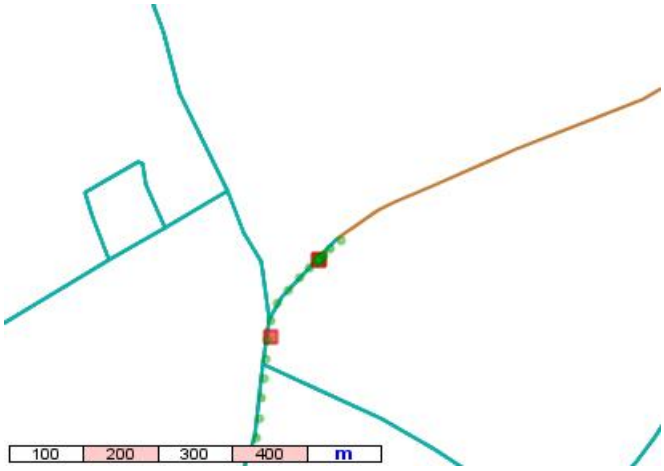


Abbildung 3-8: Darstellung der gekürzten Referenzfahrt 81 vom 13.09.2007

Fahrten am 07.11.2007

Tabelle 3-5: Überblick über die Testfahrten am 07.11.2007

Nr.	Bezeichnung	Soll-Identifizierung	Szenario	Datum/Uhrzeit	Dauer GPS FPD
1	stand 1	P / Stand	Stand	2007.11.07 / 09:12	8 min 11 min
~~2~~	~~stand 2~~	~~P / Stand~~	~~Stand~~	~~2007.11.07 / 09:31~~	~~6 min 2 sec~~
3	Bus 62 01	P / Linienbus	Linienbusfahrt Linie 62	2007.11.07 / 09:48	7 min 7 min
4	stand 3	P / Stand	Stand	2007.11.07 / 09:58	4 min 4 min
5	Bus 62 02	P / Linienbus	Linienbusfahrt Linie 62	2007.11.07 / 10:05	11 min 6 min
~~6~~	~~stand 4~~	~~P / Stand~~	~~Stand~~	~~2007.11.07 / 10:17~~	~~4 min 0 min~~
7	U-4 1	P / Straßenbahn	Straßenbahnfahrt Linie 4	2007.11.07 / 10:22	12 min 10 min
~~8~~	~~stand 5~~	~~P / Stand~~	~~Stand~~	~~2007.11.07 / 10:36~~	~~3 min 2 sec~~
9	Fußgänger 01	P / Fußgänger	Fußgänger	2007.11.07 / 10:41	7 min 12 min
~~10~~	~~stand 6~~	~~P / Stand~~	~~Stand~~	~~2007.11.07 / 11:46~~	~~3 min 1 sec~~
11	Fußgänger 02	P / Fußgänger	Fußgänger	2007.11.07 / 11:51	30 sec 3 min
12	Fußgänger 03	P / Fußgänger	Fußgänger	2007.11.07 / 12:17	9 min 3 min
13	stand7	P / Stand	Stand	2007.11.07 / 12:27	5 min 4 min
~~14~~	~~Fußgänger 04~~	~~P / Fußgänger~~	~~Fußgänger~~	~~2007.11.07 / 12:33~~	~~4 min 4 sec~~
16	U-2 1	P / Straßenbahn	Straßenbahnfahrt Linie 2	2007.11.07 / 13:01	11 min 8 min
~~17~~	~~U-2 2~~	~~P / Straßenbahn~~	~~Straßenbahnfahrt Linie 2~~	~~2007.11.07 / 13:23~~	~~13 min 0 min~~
~~50~~	~~Karlsruhe Ring_1~~	~~A / PKW~~	~~PKW Fahrt~~		
~~51~~	~~Karlsruhe Ring_2~~	~~A / PKW~~	~~PKW Fahrt~~		
~~52~~	~~Ostschleife groß~~	~~A / PKW~~	~~PKW Fahrt~~		
~~53~~	~~Ostschleife klein~~	~~A / PKW~~	~~PKW Fahrt~~		
~~54~~	~~Zickzack_1~~	~~A / PKW~~	~~PKW Fahrt~~		
~~55~~	~~Zickzack_2~~	~~A / PKW~~	~~PKW Fahrt~~	~~2007.11.07 / 12:57~~	~~29 min 0 min~~

Fahrt 7 weist große Abweichungen der Ortungspositionen von den GPS-Positionen auf. Als Problem wird eine zu optimistische Schätzung der Positionsgenauigkeiten von 100m bis 150m vermutet. Abbildung 3-9 zeigt die Situation mit Darstellung der Straßenbahnschienen in schwarz, dem Straßennetz in grau, den GPS-Positionen in grün, den Positionsschwerpunkten in rot und zwei ausgewählten Positionsschwerpunkte mit einem umgebenen Kreis mit Radius 150m zur Verdeutlichung der Abweichungen.

Abbildung 3-9: Darstellung der Fahrt 7 vom 07.11.2007 mit Kreisen von 150m Radius, die gemessene Distanz vom linken Punkt mit Kreis zu GPS beträgt ca. 400m

Die Versuchsfahrten 50 bis 55 konnten nicht aus den Mobilfunkdaten entnommen werden. Vermutlich führten häufiges Verlassen des Abis-Aufzeichnungsgebietes zu Zuordnungsproblemen der Daten und damit zu deren Verlust.

Fahrten am 19.12.2007

Tabelle 3-6: Überblick über die Testfahrten am 19.12.2007

Nr.	Bezeichnung	Soll-Identifizierung	Szenario	Datum/Uhrzeit	Dauer
1	Bahnsteig56_H1	P / Fußgänger	Gehen auf Bahnsteig KA Hbf Bahnsteig 5,6	2007.12.19 / 10:05	5 min
2	Bahnsteig56_H2	P / Fußgänger	Gehen auf Bahnsteig KA Hbf Bahnsteig 5,6	2007.12.19 / 10:05	5 min
3	Kb_H1	P / S-Bahn	S-Bahnfahrt, S31, Ka→ Bruchsal Ab 10:13 / An 10:33	2007.12.19 / 10:17	9 min
4	Kb_H2	P / S-Bahn	S-Bahnfahrt, S31, Ka→ Bruchsal Ab 10:13 / An 10:33	2007.12.19 / 10:17	9 min
5	Gk2_H1	P / RB - DB	RegioB-fahrt, RB18611, Blankenloch → Ka Ab 11:22 / An 11:32	2007.12.19 / 11:24	8 min
6	Gk2_H2	P / RB - DB	RegioB-fahrt, RB18611, Blankenloch → Ka Ab 11:22 / An 11:32	2007.12.19 / 11:24	8 min
7	K-Ettl_H1	P / S-Bahn	S-Bahnfahrt, S1, Ka→ Ettlingen Ab 12:27 / An 12:41	2007.12.19 / 12:27	14 min
8	K-Ettl_H2	P / S-Bahn	S-Bahnfahrt, S1, Ka→ Ettlingen Ab 12:27 / An 12:41	2007.12.19 / 12:27	14 min
9	Bahnhofvorplatz_H1	P / Fußgänger	Fußgänger in Ka Nähe Hbf	2007.12.19 / 13:06	7 min
10	Bahnhofvorplatz_H2	P / Fußgänger	Fußgänger in Ka Nähe Hbf	2007.12.19 / 13:06	7 min
11	Bahnsteig12_H1	P / Fußgänger	Fußgänger auf Bahnsteig	2007.12.19 / 13:21	5 min
12	Bahnsteig12_H2	P / Fußgänger	Fußgänger auf Bahnsteig	2007.12.19 / 13:21	5 min
~~13~~	~~Bahnsteig34_H2~~	~~P / Fußgänger~~	~~Fußgänger auf Bahnsteig~~	~~2007.12.19 / 13:30~~	~~5 min~~
14	Kj_1_H1	P / S-Bahn	S-Bahnfahrt, S4, Ka→ Grötzingen Ab 13:56 / An --:--	2007.12.19 / 13:58	20 min
15	Kj_1_H2	P / S-Bahn	S-Bahnfahrt, S4, Ka→ Grötzingen Ab 13:56 / An --:--	2007.12.19 / 13:58	20 min
16	Kj2_H1	P / S-Bahn	S-Bahnfahrt, S4, Grötzingen→ Jöhlingen Ab --:-- / An 14:26	2007.12.19 / 14:19	7 min
~~17~~	~~Jk_H1~~	~~P / S-Bahn~~	~~S-Bahnfahrt, S4, Jöhlingen → Ka Ab 14:30 / An 15:01~~	~~2007.12.19 / 14:36~~	~~29 min~~
18	Bahnhof_Garten_H1	P / Fußgänger	Fußgänger in Ka Nähe Hbf bis Garten	2007.12.19 / 15:33	2 min
19	Bahnhof_Garten_H2	P / Fußgänger	Fußgänger in Ka Nähe Hbf bis Garten	2007.12.19 / 15:33	2 min
20	Garten1_H1	P / Fußgänger	Fußgänger in Ka-Garten	2007.12.19 / 15:42	5 min
21	Garten1_H2	P / Fußgänger	Fußgänger in Ka-Garten	2007.12.19 / 15:42	5 min

22	Garten2_H1	P / Fußgänger	Fußgänger in Ka-Garten	2007.12.19 / 15:50	3 min
23	Garten2_H2	P / Fußgänger	Fußgänger in Ka-Garten	2007.12.19 / 15:50	3 min
24	K_s_H1	P / IC-Bahn	IC-Bahnfahrt, IC2265, Ka → Stuttgart Ab 16:06 / An 16:49	2007.12.19 / 16:08	5 min
25	K_s_H2	P / IC-Bahn	IC-Bahnfahrt, IC2265, Ka → Stuttgart Ab 16:06 / An 16:49	2007.12.19 / 16:08	5 min

© Do-iT 2009

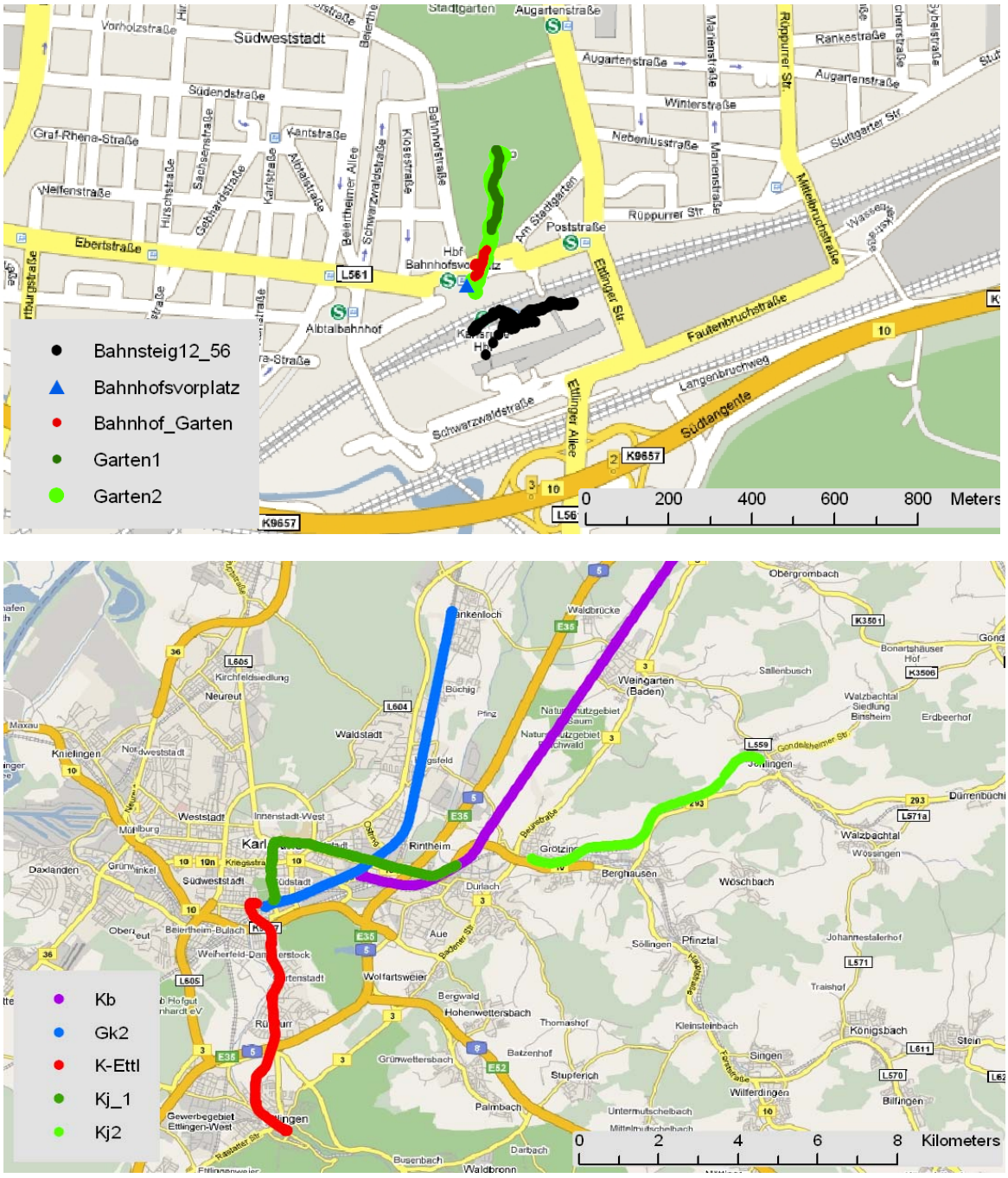


Abbildung 3-10: Trajektorien der Versuchsfahrten vom 19.12.2007

Besonderheiten für die Messfahrten vom 19.12.2007:

- Die Zeitverschiebung von einer Stunde zwischen Ortszeit und FPD-Serverzeit wurde korrigiert. Die GPS-Zeit entspricht der korrekten Ortszeit. Der FPD-Zeit folgt aus: „FPD_Zeit_korrigiert“ = „FPD_Zeit“ + 60 min
- Bei den Messfahrten 24 und 25 (IC-Bahnfahrt von Karlsruhe nach Stuttgart) sind nur für die ersten 10 Sekunden GPS-Daten vorhanden

Fahrten am 27.02.2009

Am 27. Februar 2009 wurden in Karlsruhe Versuchsfahrten zur Identifikation aktiver Verkehrsteilnehmer zu Fuß, mit dem Fahrrad, mit der Straßenbahn und mit dem Zug durchgeführt.

Tabelle 3-7: Überblick über die Testfahrten am 27.02.2008

Nr.	Bezeichnung	Soll-Identifizierung	Szenario	Datum/ Uhrzeit	Daten ID
~~1~~	~~S1~~	~~P / Stand~~	~~Frühstückstisch (Agathenstr. 33, KA), stationär~~	~~2009.02.27 / 7:00~~	
2	S2	P / Fußgänger	Fußweg zur Bahn	2009.02.27 / 7:10	101
3	S3_a1	P / Straßenbahn	Straßenbahnfahrt	2009.02.27 / 7:37	102
~~4~~	~~S3_a2~~	~~P / Straßenbahn~~	~~Straßenbahnfahrt~~	~~2009.02.27 / 7:54~~	
5	S3_b1	P / Straßenbahn	Straßenbahnfahrt	2009.02.27 / 9:28	103
6	S4	P / Fahrrad	Fahrradfahrt	2009.02.27 / 9:48	104
7	S5_a	P / Fußgänger	Gehen langgestreckter Weg und um den Block	2009.02.27 / 10:15	105
8	S5_b1	P / Fußgänger	Gehen langgestreckter Weg und um den Block	2009.02.27 / 10:43	106
~~9~~	~~S5_b2~~	~~P / Fußgänger~~	~~Gehen langgestreckter Weg und um den Block~~	~~2009.02.27 / 11:05~~	
10	S5_c1	P / Fußgänger	Gehen langgestreckter Weg und um den Block	2009.02.27 / 11:10	107
~~11~~	~~S5_c2~~	~~P / Fußgänger~~	~~Gehen langgestreckter Weg und um den Block~~	~~2009.02.27 / 11:43~~	
12	S6_1	P / DB	Zugfahrt Ka-LHS Ab 12:06 / An 12:49	2009.02.27 / 12:06	108 Kein GPS
~~13~~	~~S6_2~~	~~P / DB~~	~~Zugfahrt LHS-Ka Ab 15:08 / An 15:53~~	~~2009.02.27 / 15:08~~	

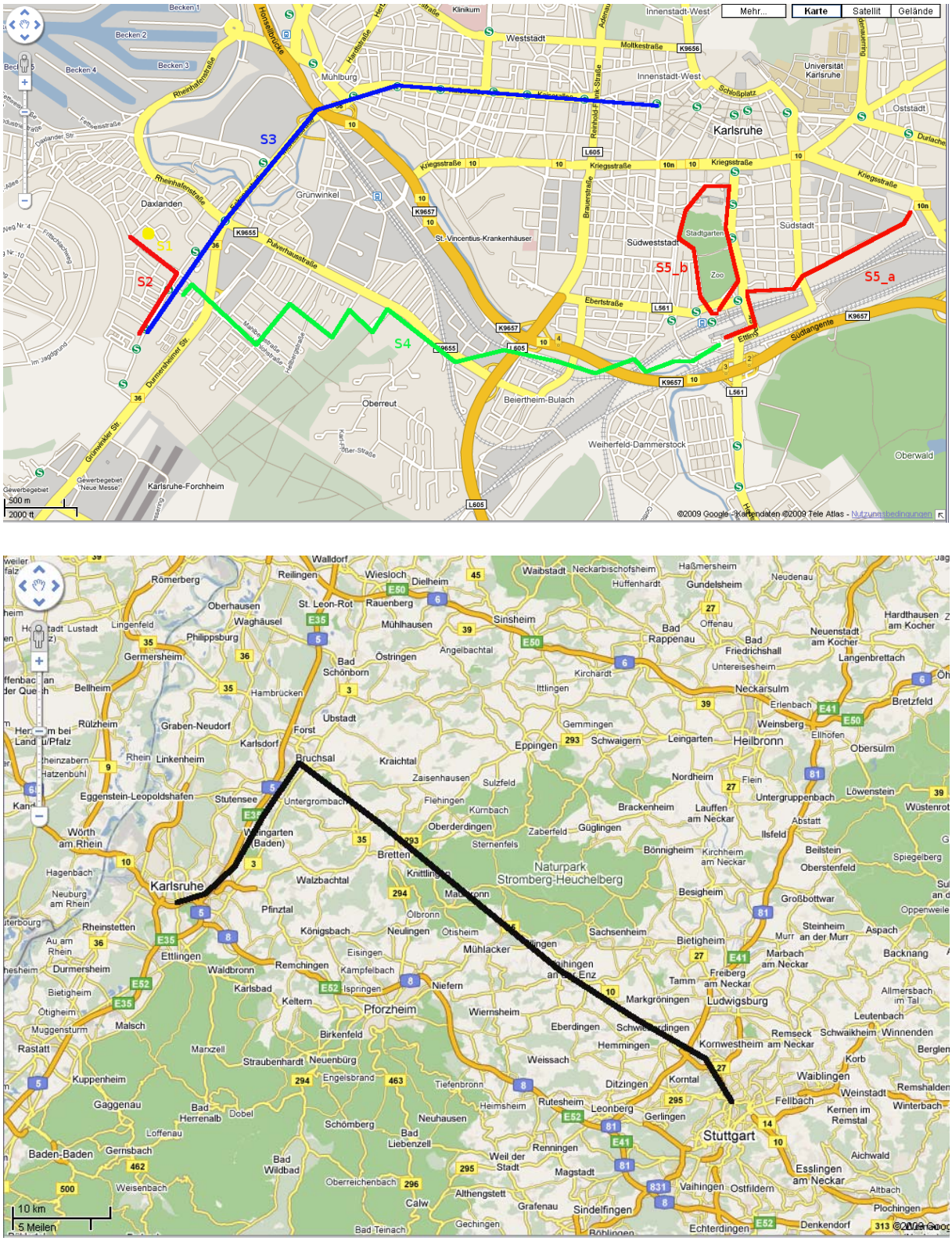


Abbildung 3-11: Trajektorien der Fahrszenarien S1-S5 (oben) und S6 (unten)

4 Analyse der Ergebnisse

In diesem Abschnitt wird zunächst ganz allgemein die Strategie der Identifizierung erläutert. Im Anschluss werden die einzelnen Identifizierungsschritte „Sektormaske“, „Bewegungsanalyse“, „Positionsschwerpunkte“, „Korrelation mit Öffentlichem Verkehr“, „Verkehrsnetzanalyse“ kurz beschrieben. Für nähere Informationen sei auf [2] verwiesen. Abschließend werden dann die Ergebnisse der einzelnen Schritte kurz zusammengefasst vorgestellt und analysiert.

4.1 Identifizierungsstrategie

Grundlegend gilt jeder MFT im Sinne der Verkehrsrelevanz zu Beginn der Identifizierung als passiv (P). Kann mit genügender Sicherheit nachgewiesen werden, dass ein ausgewählter Datensatz eines MFT als verkehrsrelevant einzustufen ist, gilt er ab dem Startzeitpunkt dieses Datensatzes als aktiv (A).

Der Grundgedanke für den Verfahrensablauf ist nicht der Nachweis der MIV-Verkehrsklasse. Vielmehr wird nach dem Ausschlussprinzip der „passiven“ Verkehrsklassen gearbeitet. Kann keine eindeutige Entscheidung getroffenen werden, also aktive und passive Verkehrsklassen sind möglich, wird der MFT als passiv eingestuft, um ggf. falsche aktiv-Identifizierungen zu unterdrücken. Eine Abwandlung dieser Strategie wird im Anschluss an die Analyse der Einzelverfahren der Identifizierung in Form verketteter MFT-Datensätze behandelt. Letztlich wird mit den Identifizierungsverfahren geprüft, ob die Merkmale der FPD-Positionen mit dem erwarteten Verhalten der Verkehrsklassen vereinbar oder nicht vereinbar sind. Daraus folgt:

- Wenn die Merkmale der FPD-Positionen mit dem Verhaltensmerkmalen der Verkehrsklasse vereinbar sind, kann die Verkehrsklasse nicht ausgeschlossen werden. Gegebenfalls wird dies durch Qualitätsmaße quantitativ bestätigt.
    - Gehört sie zur aktiven Gruppen der Verkehrsklassen, beinhaltet der übrigbleibende Suchraum (Menge der noch möglichen Verkehrsklassen) zumindest auch diese Verkehrsklasse. Es bedeutet jedoch noch nicht die Alleingültigkeit dieser aktiven Verkehrsklasse, da noch weitere ggf. konkurrierende Verkehrsklassen ebenfalls möglich sein könnten. Durch weitere Identifikationen ist der Suchraum somit weiter zu verkleinern.
    - Gehört sie zur Gruppe der passiven Verkehrsklassen, beinhaltet der Suchraum auch diese passive Verkehrsklasse. Sind nach Beendigung des letzten Identifizierungsverfahrens noch Kandidaten der aktiven und passiven Verkehrsklassen im Suchraum, kann keine eindeutige Entscheidung getroffen werden. Wie oben genannt, wird dann der MFT passiv eingestuft bzw. bei der verketteten Identifizierung (s. u. in Kapitel 4.4) nach anderen Entscheidungskriterien verfahren. Beinhaltet der Suchraum entweder passive oder aktive Verkehrsklassen, folgt die entsprechende eindeutige passive bzw. aktive Einstufung.

- Wenn die Merkmale der FPD-Positionen nicht mit den Verhaltensmerkmalen der Verkehrsklasse vereinbar sind, kann diese Verkehrsklasse ausgeschlossen werden. Die Angabe von Qualitätsmaßen ist hier unter Umständen nicht möglich, da im Grunde genommen die Untersuchung, und damit das verwendete funktionale Modell, auf einem nicht geeigneten Datensatz basieren.

Das Ausschlussprinzip passiver Verkehrsklassen kann also nicht uneingeschränkt als Bestätigung der aktiven Verkehrsklasse dienen. Es kann jedoch zur Eingrenzung und ggf. zur Eingrenzung auf die interessierende Verkehrsklasse des MIV führen. Letztlich handelt es sich aber um die Prüfung, ob ein Nicht-Ausschluss der MIV Verkehrsklasse gerechtfertigt ist.

4.2 Kurzbeschreibung der Identifizierungsschritte

4.2.1 Sektormaske

Mit dem Verfahren „Sektormaske“ wird für die Optimierung des Online-Betriebes eine Priorisierung der Mobilfunkteilnehmer vorgenommen. Eine Identifikation findet somit mittelbar nicht statt. Vielmehr wird der Parameter PZ (Priorisierungszahl) für die Priorisierung der Teilnehmer für die anschließende Auswertung ermittelt. Zur Berechnung der Priorisierungszahl werden folgende teilnehmerspezifischen Kriterien herangezogen:

1. H_{ap}: Grad der IV-Aktivität (**H**istorie): Teilnehmer war **a**ktiv = 1, **p**assiv = 0
2. Grad der Verkehrsrelevanz:
   - TAC_{Vr}: **TAC** Identifizierung für **V**erkehrs**r**elevanz (0..1)
   - P_{Sektor}: arithmetisches Mittel der Wahrscheinlichkeit (**P**) der Verkehrsrelevanz bestimmt mittels der erweiterten **Sektor**en (0..1)
3. Grad der Bewegung:
   - Anzahl unterschiedlicher Sektoren (0,1,2 ... $n_{Sektoren}$)
   - Anzahl Messelemente (0,1,2 … $n_{Messelemente}$)

Zur effizienten Abarbeitung der Rangliste wird aus den Kriterien die Priorisierungszahl (PZ) mit der Formel (4-1) abgeleitet. Je höher diese Zahl ist, desto höher ist die Priorität des Teilnehmers für die Auswertung.

$$PZ = \mathrm{int}\left(\begin{array}{ll} H_{ap} & \cdot 100\,000 \\ + TAC_{Vr} & \cdot 1000\,000 \\ + P_{Sektor} & \cdot 10\,000 \\ + \frac{n_{Sektoren}}{n^{max}_{Sektoren}} & \cdot 1000 \\ + \frac{n_{Messelemente}}{n^{max}_{Messelemente}} & \cdot 100 \end{array} \right) \qquad (4\text{-}1)$$

$$\text{mit}: n^{max}_{Sektoren} \quad \text{bislang maximale Anzahl Sektoren je Trajektorie}$$

$$n^{max}_{Messelemente} \quad \text{bislang maximale Anzahl Messwerte je Trajektorie}$$

Vermeintlich „interessante“ Teilnehmer werden für anschließenden Auswerteprozesse, beginnend mit der Ortung, dann Identifizierung etc., an den Anfang gestellt. Aufgrund der begrenzten Auswertekapazität im Online-Betrieb können zumeist nicht alle Teilnehmer abgearbeitet werden. In der Folge gelangen ggf. „interessante“ Teilnehmerdaten nicht in den Auswerteprozess, weshalb die Vermutung eines fehlerhaften Ausschlusses geschlussfolgert werden könnte. Dies wird jedoch als Priorisierungsfehler und nicht Identifizierungsfehler aufgefasst und daher in diesem Bericht nicht behandelt.

Im Offline-Betrieb werden sämtliche Teilnehmer und Teilnehmerdaten bearbeitet, unabhängig von der Priorisierung. Die Reihenfolge der abgearbeiteten Teilnehmer ist aufgrund der Priorisierung ggf. nicht mit dem zeitlichen Eintreffen der Teilnehmerdaten gleich. Die zeitlichen Beschränkungen infolge begrenzter Auswerteressourcen wirken sich auf das Ergebnis nicht aus, da die vollständige Bearbeitung abgewartet wird. Die hier vorgestellten Identifizierungsergebnisse beruhen ausschließlich auf Auswertungen im Offline-Betrieb.

4.2.2 Bewegungsanalyse

Für die Identifikation, ob der MFT steht oder sich bewegt, wird die Hypothese der Bewegung des FPD angenommen und im Anschluss überprüft, ob die Annahme zutrifft. Fand eine Bewegung des FPD statt, so soll dieses Verhalten mittels einer Signifikanzprüfung der räumlich genügend weit auseinander liegenden Positionsschwerpunkte, siehe nachfolgenden 4.2.3, nachgewiesen werde können.

4.2.2.1 Szenarien zur Bewegungserkennung mittels Positionsschwerpunkten

Die Eigenschaft einer FPD-Bewegung zeigt sich in der Signifikanz der Gesamtdistanz für den zurückgelegten Weg. Die Gesamtdistanz kann mittels Aufaddierung der Koordinatendifferenzen der Positionsschwerpunkte berechnet werden. Die Gesamtdistanz ist abhängig von der Generalisierung der Geometrie der FPD-Positionsfolge und somit auch von der Anzahl und der räumlichen Anordnung der berechneten Positionsschwerpunkte. Der Einfluss der Genauigkeiten der Einzelortungen auf die Positionsschwerpunkte steigt letztlich

mit der Anzahl eben dieser Positionsschwerpunkte, ebenso steigt der erforderliche Rechenbedarf.

Das Ziel ist daher die Repräsentation der generalisierten Positionsfolge im Sinne der Gesamtdistanzbestimmung möglichst realistisch und die Anzahl der Positionsschwerpunkte dabei möglichst gering zu halten.

Aus geometrischen Gesichtspunkten genügen für geradlinige Bewegungen zwei Positionsschwerpunkte, siehe Abbildung 4-1 Szenario „geradlinig“, zur Repräsentation. Finden Fahrten mit Hin- und Rückweg statt, muss mindestens ein weiterer Positionsschwerpunkt berechnet werden. In Abbildung 4-1 wird diese Szenario mit „gebogen 1. Art“ bezeichnet. Würden nur zwei Positionsschwerpunkte berechnet werden, könnten sie räumlich in unmittelbarer Nähe liegen, wodurch die Distanz falsch geschätzt werden würde. Bei einer angenommen Bewegung mit zweifacher Kreuzung des Weges, wie es in Abbildung 4-1 im Szenario „gebogen 2. Art“ dargestellt ist, sind bereits mindestens vier Positionsschwerpunkte erforderlich. Je komplizierter die Bewegungsformen werden, bei denen in die Nähe früherer Positionen zurückgekehrt wird oder Wege gekreuzt werden, desto mehr Positionsschwerpunkte sind zur Aufdeckung der Gesamtdistanz notwendig. Kompliziertere Wegformen als „gebogen 2. Art“ werden in Do-iT nicht betrachtet.

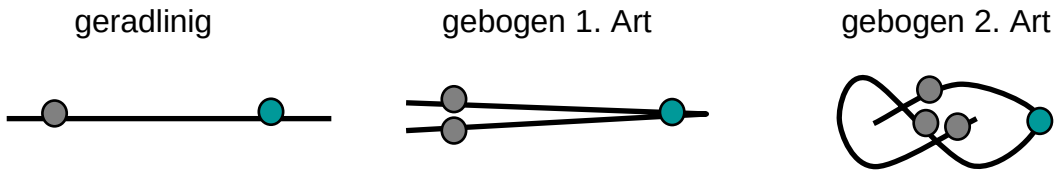


Abbildung 4-1: Szenarienformen zur Erkennung von Bewegungen

4.2.3 Positionsschwerpunkte

Folgt aus der Bewegungsanalyse, dass sich das FPD bewegt, ist die Annahme mehrerer FPD-Schwerpunkte für die Positionsfolge des FPD gerechtfertigt. In der Bewegungsanalyse wurden basierend auf den drei behandelten Szenarienformen jedoch nur maximal vier Positionen berechnet. Um die geometrische Form und den Verlauf der Geschwindigkeit repräsentativer zu bestimmen, sind besonders längere Positionsfolgen zeitlich und räumlich höher aufzulösen. Bei der Wahl der FPD-Schwerpunktanzahl sind unter anderem die folgenden Aspekte abzuwägen: je höher die FPD-Schwerpunktanzahl desto höher ist die räumliche und zeitliche Auflösung, desto mehr können sich Ungenauigkeiten auswirken, desto geringer ist die Informationsdichte und desto mehr Rechenressourcen benötigt das Clusterverfahren.

In der Bewegungsanalyse und für die Berechnung der Positionsschwerpunkte mit höherer Auflösung wird das gleiche Clusterverfahren verwendet. Es handelt sich um eine Variation

des Fuzzy-C-Means Clusterverfahren nach Gustafson-Kessel. Das FCM-Gustafson-Kessel berechnet als Repräsentanten der Cluster Schwerpunkte deren Koordinaten als die Positionsschwerpunkte der FPD-Positionen interpretiert werden. Kurz zusammengefasst arbeitet das iterative Clusterverfahren nach FCM wie folgt:

Gegeben seien die Eingangsdaten X
Wähle die Clusteranzahl $2 \le c < n$
Wähle das Abbruchkriterium ε
Initialisiere die Zugehörigkeitsmatrix U
  Wiederhole die folgende Berechnung bis mindestens ε erreicht ist:
    Berechne Zugehörigkeitsmatrix U
    Berechne Clusterzentren V

Das Abbruchkriterium der Iteration ist erreicht, wenn die Differenz der Zugehörigkeitsmatrix des vorherigen Iterationsschrittes zum aktuellen kleiner $\varepsilon \geq \left(\left\| U^{(i-1)} - U^{(i)} \right\| \right)$ ist. Die Zugehörigkeitsmatrix berechnet sich aus:

$$u_{i,k} = 1 \Big/ \sum_{j=1}^{c} \left(\frac{\left\| x_k - v_i \right\|}{\left\| x_k - v_j \right\|} \right)^{\frac{2}{m-1}} \tag{4-2}$$

und die Clusterzentren folgen mit:

$$v_i = \frac{\sum_{k=1}^{n} \left(u_{i,k}^m \cdot x_k \right)}{\sum_{k=1}^{n} u_{i,k}^m} \,. \tag{4-3}$$

Die Variation nach Gustafson-Kessel besteht in der Berechnung der Distanzen zwischen den Datenpunkten x_k und den Clusterzentren v_i. Sie werden durch eine aus den Datenpunkten und Clusterzentren berechneten Kovarianzmatrix in Form einer ellipsoidischen Verteilung mit der Mahalanobis-Distanz gewichtet:

$$d\left(x_k, v_i\right) = \sqrt{\left(x_k - v_i\right)^T \mathbf{A}_i^M \left(x_k - v_i\right)} \tag{4-4}$$

Die Matrix $\mathbf{A}_i^M$ folgt aus:

$$\mathbf{A}_i^M = \sqrt[p]{\det\left(\mathbf{S}_i\right)} * \mathbf{S}_i^{-1} \quad \text{mit} \quad \mathbf{S}_i = \frac{\sum_{k=1}^{n} u_{k,i}^m \cdot \left(x_k - v_i\right) \cdot \left(x_k - v_i\right)^T}{\sum_{k=1}^{n} u_{k,i}^m} \tag{4-5}$$

Zu beachten ist, dass die zeitlichen Abstände der berechneten Schwerpunkte ebenfalls abhängig von der räumlichen Lage sind und daher nicht konstant sind. Die Abbildung 4-2 zeigt reale Beispieldaten für die berechneten FPD-Schwerpunkte, rote Punkte mit roten

Verbindungslinien. Die zu den jeweiligen FPD-Schwerpunkten am meisten zugehörigen FPD-Positionen sind durch grüne Punkt mit grünen Verbindungslinien dargestellt.

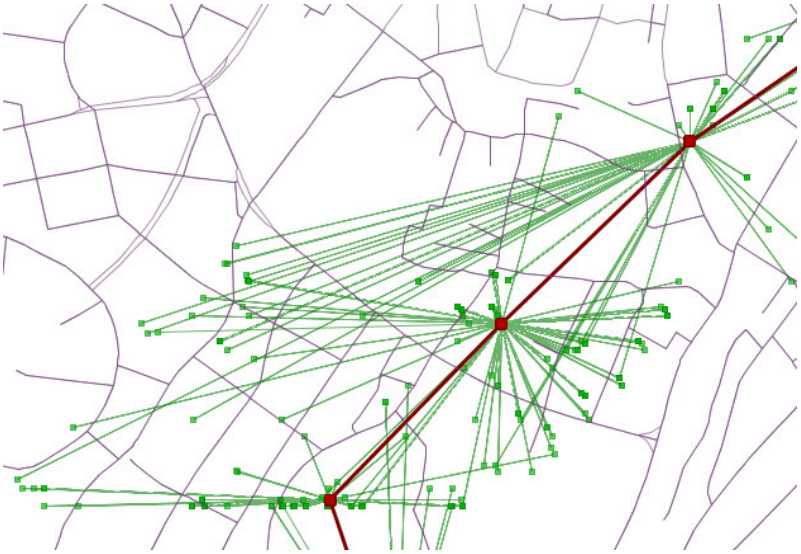


Abbildung 4-2: FPD-Positionsschwerpunkte (rot) für sich bewegende FPD mit Verbindungslinien zu den zugeordneten FPD-Positionen (grün)

4.2.4 Korrelation mit Öffentlichem Verkehr

Zur Analyse der FPD-Daten gegenüber öffentlichen Verkehren, werden die geplanten Fahrzeugpositionen (PlanÖV) auf der digitalen Karte des Lokalisierungsnetzes projiziert. Aus den Fahrplänen können hierfür die An- und Abfahrzeiten der Fahrzeuge an den Haltestellen und somit auch die Fahrzeiten zwischen den Haltestellen entnommen werden. Anhand der Fahrplandaten werden die Fahrwege geroutet, die zeitliche Abfolge an den Knotenpunkten mit An- und Abfahrtszeiten berechnet. Die Fahrwege wurden anschließend verkachelt und die planmäßigen Aufenthaltszeiten eines ÖV-Fahrzeuges innerhalb der jeweiligen Kachel prognostiziert. Die planmäßigen Aufenthaltszeiten wurden zusätzlich um Fahrplanabweichungen (Verfrühungen und Verspätungen) fuzzyfiziert (unscharfe zeitliche Zuordnung). Diese Berechnungen erfolgten in einer Vorprozessierung. Für den Online-Betrieb wurden effiziente Datenzugriffe implementiert, die anhand der zu untersuchenden Kachel der FPD-Position und der Aufenthaltszeit die ÖV-Zugehörigkeit berechnen. Durch die vorprozessierte Fuzzyfizierung der Fahrplanzeiten, ist lediglich eine lineare Interpolation der Aufenthaltszeit des FPD zwischen zwei Zugehörigkeitswerten der Fuzzyfizierung je Kachel erforderlich, vgl. nachfolgende Abbildung 4-3 zum Zeitpunkt t_i.

© Do-iT 2009

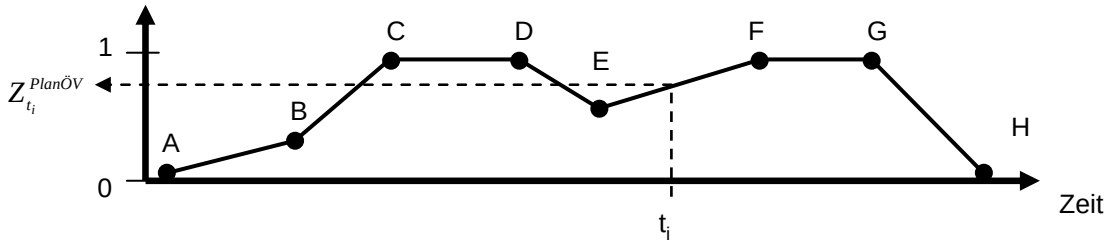


Abbildung 4-3: Interpolation des ÖV-Zugehörigkeitswertes

Wie bereits in der Tabelle 3-2 der Datenversorgung aufgeführt, stammen die Fahrplandaten aus dem Gültigkeitsraum des Jahres 2007 oder des ersten Halbjahres 2007 (DB). Versuchsfahrten die innerhalb dieses Gültigkeitsraumes stattfanden, werden auf diesen Datenbestand angewendet. Das betrifft die Versuchsfahrten 13.09.2007 und 07.11.2007 für die Linienbusdaten und U-/Straßenbahnfahrten. Die Fahrten des 19.12.2007 sind bereits außerhalb des Gültigkeitsraumes. Insbesondere befanden sich alle Versuchsfahrten für DB nicht in dem Gültigkeitsraum (Dez. 2006 bis Jun. 2007, Datenbereitstellung Nov. 2007).

Um dennoch das Verfahren der ÖV-Korrelation anzuwenden, fand eine Prognose der Fahrplandaten in die Zeiträume der Versuchsfahrten und des Demonstratorzeitraumes statt. Aufgrund der sehr heterogenen Daten sowie Uneindeutigkeiten der Betriebtagsdefinition wurde unterstellt, dass für die prognostizierten Gültigkeitszeiträume die gleichen Betriebstage (einschließlich Ferien- und Feiertage) gelten.

Verschiedene Ursachen, wie Fahrgastwechsel, Verkehrslage, Baustellen u. a., führen beim Betrieb der ÖV-Verkehre zu Abweichungen von den geplanten Zeitangaben. Derartige Fahrplanabweichungen werden in der Vorprozessierung mit angenommenen Verfrühungen und Verspätungen mittels einer Fuzzyfizierung der Kachelein- und Kachelausfahrzeiten der geplanten ÖV-Fahrzeugtrajektorien modelliert. Für Linienbusse werden Verfrühungen von 60 Sekunden und Verspätungen von bis zu 180 Sekunden, für Straßenbahnen 30 Sekunden Verfrühung und 120 Sekunden Verspätung sowie für DB Bahnen bis zu 120 Sekunden Verspätung jeweils als graduelle (monoton linear abfallend) Zugehörigkeitsfunktionen zur geplanten Ankunftszeit und Abfahrtszeit berücksichtigt, vgl. nachfolgende Abbildung 4-4.

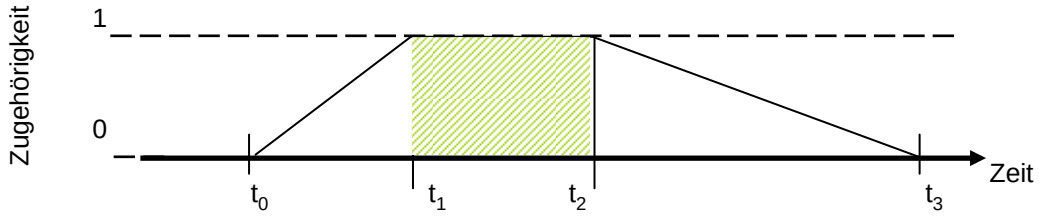


Abbildung 4-4: Schematische Darstellung der Modellierung von Fahrplanabweichungen, mit t_1-t_0=maximal Verfrühung und t_3-t_2=maximale Verspätung

Zum Identifikationsverfahren: Für den untersuchten MFT-Datensatz wird für jeden Schwerpunkt geprüft, ob eine Zugehörigkeit zum entsprechenden ÖV-Angebot besteht. Ist das für mindestens 80% der Schwerpunkte der Fall, wird die Nutzung dieses ÖV-Angebots als möglich eingestuft. Die entsprechende Verkehrsklasse und ggf. das dazugehörige Verkehrsnetz werden dann nicht ausgeschlossen. Ferner muss angenommen werde, dass die aktiven Verkehrsklassen (MIV) nicht mehr eindeutig heraustrennbar sind. Der MFT-Datensatz wird daher als passiv eingestuft. Ob zusätzlich auch eine der aktiven Verkehrsklassen möglich wäre, wird dann wegen der Untrennbarkeit nicht mehr untersucht.

4.2.5 Verkehrsnetzanalyse

Die Verkehrsnetzanalyse unterteilt sich in zwei Bereiche. Der erste analysiert auf Geometrieebene die Zugehörigkeit der berechneten Schwerpunkte einschließlich der fehlerzeigenden Figuren zu den Verkehrsnetzen. Die Abbildung 4-5 zeigt schematisch die in der Vorprozessierung ermittelte Zugehörigkeit von Verkehrsnetzen zu einer Kachel. Alle dargestellten Verkehrsnetze in dieser Abbildung besitzen einen Zugehörigkeitswert größer null.

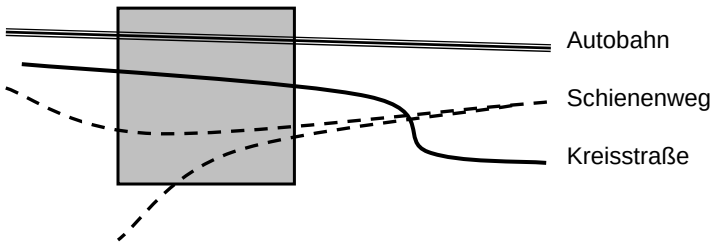


Abbildung 4-5: Schematisierte Darstellung eines diskretisierten Geometriebereiches zur effizienten Analyse der Verkehrsnetzzugehörigkeit

In der nachfolgenden Abbildung 4-6 sind in Form kleiner Kuchendiagramme die Verkehrsnetzzugehörigkeiten der Kacheln für einen Ausschnitt in Höhe Besigheim (ca. 25 km nördlich von Stuttgart) farbig visualisiert. Etwa in Bildmitte verläuft in Nord-Süd-Richtung die Bahnverbindung Heilbronn-Stuttgart, schwarze Kuchendiagrammanteile. Zur Orientierung ist die B27 in orange dünn eingezeichnet.

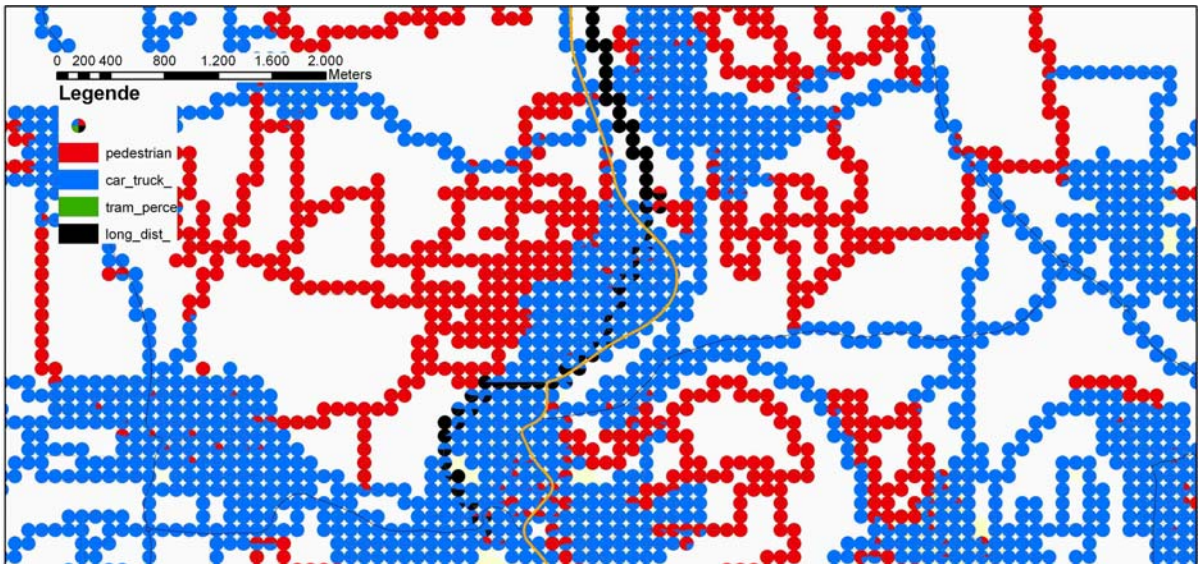


Abbildung 4-6: Kuchendarstellung der Verkehrsnetzzugehörigkeit der Kacheln, nördlich von Stuttgart in Höhe Besigheim, in dünn orange ist die B27 eingezeichnet, Kachelgröße 100x100 m^2

Nur wenn zu jedem Schwerpunkt und den dazugehörigen Kacheln mindestens ein Element des entsprechenden Verkehrsnetzes gefunden werden kann, ist auch von einer möglichen Route, also von der Machbarkeit der Befahrungsmöglichkeit auf diesem Verkehrsnetz zu rechnen. Ist dies für ein oder mehrere Verkehrsnetze nicht der Fall, so wird dieses Verkehrsnetz einschließlich der darauf angenommenen Verkehrklassen für den MFT-Datensatz ausgeschlossen.

Zu den hier eingesetzten Verkehrsnetzen und den zugeordneten Verkehrsklassen gehören:

Tabelle 4-1: Definition der Verkehrsklassen/-gruppen für die Verkehrsnetzklassifizierung

Nr.	Name der Verkehrsnetzgruppe	Beschreibung	
		Verkehrsnetz	Verkehrsklasse(n)/-gruppe
0	other	Sonstige Verkehrsnetze	keine zugeordnete Verkehrsklasse
1	pedestrian	Verkehrsnetze für Fußgänger	Fußgänger; Straßenkategorien der NAVTEQ-Verkehrsnetze
2	car_truck	Fahrzeuge des Individualverkehrs	PKW, LKW, Bus; Straßenkategorien der NAVTEQ-Verkehrsnetze
3	tram	Verkehrsnetz der ÖV Straßenbahnen	Schienennetze der Straßenbahnen Stuttgart und Karlsruhe
4	long_dist_rail	Schiene der Fern- und Regionalbahnen	Regional und Fernzüge (DB)

Die folgende Abbildung 4-7 zeigt die Programmoberfläche zur Darstellung der Verkehrsnetzanalyse während der Identifikation. Die Grafikanzeige visualisiert im oberen Teil für jeden Schwerpunkt die relativen Wahrscheinlichkeiten der Verkehrsnetze. Im Beispiel sind für 17 Schwerpunkte die Wahrscheinlichkeitswerte zu erkennen. Dargestellt in rot für das

Fußwegenetz (pedestrian), in grün für PKW und LKW (car_truck), in blau für Straßenbahnen (tram), in gelb für das Schienennetz der DB (long_dist) und in weiß für andere nicht thematisch zugeordnete Netze (other). Im Lokalisierungsnetz ist jedem Kartenelement mindestens ein Verkehrsnetz zugeordnet, weshalb keine Elemente in weiß angezeigt werden. Ob durchgängig für jeden Schwerpunkt eine Wahrscheinlichkeit größer Null eines Verkehrsnetzes vorhanden ist, kann aus der Zahlenzusammenstellung im unteren Teil, dritte Zahlenspalte von links entnommen werden. Ein negativer Wert oder Null weist auf kein oder kein durchgängiges Vorhandensein des Verkehrsnetzes hin. Eine Route kann folglich in diesen Fällen nicht gefunden werden. Ein positiver Wert größer null zeigt, dass zumindest Verkehrsnetzelemente des Verkehrsnetzes vorhanden sind. Ob topologisch und in den zeitlichen Restriktionen tatsächlich eine Route zwischen den Schwerpunkten möglich ist, wird im anschließenden Verfahren untersucht.

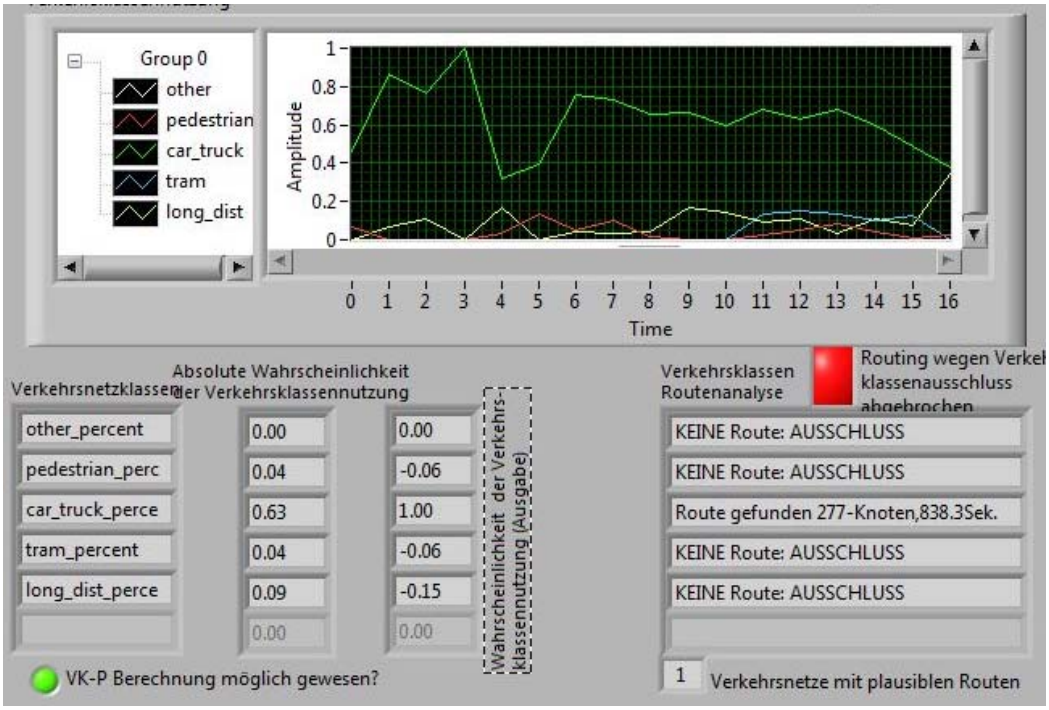


Abbildung 4-7: Programmoberfläche zur Beobachtung der Auswertung zur Verkehrsnetzanalyse während der Identifikation

Im zweiten Bereich der Verkehrsnetzanalyse wird die topologische sowie zeitliche Machbarkeit einer möglichen Befahrung analysiert. Hierfür werden für jeden Schwerpunkt sowie den fehlerzeigenden Figuren die Zeiten ab wann und bis wann sich der Teilnehmer auf den Elementen der Verkehrsnetze aufgehalten hat, zusammengestellt, vgl. Abbildung 4-8. Es wird angenommen, dass sich ein Teilnehmer maximal mit einer verkehrsnetzspezifischen Maximalgeschwindigkeit sowie ausschließlich auf topologisch verknüpften Verkehrsnetzelementen auf dem entsprechenden Verkehrsnetz fortbewegen kann. Um eine zeitlich günstigere Auflösung sowie eine bessere Zugriffsperformance zu erhalten, wurden die Verkehrsnetze zuvor verkachelt sowie an den Kachelgrenzen in Unter-Verkehrsnetze (subnet) transformiert.

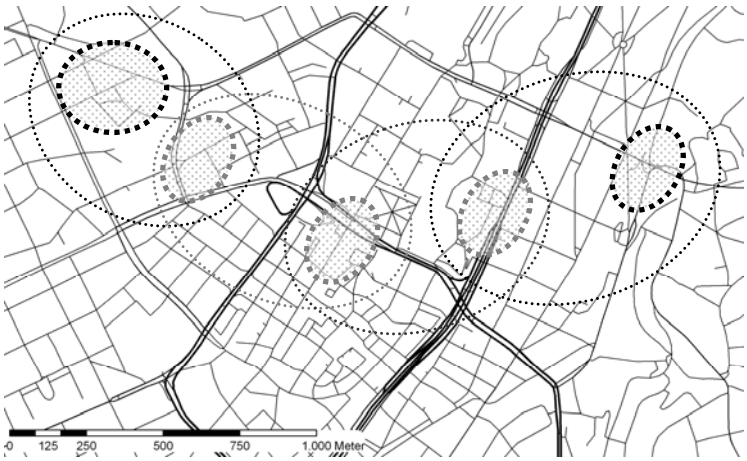


Abbildung 4-8: Mögliche Korridore einer FPD-Positionsfolge

Mit einem Routing-Algorithmus der sowohl klassisch topologische Routen bildet, als auch während des Routings die oben genannten zeitlichen Restriktionen kontrolliert, wird die prinzipielle Machbarkeit der Befahrung geprüft. Es wird hierbei versucht eine Route auf den Elementen des Verkehrsnetzes zu finden, die durch die zeitlich getrennten Korridoren verläuft, vgl. Abbildung 4-9. Es ist nicht vorrangiges Ziel die wahrscheinlichste Route zu ermitteln, alleinig die Existenz einer beliebigen, den Kriterien genügende Route genügt, um die entsprechenden Verkehrsklassen des behandelten Verkehrsnetzes nicht auszuschließen. Im Umkehrschluss erfolgt natürlich bei Nicht-Existenz einer Route der Ausschluss. Aus der Definition heraus zwischen zeitlich getrennten Korridoren eine Route zu suchen, ist es ebenso möglich Trajektorien mit zwischenzeitlich stehendem MFT oder auch Routen mit wiederholtem Aufenthalt auf bereits befahrenen Verkehrsnetzelementen zu berechnen.

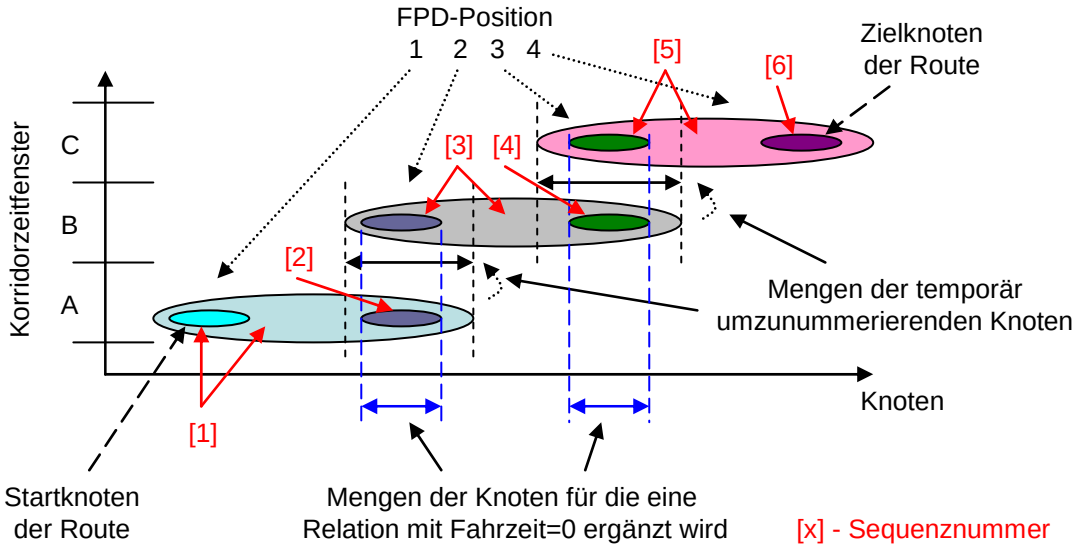


Abbildung 4-9: Topologische und zeitliche Darstellung des Routings mit Zeitrestriktionen

Wie bereits im Bericht zu AP4.5 [3] „Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren“ (Oktober 2008) hingewiesen wurde, sind erhebliche Abweichungen der FPD-Ortung am Anfang und zum Teil auch am Ende der Positionsfolgen festzustellen. Infolgedessen können für die Identifizierung mittels topologisch-zeitlicher Routen-Analyse ggf. auf keinem der untersuchten Verkehrsnetz Routen gefunden werden, obwohl eine Bewegung eindeutig nachgewiesen wurde. Ursächlich werden grobe und systematische Positionssprünge vermutet. Zur Verringerung des so resultierenden Ausschlusses evtl. aktiver Trajektorien, wurde in den Identifizierungsablauf zusätzlich folgender Ansatz implementiert:

Wurde für eine Trajektorie 1. eine Bewegung nachgewiesen, besteht 2. keine signifikante ÖV-Korrelation und es konnte 3. keine Route auf einem der Verkehrsnetze gefunden werden, besteht jedoch 4. für mindestens 80% der Positionsschwerpunkte eine Zugehörigkeit zum Verkehrsnetz „car-truck“, dann kürze die Trajektorie. Die Kürzung erfolgt alternierend erst am Anfang, dann am Ende, am Anfang usw.. Wurde eine Kürzung durchgeführt, startet beginnend nach der Schwerpunktbestimmung der Identifizierungsprozess für den MFT-Datensatz erneut. Es wird solange mit der Kürzung fortgefahren bis ein Nachweis über die Identifizierung erbracht wurde oder maximal 20% der Trajektorie abgeschnitten wurden. Endet diese Prozedur ohne den Nachweis einer den zeitlichen und topologischen Kriterien genügenden Route, wird der Datensatz passiv eingestuft.

4.3 Vergleich mit GPS-Referenzdaten und empirische Analyse

Für alle hier genannten Identifizierungsverfahren gilt, dass sie nur ausgeführt werden, wenn der (Mobilfunkteilnehmer) MFT nicht bereits durch ein vorhergehendes Identifizierungsverfahren eindeutig passiv eingestuft worden ist. Wird erstmalig ein Datensatz eines MFT untersucht, gilt die passiv-Annahme. Erst wenn ein aktives (MIV) Verhalten nachgewiesen werden konnte, werden auch nachfolgende ggf. stehende MFT-Datensätze als Trajektorien an die folgenden Do-iT Prozesse weitergegeben. Diese Einschränkung wurde aufgrund der geringen Positionsgenauigkeit der Mobilfunkdaten eingeführt. MFT, die sich bspw. tatsächlich auf einer Autobahn befinden und sich bereits im Stau befindend erstmalig telefonisch melden, werden hierdurch ggf. nicht als aktiv erkannt, insofern in räumlicher Nähe auch Verkehrsnetze passiver Verkehrsteilnehmerklassen vorhanden sind. Eine eindeutige Zuordnung auf welchem Verkehrsnetz sich der MFT befindet, ist aufgrund des Genauigkeitsniveaus ggf. nicht möglich. Konnte hingegen durch einen vorherigen Datensatz des MFT seine aktive Verkehrsrelevanz festgestellt werden, wird der MFT, auch stehend, nicht ausgeschlossen und dessen Daten an die nachfolgenden Analysen als aktiver MFT weitergegeben. Näher wird dieses Thema am Ende der nun folgenden Einzelverfahren der Identifikation behandelt.

Wenn nicht anders angegeben, wird den GPS-Positionen in den folgenden Ausführungen eine Genauigkeit von 50 m unterstellt. In vielen Fällen ist die tatsächliche Genauigkeit sicherlich besser, jedoch zeigen sich in den Daten vereinzelnd auch erheblich größere Ab-

weichungen. Um die GPS-Daten in gleicher Weise wie die MFT-Ortungen der Identifizierung zuzuführen und Überprüfungen jeder einzelnen GPS-Koordinate durch eine weitere Referenz zu vermeiden, wurde diese pauschale Definition getroffen. Der Wert 50 m soll Extremwerten gerecht werden. Angemerkt sei, dass die GPS-Positionen im Vergleich zur durchschnittlichen Genauigkeit der FPD-Positionen noch ca. um einen Faktor 10, wie unten noch gezeigt wird, besser sind und durchaus als Referenz dienen können.

4.3.1 Bewegungsanalyse

Grundidee der Bewegungsanalyse ist die Bestimmung der Bewegungswahrscheinlichkeit auf Basis der stochastischen Signifikanzgrenze der Koordinatendifferenzen zwischen den Schwerpunkten. Die Schwerpunkte selbst werden mit einem Fuzzy-C-Means Clusterverfahren nach Gustafson-Kessel berechnet, vgl. hierzu Abschnitt 4.3.2. Wird die Sicherheitswahrscheinlichkeit P der Bewegung überschritten, gilt der Teilnehmer / Teilnehmerdatensatz als sich bewegend (bewegt). Wird keine Bewegung festgestellt und der MFT war bislang passiv, so wird ohne die Ausführung der anderen Identifikationsverfahren der MFT als passiv eingestuft. War es zuvor aktiv, könnte eine Stausituation vorliegen, weshalb die Einstufung zwar „stehend“, aber nicht passiv vorgenommen wird. Mit den nachfolgenden Verfahren muss die Aufrechterhaltung der bislang aktiven Einstufung geprüft werden.

Die Berechnung der Korrektheit der Bewegungsanalyse, erfolgt über das Richtigkeitsmaß $RM = 1 - \frac{n}{N}$, mit der Anzahl n unkorrekter Einheiten und der Anzahl N der Grundgesamtheit des ausgewerteten Datenbestandes. Unkorrekte Einheiten setzen sich bei stehenden Szenarien aus der Anzahl sich dennoch als bewegt identifizierten zusammen. Entsprechend werden bei in Bewegung befindlichen Szenarien die Einstufungen „stehend“ als unkorrekt gezählt. Die Ergebnisse der Versuchsfahrten zur Bewegungsanalyse sind in den nachfolgenden Tabellen zusammengefasst. In der ersten Spalte sind jeweils die betrachteten Szenarien und in der zweiten Spalte die Kategorisierung der Soll-Identifikation mit „nicht bewegt“ und „bewegt“ aufgeführt. Die Soll-Identifikation gibt vor, welches die korrekte Einheit ist. Im Stand-Szenario somit „nicht bewegt“, in allen anderen Szenarien „bewegt“. In Spalte drei sind die Anzahlen der Datensätze zu den jeweiligen Szenarien zusammengefasst. In dieser sowie in allen weiteren Spalten bedeutet die erste Zahl die Anzahl bzw. den Wert der mit GPS auswertbaren Datensätze und die zweite Zahl diejenigen mit Mobilfunkdaten (FPD). Die Anzahl der Versuchsfahrten von GPS und FPD kann unterschiedlich sein, da nicht bei jedem Szenario genügender GPS-Empfang war oder auch unvollständige FPD-Daten vorliegen können. Insofern eine eindeutige Zuordnung der Daten bspw. aus Handschriftprotokollen möglich war, wurden die vorhandenen Daten der Auswertung zugeführt.

Die Bewegungsanalyse wurde mit den drei Sicherheitswahrscheinlichkeiten 50%, 68% und 80% durchgeführt. Die 68 % stehen dabei unter der Annahme einer normalverteilten Grundgesamtheit für einen Grenzwert in Höhe der Standardabweichung σ. Die beiden anderen Prozentwerte haben sich aus empirischen Analysen ergeben. Die Anzahl der richtig erkannten Bewegungseinstufungen sind in den letzten drei Spalten zusammengefasst.

Zur Korrektheitsberechnung der letzten drei Zeilen: Die obere Zeile zeigt die Zusammenfassung für die Szenarien separiert nach Soll-Identifikation „nicht bewegt“. Der Wert der Korrektheit RM_{Stand} gibt die im Stand „als stehend“ korrekt identifizierten Szenarien an. Die folgende Zeile fasst die Szenarien mit der Soll-Identifikation „bewegt“ mit RM_{bewegt} zusammen. In der letzten Spalte ist übergreifend für alle Soll-Identifikationen, der berechnete Korrektheitswert RM_{gesamt} angegeben.

Für die folgenden Tabellen sei angemerkt, dass in der obersten linken Zelle der Versuchstag eingetragen ist und das Symbol „-“ nicht an diesem Versuchstag durchgeführt oder auch nicht berechenbar bedeutet.

Tabelle 4-2: Korrektheitsmaße der Bewegungsanalyse für die Referenzfahrten des 13.09.2007

Szenario (13.09.2007)	**Soll-Identifikation**	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD		
Sicherheitswahrscheinlichkeit →			**50%**	**68%**	**80%**
Stand	nicht bewegt	8 / 8	8 / 2	8 / 4	8 / 4
Fußgänger	bewegt	- / -	- / -	- / -	- / -
Radfahrer	bewegt	- / -	- / -	- / -	- / -
PKW	bewegt	44 / 41	44 / 41	44 / 41	44 / 37
Linienbus	bewegt	- / -	- / -	- / -	- / -
Straßenbahn	bewegt	- / -	- / -	- / -	- / -
Nah/Fern DB	bewegt	- / -	- / -	- / -	- / -
Σ		52 / 49	52 / 43	52 / 41	52 / 41
	Korrektheit RM_{Stand}=	**8 / 8**	**1.00 / 0.25**	**1.00 / 0.50**	**1.00 / 0.50**
	Korrektheit RM_{bewegt}=	**44 / 41**	**1.00 / 1.00**	**1.00 / 1.00**	**1.00 / 0.90**
	Korrektheit RM_{gesamt}=	**52 / 49**	**1.00 / 0.88**	**1.00 / 0.88**	**1.00 / 0.84**

Tabelle 4-3: Korrektheitsmaße der Bewegungsanalyse für die Referenzfahrten des 07.11.2007

Szenario (07.11.2007)	**Soll-Identifikation**	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD		
Sicherheitswahrscheinlichkeit →			**50%**	**68%**	**80%**
Stand	nicht bewegt	3 / 3	3 / 1	3 / 1	3 / 1
Fußgänger	bewegt	3 / 3	2 / 1	2 / 0	2 / 0
Radfahrer	bewegt	- / -	- / -	- / -	- / -
PKW	bewegt	- / -	- / -	- / -	- / -
Linienbus	bewegt	2 / 2	2 / 2	2 / 2	2 / 2
Straßenbahn	bewegt	2 / 2	2 / 2	2 / 2	2 / 2
Nah/Fern DB	bewegt	- / -	- / -	- / -	- / -
Σ		10 / 10	9 / 6	9 / 5	9 / 5
	Korrektheit RM_{Stand}=	**3 / 3**	**1.00 / 0.33**	**1.00 / 0.33**	**1.00 / 0.33**
	Korrektheit RM_{bewegt}=	**7 / 7**	**0.86 / 0.71**	**0.86 / 0.57**	**0.86 / 0.57**
	Korrektheit RM_{gesamt}=	**10 / 10**	**0.90 / 0.60**	**0.90 / 0.50**	**0.90 / 0.50**

Tabelle 4-4: Korrektheitsmaße der Bewegungsanalyse für die Referenzfahrten des 19.12.2007

Szenario (19.12.2007)	**Soll-Identifikation**	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD		
Sicherheitswahrscheinlichkeit →			**50%**	**68%**	**80%**
Stand	nicht bewegt	- / -	- / -	- / -	- / -
Fußgänger	bewegt	12 / 12	10 / 8	10 / 7	10 / 7
Radfahrer	bewegt	- / -	- / -	- / -	- / -
PKW	bewegt	- / -	- / -	- / -	- / -
Linienbus	bewegt	- / -	- / -	- / -	- / -
Straßenbahn	bewegt	- / -	- / -	- / -	- / -
Nah/Fern DB	bewegt	9 / 11	9 / 11	9 / 11	9 / 11
Σ		21 / 23	19 / 19	19 / 18	19 / 18
Korrektheit RM_{Stand}=		**- / -**	**- / -**	**- / -**	**- / -**
Korrektheit RM_{beweat}=		**21 / 23**	**0.90 / 0.83**	**0.90 / 0.78**	**0.90 / 0.78**
Korrektheit RM_{gesamt}=		**21 / 23**	**0.90 / 0.83**	**0.90 / 0.78**	**0.90 / 0.78**

Tabelle 4-5: Korrektheitsmaße der Bewegungsanalyse für die Referenzfahrten des 27.02.2009

Szenario (27.02.2009)	**Soll-Identifikation**	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD		
Sicherheitswahrscheinlichkeit →			**50%**	**68%**	**80%**
Stand	nicht bewegt	- / -	- / -	- / -	- / -
Fußgänger	bewegt	4 / 4	4 / 3	4 / 3	4 / 3
Radfahrer	bewegt	1 / 1	1 / 1	1 / 1	1 / 1
PKW	bewegt	- / -	- / -	- / -	- / -
Linienbus	bewegt	- / -	- / -	- / -	- / -
Straßenbahn	bewegt	2 / 2	2 / 2	2 / 2	2 / 2
Nah/Fern DB	bewegt	0 / 1	0 / 1	0 / 1	0 / 1
Σ		7 / 8	7 / 7	7 / 7	7 / 7
Korrektheit RM_{Stand}=		**- / -**	**- / -**	**- / -**	**- / -**
Korrektheit RM_{beweat}=		**7 / 8**	**1.00 / 0.88**	**1.00 / 0.88**	**1.00 / 0.88**
Korrektheit RM_{gesamt}=		**7 / 8**	**1.00 / 0.88**	**1.00 / 0.88**	**1.00 / 0.88**

Tabelle 4-6: Zusammenfassung der korrekt erkannten Bewegungseinstufung aller Referenzdaten

Erfassungs-tag	**nicht bewegt GPS / FPD**				**bewegt GPS / FPD**			
Sicherheits-wahr-scheinlichkeit	Soll	50%	68%	80%	Soll	50%	68%	80%
13.09.2007	8 / 8	8 / 2	8 / 4	8 / 4	44 / 40	44 / 40	44 / 40	44 / 36
07.11.2007	3 / 3	3 / 1	3 / 1	3 / 1	7 / 7	6 / 5	6 / 4	6 / 4
19.12.2007	0 / 0	- / -	- / -	- / -	21 / 23	19 / 19	19 / 18	19 / 18
27.02.2009	0 / 0	- / -	- / -	- / -	7 / 8	7 / 7	7 / 7	7 / 7
Summe:	11 / 11	11 / 3	11 / 5	11 / 5	79 / 78	76 / 71	76 / 69	76 / 65
Korrektheitsgrad GPS:		**1.00**	**1.00**	**1.00**		**0.96**	**0.96**	**0.96**
FPD:		**0.27**	**0.45**	**0.45**		**0.91**	**0.88**	**0.83**

Zwischenfazit: Aus den Ergebnissen kann geschlussfolgert werden, dass im Falle eines tatsächlichen Standes 100% aller mit GPS erfassten Fahrten korrekt erkannt wurden, jedoch nur maximal 45% aller FPD-Daten. 55% der sich im Stand befindlichen FPD-Daten werden somit als bewegt eingestuft. Festgestellt werden muss aber auch, dass wegen der geringen Grundgesamtheit von 11 diese Auswertung statistisch eher als unsicher zu werten ist.

In der Analyse der sich tatsächlich in Bewegung befindlichen FPD wurden bei den GPS-Daten 96% korrekt identifiziert. Die genauere Betrachtung der falsch erkannten Szenarien zeigt, dass es sich hierbei um Fußgänger in dichter Bebauung und sogar innerhalb von Gebäuden handelt. Der verwendete GPS-Empfänger liefert in derartigen Umgebungssituationen ungenügende Positionsergebnisse. Die angenommene Genauigkeit von 50 m wird den erfassten GPS-Positionen bei diesen Messungen nicht repräsentativ, weshalb letztlich Ortungsfehler als Ursache für die fehlerhafte Identifikation angenommen werden können. Bei den FPD Analysen zeigt sich eine Korrektheit von 91% bei 50%, 88% bei 68% und 83% bei 80% Bewegungswahrscheinlichkeit. D.h. die Erkennung einer Bewegung von FPD, die sich tatsächlich in Bewegung befinden, ist im Vergleich zu stehenden FPD deutlich besser. Bei dieser Bewertung ist jedoch die oben genannte statische Unsicherheit bei „nicht bewegt“ zu beachten. Die Abbildung 4-10 verdeutlicht die Verteilung der Korrektheitsmaße an einem Kuchendiagramm für eine angenommene Sicherheitswahrscheinlichkeit von 68%.

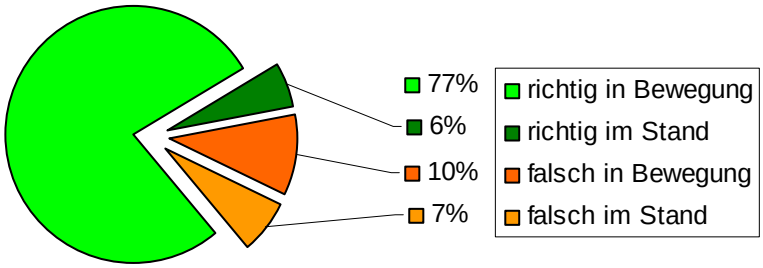


Abbildung 4-10: Kuchendiagramm zur Verteilung der Korrektheit für die Bewegungsanalyse bei der Sicherheitswahrscheinlichkeit von 68%

Für die weiteren Analysen wird eine Sicherheitswahrscheinlichkeit für die Bewegung von 68% verwendet. Bei diesem Wert wird der maximale Wert (45%) der Korrektheit der „nicht bewegten“ FPD erreicht und es werden gleichzeitig die zweit-meisten FPD (88%) in Bewegung richtig erkannt.

4.3.2 Positionsschwerpunkte

Im Projektbericht AP4-5 [3] wurden bereits Untersuchungen bezüglich der Reduktion von Ortungsberechnungen vorgestellt, deren Hauptziel die Reduktion von Rechenzeiten war. Die ursprüngliche, aus der Abis-Datenerfassung resultierende Frequenz zur Positionsberechnung beträgt ca. 2 Hz (dt=470ms). Basierend auf der Annahme von zufälligen Fehlern

wurden mit der Ausdünnung auf jeden 5. beziehungsweise jeden 10. Messwert verschiedene Genauigkeitsuntersuchen bei der Routenbestimmung vorgenommen.

Die Rechenzeitoptimierung steht auch bei dem für die Identifizierung verwendeten Verfahren der Datenreduktion im Vordergrund. Zudem fließen aber auch Aspekte der räumlichen und zeitlichen Punktverteilungen sowie die Ableitung von Genauigkeiten für die resultierenden Positionen mit ein. Es wird ein FCM Clusterverfahren nach Gustafson-Kessel, vgl. [4], verwendet, das die Ortungen des Signalstärken-Matchings, siehe AP4-5 [3], clustert. Die Schwerpunkte dieser Cluster werden als diejenigen Positionen interpretiert, die die Cluster repräsentieren und hier als Zielgröße angesehen werden. Im weiteren Verlauf wird daher von Schwerpunkten gesprochen.

Wie alle FCM Clusterverfahren beinhaltet auch das FCM-Gustafson-Kessel einen iterativen Berechnungsweg. Die Menge der Eingangsdaten (Anzahl der Ortungen) sowie die Anzahl der resultierenden Schwerpunkte, sind maßgebliche Faktoren für die Berechnungszeit. Auf die Menge der Ortungen soll kein Einfluss genommen werden, jedoch wird die Anzahl der Schwerpunkte je auszuwertenden Datensatz auf maximal 25 begrenzt. Zudem wird ein durchschnittlicher Schwerpunktabstand von 30 Sekunden angestrebt. Mit diesen Einstellungen können Trajektorien von bis zu ca. 12 Minuten mit der durchschnittlichen zeitlichen Auflösung von 30 Sekunden bestimmt werden. Bei zeitlich längeren Trajektorien ändert sich die durchschnittliche Auflösung entsprechend, die Anzahl der Cluster erreicht maximal 25. Kürzere Trajektorien werden mit der gerundeten Clusteranzahl aus 30 Sekunden dividiert durch Trajektoriendauer berechnet, jedoch mindestens 2 Cluster.

Der angestrebte zeitliche Schwerpunktabstand von 30 Sekunden folgt aus der Abschätzung der Positionsgenauigkeit von 100 m und schlechter sowie der Annahme, dass Bewegungen von mindestens 10 km/h (ca. 2,8 m/s) aufgelöst werden sollen. Bei diesen Werten lässt sich ein zeitlicher Abstand von ca. $100m \Big/ 2{,}8\frac{m}{s} \approx 35s$ berechnen. Da es sich hier nur um grobe Schätzung handelt, wurde ein Wert von 30 Sekunden gewählt. Wie unten noch gezeigt wird, werden meist schlechtere Positionsgenauigkeiten erzielt, weshalb letztlich auch größere Zeitabstände gerechtfertigt wären.

Vergleich GPS mit den Schwerpunkten des FCM-Gustafson-Kessel

Im Folgenden sollen die resultierenden Ergebnisse der Datenreduktion mit dem FCM-Gustafson-Kessel mit den Daten der oben genannten Referenzfahrten näher betrachtet werden. Für geometrische Vergleiche werden die GPS-Positionen der Referenzfahrten verwendet. Ausgangspunkt der Untersuchung ist ein Vergleich eben dieser GPS-Positionen mit den daraus berechneten Schwerpunkten. Für die GPS-Positionen wird, wie bereits oben erwähnt, eine Positionsgenauigkeit der Lagekoordinaten von 50 m je Position zur Berücksichtigung der größtmöglichen erwartbaren GPS-Ortungsgenauigkeiten als Pauschalwert angenommen.

Als vergleichendes Maß wird die metrische Distanz zwischen den Schwerpunkten und den jeweils zeitlich dichtesten GPS-Positionen berechnet. Dabei handelt es sich ausschließlich um Referenzfahrten mit Soll-Identifikation „bewegt“. Das Histogramm in Abbildung 4-11 für die Abweichungen der Schwerpunkte aus GPS-Positionen sowie denen aus MFT-Ortungen in Abbildung 4-12 vermitteln einen Eindruck über die Verteilung der Schwerpunktgenauigkeiten.

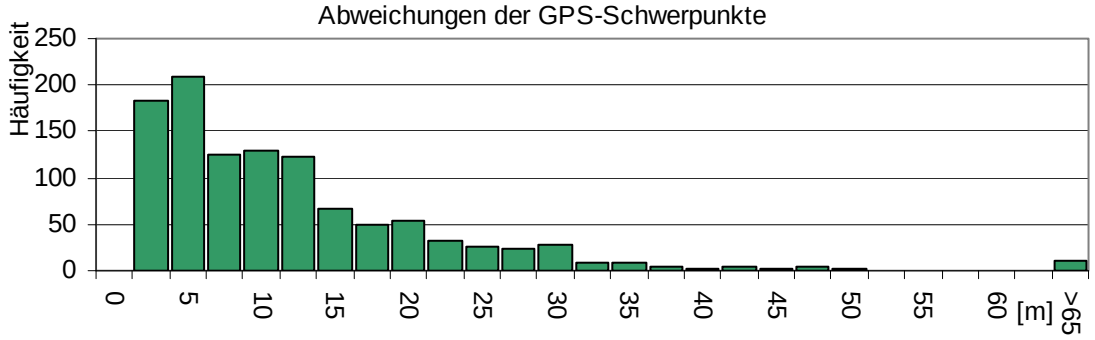


Abbildung 4-11: Histogramm für räumliche Abweichungen der GPS-Schwerpunkte aus den Referenzdaten zu den zeitlich nächsten GPS-Referenzpositionen

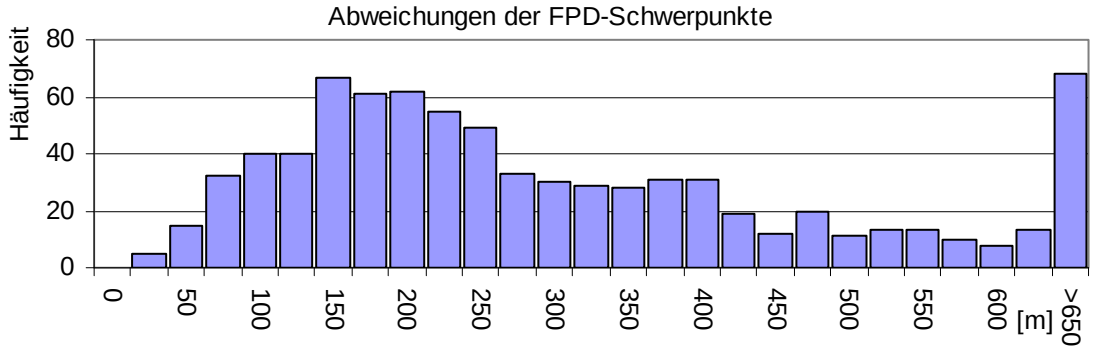


Abbildung 4-12: Histogramm für räumliche Abweichungen der FPD-Schwerpunkte zu zeitlich nächsten GPS-Referenzpositionen

Zu beachten sind die Skalen der metrischen Abweichungen zwischen Abbildung 4-11 und Abbildung 4-12. Es liegt ein Faktor 10 dazwischen. Die oben veranschlagte GPS-Positionsgenauigkeit von 50 m zeigt sich auch in den Abweichungen der Positionsschwerpunkte aus den GPS-Referenzdaten als taugliche Annahme, wenn auch teils als zu pessimistisch. Im Vergleich zu den Abweichungen aus den FPD-Positionsschwerpunkten (Abbildung 4-12) erkennt man deutlich, dass die GPS-Referenzmessungen und die daraus abgeleiteten Positionsschwerpunkte (Abbildung 4-11) mindestens um einen Faktor 10 besser bestimmt werden können.

Mit Blick auf die folgende Abbildung 4-13 erkennt man aber auch, dass durch die Positionsschwerpunktbildung keine Verbesserung der geometrischen Genauigkeit der FPD-Positionen mit dem hier eingesetzten Clusterverfahren erwartet werden kann. Eine Verringerung der Genauigkeit folgt letztlich aber auch nicht. In der Grafik sind die relativen Häufigkeiten

der metrischen Abweichungen zwischen den FPD-Positionen aus dem Signalstärken-Matching und den zeitlich dichtesten (wegen der Taktrate von 0.5 Hz bei der GPS-Aufzeichnung treten maximale Zeitdifferenzen zwischen GPS-Position und FPD-Position von ca. 1 Sekunde auf) GPS-Referenzpositionen aufgezeigt. Bestätigt wird dieser Sachverhalt durch die diskreten Verteilungsfunktionen der FPD-Positionen im Vergleich zu den FPD-Schwerpunkten in Abbildung 4-14. Die Genauigkeit für $1 \cdot \sigma$ (68,27%) ist für beide Reihen bei etwa 400 m.

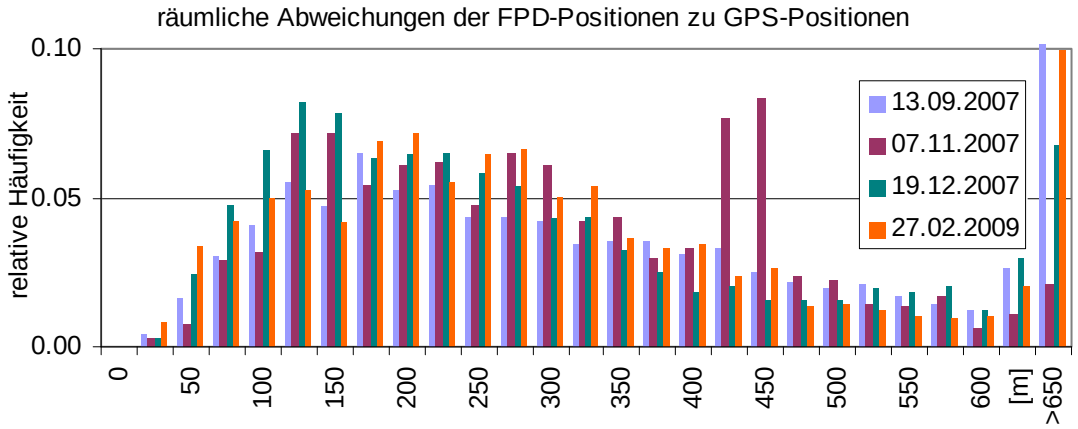


Abbildung 4-13: Relative Häufigkeiten der Abweichungen zwischen FPD-Positionen zu den GPS-Referenzpositionen

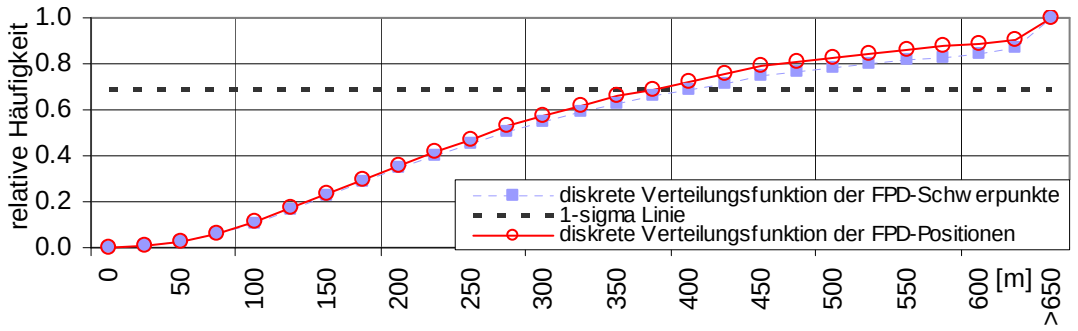


Abbildung 4-14: Diskrete Verteilungsfunktion der Schwerpunkte

Dieser Grafiken liegen folgende Messanzahlen zugrunde: 13.09.2007 → 39 309, 07.11.2007 → 4 064, 19.12.2007 → 13 401, 27.02.2009 → 12 495.

Zwischenfazit: Es kann festgehalten werden, dass bei etwa gleichbleibender geometrischer Genauigkeit eine erhebliche Verringerung der FPD-Positionsmenge erzielt werden kann. In der folgenden Tabelle 4-7 sind die Anzahl der FPD-Positionen und die resultierenden Schwerpunkte für die Messtage zusammengefasst.

Tabelle 4-7: Zahlenübersicht zur Datenreduktion durch Bildung von Schwerpunkten

Messtag	Anzahl		Datenreduktion auf: [%]
	FPD-Positionen	Schwerpunkte	
19.12.2007	4 274	69	1.61
27.02.2009	3 634	56	1.54
13.09.2007	34 118	609	1.78
11.07.2007	2 463	61	2.48

Aus den 44 489 MFT-Einzelortungen der Referenzfahrten in Bewegung entstanden 795 Schwerpunkte. Das entspricht einer durchschnittlichen Datenreduzierung auf 1,79%. Die zu verarbeitende Datenmenge der nachfolgenden Identifizierungsverfahren ist somit wesentlich geringer als ohne Datenreduktion, weshalb eine effizientere Bearbeitung erfolgen kann.

4.3.3 Korrelation mit Öffentlichem Verkehr

Mit der Analyse zur Korrelation mit dem ÖV startet im Identifizierungsprozess das erste Verfahren zur Prüfung, ob eine oder mehrere Verkehrsklassen von den weiteren Untersuchungen ausgeschlossen werden können. Entsprechen die Merkmale (Ort und Zeit) der Schwerpunkte mindestens einer ÖV-Klasse, dann kann dieser Datensatz des MFT nicht mehr eindeutig vom ÖV getrennt werden. Der Datensatz wird folglich als passiv eingestuft. Zeigt keine der untersuchten ÖV-Klassen eine genügende Merkmalsübereinstimmung, dann können und werden diese Verkehrsklassen für die weiteren Prüfungen ausgeschlossen. Der Suchraum der Verkehrsklassen reduziert sich um die ÖV-Klassen mit ungenügender Merkmalsübereinstimmung.

Das Kriterium der Merkmalsübereinstimmung wurde wie folgt definiert: Eine Verkehrsklasse aus der Gruppe ÖV (Linienbus, Straßen-/U-Bahn, DB-AG=Fernbahn) gilt als nachgewiesen, wenn für mindestens 80% der Schwerpunkte einer ÖV-Klasse (KVV Linienbusse, KVV Straßenbahnen, VVS Linienbusse, VVS Straßen-/U-Bahnen, DB-Nah-/Fernbahnen) eine Zugehörigkeit größer null ermittelt wird. Dieser Grenzwert wurde im Rahmen der vorhandenen geringen Stichprobe empirisch gewählt und ist in zukünftigen Arbeiten zu überprüfen. Wie die nachfolgenden Untersuchungen zeigen, können mit diesem Grenzwert in jedem Fall nachvollziehbare Ergebnisse erzielt werden. Zusätzlich zu den Schwerpunkten werden die ebenfalls aus der Clusterberechnung abgeleiteten Fehlerellipsen einschließlich der Kacheln mit einer Aufenthaltswahrscheinlichkeit von 5% oder höher, vgl. Abbildung 4-15, benötigt. Wie oben gezeigt, entsprechen die geometrischen Genauigkeiten der Schwerpunkte etwa denen der FPD-Positionen des Signalstärken-Matchings. Die Größen der Fehlerellipsen der Schwerpunkte entsprechen daher in etwa denen der zugrunde liegenden FPD-Positionen. Sobald für einen Schwerpunkt mindestens eine Kachel mit einer Aufenthaltswahrscheinlichkeit von mindestens 5% eine ÖV-Zugehörigkeit größer null festgestellt wird, kann eine ÖV-Zugehörigkeit nicht mehr ausgeschlossen werden, ein solcher Schwerpunkt zählt dann als zugehörig zur ÖV-Klasse.

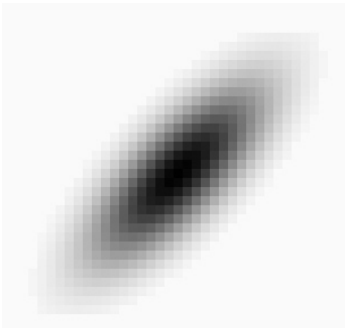


Abbildung 4-15: Beispiel für die Aufenthaltswahrscheinlichkeiten größer gleich 5% in den Kacheln eines Schwerpunktes; je dunkler die Kachel dargestellt ist, desto höher ist die Aufenthaltswahrscheinlichkeit

In den nun folgenden empirischen Nachweisen werden die Versuchsfahrten mit dem Merkmal „bewegt“ bei 68% Sicherheitswahrscheinlichkeit auf Korrelation mit dem öffentlichen Verkehr analysiert. Ziel des Verfahrens ist es bei Fahrten im ÖV zu zeigen, dass eine Zugehörigkeit zur Gruppe der ÖV-Klassen besteht. Es sei angemerkt: Eine Einstufung „ÖV kann nicht ausgeschlossen werden“ bedeutet nicht, dass die Verkehrsklasse MIV auszuschließen ist! Sie ist in diesem Fall jedoch nicht mehr eindeutig vom ÖV trennbar. Ist das Resultat, dass für keine der untersuchten ÖV-Klassen die Kriterien erfüllt werden, kann weiterhin die „aktive“ MIV-Verkehrsklasse möglich sein und ggf. noch als einzig mögliche identifiziert werden. Sie muss es jedoch nicht zwingend.

Die zusammenfassenden Tabellen der Versuchsfahrten sind ähnlich denen der Bewegungsanalyse aufgebaut. Zur besseren Übersicht wurden jedoch die einzelnen Fahrten bereits in den ÖV-Klassen Linienbus, U-/Straßenbahn und Nah/Fern DB in der ersten Spalte gruppiert. Die zweite Spalte zeigt die Anzahl der für die mit GPS und die mit Mobilfunkdaten (FPD) durchgeführten Fahrten. In der dritten und vierten Spalte folgen schließlich die Anzahl der richtig erkannten Szenarien und die Berechnung des Korrektheitsmaßes. Die unterste Zeile zeigt das berechnete Maß der Korrektheit über alle ÖV-Klassen, es schließt die „nicht ÖV“ Fahrten des Versuchstages in die Grundgesamtheit mit ein. Die maximale Korrektheit der „nicht ÖV“ Gruppe ist dann erreicht, wenn keines dieser Szenarien einer ÖV-Klasse zugeordnet wird. Der Wert in dieser Zeile bedeutet also, dass „nicht zu ÖV zugehörig“ richtig erkannt wurde. Es können bei einem Szenario verschiedene ÖV-Klassen gleichzeitig nachgewiesen (Linienbus und Straßenbahn) werden. Bei der Berechnung des Korrektheitsmaßes je ÖV-Klasse (Spalte vier) wird jeder Nachweis einzeln gezählt. Bei der Berechnung der Korrektheit „gesamt“ wird das entsprechende Szenario nur einmal gezählt, sobald eine beliebige ÖV-Klasse nachgewiesen wurde (unterste Zeile).

Tabelle 4-8: Korrektheitsmaße der ÖV-Korrelation für die Referenzfahrten des 13.09.2007

Szenario (13.09.2007)	**Anzahl Soll** GPS / FPD	**Ist Anzahl erkannt** GPS / FPD	**Korrektheit** RM
Linienbus	- / -	- / -	- / -
Straßenbahn	- / -	- / -	- / -
Nah/Fern DB	- / -	- / -	- / -
Nicht ÖV	44 / 45	44 / 30	1.00 / 0.67
Σ	44 / 45	44 / 30	
		Korrektheit RM_{gesamt}=	**1.00 / 0.67**

Anmerkung zum 13.09.2007: An diesem Tag wurden keine ÖV-Referenzfahrten durchgeführt. Dennoch wurden für die FPD-Daten 15 Fahrten als ÖV eingestuft.

Tabelle 4-9: Korrektheitsmaße der ÖV-Korrelation für die Referenzfahrten des 07.11.2007

Szenario (07.11.2007)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Linienbus	2 / 2	2 / 2	1.00 / 1.00
Straßenbahn	2 / 2	0 / 2	0.00 / 1.00
Nah/Fern DB	- / -	- / -	- / -
Nicht ÖV	2 / 1	1 / 1	0.50 / 1.00
Σ	6 / 5	3 / 5	
		Korrektheit RM_{gesamt}=	**0.50 / 1.00**

Tabelle 4-10: Korrektheitsmaße der ÖV-Korrelation für die Referenzfahrten des 19.12.2007

Szenario (19.12.2007)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Linienbus	- / -	- / -	- / -
Straßenbahn	- / -	- / -	- / -
Nah/Fern DB	9 / 11	7 / 5	0.78 / 0.45
Nicht ÖV	10 / 7	4 / 0	0.40 / 0.00
Σ	19 / 18	11 / 5	
		Korrektheit RM_{gesamt}=	**0.58 / 0.27**

Anmerkung zum 19.12.2007: Alle falsch erkannten „nicht ÖV“ sind Fußgängerszenarien auf Bahnsteigen oder in direkter Nähe zum Hauptbahnhof Karlsruhe.

Tabelle 4-11: Korrektheitsmaße der ÖV-Korrelation für die Referenzfahrten des 27.02.2009

Szenario (27.02.2009)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Linienbus	- / -	- / -	- / -
Straßenbahn	2 / 2	2 / 2	1.00 / 1.00
Nah/Fern DB	- / 1	- / 0	- / 0.00
Nicht ÖV	5 / 4	3 / 2	0.60 / 0.75
Σ	7 / 7	5 / 4	
		Korrektheit RM_{gesamt}=	**0.71 / 0.57**

Tabelle 4-12: Korrektheitsmaße der ÖV-Korrelation für alle Referenzfahrten

Szenario (alle)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Linienbus	2 / 2	2 / 2	1.00 / 1.00
Straßenbahn	4 / 4	2 / 4	0.50 / 1.00
Nah/Fern DB	9 / 12	7 / 5	0.78 / 0.42
Σ	15 / 18	11 / 11	
		Korrektheit $RM_{gesamt,\ ÖV}$=	**0.73 / 0.61**
Nicht ÖV	61 / 57	52 / 33	
		Korrektheit $RM_{gesamt,\ nicht\ ÖV}$=	**0.85 / 0.58**
Σ	76 / 75	63 / 44	
		Korrektheit RM_{gesamt}=	**0.83 / 0.59**

Wie in der Kartennachbereitung im Kapitel 3 bereits angesprochen, wurden Verkehrsnetzelemente der Schienenwege zwar topologisch nacherfasst, jedoch sind vereinzelt Elemente weder in den NAVTEQ-Daten noch durch die VuV Einpflege geometrisch genügend auflösend. Gravierend wirkt sich dieser Umstand bspw. bei den S-Bahn Fahrten zwischen Karlsruhe und Ettlingen aus, vgl. Abbildung 4-16. Dargestellt sind in schwarz die NAVTEQ-Schienendaten, in rot die VuV-Schienendaten, in orange das Autobahnnetz, in dunkelgrün die GPS-Positionen, in hellgrün die FPD-Ortungen und im Hintergrund in rosa das Abis-Datenaufzeichnungsgebiet vom 19.12.2007.

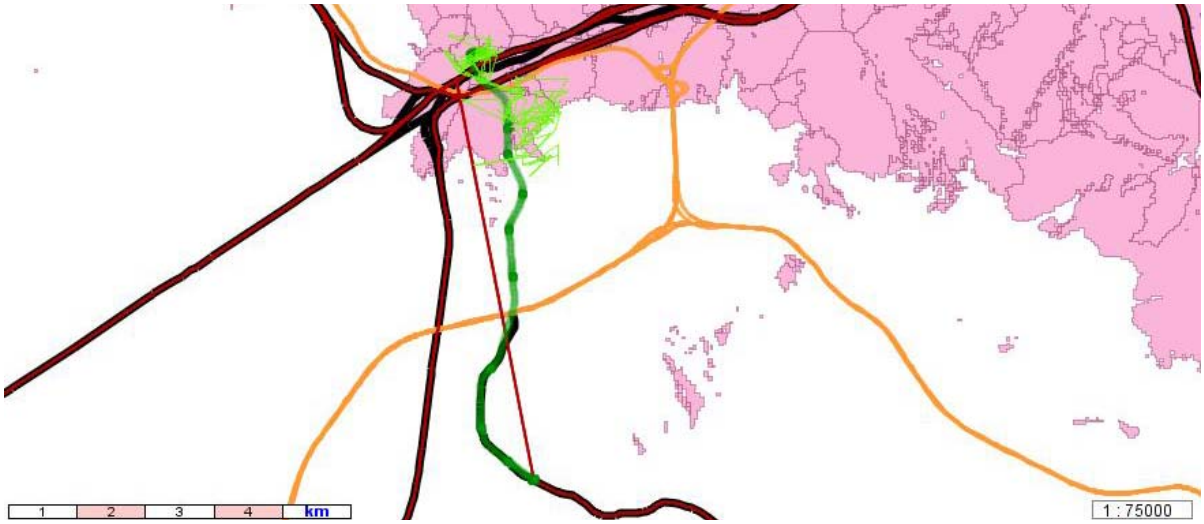


Abbildung 4-16: Referenzfahrt 7 vom 19.12.2007, S-Bahn Karlsruhe-Ettlingen

Deutlich erkennt man, dass sowohl GPS als auch FPD nur in der Nähe des Hauptbahnhofes Karlsruhe (Bild oben Mitte, Beginn der grünen Trajektorien, Verlauf: Nord→Süd) mit Schienendaten und den ÖV-Fahrplandaten korreliert werden können. Etwa 2 km südlich des Hbf Ka existieren für ca. 4 km keine NAVTEQ-Geometrien. Die rot markierte Bahnstrecke der VuV-Digitalisierung weicht in Westrichtung deutlich sowohl von den FPD als

auch von den GPS-Positionen ab. In beiden Fällen kann aufgrund der zu geringen räumlichen Auflösung der Schienendaten die ÖV-Korrelation keine Identifizierung durchführen.

Zwischenfazit: Das Verfahren zur Korrelation der FPD mit ÖV-Verkehrklassen ist maßgeblich von der geometrischen Korrektheit der Kartenelemente und von der Genauigkeit der Fahrplandaten, die hier nicht näher betrachtet wurde, abhängig. Es kann jedoch festgehalten werden: Von den 18 Referenzfahrten mit FPD-Erfassung wurden 11 erkannt. D.h. die Erkennungsrate beträgt 61% bei FPD und 11/15=73% mit GPS. Die nicht erkannten ÖV-Fahrten betreffen bei FPD vollständig die Nah/Fern DB Versuchsfahrten. Es kann erwartet werden, dass bei Vervollständigung der Schienengeometriedaten ein deutlicher Anstieg der Korrektheit beobachtet werden kann. Der hier anvisierte Nachweis für einen Ausschluss von FPD-Datensätzen in ÖV konnte jedoch besonders für Linienbusse und Straßenbahnen gezeigt werden.

Wenn ein FPD-Datensatz mit ÖV Soll-Identifikation mit dem Korrelationsverfahren erkannt wird, dann handelt es sich (bei Verwendung des Do-iT Lokalisierungsnetzes in der vorliegenden Version und den FPD Ortungen der untersuchten Szenarien) zu 59% tatsächlich um ein ÖV-Szenario. Mit GPS wurden 83% erreicht. Interpretiert man die Fußgängerszenarien in der Nähe und im Hbf Karlsruhe nicht als falsch erkannt, dann werden 68% (= (44+7) / 75) der Versuchsfahrten richtig identifiziert. Bei GPS sind es dann 91% (= (63+6) / 76). Es sei aber auf den sehr geringen und für verallgemeinernde Statistiken nicht repräsentativen Stichprobenumfang der ÖV-Versuchsfahrten hingewiesen! Des Weiteren wurden bei den hier gezeigten Analysen die modellierten Fahrplanabweichungen nicht näher betrachtet. Größere temporäre Fahrplanabweichungen könnten ggf. auch Ursache der Fehlidentifizierungen bei den mit GPS erfassten Straßenbahnfahrten sein.

4.3.4 Verkehrsnetzanalyse

Wie bereits oben im Abschnitt 4.2.5 beschrieben, handelt es bei der Verkehrsnetzanalyse um zwei getrennte Verfahren. Das erste Verfahren prüft ausschließlich anhand der räumlichen Lage der Schwerpunkte, ob für jeden Schwerpunkt Verkehrsnetzelemente der zu untersuchenden Verkehrsklassen vorhanden sind. Mit der Aufenthaltswahrscheinlichkeit der Kacheln wird schließlich eine Wahrscheinlichkeit für den Aufenthalt auf den Verkehrsnetzelementen abgeleitet. Im zweiten Verfahren werden zusätzlich topologische und zeitliche Kriterien hinzugenommen.

Im Folgenden werden die empirischen Ergebnisse der Bestimmung der Verkehrsnetzwahrscheinlichkeit zusammengestellt. Es werden alle diejenigen Fahrten ausgewertet, für die nach der Korrelation mit dem ÖV der Suchraum die MIV Verkehrsklasse noch enthält. Als entscheidend für die Korrektheitsmaßbestimmung wird definiert, dass für die untersuchte Verkehrsklasse das entsprechende Verkehrsnetz mit eingeschlossen ist. Für Fahrten mit der Soll-Identifikation PKW/LKW sind Einschlüsse mit dem Straßennetzes (car_truck) korrekt. Das Erkennen weiterer Verkehrsnetze wird nicht als falsch gewertet, da auch gleiche Verkehrsnetzkanten unterschiedliche Widmungen haben können. Hingegen können Refe-

renzfahrten mit dem Fahrrad nicht ausgewertet werden, da im Lokalisierungsnetz keine Widmung für Fahrrad-Verkehrsnetze vorhanden ist. Derartige Referenzfahrten werden daher in den folgenden Tabellen nicht berücksichtigt.

Analog zum Aufbau der Ergebnistabellen in den vorherigen Identifikationsverfahren beinhalten die folgenden Tabellen in Spalte eins das Szenario, in Spalte zwei die Anzahl der betrachteten Datensätze, in Spalte drei die Anzahl der korrekt erkannten Datensätze und in Spalte vier das Korrektheitsmaß für das Szenario. Die Versuchsfahrtszenarien wurden der besseren Übersicht wegen bereits in Fußweg, Straße, Straßenbahn und dem Schienennetz der DB AG mit Nah/Fern DB zusammengefasst. In der letzten Zeile findet sich das über alle Szenarien berechnete Korrektheitsmaß.

Tabelle 4-13: Korrektheitsmaße der Verkehrsnetzzugehörigkeit für die Referenzfahrten des 13.09.2007

Szenario (13.09.2007)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Fußweg	- / -	- / -	- / -
Straße	44 / 30	44 / 30	1.00 / 1.00
Straßenbahn	- / -	- / -	- / -
Nah/Fern DB	- / -	- / -	- / -
Σ	44 / 30	44 / 30	1.00 / 1.00
		Korrektheit RM_{gesamt}=	**1.00 / 1.00**

Tabelle 4-14: Korrektheitsmaße der Verkehrsnetzzugehörigkeit für die Referenzfahrten des 07.11.2007

Szenario (07.11.2007)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Fußweg	1 / -	0 / -	0.00 / -
Straße	- / -	- / -	- / -
Straßenbahn	2 / -	2 / -	1.00 / -
Nah/Fern DB	- / -	- / -	- / -
Σ	3 / -	2 / -	
		Korrektheit RM_{gesamt}=	**0.67 / -**

Tabelle 4-15: Korrektheitsmaße der Verkehrsnetzzugehörigkeit für die Referenzfahrten des 19.12.2007

Szenario (19.12.2007)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Fußweg	4 / -	0 / -	0.00 / -
Straße	- / -	- / -	- / -
Straßenbahn	- / -	- / -	- / -
Nah/Fern DB	2 / 3	2 / 2	1.00 / 0.67
Σ	6 / 3	2 / 2	
		Korrektheit RM_{gesamt}=	**0.33 / 0.67**

Tabelle 4-16: Korrektheitsmaße der Verkehrsnetzzugehörigkeit für die Referenzfahrten des 27.02.2009

Szenario (27.02.2009)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Fußweg	2 / 1	0 / 0	0.00 / 0.00
Straße	- / -	- / -	- / -
Straßenbahn	- / -	- / -	- / -
Nah/Fern DB	- / 1	- / 1	- / 1.00
Σ	2 / 2	0 / 1	
		Korrektheit RM_{gesamt}=	**0.00 / 0.50**

Tabelle 4-17: Zusammenfassung der korrekt erkannten Verkehrsnetzzugehörigkeiten aller Versuchstage

Szenario (alle)	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM
Fußweg	7 / 1	0 / 0	0.00 / 0.00
Straße	44 / 30	44 / 30	1.00 / 1.00
Straßenbahn	2 / -	2 / -	1.00 / -
Nah/Fern DB	2 / 4	2 / 3	1.00 / 0.75
Σ	55 / 35	48 / 33	
		Korrektheit RM_{gesamt}=	**0.87 / 0.94**

Zwischenfazit: Die entsprechenden Verkehrsnetze der Szenarien werden zu 87% bei GPS und zu 94% mit den FPD Daten korrekt identifiziert. Maßgeblich unkorrekt sind die Szenarien der Fußwege und der Nah/Fern DB Schienenwege. Das erklärt auch die geringe Korrektheitsrate bei GPS, da hier mehr dementsprechende Fahrten eingingen. Die Abbildung 4-17 veranschaulicht für GPS und Abbildung 4-18 für FPD die Verteilung der richtig und falsch identifizierten Verkehrsnetze. Angemerkt sei, dass die Identifizierung der Nah/Fern DB Schienen nur wegen des deutlich größeren Beobachtungsraumes mit ca. 1% zugeordnet wurden, vgl. Abbildung 4-16. Eine Ursache für die unkorrekten Zuordnungen könnte die Datenquelle des Lokalisierungsnetzes sein. Es basiert auf den NAVTEQ-Daten, deren vorrangiges Ziel die Abdeckung und Widmung der Straßen ist. Fußwege und die bereits oben angesprochenen Schienennetze sind bislang eher aus Visualisierungsgründen in den Daten vorhanden, was sich schließlich hier durch die geringe Identifizierungsrate bestätigt.

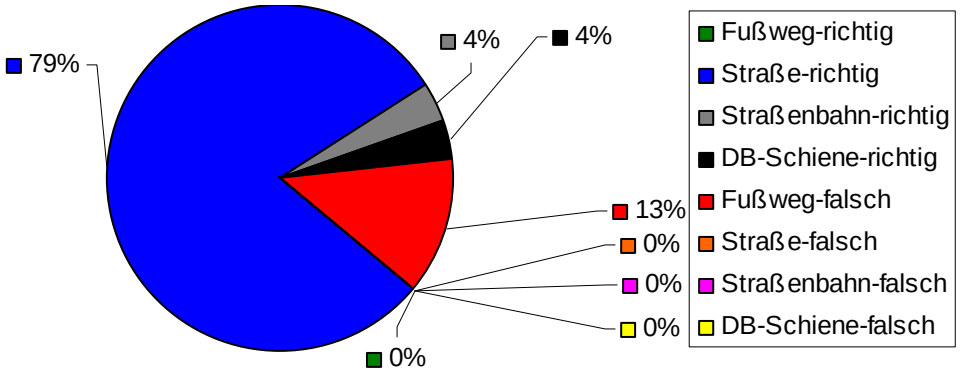


Abbildung 4-17: Verteilung für GPS der richtig und falsch identifizierten Verkehrsnetze mit der Analyse zur räumlichen Verkehrsnetzwahrscheinlichkeit

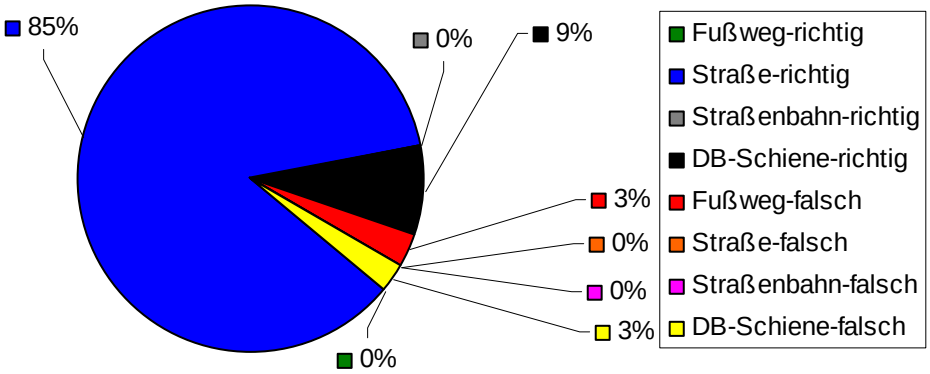


Abbildung 4-18: Verteilung für FPD der richtig und falsch identifizierten Verkehrsnetze mit der Analyse zur räumlichen Verkehrsnetzwahrscheinlichkeit

Im zweiten Verfahren wird nun geprüft, ob zwischen den Schwerpunkten innerhalb der Zeitdifferenzen der Schwerpunkte eine Route gefunden werden kann. Wird eine Route gefunden, muss von der prinzipiellen Möglichkeit ausgegangen werden, dass sich prinzipiell auf dem verwendeten Verkehrsnetz mit den Merkmalen (Positionen und Zeit) der Schwerpunkte fortbewegt werden kann. Die zum untersuchten Verkehrsnetz zugehörigen Verkehrsklassen können schließlich, sofern sie noch im Suchraum vorhanden sind, nicht ausgeschlossen werden. Kann keine Route gefunden werden, dann ist wiederum der Ausschluss der entsprechenden Verkehrsklasse begründet. Die Gründe, warum keine Route gefunden werden kann, sind unterschiedlicher Natur. Im Folgenden seien anhand einiger ausgewählter Referenzfahrten wesentliche Ursachen an Beispielen näher erläutert.

Das erste Beispiel entstammt der Referenzfahrt 107 vom 27.02.2009. In den beiden Karten der Abbildung 4-19 sind die GPS-Referenzdaten in grün und die daraus berechneten Schwerpunktpositionen in rot eingezeichnet. In der linken Karte ist das selektierte Straßennetz im Hintergrund (blau) sowie rechts dem selektierten Fußwegenetz der Fußwegroute (gelb) dargestellt. Beide Netzausschnitte folgen aus den Korridordefinitionen jeweils zwischen zwei Schwerpunkten, vgl. Abschnitt 4.2.5.

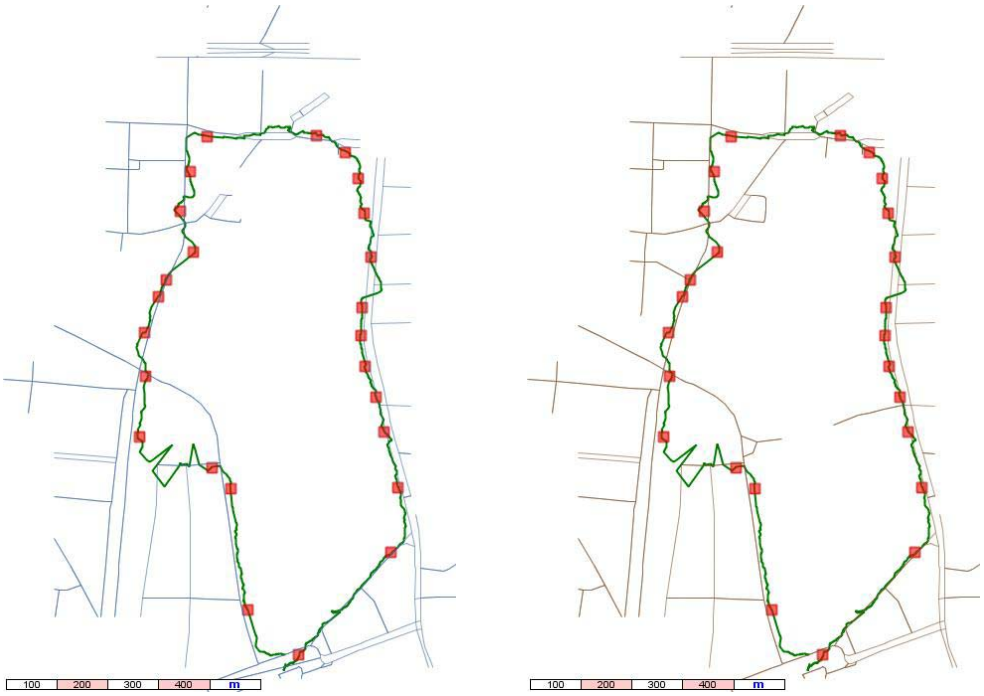


Abbildung 4-19: links Straßennetz, rechts Fußwegenetz der Referenzfahrt 107 bei 50 m Positionsgenauigkeit, keine Route wegen fehlender topologischer Verbindungen

Man erkennt in den Karten jeweils oben rechts, dass die Netze topologisch entlang der GPS-Route nicht miteinander verbunden sind. Ursachen können sein, dass die Netze, besonders das Fußwegenetz, nicht vollständig sind oder die veranschlagte Genauigkeit der Schwerpunkte zu optimistisch ist. Schließlich erfolgt die Selektion der Verkehrsnetze auf Basis der zugehörigen Kacheln. Ob eine Kachel zur Auswahlliste gehört, wird durch den Wert der Aufenthaltswahrscheinlichkeit innerhalb der Kachel, bezogen auf den geometrischen Schwerpunkt der Kachel, bestimmt. Ist der Wert mindestens 5%, ist die entsprechende Kachel Teil der Liste. Die Kantenlängen der Kacheln betragen 100 m. Die passiv-Einstufung dieses Datensatzes ist somit Folge der Verkehrsnetzselektion, bestehend aus einem topologisch nicht zusammenhängenden Straßennetz (und damit fehlerhaft).

Im nächsten Beispiel wurden die Positionsgenauigkeiten von 50 m, sonst Standard für alle sonstigen GPS-Berechnungen in diesem Bericht, auf 150 m verschlechtert. Die GPS-Positionen sind wieder mit grün und Schwerpunkte in rot eingezeichnet. Die linke Karte der Abbildung 4-20 zeigt das selektierte Straßennetz im Hintergrund und die Route für PKW in blau, die rechte Karte das selektierte Fußwegenetz im Hintergrund und der Fußwegroute in gelb.

Abbildung 4-20: links Straßennetz, rechts Fußwegenetz der Referenzfahrt 107 bei 150 m Positionsgenauigkeit, jeweils wurde eine Route gefunden

In beiden Verkehrsnetzen konnte eine zeitlich sowie topologisch mögliche Route gefunden werden. Die passiv-Einstufung erfolgt hier, da die Verkehrsrelevanz des Datensatzes nicht eindeutig geklärt werden konnte.

Eine weitere Möglichkeit für einen ggf. unrichtigen Ausschluss eines Verkehrsnetzes zeigt Abbildung 4-21. Die Daten stammen aus der Versuchsfahrt 24 in einem IC am 19.12.2007. Dargestellt sind in schwarz das Schienennetz der DB und die Startknoten im Bild unten links. In grün sind die Positionen des Signalstärken-Matchings und in rot die Schwerpunktpositionen gezeichnet. GPS-Koordinaten konnten wegen Signalabschattung und Signalstörung nur für die ersten Sekunden, jedoch nicht repräsentativ für die Referenzfahrt aufgezeichnet werden.

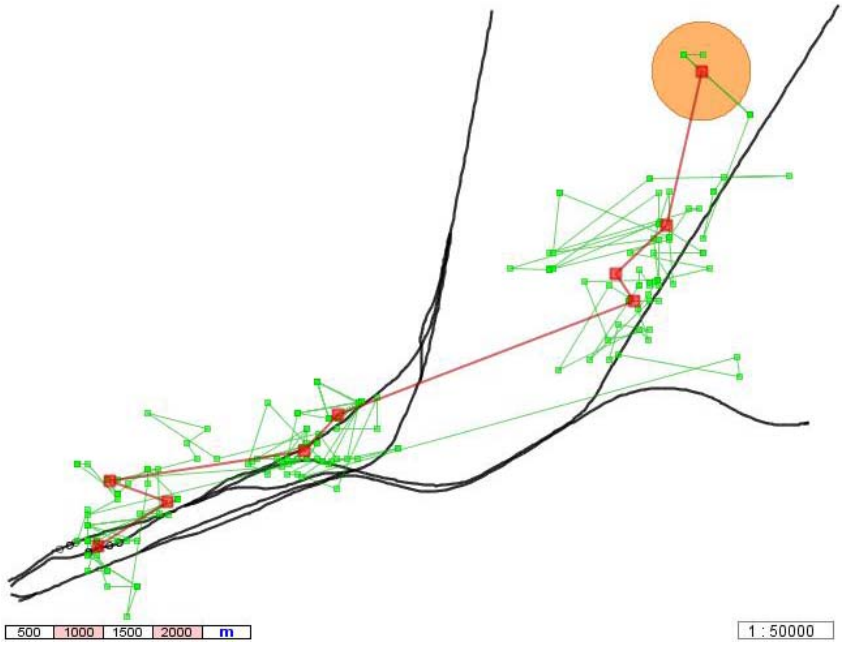


Abbildung 4-21: Verkehrsnetzausschluss wegen fehlender Zielknoten (Versuchsfahrt von 19.12.2007 Nr. 24)

Der rote Kreis im Bild oben rechts symbolisiert einen 500 m Radius um den letzten Schwerpunkt. Deutlich ist zu erkennen, dass das Schienennetz nicht innerhalb dieses Kreises liegt. Das nächste Schienenelement zu diesem letzten Schwerpunkt ist ca. 800 m und damit deutlich weiter entfernt als die geschätzte Positionsgenauigkeit der letzten Positionen des Datensatzes von 175 m bis 375 m. Die Folge ist, dass für diesen Datensatz zwar Startknoten (schwarze Punkte) und auch Korridore für die Schwerpunkte gefunden werden, jedoch keine Zielknoten. Auch eine versuchsweise Verschlechterung der Genauigkeit auf 500 m genügte nicht. Erst ab ca. 700 m ist der Suchbereich genügend groß, um Verkehrsnetzelemente des Schienennetzes und somit Zielknoten zu ermitteln. Obwohl für den Datensatz 24 vom 19.12.2007 durchgängig Zugehörigkeiten zu Straßenelementen vorhanden sind, konnte keine Route innerhalb der Zeitrestriktionen gefunden werden. Der Teilnehmer wurde daher als passiv eingestuft. Sinnvoll untermauert wäre das Ergebnis in diesem DB-IC-Fahrt Szenario, wenn eine DB-Route gefunden würde. Schließlich schlug die Routensuche aufgrund der nicht repräsentativen Genauigkeitsschätzung der FPD Positionen fehl. Die in der Grafik erkennbaren systematischen Abweichungen wurden auch auf die Schwerpunkte übertragen.

In den genannten Beispielen erfolgte jeweils die geforderte Einstufung des Datensatzes als passiv. Die Ursachen hierfür basieren jedoch auf verschiedenen Kriterien: Topologie, Eindeutigkeit und Positionsgenauigkeit. Eine weitere passiv-Einstufung wäre möglich, wenn die Bewegungsgeschwindigkeit höher ist, als es die plausible Maximalgeschwindigkeit innerhalb des Verkehrsnetzes es gestatten würde. Indirekt wurde dieser Fall im letzten Beispiel gezeigt, da durch die hohe IC Bewegungsgeschwindigkeit keine Route auf dem Straßennetz den Bedingungen der Zeitrestriktionen genügte.

Für den empirischen Nachweis soll gelten, dass für die aktive Verkehrsklassengruppe des MIV (LKW / PKW / Reisebus / Motorrad) eine korrekte Erkennung vorliegt, sobald auf dem Straßennetz eine Route für die entsprechende Referenzfahrt gefunden wird. Wie bereits beim vorherigen Verfahren genannt, können Verkehrsnetze mehrere Widmungen haben, weshalb der Ausschluss von Routen anderer Verkehrsklassen hier nicht als Unkorrektheit verwendet werden kann. Vielmehr ist das Vorhandensein von Routen bei unterschiedlichen Verkehrsklassen eine Frage der Eindeutigkeit. Dieser Aspekt wird im anschließenden Abschnitt 4.3.5 behandelt. Zu den Referenzfahrten sei gesagt, dass ausschließlich am 13.09.2007 Fahrten mit aktiver Soll-Identifikation durchgeführt wurden. Alle anderen Versuchstage dienten dem Nachweis passiver Identifikationen. Der Nachweis von Routen auf Verkehrsnetzen passiver Identifikation kann zum Ausschluss aktiver Verkehrsklassen oder ggf. zur Bestimmung der uneindeutige Identifizierung dienen.

Tabelle 4-18: Zusammenfassung der korrekt erkannten Verkehrsnetze aller Versuchstage mittels der zeitlich topologischen Routensuche

Versuchstag	**Szenario**	**Anzahl Soll** GPS / FPD	**Ist Anzahl richtig erkannt** GPS / FPD	**Korrektheit** RM GPS / FPD
13.09.2007	Fußweg	- / -	- / -	- / -
	Straße	44 / 29	44 / 19	1.00 / 0.66
	Straßenbahn	- / -	- / -	- / -
	Nah/Fern DB	- / -	- / -	- / -
07.11.2007	Fußweg	- / -	- /	- / -
	Straße	- / -	- / -	- / -
	Straßenbahn	2 / -	0 / -	0.00 / -
	Nah/Fern DB	- / -	- / -	- / -
19.12.2007	Fußweg	- / -	- / -	- / -
	Straße	- / -	- / -	- / -
	Straßenbahn	- / -	- / -	- / -
	Nah/Fern DB	2 / 3	0 / 0	0.00 / 0.00
27.02.2009	Fußweg	2 / 1	0 / 0	0.00 / 0.00
	Straße	- / -	- / -	- / -
	Straßenbahn	- / -	- / -	- / -
	Nah/Fern DB	- / 1	- / 0	- / 0.00
alle Σ	Fußweg	2 / 1	0 / 0	**0.00 / 0.00**
alle Σ	Straße	44 / 29	44 / 19	**1.00 / 0.66**
alle Σ	Straßenbahn	2 / -	0 / -	**0.00 / -**
alle Σ	Nah/Fern DB	2 / 4	0 / 0	**0.00 / 0.00**
	Σ	50 / 34	44 / 19	
			Korrektheit RM_{gesamt}=	**0.88 / 0.56**

Anmerkung für GPS: Für die beiden Straßenbahnfahrten am 07.11.2007 wurde jeweils eine Route auf dem Straßennetz (car_truck) bestätigt. Die beiden Fußgängerszenarien am 27.02.2009 wurden jeweils mit einer Route auf dem Straßennetz (car_truck) bestätigt.

Anmerkung für FPD: Für die S-Bahnfahrt von Karlsruhe nach Ettlingen am 19.12.2007 wurde zwar keine Nah/Fern DB Route jedoch eine Route auf dem Straßennetz (car_truck) gefunden. Am 27.02.2009 wurde die Fahrradfahrt ebenfalls mit einer Route auf dem Straßennetz (car_truck) bestätigt. Ein eigenes Fahrradverkehrsnetz existiert im Lokalisierungsnetz nicht.

Zwischenfazit: Für alle GPS-Referenzdaten mit aktiver Soll-Identifikation des Straßennetzes (car_truck) wurden Routen mit Einhaltung der zeitlichen und topologischen Restriktionen der Schwerpunkte auf dem Straßennetz gefunden. Bei den FPD Daten konnten für 66% der Fahrten entsprechende Route nachgewiesen werden. Die restlichen 34% Fahrten wurden aus Sicht der Soll-Identifikation fehlerhaft als passiv eingestuft. Für die anderen Soll-Identifikationen Fußweg, Straßenbahn und Nah/Fern DB konnten keine bestätigenden Routen gefunden werden. Der Stichprobenumfang mit 6 bei GPS und 5 bei FPD ist als sehr gering einzustufen, um Schlussfolgerungen daraus abzuleiten. Ein Teil bspw. der ÖV-Fahrten wird bereits bei der ÖV-Korrelation gefiltert und gelangt schließlich nicht zur Routenanalyse. Festzuhalten ist, dass nach der Grundidee alle ÖV-Fahrten herauszufiltern sind. Zur Untersuchung der Routensuche mit Zeit- und Topologierestriktionen auf ÖV-Verkehrsnetzen, wurde testweise die ÖV-Analyse für die FPD-Fahrten am 07.11.2007 in einer separat durchgeführten Auswertung abgeschaltet. FPD mit zeitlichen und räumlichen ÖV-Merkmalen werden somit in der ÖV-Analyse absichtlich nicht herausgefiltert. Als Folge wurde festgestellt, dass neben einer möglichen MIV-Route jeweils auch eine Straßenbahnroute gefunden wurde. Eine ÖV-Korrelation kann infolgedessen nicht mehr ausgeschlossen werden, bestätigt durch Vorhandensein einer ÖV-Route mittels des Verfahrens zur Routensuche mit Zeit- und Topologierestriktionen. Damit konnte die Funktionsfähigkeit prinzipiell bestätigt werden. Sie wird jedoch bei räumlich erheblich abweichenden Verkehrsnetzgeometrien sowie bei großen Ortungsfehlern eingeschränkt. Abschließend errechnet sich für die in Tabelle 4-18 zusammengestellten Versuchfahrten ein Gesamtkorrektheitsmaß für die Routensuche bei GPS von 88% und für FPD von 56%. Man könnte zusätzlich aus der Tabelle 4-18 ableiten, dass 100% der gefundenen Routen zu den aktiven Verkehrsklassen gehören. Aufgrund des geringen Stichprobenumfangs der nicht aktiven Verkehrsklassen ist diese Aussage jedoch vage und statistisch eher unsicher.

4.3.5 Analyse des Gesamtverfahrens

Mit den zuvor untersuchten Identifikationsverfahren wurden Übereinstimmungen von Verkehrsklassenmerkmalen zu den FPD-Datensätzen und als Vergleich auch zu den GPS-Daten gesucht. Der Suchraum der infrage kommenden Verkehrsklassen wird mit den Identifikationsverfahren reduziert. In diesem Abschnitt wird die Identifikation von aktiven Verkehrsteilnehmern als Gesamtverfahren untersucht. Die Referenzfahrten werden, wie bisher auch, separat behandelt. Jede Fahrt entspricht somit einem eigenen virtuellen Teilnehmer und damit einem abgeschlossenen Szenario. Kann der Suchraum nach dem letzten Verfahren auf die aktiven Verkehrsklassen eingegrenzt werden, gelten der Datensatz und damit der virtuelle Teilnehmer als aktiv erkannt. Andernfalls erfolgt die passiv-Einstufung.

Erfolgt die Identifizierung uneindeutig, wird zur Reduzierung ggf. fehlerhaft aktiv eingestufter Datensätze die passiv-Einstufung vorgenommen. Unter Umständen ist zu erwarten, dass auf Kosten der sichereren aktiv-Einstufung, in Wirklichkeit aktive Datensätze zu Unrecht ausgeschlossen werden. Die Datenmenge für eine nachfolgende Verkehrslageschätzung verringert sich ggf. unberechtig. Durch dieses Vorgehen kann jedoch der sich mitunter gravierend auswirkende Fehler von aktiv eingestuften Datensätzen, die in Wirklichkeit passiv werden, verringert werden.

In der folgenden Tabelle 4-19 werden die Anzahlen der nach Soll-Identifikation aktiven und passiven Szenarien je Versuchstag zusammengefasst, sowie die Anzahlen der jeweils korrekt erkannten Einstufungen in Spalte drei angegeben. Die Korrektheitsmaße folgen je Szenario in der letzten Spalte und über alle Versuchstage hinweg in der letzten Zeile. In der letzten Spalte stehen hochgestellte Nummern, auf die im nachfolgenden Text referenziert wird.

Tabelle 4-19: Zusammenstellung der Identifikationsergebnisse der nach Szenarien separierten Datensätze

Szenario	**Anzahl** GPS / FPD	**Ist Anzahl (Zeit) korrekt** GPS / FPD	**Ist Anzahl (Zeit) unkorrekt** GPS / FPD	**Korrektheit** ***RM***
13.09.2007 Soll: aktiv	44 / 41	44 (28560) / 20 (11292)	0 (0) / 21 (9573)	1.00 / 0.49
Soll: passiv	8 / 8	8 (3730) / 7 (1877)	0 (0) / 1 (618)	1.00 / 0.88
Summe:	52 / 49			
07.11.2007 Soll: aktiv	- / -	- / -	- / -	- / -
Soll: passiv	10 / 10	8 (3090) / 10 (3574)	2 (1345) / 0 (0)	0.80 / 1.00
Summe:	10 / 10			
19.12.2007 Soll: aktiv	- / -	- / -	- / -	- / -
Soll: passiv	23 / 23	19 (10298) / 22 (7537)	4 (1308)/ 1 (228)	1.00 / 0.96
Summe:	23 / 23			
27.02.2009 Soll: aktiv	- / -	- / -	- / -	- / -
Soll: passiv	7 / 8	5 (5852) / 6 (5091)	2 (2669) / 2 (2044)	0.71 / 0.75
Summe:	7 / 8			
Σ gesamt aktiv	44 / 41	korrekt aktiv 44 / 20	unkorrekt aktiv 8 (0) / 4 (0)	**1.00**[1)] **/ 0.49**[2)]
Σ gesamt passiv	48 / 49	korrekt passiv 40 / 45	unkorrekt passiv 0 / 21	**0.83**[3)] **/ 0.92**[4)]
Σ gesamt	92 / 90	84 / 65	8 / 25	
			Korrektheit RM_{gesamt}**=**	**0.91**[5)] **/ 0.72**[6)]

Zwischenfazit: Wenn ein MFT-Datensatz tatsächlich aktiv ist, dann wird er mit GPS zu 100% richtig als aktiv erkannt, vgl. Markierung [1)]. Bei den FPD sind es 49% korrekt als aktiv identifizierte aktive MFT-Datensätze[2)]. Ist der MFT-Datensatz passiv, dann wird er mit GPS zu 83%[3)] und mit FPD zu 92%[4)] richtig als passiv eingestuft. Daraus folgt, dass zu 8% (=1-0.92[4)]) ein tatsächlich passiver MFT-Datensatz fehlerhaft als aktiv identifiziert wird. Allerdings ist auch festzuhalten, dass 51% (=1-0.49[2)]) der aktiven MFT-Szenarien nicht als aktiv

identifiziert wurden. Dadurch bedingt, werden nur etwa die Hälfte der aus den Mobilfunkdaten möglichen aktiven Trajektorien an die Verkehrslageschätzung weitergereicht. Die Chance, dass ein beliebiger Datensatz eines aktiven MFT auch korrekt, sowohl passiv als auch aktiv, identifiziert wird, beträgt 72%[6)] bei FPD und 91%[5)] bei GPS. Und weiter: Ein beliebiger als aktiv identifizierter MFT-Datensatz der FPD-Datensätze, wurde schließlich mit einer Wahrscheinlichkeit von nur 8% falsch erkannt.

Die folgende Abbildung 4-22 zeigt die Verteilungen der korrekt aktiv und passiv erkannten sowie die jeweiligen Anteile der unkorrekt identifizierten Datensätzen. In Abbildung 4-22 a) sind die jeweilige Anzahl der korrekt bzw. unkorrekt identifizierten Szenarien der GPS-Messungen und in Abbildung 4-22 b) der FPD Auswertung dargestellt.

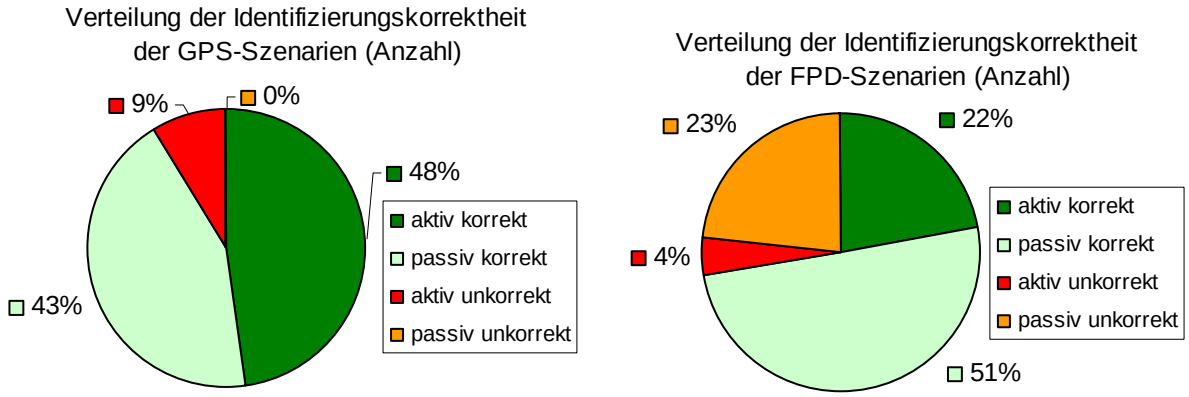


Abbildung 4-22: Darstellung der korrekt bzw. unkorrekt identifizierten Szenarien links der GPS-Daten, rechts der FPD-Daten

Die folgenden beiden Darstellungen der Abbildung 4-23 berücksichtigen neben der Anzahl der Szenarien auch deren zeitliche Dauer. Das heißt die Prozentangaben beziehen sich auf die Dauer der identifizierten Gespräche der MFT. Die einzelnen Anteile verschieben sich nur unwesentlich.

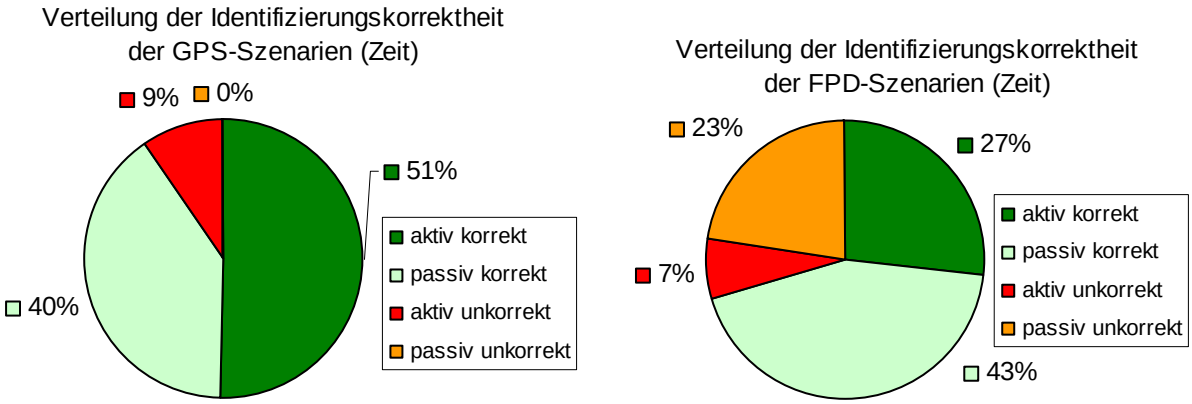


Abbildung 4-23: Darstellung der korrekt bzw. unkorrekt identifizierten Szenarien mit Berücksichtigung der Messzeiträume links der GPS-Daten, rechts der FPD-Daten

Abschließend kann man feststellen, dass etwa die Hälfte der telefonierenden im verkehrlichen Sinne aktiven Verkehrsteilnehmer erkannt werden. Das ist bei Hochrechnungen des Verkehrsaufkommens zu berücksichtigen.

4.4 Analysen der verketteten Identifizierung von MFT-Datensätzen

Die bisher durchgeführten Analysen haben jeden Datensatz als Einzelereignis eines MFT betrachtet. Zum einen generiert ein Mobilfunkteilnehmer zumeist mehrere Datensätze pro Tag, zum anderen kann durch die Einzelverarbeitung „Identifikationswissen“ z.B. vorheriger Auswertungen nicht genutzt werden. Datensätze mit Identifikation „stehend“ werden infolgedessen immer passiv eingestuft. Ebenso erfolgt die passiv-Einstufung für Datensätze, bei denen mehrere Routen auf mehreren ggf. konkurrierenden Verkehrsnetzen (Straßenbahn, Straße) möglich sind.

Um auch Datensätze mit uneindeutiger Verkehrsklassenidentifizierung oder von stehenden MFT zu identifizieren, bei denen von einem vorherigen Datensatz bereits ein aktives Verhalten bekannt ist, werden die einzelnen MFT-Datensätze im Folgenden als verkettet betrachtet. Bei der Annahme, dass sich das Verkehrsklassenverhalten des MFT nicht geändert hat, kann bspw. bei den einleitend genannten Datensätzen die vorherige aktiv Einstufung übernommen werden. Zu beachten ist aber auch, dass durch die Verkettung eine aktive Identifikation z.B. bei Identifizierungsuneindeutigkeiten zu Unrecht übernommen wird, wenn zwischenzeitlich in eine passive Verkehrsklasse gewechselt wurde. Das verkehrsteilnehmer-spezifische Verhalten für Verkehrsklassenwechsel muss, besonders bei anonymisierten Datensätzen, als unbekannt angesetzt werden. Intermodale Knotenpunkte könnten einen derartigen Wechsel ggf. räumlich eingrenzen, sie sind jedoch nicht für alle hier betrachteten Wechsel verfügbar bzw. praktikabel realisierbar. Als Beispiel sei ein Wechsel von PKW zu Fußgänger genannt, der auf allen Straßenwegen möglich wäre.

Um dennoch die „Erfahrungswerte“ aus der vorherigen Identifikationen des MFT zu nutzen, wurde der Parameter $dt_{\max\,active}$ für die maximal zulässige Zeitdifferenz zur Übernahme einer früheren, eindeutig identifizierten aktiv-Einstufung bei uneindeutiger oder „stehender“ Identifikation eingeführt. Der konstante Parameter $dt_{\max\,active}$ wird der Komplexität des Verkehrsteilnehmerverhaltens sicherlich nicht vollends gerecht, jedoch wird die Möglichkeit gegeben, stehende MFT in Regionen mit räumlich uneindeutiger Verkehrsnetzzugehörigkeit oder uneindeutige Ergebnisse der zeitlich-topologischen Routensuche identifizieren zu können. Bspw. sind damit MFT-Datensätze in Verkehrsstaus als aktiv identifizierbar, obwohl Fußwege oder Bahntrassen in räumlicher Nähe verlaufen, vgl. Verkehrsnetzzugehörigkeit in Abschnitt 4.3.4. Die Auswirkungen der verketteten Identifizierung können wegen dem unbekannten MFT-Verkehrsklassenverhaltens der anonymisierten Daten nur anhand der Versuchsfahrten betrachtet werden. Die Übertragbarkeit dieser Statistiken auf die restlichen anonymen MFT muss offen bleiben.

Aufgrund der Verkettung der Datensätze werden zur Bestimmung der Korrektheitsmaße die Szenarien nicht, wie bisher geschehen, getrennt betrachtet und deren Korrektheit bestimmt. Vielmehr werden die aktiv/passiv-Einstufungen der Zeiträume der Soll-Identifizierung mit denen der Ist-Identifizierung verglichen. Stimmen die Soll- und Ist-Identifizierung überein, wird der Zeitraum als korrekt und bei nicht Übereinstimmung als unkorrekt gewertet. Als Maß der Korrektheit wird das Richtigkeitsmaß $RM^{Soll:A}$ (aktive Szenarien werden korrekt als aktiv identifiziert) definiert, bei dem die Zeiträume der zu unrecht passiv eingestuften Szenarienzeiten $t_{Ist:P}^{Soll:A}$ ermittelt werden:

$$RM^{Soll:A} = 1 - \frac{t_{Ist:P}^{Soll:A}}{T} \tag{4-6}$$

Analog dazu folgt das Maß der Korrektheit des Richtigkeitsmaßes $RM^{Soll:P}$ (passive Szenarien werden als passiv identifiziert), für die zu unrecht aktiv eingestuften Szenarienzeiten $t_{Ist:A}^{Soll:P}$:

$$RM^{Soll:P} = 1 - \frac{t_{Ist:A}^{Soll:P}}{T} \tag{4-7}$$

Die Grundgesamtheit beider Korrektheitsmaße basiert auf der Summe T aller betrachteten Szenarienzeiten.

Anmerkung: Die Versuchsfahrten vom 13.09.2007 fanden vollständig als MIV statt. Die Fahrten 121 bis 152 (Standphasen) sind zwar keine Fahrten in Bewegung, eine Verkehrsklassenänderung findet im Sinne der Versuchsfahrten jedoch ebenfalls nicht statt, da sie ursprünglich einen Verkehrsstau simulieren sollten. Für alle Fahrten des 13.09.2007 gilt somit die Soll-Identifikation aktiv. Bei den Versuchsfahrten vom 07.11.2007, 19.12.2007 und 27.02.2009 wurden ausschließlich passive Datensätze erzeugt. Änderungen der Verkehsteilnehmereinstufung fanden hier ebenfalls nicht statt. Infolgedessen können mit den Versuchsfahrtdaten letztlich nur falsch erkannte Einstufungen, aktiv anstatt passiv oder passiv anstatt aktiv, als Auswirkungen der verketteten Identifizierung beobachtet werden.

Die Referenzfahrten wurden aus verkehrstechnischen Gründen teilweise nicht auf den Strecken der Szenariendefinitionen durchgeführt. Markant zeigt sich das an den Standszenarien des 13.09.2007, die einen Stau auf einer Autobahn simulieren sollten. Abbildung 4-24 zeigt, dass die Standszenarien zwar in der Nähe der Autobahn, jedoch ebenso in direkter Nachbarschaft zu einer Linienbusverbindung zwischen Grötzingen und Hagsfeld (Ost-West-Richtung) sowie einer Straßenbahn- und DB Bahnverbindung zwischen Weingarten und Durlach (Nord-Süd-Richtung) stattfanden. Unbeabsichtigt für die Referenzfahrten war die Korrelation zu eben diesen ÖV-Verkehren.

Abbildung 4-24: Ausschnittsvergrößerung der Abbildung 3-4 (etwa Mitte), Standszenarien symbolisiert als schwarzes Dreieck im Bild oben

Infolgedessen kommt es zu den Zeiten der Standszenarien (Herdweg L604) zu einer unbeabsichtigten Korrelation mit einem Linienbus (KVV Linie 31) und Straßenbahn (KVV S4 bzw. S5)

Zwischenfazit: Mit den durchgeführten Referenzfahrten ist direkt das Verhalten der Identifizierung bei Verkehrsklassenwechsel aufgrund fehlender Verkehrklassenwechsel aktiv ←→ passiv nicht überprüfbar. Indirekt werden die Auswirkungen nach den falsch identifizierten Messzeiträumen untersucht. Außerdem wird für die MIV Fahrten des 13.09.2007 zum Zweck der Nachweisführung der prinzipiellen Funktionsfähigkeit das Identifikationsverfahren ÖV-Korrelation abgeschaltet.

In den nachfolgenden Grafiken werden für die Versuchsfahrten des 13.09.2007 die Zeiten der korrekt und der unkorrekt identifizierten Zeiten angegeben. Die Soll-Identifizierung ist in grün dargestellt und ist für alle Zeiträume definitionsgemäß aktiv (jeweils linke Linie). Die ermittelte Ist-Identifizierung ist blau gezeichnet (jeweils rechte Linie). Zwischen diesen beiden Linien sind zur Verdeutlichung die von den Soll-Identifizierungen abweichenden Zeiträume rot markiert (jeweils mittlere Linie). In der folgenden Abbildung 4-25 sind die Ergebnisse für das Extrem $dt_{\max\,active} = 0\,\min$, ohne Übernahme von aktiv-Einstufungen dargestellt.

Die Abbildung 4-25 a) zeigt die Identifizierung der GPS-Daten, b) und c) die Identifizierungen der FPD-Daten des Handys 1 und des Handys 2. Gelb hinterlegt sind die Zeiträume der Standphasen.

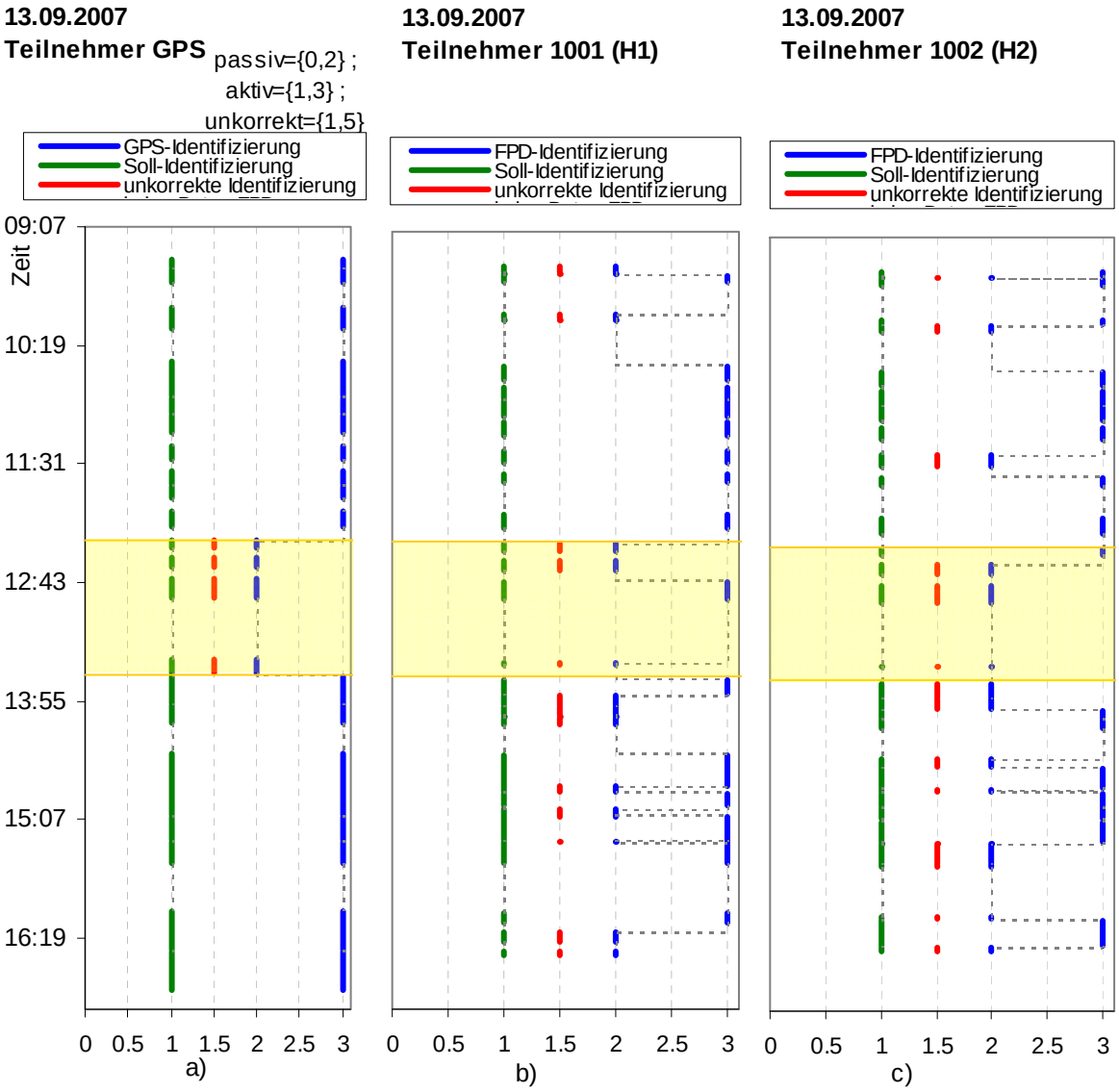


Abbildung 4-25: Zeitlicher Vergleich korrekter und unkorrekter aktiv/passiv-Einstufung für die Versuchsfahrten des 13.09.2007, a) GPS, b) MFT Handy 1 und c) MFT Handy 2 jeweils ohne Übernahme vorhergehender Identifizierungseinstufungen ($dt_{\max\,active} = 0\,\min$).

In Abbildung 4-25 a) sind deutlich die passiv-Einstufungen der Standphasen in den gelb hinterlegten Zeiträumen von ca. 12:17 Uhr bis ca. 13:39 Uhr zu erkennen. Im gleichen Zeitraum lassen sich auch bei den FPD-Daten von Handy 1 und Handy 2 die passiv-Einstufungen, bis auf eine fehlerhafte aktiv-Einstufung bei Handy 1, der Standszenarien wiederfinden. Deutlich davon unterscheiden sich die Ergebnisse bei denen ein Übernahmezeitraum für die aktiv-Einstufungen von bis zu 30 Minuten ($dt_{\max\,active} = 30\,\min$) zugrunde lag, dargestellt in den Grafiken der nachfolgenden Abbildung 4-26.

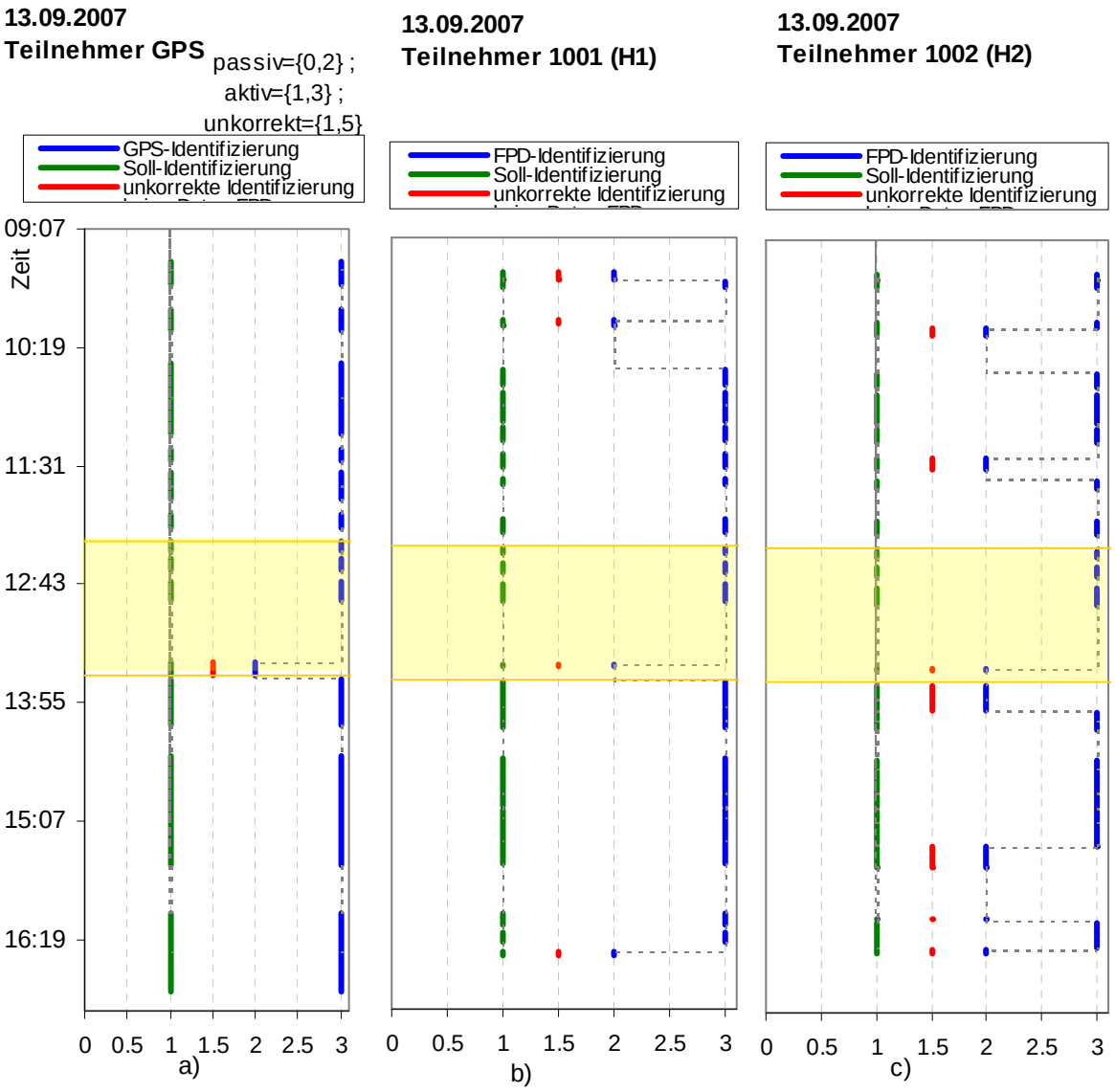


Abbildung 4-26: Zeitlicher Vergleich korrekter und unkorrekter aktiv/passiv-Einstufung für die Versuchsfahrten des 13.09.2007, a) GPS, b) MFT Handy 1 und c) MFT Handy 2 jeweils mit Übernahme einer vorhergehenden aktiv Einstufung von bis zu 30 Minuten ($dt_{\max\,active} = 30\,\min$).

Bis auf die letzte Standphase, die mit 39 Minuten etwa 9 Minuten länger als die maximale Übernahmezeit von 30 Minuten zeitlich entfernt ist, wurden alle Standphasen sowohl bei GPS als auch bei FPD als aktiv eingestuft. Die prinzipielle Funktionstüchtigkeit der Verkettung von Datensätzen zur Übernahme von aktiv-Einstufungen von früheren Datensätzen eines MFT ist damit nachgewiesen.

4.5 Statistiken zur 24h Auswertung des 19.12.2007

Mit der Analyse der Identifizierung von Mobilfunkdaten eines gesamten Tages soll im Folgenden die praxisgerechte Einsatztauglichkeit der hier vorgestellten Identifizierungsverfahren nachgewiesen werden. Als Datenbasis wurde die Mobilfunkdatenaufzeichnung der Messfahrt vom 19.12.2007 ausgewählt. Die Auswertung betrifft die LAC 29453 Karlsruhe für einen Zeitraum von 24 Stunden. Die Identifikation der anonymisierten MFT-Datensätze erfolgte zur Untersuchung der Auswirkung der Übernahme der aktiv-Einstufungen für unterschiedliche Zeitdifferenzen $dt_{\max active}$ verkettet. In der Tabelle 4-20 sind einige Eckdaten der Datenbasis, sowie Grundgesamtheiten für Durchschnittwerte, Datenraten und weiteren unten aufgeführten Statistiken zusammengefasst.

Tabelle 4-20: Statistiken der Tagesdatenauswertung vom 19.12.2007

Merkmal	**Wert**
Datum: [dd.mm.yyyy]	19.12.2007
Region:	Karlsruhe, LAC 29453
Erster Positionseintrag: [hh:mm:ss.sss]	00:00:00.006
Letzter Positionseintrag: [hh:mm:ss.sss]	23:59:59.991
Gesamtanzahl Datentelegramme (Abis- und A-Schnittstelle):	60 238 421
Datentelegramme (nur A-Schnittstelle):	4 085 422
Datentelegramme (nur Abis-Schnittstelle):	56 152 999
Gesamtanzahl MFT	833 244
Gesamtanzahl Telefonate (calls) (Abis- und A-Schnittstelle):	636 739
Telefonate (calls) (nur A-Schnittstelle):	635 461
Telefonate (calls) (nur Abis-Schnittstelle):	586 325
Nach Ortung mittels „Signalstärken-Matching“	
Anzahl Telefonate (calls) mit mindestens einer Position:	172 061
Anzahl der georteten Positionen:	36 509 929
Gesamtanzahl MFT:	208 266
MFT mit mindestens einer Position:	173 908
MFT mit mindestens 25 Positionen (Grenzwert für Bewegungsanalyse und Schwerpunktbestimmung):	129 028
MFT mit identifizierter aktiver MIV-Verkehrsrelevanz	13 761
MFT mit identifizierter passiver MIV-Verkehrsrelevanz	194 505

Die nachfolgende Abbildung 4-27 zeigt die Häufigkeiten der Mobilfunkdatentelegramme der Do-iT-Abis-Schnittstelle. Die Klassenbreite beträgt 1 Minute. Etwa gegen 12:30 Uhr ist ein markanter Abfall der Telegrammrate erkennbar, der auf ein versehentliches Abschalten der Datenaufzeichnung seitens der TMD zurückzuführen ist. Um 14:33 Uhr wurde die maximale Rate von 82 835 Datentelegrammen je Minute (Tel/m) erreicht. Der Mittelwert beträgt 41 774.2 Tel/m, der Median 46 745 Tel/m.

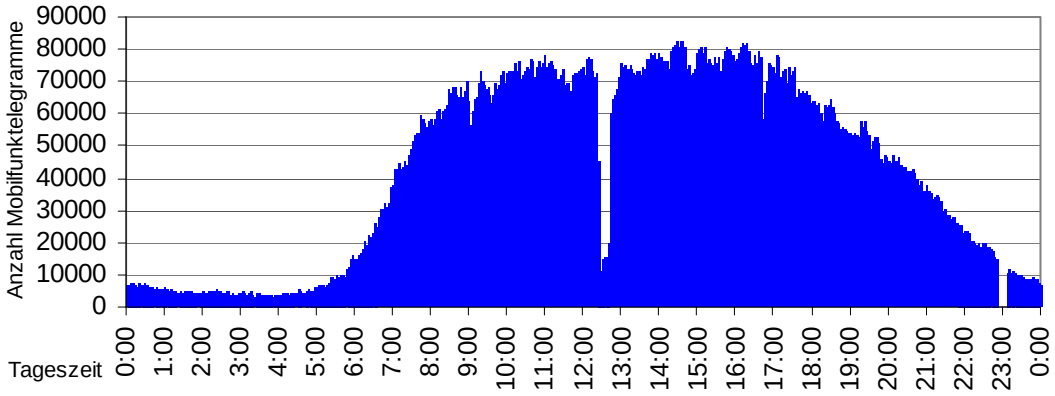


Abbildung 4-27: Häufigkeiten der Mobilfunktelegramme der Do-iT-Abis-Schnittstelle der LAC 29454 am 19.12.2007, Klassenbreite 1 Minute

Die Daten wurden mit dem Ortungsverfahren „Signalstärken-Matching“ aus AP4-5 [3] geortet und anschließend den Identifizierungsverfahren zugeführt. Wie bereits einleitend genannt, wurde die Identifizierung für unterschiedliche $dt_{\max\,active}$ durchgeführt. Ausgewählte Ergebniswerte können der Tabelle 4-21 entnommen werden. Gelb hinterlegt ist die bereits in Kapitel 4.4 ausgewählte aktiv-Übernahmezeit $dt_{\max\,active} = 30\,\min$. Für die Identifizierungsdurchläufe der 24 Stundenauswertung wurden Berechnungszeiten von jeweils ca. 15,5 Stunden benötigt, womit die prinzipielle Einsatztauglichkeit der Identifizierungsverfahren für den Echtzeitbetrieb gezeigt werden konnte. Die Daten des zugrundeliegenden Rechners sind unten in Tabelle 4-21 näher spezifiziert. Mit $dt_{\max\,active} = 30\,\min$ wird durchschnittlich ca. alle 3 Sekunden (0.3579 je Sekunde) eine aktive Trajektorie für die LAC 29453 ermittelt.

Tabelle 4-21: Statistiken der Identifikationsergebnisse der Tagesdatenauswertung vom 19.12.2007

$dt_{\max\,active}$ [Minuten]	Prozessierungs-dauer [hh:mm:ss]	Anzahl aktiv identifizierter Trajektorien	durchschn. bearbeitete MFT je Sekunde	durchschn. aktive Trajektorien je Sekunde	durchschn. Anzahl Trajektorien je Teilnehmer
0	15:30:24	15 650	0.2465	0.2803	0.0751
1	15:24:57	15 656	0.2480	0.2821	0.0752
2	15:24:10	16 405	0.2482	0.2959	0.0788
5	15:24:13	17 388	0.2482	0.3136	0.0835
15	15:23:02	18 711	0.2485	0.3379	0.0898
30	15:24:30	19 854	0.2481	0.3579	0.0953
60	15:24:13	21 138	0.2482	0.3812	0.1015
120	15:22:57	22 562	0.2485	0.4074	0.1083
86400	15:26:08	22 758	0.2476	0.4096	0.1093

Eine grafische Darstellung der Abhängigkeit der Trajektorienanzahl von der maximal zulässigen Zeitdifferenz $dt_{\max\,active}$ gibt Abbildung 4-28. Wie bereits im Kapitel 4.4 erwähnt, muss bei Verwendung einer höheren Zeitdifferenz $dt_{\max\,active}$ auch eher mit nicht entdeckten Verkehrsklassenwechsel aktiv → passiv gerechnet werden. Zu erkennen ist, dass eine hohe

Zuwachsrate der Trajektorienanzahl bereits im Bereich von 0 bis 30 Minuten erreicht wird und damit in diesem Bereich auch mit den höchsten Gewinnen an zusätzlichen Verkehrslageinformationen bei ggf. vertretbarer Menge an Falscheinstufungen gerechnet werden kann.

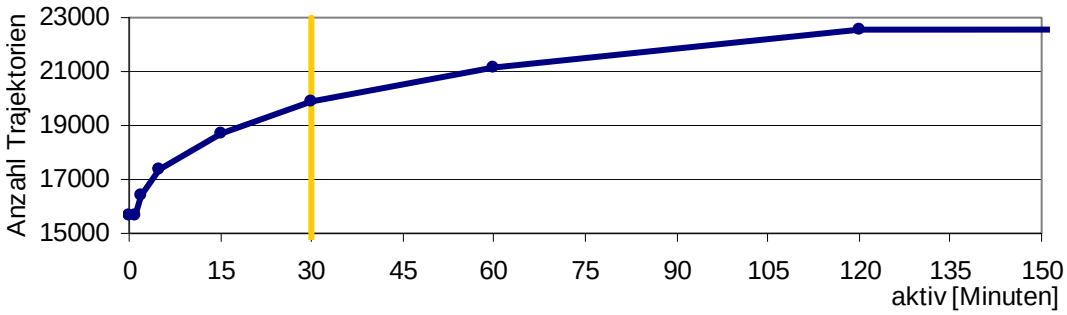


Abbildung 4-28: Abhängigkeit der Trajektorienanzahl von der maximal zulässigen Zeitdifferenz für die Übernahme von aktiv-Einstufungen vorhergehender MFT-Datensätze

Die Tagesauswertung für den 19.12.2007 wurde ebenfalls mit einer weiteren Identifikationsbedingung durchgeführt, bei der zusätzlich die Trajektorien aktiv eingestuft wurden, bei denen mit nachgewiesenem Stand eine alleinige Zugehörigkeit zu Straßennetz car_truck ermittelt wurde. Insofern weder die ÖV-Korrelation noch die Verkehrsnetzanalyse die MIV-Verkehrsklasse ablehnten, wurden diejenigen Trajektorien als aktiv eingestuft. Bereits bei der Identifikation mit $dt_{\max\,active} = 0\,\min$ wurden 66 067 aktive Trajektorien gefunden. Das entspricht im Vergleich zur sonst eingesetzten Bedingungsregel (Identifizierung von mindestens einer Route auf dem Straßennetz) etwa der 4-fachen Anzahl der aktiven Trajektorien. Obwohl die Rate der dadurch aktiv identifizierten Trajektorien deutlich höher ist, fehlt den Daten der Nachweis der Verkehrsbeteiligung. Es wird vermutet, dass es sich in den meisten Fällen um stehende MFT abseits von Schienenwegen handelt, was alleinig die Verkehrsbeteiligung jedoch noch nicht genügend bestätigen lässt. Dieser Art der Identifikation wurde schließlich zu Gunsten des besseren Nachweises der Verkehrsbeteiligung, jedoch auch zu Lasten der geringeren Anzahl aktiver Trajektorien nicht weiter verfolgt.

4.5.1 Statistik zur tageszeitlichen Trajektorienverteilung

Aus der Abbildung 4-27 konnten bereits unterschiedliche von der Tageszeit abhängige Datentelegrammhäufigkeiten gelesen werden. In der nachfolgenden Abbildung 4-29 zeigt sich ein ähnliches Verhalten für die aktiven Trajektorien. In den Mittags- und Nachmittagsstunden werden die höchsten Trajektorienraten erzielt. Mit Berücksichtigung von $dt_{\max\,active} = 30\,\min$ werden maximale Raten um 10:30 Uhr von 847 Trajektorien je 30 Minuten und um 15:00 Uhr von 843 Trajektorien je 30 Minuten erreicht. Etwa um 12:30 Uhr zeigt sich wieder ein Abfallen der Rate, die vermutlich Ursache der Unterbrechung der Datenaufzeichnung ist. Deutlich fallen die höheren Trajektorienraten in den Mittagsstunden für $dt_{\max\,active} = 30\,\min$ (orange) gegenüber denen für $dt_{\max\,active} = 0\,\min$ (blau) oder auch

$dt_{\max\,active} = 5\,\min$ (grün) auf. Abhängigkeiten einer höheren Verkehrsstaurate zu diesen Zeiten wurden nicht untersucht. Ein Nachweis von Abhängigkeiten zwischen der höheren Trajektorienrate und ggf. einer Verkehrsstaurate könnte die Wahl des Wertes von $dt_{\max\,active}$ untermauern.

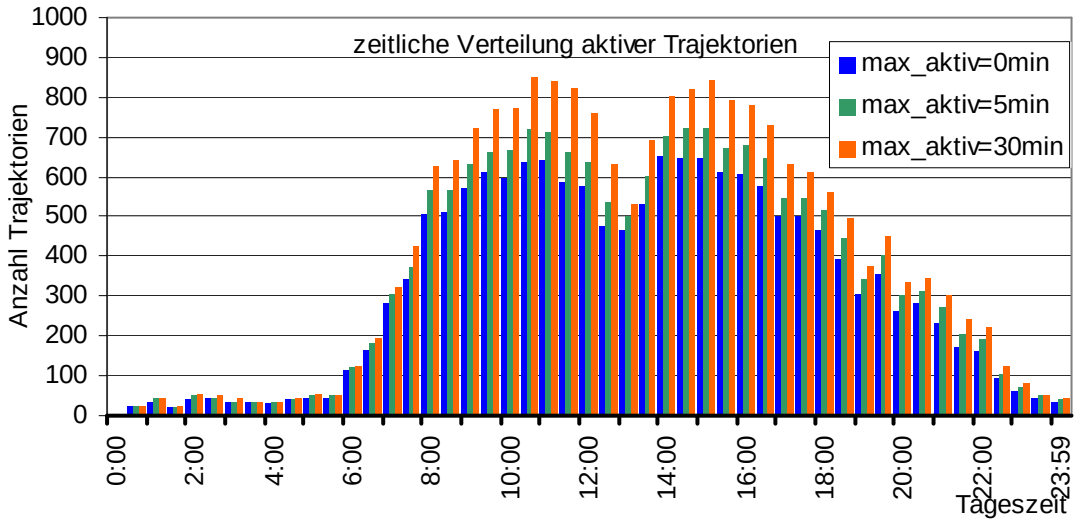


Abbildung 4-29: Zeitliche Verteilung der aktiven Trajektorien des 19.12.2007 für unterschiedliche $dt_{\max\,active}$, Klassenbreite 30 Minuten

Es fällt auf, dass im Vergleich zum Einbruch der Telegrammrate in Abbildung 4-27 der Einbruch der Trajektorienrate in Abbildung 4-29 geringer ausfällt. Zur Erzeugung des Histogramms der Abbildung 4-29 wurden die Mittelwerte aus Anfangszeit und Endzeit der Trajektorien verwendet, sodass keine „scharfe“ zeitliche Zuordnung der Trajektorienanzahl zu spezifischen Zeitpunkten wiedergegeben wird. Außerdem wurde eine größere Klassenbreite in Abbildung 4-29 von 30 min gewählt im Vergleich zu 1 min in Abbildung 4-27, weshalb zusätzlich eine Art Verwischungseffekt zu erwarten ist. Des Weiteren lässt sich diese Eigenheit auf die Verkettung zurückführen.

4.5.2 Flächenabdeckung und räumliche Verteilung

Ein wesentliches Kriterium für den Einsatz von FPD ist die Möglichkeit flächendeckende Verkehrslageinformationen daraus generieren zu können. Die nachfolgenden Bilder zeigen daher die räumliche Verteilung der aus den Ortungsdaten des „Signalstärken-Matchings“ abgeleiteten Positionsschwerpunkte. Zur Orientierung sei genannt, dass sich ca. 4 km nordwestlich der Bildmitten das Zentrum von Karlsruhe und ca. 8 km südlich davon das Zentrum Ettlingens befindet, vgl. jeweils die Maßstabsleiste in den Abbildungen unten links.

Je Kachel (100 m x 100 m) wird die Häufigkeit der Positionsschwerpunkte bestimmt und farblich wiedergegeben. Je dunkler bzw. kräftiger der Farbton der Kachel ist desto mehr Positionsschwerpunkte beinhaltet sie. In gelb sind die Autobahnen A5 und A8, in rot die Bundesstraßen und in dünnen schwarzen Linien sind ausgewählte Landes-, Kreis- und

Gemeindeverbindungsstraßen dargestellt. Die kräftig schwarzen Linien zeigen die DB-Bahn Schienen. Im Hintergrund der Bilder sind die Bediengebiete der Mobilfunkzellen des sogenannten Bestserver-Plots der TMD in rosa dargestellt. Auffällig ist, dass nicht in allen Bereichen des Bestserver-Plots Positionsschwerpunkte zu finden, aber wiederum zahlreiche Positionsschwerpunkte außerhalb vorhanden sind. Ursachen können in den abweichenden Signalausbreitungen gegenüber der Bestserver-Plot-Darstellung gefunden werden. Der Bestserver-Plot zeigt nur die Regionen einer stationären Mobilfunkstation, in denen die Station mit maximaler Empfangsstärke vermutet wird. In anderen Regionen kann sie dennoch bspw. als zweit, dritt usw. ebenfalls empfangbar sein, wie anscheinend westlich des Bestserver-Plots häufig der Fall ist. Man kann daraus schlussfolgern, dass der Bestserver-Plot zwar für Visualisierungen gut geeignet ist, jedoch für die Positionsbestimmung von MFT deutliche Genauigkeitsgrenzen auferlegt. Die Positionsbestimmung im vorliegenden Bericht stützt sich nicht auf diesen Plot, weshalb folgend ausschließlich die Visualisierungsvorteile genutzt werden können.

In der ersten Abbildung 4-30 sind die Häufigkeiten der Positionsschwerpunkte sowohl für passive als auch für aktive Trajektorien dargestellt. Man kann daraus die räumliche Verteilung sowie die Menge der Telefonate ableiten. Zeitliche Variationen können hingegen nicht entnommen werden.

Die anschließende Abbildung 4-31 zeigt ausschließlich die Positionsschwerpunkte und deren Häufigkeit für aktive Trajektorien. Man erkennt, dass besonders entlang der Hauptverkehrsachsen weiterhin eine hohe Dichte und Häufigkeit der Positionsschwerpunkte zu finden sind. Die Dichte und Häufigkeit in der Nähe zu Schienenwegen ist im Vergleich geringer, was sich besonders südlich und südwestlich von Ettlingen zeigt. Wenn auch geringer als in Nähe der Autobahnen und Bundesstraßen, finden sich bei zahlreichen Landes-, Kreis- und Gemeindeverbindungsstraßen Positionsschwerpunkte aktiver Trajektorien. Die Tauglichkeit, dass FPD in der Fläche möglich ist und diese FPD prinzipiell auch als aktive Verkehrsteilnehmer gefunden werden können, ist damit gezeigt.

© Do-iT 2009

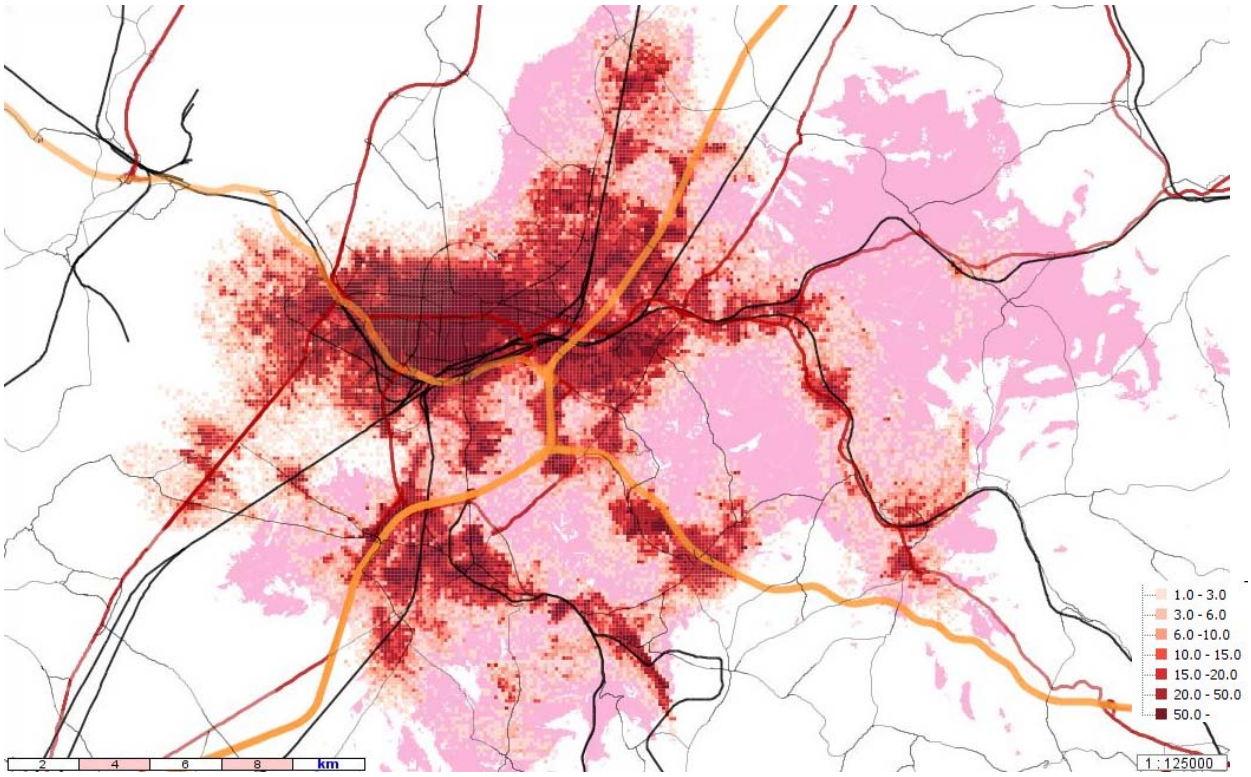


Abbildung 4-30: Räumliche Verteilung und Häufigkeit (vgl. Legende rechts unten im Bild) aller Positionsschwerpunkte von aktiven und passiven Trajektorien

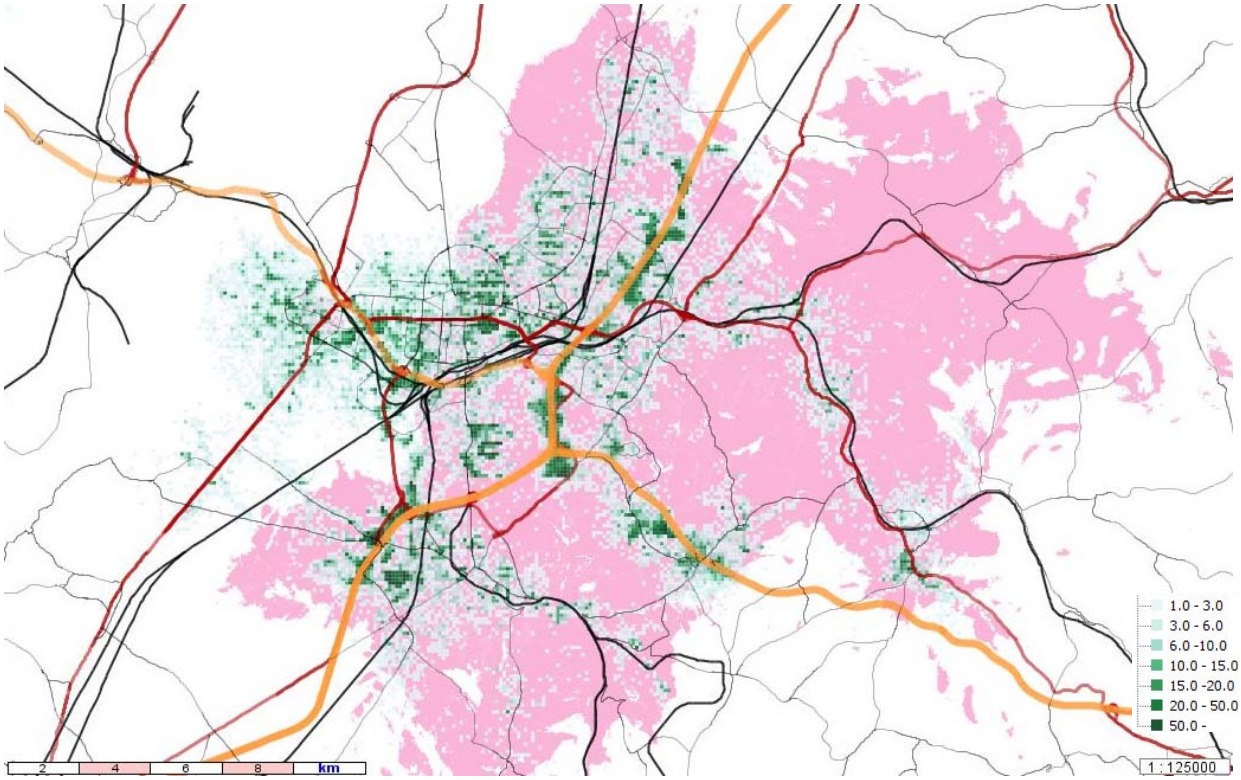


Abbildung 4-31: Räumliche Verteilung und Häufigkeit (vgl. Legende rechts unten im Bild) der Positionsschwerpunkte der aktiv identifizierten Trajektorien

Als Nachweis für die aktiv-Einstufung eines Datensatzes, wird in der Verkehrsnetzanalyse eine Route gesucht, die den zeitlichen und topologischen Kriterien aus den Positionsschwerpunkten genügt. Einzige Ausnahme sind bei der verketteten Identifikation die aktiv eingestuften stehenden Datensätze. Es sei daran erinnert, dass in der Verkehrsnetzanalyse das Finden einer beliebigen den Kriterien entsprechenden Route als Nachweis genügt. Aus Effizienzgründen endet die Routensuche, wenn eine erste Route aus einer ggf. zufälligen Start-Zielknotenkombination gefunden wird. Diese Route gehört zur Menge der „möglichen Routen“. Eine Wahrscheinlichkeit für die Wahl dieser Route durch den MFT kann daraus jedoch nicht abgeleitet werden. Außer dass die Wahrscheinlichkeit der gefundenen Route größer null ist. Man kann jedoch annehmen, dass die „wahre“ Route, wenn nicht mit der gefundenen identisch, zumindest in räumlicher Nähe zu finden ist. Aus dieser Überlegung heraus, kann die nachfolgende Darstellung der Kantenhäufigkeiten in Abbildung 4-32 mindestens als regionale Eingrenzung der Routen der aktiven Trajektorien verstanden werden. Für eine detailliertere Darstellung sei auf [5] verwiesen.

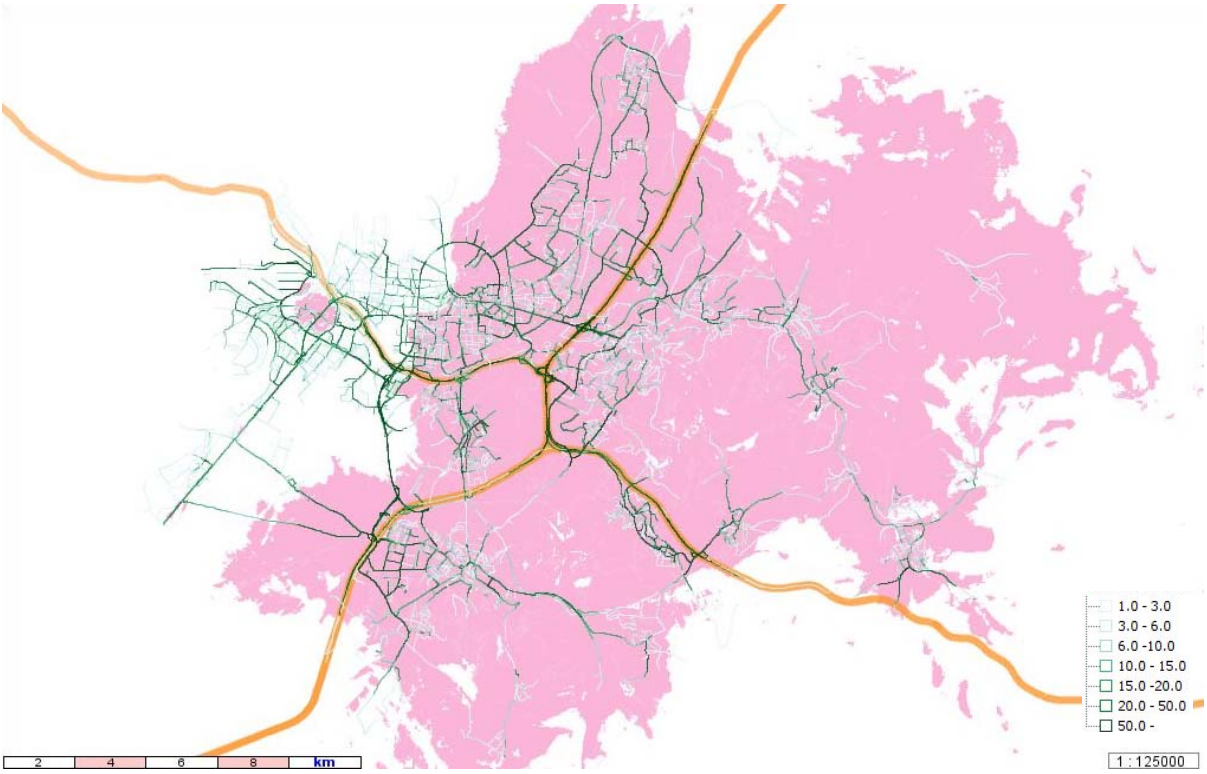


Abbildung 4-32: Häufigkeiten der sub-edges (vgl. Legende rechts unten im Bild) der erstgefundenen Routen der Verkehrsnetzanalyse für die aktiven Trajektorien

Zur besseren Übersicht sind nur die Autobahnen in gelb, der Bestserver-Plot der TMD in rosa und die Kanten der aktiven Trajektorien in der nach Häufigkeit klassifizierten Farbgebung, vgl. auch Legende unten rechts im Bild, dargestellt. Die Verkehrsnetzanalyse verwendet eine höhere räumliche Auflösung, als sie durch die Kanten der digitalen Straßenkarte gegeben ist. Die Kanten werden in sogenannte sub-edges (Unterkanten) aufgeteilt, wenn sie eine Kachelgrenze überschreiten. Die Kacheln selbst haben eine Größe

von 100m x 100m, wodurch letztlich eine räumliche Auflösung der Kanten von mindestens derselben Größe folgt. Die maximal vorhandene Häufigkeit einer sub-edge für die 24h-Auswertung des 19.12.2007 betrug 1428. Auf den Autobahnen und Bundesstraßen finden sich erwartungsgemäß hohe Häufigkeiten. Aber auch abseits dieser Verkehrswege sind „in der Fläche“ zahlreiche Kanten mit hoher Frequentierung zu finden. Die Einsatzmöglichkeit der FPD zur Verkehrslageerfassung in diesen Regionen sei hiermit erneut bestätigt.

4.5.3 Rechengeschwindigkeit

Zu Beginn des Kapitels wurde bereits auf die Identifizierungsdauer der gesammelten Mobilfunkdaten aus 24 Stunden von ca. 15,5 Stunden hingewiesen. Dieser Wert gilt nahezu unabhängig vom gewählten $dt_{\max active}$. Die folgenden Statistiken beziehen sich auf $dt_{\max active} = 0$. Die Berechnung einer aktiven Trajektorie benötigt im Mittel ca. 1.17 Sekunden, der Median liegt bei 0.52 Sekunden. Eine Verteilung der Berechnungszeiten der aktiven Trajektorien kann der Abbildung 4-33 entnommen werden.

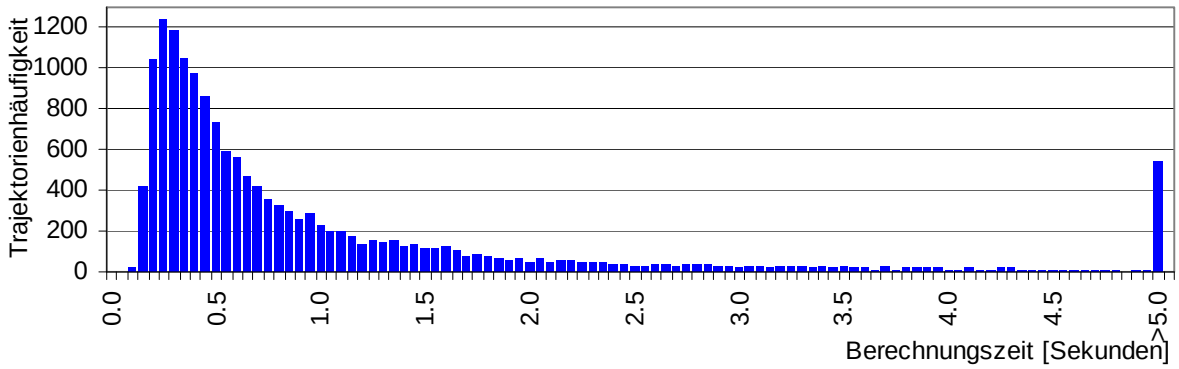


Abbildung 4-33: Trajektorienhäufigkeiten in Abhängigkeit ihrer Berechnungsdauer

Die nachfolgende Abbildung 4-34 zeigt grafisch die Anteile der jeweiligen Identifizierungsverfahren an der Berechnungsdauer der aktiven Trajektorien. Auffällig ist, dass häufig 80% und mehr der Berechnungsdauer durch das Routing mit zeitlichen und topologischen Kriterien der Verkehrsnetzanalyse entfallen. Im Vergleich zu den anderen Identifizierungsverfahren sind hier die größten rechen- und programmiertechnischen Zeitvorteile durch zukünftige Optimierungen zu erwarten. Eine weitere Erhöhung der notwendigen Rechenzeit ergäbe sich, wenn anstatt des Abbruchs des Routings nach Finden einer ersten Route, die Suche nach der kürzesten Route erfolgen würde, wie es bspw. im Map-Aiding, vgl. Abbildung 1-1, nach der Identifikation aktiver Verkehrsteilnehmer geschieht. Das Verfahren ist bereits implementiert, wurde hier jedoch nicht untersucht. Aus den Positionsschwerpunkten kann für die umgebenen Verkehrsnetzelemente eine Aufenthaltwahrscheinlichkeit abgeleitet werden. Daraus wiederum könnten Wahrscheinlichkeiten für gefundene Routen berechnet oder gezielt Routen mit hohen Aufenthaltswahrscheinlichkeiten gesucht werden. Diese Verfahren wurden bislang nicht implementiert, sodass sie ggf. in Folgearbeiten zu untersuchen sind.

Zwischenfazit: Es kann festgehalten werden, dass die in Do-iT geforderten zeitlichen Kriterien für den Online-Einsatz der Identifikation aktiver Verkehrsteilnehmer mit ca. 15,5 h Berechnungsdauer je 24 h Mobilfunkdaten im Abis-Demonstrator Karlsruhe eingehalten werden.

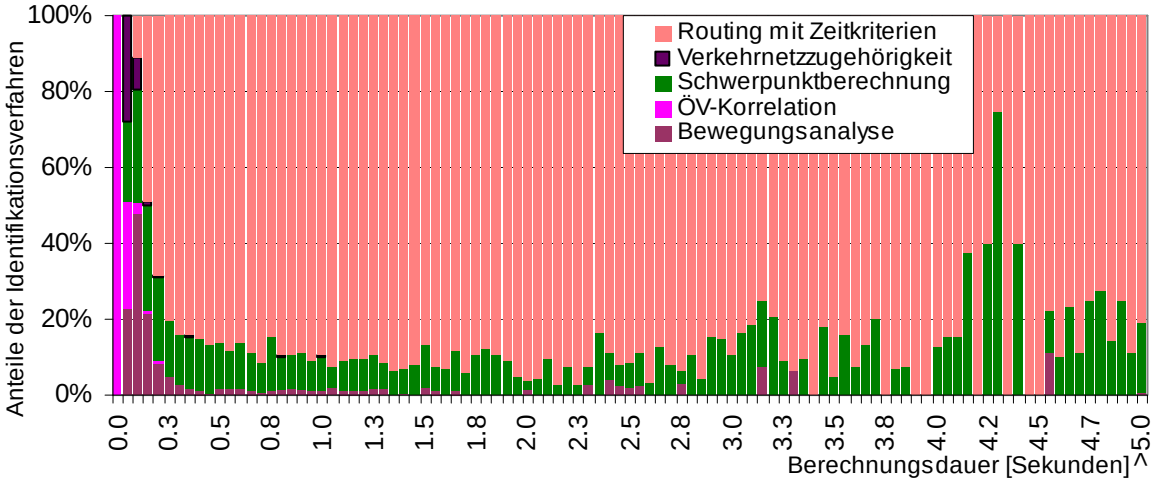


Abbildung 4-34: Anteile der Identifikationsverfahren an der Berechnungsdauer der aktiven Trajektorien in Abhängigkeit zur Gesamtberechnungsdauer, aufgeschlüsselt nach den einzelnen Identifizierungsverfahren

In der nachfolgenden Tabelle 4-22 sind wesentliche Eckdaten des Rechners zusammengefasst, der für die Berechnungen und Auswertungen in diesem Bericht eingesetzt wurde. Es handelt sich um einen handelsüblichen Desktop-PC aus dem Jahr 2008. Die Zeitmessungen basieren auf programmtechnische Ablesungen der systemeigenen Uhr.

Tabelle 4-22: Daten des verwendeten Rechners für die Identifizierungsberechnungen des vorliegenden Berichtes

Merkmal	**Wert**
Prozessortyp	Intel® Core™2 Quad Q6600
Prozessortaktfrequenz	2.40 GHz
Arbeitsspeichertyp	DDR2 800MHz PC2-6400
Arbeitsspeichergröße	4.0 GB
Betriebssystem	Microsoft® Windows™Vista 64-Bit
Datenbanksoftware	MySQL® 5.1.24 rc-community MySQL® Client Version 5.1.11
Programmier-/Entwicklungsumgebung	National Instruments® LabView™ 8.5

5 Zusammenfassende Qualitätsbewertung

Der vorliegende Bericht zur Qualitätsbewertung der Identifikation von aktiven Verkehrsteilnehmern im Projekt Do-iT behandelt die Identifikation zur Auswertung der Abis-Mobilfunkdaten des Demonstrators in Karlsruhe. Es wurden insgesamt vier Versuchstage mit 89 Referenzfahrten ausgewertet. Neben den Ortungsdaten aus den Abis-Mobilfunkdaten, vgl. dazu Bericht des AP 4-5, liegen in den Referenzfahrten aufgezeichnete GPS-Daten zur vergleichenden Auswertung zu Grunde. Die Verfahren zur Identifizierung Mobilfunkteilnehmer (MFT) bestehen im Einzelnen aus der „Bewegungsanalyse“ zur Erkennung des Bewegungszustandes, der „ÖV-Korrelation“ zur Erkennung von ÖV-Verkehren, der Verkehrsnetzzugehörigkeit zur Prüfung der infrage kommenden Verkehrsnetze auf Geometrieebene und einer Routensuche mit zeitlich-topologischen Restriktionen. Die Ergebnisse der Verfahren sowie deren Qualitätsbewertung werden im Folgenden zusammengefasst.

In der Bewegungsanalyse wird der Bewegungszustand der MFT anhand einer Signifikanzprüfung des zurückgelegten Weges mit einer Sicherheitswahrscheinlichkeit geprüft. Als Sicherheitswahrscheinlichkeit stellte sich ein Wert von 68% als bester Kompromiss heraus. Mit diesem Wert konnten im Falle eines unbewegten Soll-Zustands 100% aller mit GPS und 45% aller FPD-Daten korrekt erkannt werden. Befand sich der MFT tatsächlich in Bewegung wurden bei den GPS-Daten 96% und bei den FPD-Daten 88% korrekt erkannt. Maßgebliche Einflüsse bei den unkorrekt erkannten Daten sind auf die Fußgängerszenarien und bei GPS auf systembedingte Ortungsfehler in dichter Bebauung und Erfassung in Gebäuden zurückzuführen.

Mit dem Ziel die Effizienz der Identifikationsverfahren zu steigern, wurde zur Reduktion der FPD-Positionsmenge ein FCM-Clusterverfahren nach Gustafson-Kessel implementiert. Mit dem Clusterverfahren werden Schwerpunktkoordinaten einschließlich Fehlerellipsen berechnet, woraus Aufenthaltsregionen der MFT mit Wahrscheinlichkeitsmaßen ableitbar sind. Es wurde gezeigt, dass bei etwa gleichbleibender geometrischer Genauigkeit eine Reduktion der Positionsmenge auf 1,79% erreicht wird.

Zur Erkennung von aktiven MFT, im Rahmen von Do-iT ist das die Erkennung von MIV (Motorisierter Individual-Verkehr), sind die nicht eindeutig vom ÖV (Öffentlichen Verkehr) trennbaren MFT auszuschließen. Hierfür werden die ÖV-Verkehre: Linienbusse, Straßenbahnen sowie die Bahnen der DB AG im Identifikationsverfahren ÖV-Korrelation berücksichtigt. Fahrplanabweichungen werden durch eine Fuzzyfizierung der Fahrplandaten modelliert. Die Ergebnisdaten dieser Fuzzyfizierung wurden für den Online-Betrieb vorprozessiert und getrennt nach ÖV-Verkehr in Form effizienter Datenstrukturen bereitgestellt.

In den Analysen konnte für die GPS-Daten eine 73%-ige und für die FPD-Daten ein 61%-ige Erkennungsrate der ÖV-Szenarien nachgewiesen werden. Wurde eine ÖV-Korrelation nachgewiesen, dann handelte es sich bei GPS zu 91% und bei FPD zu 68% tatsächlich um Fahrten in einem ÖV-Fahrzeug oder um Fußgänger im Bahnhof. Das heißt eine zu Unrecht

vorgenommene ÖV-Einstufung, und der damit verbundene Ausschluss der FPD-Daten, geschah nur zu 9% bei GPS und zu 32% bei FPD. Maßgebliche Fehler bei der Erkennungsrate sowohl bei GPS als auch bei FPD konnten auf fehlerhafte Geometrien der Schienenwege zurückgeführt werden. Nicht näher untersucht wurden die Auswirkungen der modellierten Fahrplanabweichungen mittels Fuzzyfizierung. Die unkorrekte Erkennung von zwei Straßenbahnszenarien mit GPS könnten hiermit u. U. erklärt werden. Es ist zu erwähnen, dass der Stichprobenumfang für verallgemeinernde statistische Aussagen insgesamt als gering einzustufen ist.

Die eingeschränkte Korrektheit der Schienendaten sowie unvollständige Fußwegenetze sind ebenfalls maßgebliche Ursachen der unkorrekten Einstufungen bei der Verkehrsnetzzugehörigkeit. In diesem Identifikationsverfahren wurden 87% bei GPS und 94% der Versuchsszenarien mit den FPD-Daten korrekt identifiziert. Die im Do-iT-Lokalisierungsnetz integrierte Straßenkarte enthält kein vollständiges Fußwegenetz und bedingt dadurch die Fehlerrate von 13% bei GPS, die vollständig auf unkorrekt erkannte Fußgängerszenarien zurückzuführen ist. Das Identifikationsverfahren Verkehrsnetzzugehörigkeit dient vorrangig dazu dem nachfolgenden Verfahren der Routensuche mit zeitlich-topologischen Restriktionen die prinzipielle Möglichkeit eine Route zu bestätigen. Und zwar kann nur dann von einer möglichen Route ausgegangen werden, wenn zu jedem Schwerpunkt, einschließlich Fehlerfigur, mindestens ein Element des zu untersuchenden Verkehrsnetzes gefunden wird. Ist das nicht der Fall, kann bereits aus topologischen Gründen keine Route gefunden und die entsprechende Verkehrsklasse ausgeschlossen werden. Für die GPS-Daten, die als Referenz aktiver MFT aufgezeichnet wurden, konnten alle Routen des MIV gefunden werden. Bei den FPD-Daten gelang dies für 66% aller Fahrten. Die restlichen 34% Fahrten wurden fehlerhaft als passiv eingestuft, da für sie keine den zeitlichen oder topologischen Kriterien genügende Route gefunden werden konnte. D.h. 34% der aktiven MFT werden durch dieses Verfahren ggf. zu Unrecht ausgeschlossen.

Die aktiv-Einstufung eines MFT erfolgte bislang nur, wenn dies mindestens durch Vorhandensein einer Route auf dem Straßennetz belegt werden konnte. Eine Bewegung des MFT wird hiermit implizit vorausgesetzt. Besonders stehende MFT sind jedoch wesentlich zur Stauerkennung. Mit der Annahme, dass ein als aktiv identifizierter MFT keinen Verkehrsklassenwechsel vornahm, können nachfolgende ggf. als stehend erkannte Datensätze ebenfalls aktiv eingestuft werden. Das verkehrstechnische Verhalten anonymer MFT ist als komplex und ohne Zusatzinformation nicht modellierbar hinzunehmen. Es wurde daher mittels einer maximal zulässigen Zeitdifferenz $dt_{\max\,active}$ ein Zeitraum definiert, innerhalb dessen angenommen wird, dass der bislang aktive MFT keinen Verkehrsmittelwechsel vornahm. Anhand der Referenzfahrten konnte die Funktionsfähigkeit der Algorithmen gezeigt werden, jedoch fehlten für eine untermauernde Analyse Referenzfahrten mit erzwungenen Verkehrsklassenwechseln. Die Verkettung der FPD-Datensätze ermöglicht eine Verbesserung der Korrektheit für unbewegte aktive MFT. Zur Reduzierung von unentdeckten Verkehrsklassenwechseln kann der Parameter $dt_{\max\,active}$ optimiert werden. Eine empirisch begründete Wahl von $dt_{\max\,active}$ wurde anhand von Auswertungen eines 24-Stundendatensatzes gezeigt.

Dieser 24-Stundendatensatz diente ebenfalls für den Nachweis der Online-Fähigkeit der Identifikationsverfahren. Als Datenbasis fand eine Aufzeichnung des 19.12.2007 des Abis-Demonstrators des Do-iT Projektes in Karlsruhe Einsatz. Statistisch wurden die Auswirkungen verschiedener Werte der Zeitdifferenz $dt_{\max active}$ gezeigt, sowie die räumliche Abdeckung der möglichen Trajektorien und deren Häufigkeiten grafisch dargestellt. Für die vollständige Identifikation dieses Datensatzes wurde eine Berechnungszeit von ca. 15,5 h benötigt. Bei der Wahl von $dt_{\max active} = 30$ Minuten wurden für 208 266 MFT mit 172 061 Telefonaten, für die jeweils mindestens eine Position vorhanden waren, 19 854 aktive Trajektorien identifiziert. Anhand von Kartendarstellungen wurde die Generierung von Verkehrslageinformationen auch abseits der Hauptverkehrsachsen, in der Fläche gezeigt und die Tauglichkeit von FPD in diesen Regionen sowie mit dem Einsatz der hier untersuchten Verfahren zur Identifikation aktiver Verkehrsteilnehmer bestätigt. Bei weiterer programmtechnischer Optimierung kann mit einer deutlichen Verringerung der Berechnungsdauer gerechnet werden. Aufgrund der für das Projekt genügenden Berechnungszeiten, wurde die Skalierbarkeit der Algorithmen nicht näher untersucht. Angesichts der auf einzelne FPD-Datensätze orientierten Auswertestrategie, sind sowohl die Skalierbarkeit und als auch dezentrale Auswertungen durchaus möglich.

Literatur

[1] AP10-1: Laufer, R.: *Evaluierungskonzept, interner Do-iT Bericht AP10-1*, September 2008.

[2] AP3-4: Schollmeyer, R.: *AP3 Bericht zur Identifizierung aktiver Verkehrsteilnehmer*, interner Do-iT-Bericht AP3, 2008.

[3] AP4-5: Czommer, R.: *Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren*, interner Do-iT-Bericht AP4, 2008.

[4] F. Höppner, F. Klawonn, R. Kruse, and T. Runkler. *Fuzzy Cluster Analysis, Methods for classification, data analysis and image recognition*. Wiley, 2000.

[5] AP10-2: Laufer, R. et al.: Evaluierungsergebnisse, interner Do-iT Bericht AP 10-2, 2009.

6 Anlage

Das Softwareprogramm zur Durchführung der Identifikationsberechnung wurde in LabView™ (National Instruments®) programmiert. Zur Verwaltung, für Einstellungen und zur Überwachung der Berechnungsdurchführung im Betrieb besitzt das Softwareprogramm Bildschirmein- und ausgaben mit teilweise grafischer Darstellung. Die folgende Abbildung 6-1 zeigt die Startoberfläche zur Definition und Überwachung der wesentlichsten Dateneinund Datenausgaben. Die Identifikation wird mit dem kleinen weißen Pfeil, oben links im Bild, gestartet und durchläuft dann selbstständig aller verfügbaren MFT-Datensätze.

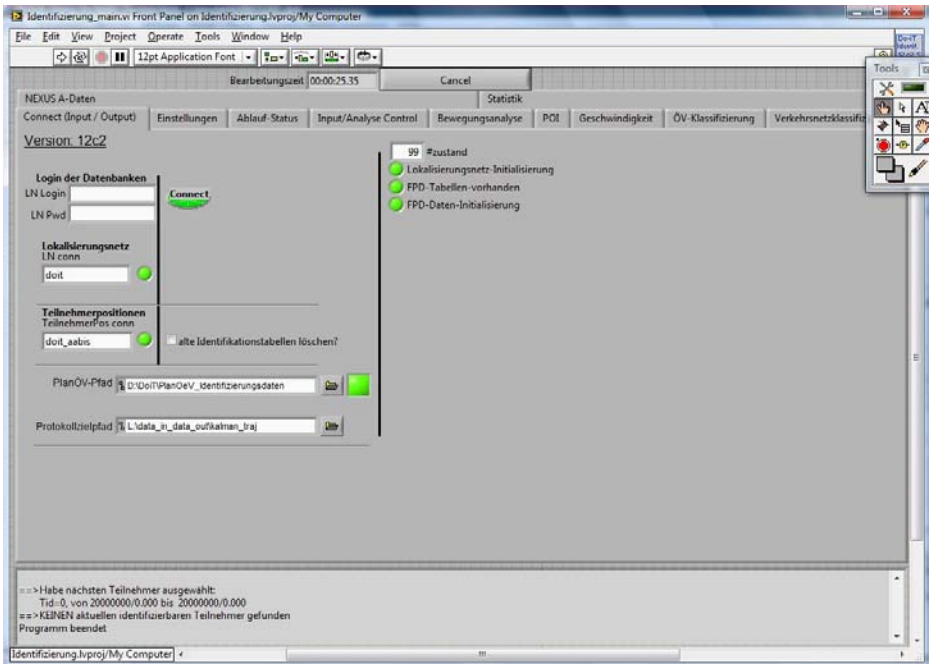


Abbildung 6-1: Bildschirmoberfläche zur Eingabe von Ein- und Ausgabepfaden

In der Menülasche „Einstellungen“, siehe Abbildung 6-2, werden die Datenbanktabellen konfiguriert und es können für Laboruntersuchungen Simulationseinstellungen vorgenommen werden.

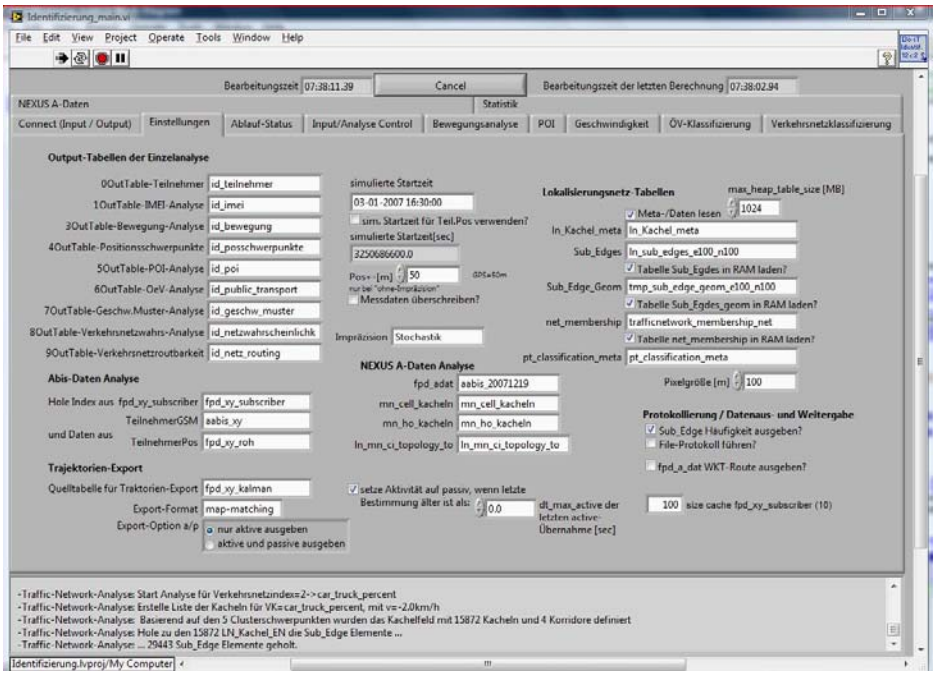


Abbildung 6-2: Bildschirmoberfläche zur Konfiguration der Datenbanktabellen und Einstellungen für Laboruntersuchungen

Die nachfolgende Abbildung 6-3 zeigt die Überwachungsoberfläche während der Identifikationsberechnungen. Es können bei Bedarf für Laboruntersuchungen einzelne Identifikationsverfahren ein- oder ausgeschaltet werden (Bildmitte), sowie der Staus zum Berechnungsfortschritt beobachtet werden (im Bild links, oben und unten rechts).

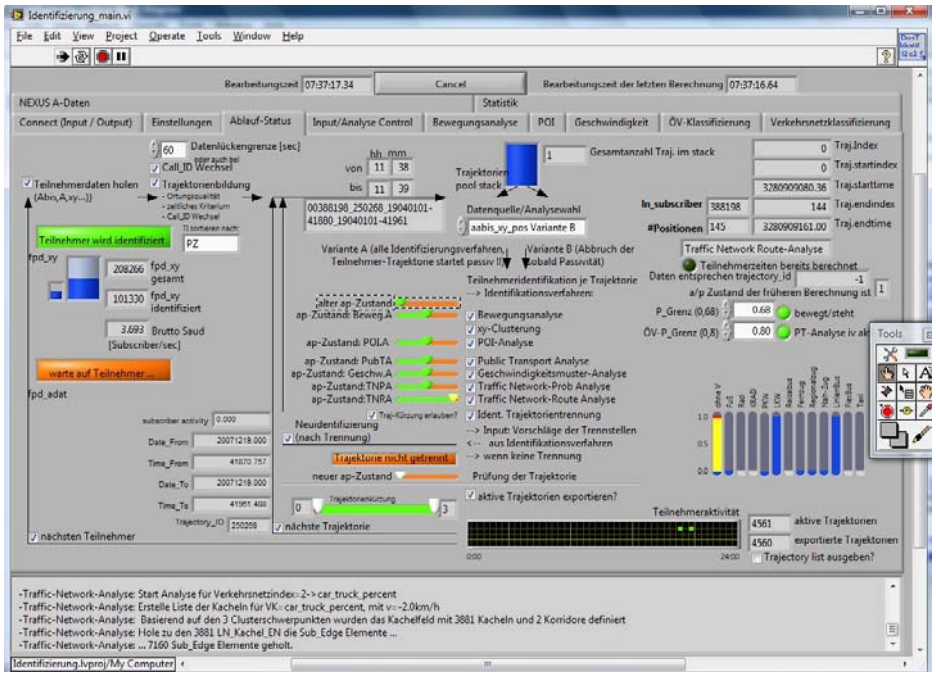


Abbildung 6-3: Bildschirmoberfläche mit Statusinformationen der Identifikation im Betrieb

Die folgenden vier Abbildungen zeigen die Statusoberflächen der Einzelverfahren „Bewegungsanalyse“ und „ÖV-Korrelation“ in Abbildung 6-4 sowie der „Verkehrsnetzzugehörigkeit“ und des „Routings mit zeitlichen und topologischen Restriktionen“ in Abbildung 6-5.

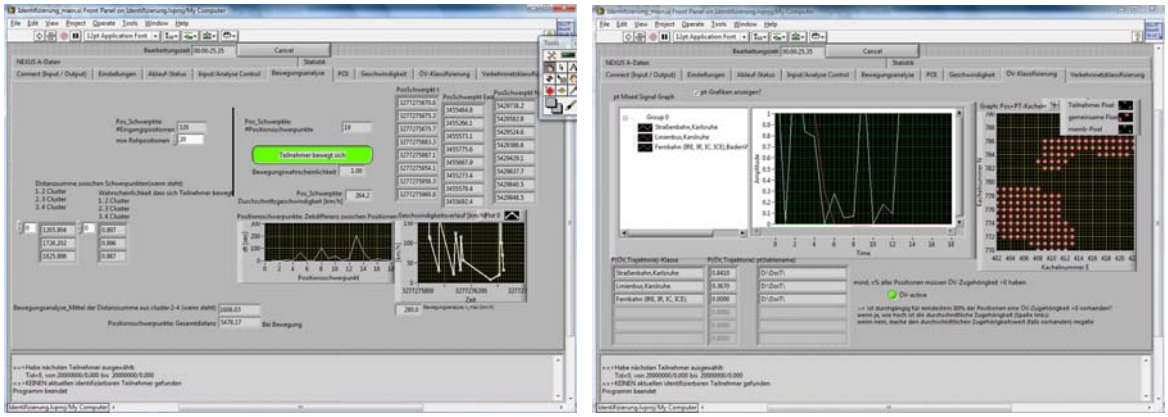


Abbildung 6-4 : Überwachungsbildschirmoberflächen mit links der „Bewegungsanalyse" und rechts der „ÖV-Korrelation"

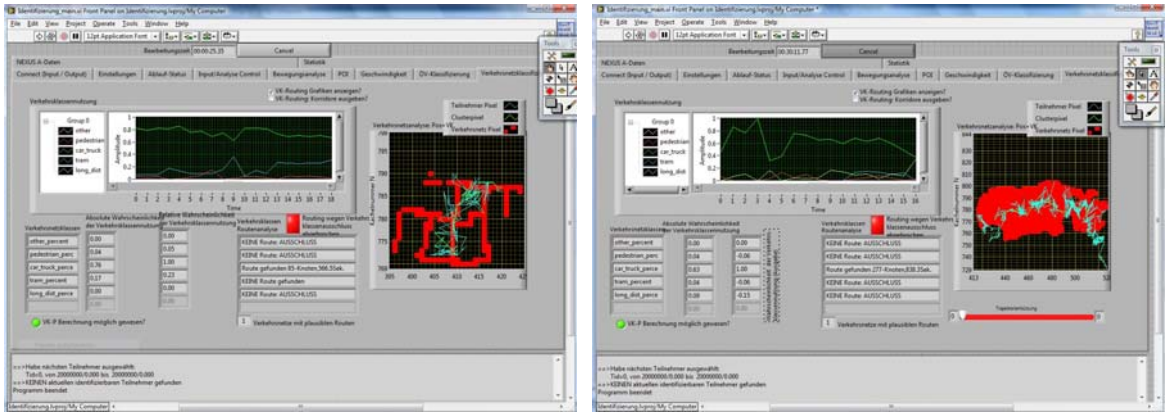
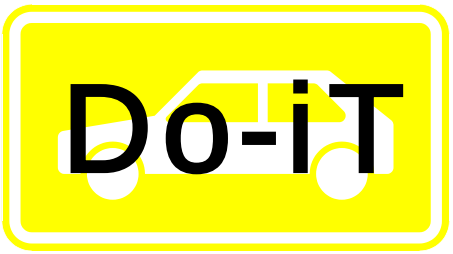


Abbildung 6-5: Zwei Beispiele für die Überwachungsbildschirmoberflächen der „Verkehrsnetzzugehörigkeit" jeweils linke Bildhälfte und des „Routings mit zeitlichen und topologischen Restriktionen“ jeweils auf der rechten Bildhälfte

Bericht zum AP 8-7

Datenoptimierung für integrierte Telematik

Do-iT



Förderkennzeichen: 19 B 5007

Nachweis der Wirksamkeit der

Qualitätssicherungsmaßnahmen (V1.0)

Mai 2009

Ralf Laufer, Stefan Schwartz

Stadt Karlsruhe

Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen

Dokument-Informationen

Version	Datum	Bearbeiter	Bemerkung
0.1	**11.02.09**	**Laufer (IAGB)**	• Erstellung der Berichtvorlage
0.2	**06.03.09**	**Laufer (IAGB)**	• Einpflegen des Teilberichts von Herrn Schwartz aus Karlsruhe
1.0	**30.04.09**	**Laufer (IAGB)**	• Einpflegen des Teilberichts des IAGB • Fertigstellung des Berichts
1.1	**29.05.09**	**Laufer (IAGB)**	• Review: Einarbeitung der Änderungen

Inhalt

1 Einleitung 4

2 Nachweis der Wirksamkeit der QS-Maßnahmen 6

2.1 Nachweis der Wirksamkeit der QS-Maßnahmen in Karlsruhe 6

2.1.1 FMEA-Analyse 7

2.1.2 Qualitätsparameter 12

2.2 Nachweis der Wirksamkeit der QS-Maßnahmen bei der Generierung von FPD-Trajektorien 16

2.2.1 FPD-Server 16

2.2.2 Berechnung von Positionsfolgen aus Mobilfunkdaten 18

2.2.3 Berechnung von FPD-Trajektorien 32

2.2.4 QS durch Protokollierung der Tagesergebnisse 38

Literatur 39

1 Einleitung

Dieser Bericht dient dem Nachweis der Wirksamkeit der im Rahmen des Projektes Do-iT umgesetzten Qualitätssicherungsmaßnahmen (QS-Maßnahmen). Da der Hauptteil des Projekts aus der Entwicklung der Algorithmen zur Generierung von Trajektorien aus Mobilfunkdaten (MFD) bestand, dienten die Maßnahmen zur Qualitätssicherung nicht nur der Sicherung eines gewissen vorhandenen Qualitätsniveaus sondern zunächst erst einmal dem Erreichen dieses Niveaus. Als Grundlage wurde dazu der Bericht *Qualitätssicherungsmaßnahmen (V1.0)* zum AP8-5, der als Bestandteil des ZMS3 erstellt wurde herangezogen (vgl. Laufer, Schwartz 2008). Die darin vorgestellten Maßnahmen beziehen sich zum einen auf die Qualitätsverbesserung und -sicherung der FPD-Trajektorien, die mit den im Projekt entwickelten Algorithmen des IAGB generiert werden. Zum anderen wurden auch Maßnahmen zur Sicherung der Datenqualität von Stationären Verkehrsdatenerfassungssystemen (SES) vorgestellt, die bei dem Projektpartner Stadt Karlsruhe im Einsatz sind. Diese SES dienen in Do-iT als Referenzsysteme zur Evaluierung der FPD-Trajektorien, daher ist eine geeignete Datenqualität auch für das Projekt Do-iT von großer Bedeutung. Da die SES, die dem Projekt zur Verfügung standen bereits seit längerem im online-Betrieb sind, beschränkten sich die Maßnahmen in diesem Fall tatsächlich auf die reine Qualitätssicherung sowie die weitere Verbesserung einzelner Parameter.

Die aus dem Katalog gewählten Maßnahmen zur Sicherung der Qualität der generierten FPD-Trajektorien haben sich an der restlichen Projektlaufzeit orientiert. Daher wurden nur Maßnahmen ergriffen, die noch innerhalb der Projektlaufzeit umgesetzt und deren erste Erfolge bis zum Projektende nachweisbar waren. Bei den stationären Verkehrsdatenerfassungssystemen der Stadt Karlsruhe hingegen wurden die Maßnahmen in kurz-, mittel- und langfristig wirksame Maßnahmen unterteilt. Diese Systeme waren und sind ein fester Bestandteil der Verkehrssteuerung in Karlsruhe und werden auch in naher Zukunft einen wesentlichen Teil zur Informationsgewinnung beitragen. Daher ist hier eine Abschätzung und Umsetzung von längerfristig wirksamen Maßnahmen zur Sicherung der Datenqualität unabhängig von einer möglichen Integration von Mobilfunkdaten im Verkehrsrechner sinnvoll.

Viele der beschriebenen Maßnahmen zum Erreichen eines notwendigen Qualitätsniveaus der berechneten Trajektorien aus Mobilfunkdaten wurden bereits vor Fertigstellung des Berichts zu den Qualitätssicherungsmaßnahmen während der Verfahrensentwicklung umgesetzt. Da sie einen wesentlichen Teil der durchgeführten Qualitätssicherungsmaßnahmen darstellen, werden Sie in diesem Bericht ebenfalls aufgenommen. Soweit es im Nachhinein noch möglich war, wurde der Erfolg der während der Softwareentwicklung durchgeführten Maßnahmen durch Vergleiche geeigneter Kennzahlen nachgewiesen. Aufgrund des teilweise enormen Zeitaufwands zur erneuten Prozessierung von Rohdaten mit älteren Softwareversionen war dies jedoch nicht in allen Fällen möglich.

Die im Bericht dargestellten Qualitätsmodelle für Verkehrsinformationen, Positionsfolgen und FPD-Trajektorien basieren auf dem für Do-iT definierten Qualitätsmodell welches in (Laufer u. Wiltschko 2006) dargestellt ist.

2 Nachweis der Wirksamkeit der QS-Maßnahmen

Im Folgenden werden zunächst die durchgeführten Maßnahmen zur Sicherung der Datenqualität der SES dargestellt. Der Nachweis deren Wirksamkeit erfolgt hierbei auf zwei Arten. Zunächst werden die im Rahmen der durchgeführten FMEA[1] ermittelten Risikoprioritätszahlen vor und nach der Einführung der Maßnahmen gegenübergestellt. Anschließend wird der Nachweis auf Basis des gegebenen Qualitätsmodells für Verkehrskenngrößen durch Auswahl geeigneter Qualitätsparameter geführt (vgl. Laufer 2008).

Im zweiten Abschnitt wird die Qualitätssicherung bei der Generierung der FPD-Trajektorien beschrieben. Der Nachweis erfolgt hier jedoch nur auf Basis ausgewählter Qualitätsparameter, da während der Softwareentwicklung keine Möglichkeit bestand, weitere Analysen wie z. B. eine FMEA durchzuführen, die nun als Referenz herangezogen werden könnten.

2.1 Nachweis der Wirksamkeit der QS-Maßnahmen in Karlsruhe

Im Rahmen des Arbeitspaketes wurden im Spätjahr 2008 diverse Qualitätssicherungsmaßnahmen für die stationären Erfassungssysteme (Induktionsschleifen und Traffic Eyes[2]) in Karlsruhe entwickelt und umgesetzt.

Der Katalog umfasst sowohl kurzfristige Maßnahmen, die noch während der Projektlaufzeit umgesetzt wurden, und deren Erfolg im Rahmen dieses Projektes nachweisbar ist, als auch mittel- und langfristige Maßnahmen, deren Wirkung sich erst im Laufe der Zeit entfalten werden (vgl. Laufer, Schwartz, 2008).

Bei der Entwicklung des Maßnahmenkatalogs hat sich gezeigt, dass kurzfristig im Wesentlichen die Erkennungs- und Behebungszeiten von Fehlern oder Ausfällen optimiert werden können und mittel- bis langfristig auch eine Reduzierung der Fehler- oder Ausfallhäufigkeit zu erwarten ist. Der Nachweis der Wirksamkeit der entwickelten QS-Maßnahmen und eine Einschätzung deren Wirkungsumfänge wird im Rahmen dieses Berichtes in zwei voneinander unabhängigen Verfahren nachgewiesen. Die Grundlage für die Bewertung bildet einerseits das in den Arbeitspaketen 8-3 und 8-4 angepasste Verfahren FMEA (Fehlermöglichkeits- und Einflussanalyse), welches im Kern eine Fehler- und Ausfallwahrscheinlichkeit aller Systemkomponenten analysiert und im Ergebnis eine Quantifizierung des Risikos in Form von Risikoprioritätszahlen ermittelt. Der Wirksamkeitsnachweis der QS-Maßnahmen kann daher in Form einer kurz-, mittel-

[1] FMEA: Fehler-Möglichkeits- und Einflussanalyse; es handelt sich dabei um ein häufig eingesetztes Werkzeug der Qualitätsplanung zur Analyse möglicher Fehlerquellen

[2] Traffic Eye Universal (TEU): Dabei handelt es sich um einen passiven Infrarotsensor der Firma Siemens zur Erfassung von Verkehrskenngrößen

und langfristigen prozentualen Risikoreduzierung erfolgen. Andererseits erlaubt das in AP 8-4 definierte Qualitätsmodell für Verkehrskenngrößen eine quantitative Bewertung der Maßnahmenwirksamkeit auf die konkreten Qualitätsparameter Verfügbarkeit, Aktualität, Vollständigkeit, Konsistenz, Korrektheit und Genauigkeit, indem alle Parameter sowohl vor, als auch nach Einführung der QS-Maßnahmen für einen längeren Zeitraum betrachtet, ausgewertet und vergleichend gegenübergestellt werden.

2.1.1 FMEA-Analyse

Die Fehlermöglichkeits- und Einflussanalyse (FMEA) wurde in AP 8-3 als geeignetes Analyseverfahren zur Bewertung der Datenqualität stationärer Erfassungssysteme eingeführt (vgl. Laufer 2007). Es wurde ein Bewertungsschema definiert, welches eine problemangepasste Bewertung der Auftretenswahrscheinlichkeit (A), der Bedeutung (B) sowie der Entdeckungs- und Behebungsmöglichkeit (E) eines Ereignisses ermöglicht und als Ergebnis das Risiko als Produkt der drei Bewertungsparameter (sog. Risikoprioritätszahl) beinhaltet. Für den Zeitraum vor Einführung der Qualitätssicherungsmaßnahmen wurde die FMEA für alle Elemente des Datenerhebungsprozesses in AP 8-4 durchgeführt und dokumentiert (vgl. Laufer 2008).

In der folgenden Tabelle ist für jede quantifizierbare QS-Maßnahme die zu erwartende Wirkung auf die einzelnen Bewertungsparameter sowie der Zeithorizont angegeben, zu dem sich die Maßnahme auswirkt. Kurzfristige Maßnahmen sind definitionsgemäß noch innerhalb der Projektlaufzeit umsetzbar und als solche in der Spalte *Nr.* grau hinterlegt.

Tabelle 1: QS-Maßnahmen mit zu erwartender Wirkung

Nr.	System	Bauteil	QS-Maßnahme	Ziel	Wirkung (A/B/E)*	Zeithorizont
1.1	Schleife (RMQ)	Induktionsschleife	Gute Vorplanung von Baumaßnahmen, Baustellen- und Ereignismanagement "BEM" ist derzeit in der ersten Betriebsphase	Reduzierung der Auftretenshäufigkeit durch rechtzeitiges Hinweisgeben auf Schleifenkabel bzw. Reduzierung der Entdeckungs- und Behebungszeit	Schleife: A 6 → 5 E 6→ 5 Zuleitung: A 4 → 3 E 6→ 5	kurzfristig
1.2	Schleife (RMQ)	Muffe	Verwendung von Muffen hoher Qualität; wird seit ca. drei Jahren in Karlsruhe praktiziert in Form von gelverfüllten witterungsbeständigen Muffen ("Gelsnap").	Mittel- bis langfristige Reduzierung der Auftretenshäufigkeit zu erwarten	A 5 → 4,5 A 5 → 4	mittel-/langfristig
1.3	Schleife (RMQ)	Zuleitung	Verwendung moderner (fettverpresster) Kabel; wird seit ca. zehn Jahren in Karlsruhe praktiziert.	Mittelfristige Reduzierung der Auftretenshäufigkeit zu erwarten	A 3→ 2,5	mittelfristig
1.4	Schleife (RMQ)	Steuergerät	Kontinuierliche Modernisierung des Gerätebestandes, insbesondere der "abgekündigten" Gerätetypen, bei denen eine Ersatzteilvorhaltung nicht mehr gewährleistet ist.	Mittel- und langfristige Reduzierung der Auftretenshäufigkeit zu erwarten	A 4→ 3 A 3→ 2	mittel-/langfristig

Nr.	System	Bauteil	QS-Maßnahme	Ziel	Wirkung (A/B/E)*	Zeit-horizont
1.5	Schleife (RMQ)	Spannungs-versorgung	Standardmäßiger Einbau einer USV zur unabhängigen Spannungsversorgung bei Netzausfall	Mittel- und langfristige Reduzierung der Auftretenshäufigkeit zu erwarten	A 4→ 3,5 A 4→ 3	mittel-/langfristig
1.6	Schleife (RMQ)	Steuerkabel	Kabelkataster	Deutliche Reduzierung der Behebungszeit bei Defekten einzelner Adern durch Umrangieren auf Ersatzaderpaar	Bruch, Kurzschluss und sporadischer Kurzschluss: E -1 E -2*	kurz-/mittelfristig
2.3	TEU	Infrarotdetek-tor	Stringente Vorgaben zur Plausibilisierungskontrolle in Concert-Software	Unplausible Daten werden vom System verworfen (Concert: Status grau), so dass eine deutliche Reduzierung der Erkennungszeit zu erwarten ist	E 7→ 6	kurzfristig
2.4	TEU	Infrarotdetek-tor	Tournusmäßige Plausibilitäts-kontrolle der Messwerte, insbesondere der Geschwindigkeit, in Bezug zu historischen Daten, 1x wöchentlich. Einrichten einer Messwert-kontrolldatei	Deutliche Reduzierung der Erkennungszeit	E 9→ 7	kurzfristig
2.5	TEU	GPRS-Server	Automatischer Restart des Softwareprozesses sonntags 12:00 Uhr	Reduzierung der Behebungs-zeit bei Ausfall am Wochenende	E 4→ 3	kurzfristig
2.6	TEU	TEU-Gesamtsys-tem	Ausarbeitung, Abstimmung, Bekanntmachung und Vorhalten von Aktionsplänen	Reduzierung der Behebungs-zeit insbesondere wenn Absperrmaßnahmen zur Fehlerbeseitigung zu organisieren sind	E -1 für alle Teile der TEU-Außenanlage	kurzfristig
3.1	Schleife / TEU	Concert-Rechner	Automatischer Restart der Verarbeitungsprozesse sonntags 12:00 Uhr	Reduzierung der Behebungs-zeit bei Ausfall am Wochenende	E 4→ 3	kurzfristig
3.2	Schleife / TEU	Gesamtsys-tem	Tournusmäßige visuelle Kontrolle des Gesamtsystems auf Funktionstüchtigkeit aller Messwerte 2x täglich in Abhängigkeit von den vereinbarten Reaktionszeiten der Servicekräfte	Reduzierung der Erkennungs- und Behebungszeiten	E -0,5 für alle ausfallrelevanten Störungen, die nicht anderweitig schneller erkannt werden (z.B. Ausfall LSA-Steuergerät) **	kurzfristig

* Bezeichnung entsprechend der durchgeführten System-FMEA: A=Auftrittswahrscheinlichkeit, B=Bedeutung E=Entdeckung und Behebung eines Fehlers

** Eine Entdeckungs- und Behebungszeit der Stufe 3 (entspr. ½ Tag) oder geringer wird bei Schleifen und Concert- bzw. Verkehrsrechner nicht weiter reduziert, bei TEU-Außenanlage eine Entdeckungs- und Behebungszeit der Stufe 4 (entspr. 1 Tag) oder geringer, da in diesen Fällen bereits die minimal erreichbaren Entdeckungs- und Behebungszeiten erreicht wurden.

Für die Bewertung der einzelnen QS-Maßnahmen wurde das früher entwickelte Bewertungsschema zur Schwere und zur Entdeckungs- und Behebungszeit nochmals geringfügig modifiziert und die ursprünglich verbale Stufendefinition mit konkreten und realistischen Zahlenwerten untermauert. Auf diese Weise ließ sich die Wirkungsquantifizierung unabhängig von subjektiven Betrachtungsweisen durchführen. Insofern ergeben sich auch innerhalb der Bestandsbewertung nach FMEA Abweichungen gegenüber der in AP 8-4 (vgl. Laufer 2008) durchgeführten Analyse, in Summe ist aber eine wesent-

lich genauere Einschätzung des Risikos bzw. des Wirksamkeitsnachweises der QS-Maßnahmen möglich. Bezogen auf die einzelnen Systembauteile ergeben sich die nachfolgend aufgeführten Bestandsrisiken sowie die kurz-, mittel- und langfristig erwartbaren Risiken nach Einführung der QS-Maßnahmen, wobei erzielbare Qualitätsverbesserungen grau hinterlegt sind:

Tabelle 2: FMEA mit und ohne QS-Maßnahmen

					QS-Maßnahmen					
			Bestand*		kurzfrist		mittelfrist		langfrist	
Bauteil	**potentielle Ausfallart**	**mögliche Ursache**	**A B E**	**RPZ**	**A B E**	**RPZ**	**A B E**	**RPZ**	**A B E**	**RPZ**
Induktionsschleife	Bruch / Abriß	mechanische Belastung; Baumaßnahmen, Alterung, Witterung	6 2 6	72	5 2 4,5	45	5 2 4,5	45	5 2 4,5	45
	Kurzschluss		4 2 7	56	4 2 6,5	52	4 2 6,5	52	4 2 6,5	52
	sporadischer Kurzschluss		4 1 8	32	4 1 7,5	30	4 1 7,5	30	4 1 7,5	30
Muffe	Bruch / Abriß	mechanische Belastung; Baumaßnahmen, Alterung, Witterung	5 2 6	60	5 2 5,5	55	4,5 2 5,5	49	4 2 5,5	44
	Kurzschluss		3 2 7	42	3 2 6,5	39	3 2 6,5	39	3 2 6,5	39
	sporadischer Kurzschluss		3 1 8	24	3 1 7,5	22	3 1 7,5	22	3 1 7,5	22
Zuleitung vom Kleinschacht zum Schaltkasten	Bruch / Abriß	mechanische Belastung; Baumaßnahmen, Alterung, Witterung	4 2 6	48	3 2 4,5	27	3 2 4,5	27	3 2 4,5	27
	Kurzschluss		3 2 7	42	3 2 6,5	39	2,5 2 6,5	32	2,5 2 6,5	32
	sporadischer Kurzschluss		3 1 8	24	3 1 7,5	22	2,5 1 7,5	19	2,5 1 7,5	19
Auswertebaugruppe	defekt / sporadisch defekt	Blitzschlag / Kurzschluss	2 3 4	24	2 3 3,5	21	2 3 3,5	21	2 3 3,5	21
		Alterung/ Qualitätsmangel	3 3 4	36	3 3 3,5	31	3 3 3,5	31	3 3 3,5	31
Steuergerät	defekt	Alterung, Witterung, mech. Belastung, Vandalismus, Unfall, Qualitätsmangel	4 4 6	96	4 4 6	96	3 4 6	72	2 4 6	48
		Blitzschlag	2 4 5	40	2 4 5	40	2 4 5	40	2 4 5	40
	Abschaltung der Signalgeber, jedoch nicht systemrelevant	Abschaltung der Signalgeber in der Nebenzeit	8 0 1	0	8 0 1	0	8 0 1	0	8 0 1	0
		Abschaltung der Signalgeber wegen Rotlampendefekt	8 0 1	0	8 0 1	0	8 0 1	0	8 0 1	0
Spannungsversorgung	defekt	Alterung, Witterung, mech. Belastung, Qualitätsmangel	4 4 2	32	4 4 2	32	3,5 4 2	28	3 4 2	24
		Blitzschlag	2 4 3	24	2 4 3	24	2 4 3	24	2 4 3	24
Steuerkabel	Bruch	Alterung, Witterung, mech. Belastung, Vandalismus, Qualitätsmangel	4 5 6	120	4 5 4,5	90	4 5 3,5	70	4 5 3,5	70
	Kurzschluss		5 5 7	175	5 5 5,5	138	5 5 4,5	113	5 5 4,5	113
	sporadischer Kurzschluss		6 4 8	192	6 4 6,5	156	6 4 5,5	132	6 4 5,5	132

			QS-Maßnahmen							
			Bestand*		kurzfrist		mittelfrist		langfrist	
Bauteil	**potentielle Ausfallart**	**mögliche Ursache**	**ABE**	**RPZ**	**ABE**	**RPZ**	**ABE**	**RPZ**	**ABE**	**RPZ**
ZMO-Schnittstelle zum VR-Steuergerät	defekt	Alterung, mangelnde Qualität	343	36	343	36	343	36	343	36
	sporadisch defekt		334	36	333,5	31	333,5	31	333,5	31
Rahmen	Netzteil defekt	Alterung, mangelnde Qualität, Blitzschlag	365	90	365	90	365	90	365	90
	Elektronik defekt		265	60	265	60	265	60	265	60
Verkehrsrechner	defekt	Ausfall einer Baugruppe / Systemkomponente, Kurzschluss	374	84	374	84	374	84	374	84
	Absturz	Überlastung, Softwarefehler	673	126	673	126	673	126	673	126
	Fehlfunktion	Virenbefall, Softwarefehler	275	70	275	70	275	70	275	70
Netzwerkkabel zum Concert-Rechner	Unterbrochen	mech. Belastung, Stecker gezogen	173	21	173	21	173	21	173	21
TEU-Solarpanel	keine Stromerzeugung	Solarpanel defekt	246	48	244,5	36	244,5	36	244,5	36
		durch Blätter, Schnee oder sonstigen Schmutz verdeckt	235	30	234	24	234	24	234	24
TEU-Verbindungskabel Solarpanel-Auswerteeinheit	Kurzschluss	Alterung, Witterung, Vandalismus, Qualitätsmangel	145	20	144	16	144	16	144	16
	Abriß		145	20	144	16	144	16	144	16
	Wackelkontakt		136	18	134,5	13	134,5	13	134,5	13
TEU-GPRS-Modul	GPRS-Modul defekt	Qualitätsmangel, mechanische Belastung	346	72	344,5	54	344,5	54	344,5	54
	Stecker lose	mangelhafte Montage	145	20	144	16	144	16	144	16
	SIM-Karte defekt oder gesperrt	Kontakte oxidiert bzw. nicht bezahlte Rechnung	245	40	244	32	244	32	244	32
TEU-Infrarotdetektor	Infrarotdetektor defekt	Qualitätsmangel, mechanische Belastung	225	20	224	16	224	16	224	16
	Infrarotdetektor verdreht, verkippt oder verdreht und verkippt	keine Messdaten	225	20	224	16	224	16	224	16
		sehr falsche Messdaten	327	42	324	24	324	24	324	24
		falsche Messdaten	329	54	326	36	326	36	326	36
TEU-Verbindungskabel Infrarotdetektor-Auswerteeinheit	Kurzschluss	Alterung, Witterung, Vandalismus, Qualitätsmangel	125	10	124	8	124	8	124	8
	Abriß		225	20	224	16	224	16	224	16
	Wackelkontakt		116	6	114	4	114	4	114	4
TEU-Auswerteeinheit	Auswerteeinheit defekt	Alterung, Witterung, Vandalismus, Qualitätsmangel	345	60	344	48	344	48	344	48
GPRS-Netz	keine Internetverbindung möglich	Überlastung, sonstige Probleme des Netzbetreibers	384	96	384	96	384	96	384	96

					QS-Maßnahmen					
			Bestand*		kurzfrist		mittelfrist		langfrist	
Bauteil	**potentielle Ausfallart**	**mögliche Ursache**	**A B E**	**RPZ**	**A B E**	**RPZ**	**A B E**	**RPZ**	**A B E**	**RPZ**
	lokal keine Internetverbindung möglich	Überlastung, Ausfall einer Zelle, sonstige Probleme des Netzbetreibers	4 5 4	80	4 5 4	80	4 5 4	80	4 5 4	80
Zuleitung zum DSL-Modem (bis Telefondose)	Unterbrechung	Stecker gezogen	1 8 3	24	1 8 3	24	1 8 3	24	1 8 3	24
		Alterung	1 8 4	32	1 8 3,5	28	1 8 3,5	28	1 8 3,5	28
		mechanische Belastung (z. B. Baumaßnahme)	1 8 4	32	1 8 3,5	28	1 8 3,5	28	1 8 3,5	28
DSL-Modem	defekt	Alterung, mangelnde Qualität	2 8 4	64	2 8 3,5	56	2 8 3,5	56	2 8 3,5	56
GPRS-Server	defekt / Absturz	Alterung, Softwarefehler	4 8 4	128	4 8 3	96	4 8 3	96	4 8 3	96
Concert-Rechner	defekt	Ausfall einer Baugruppe / Systemkomponente, Kurzschluss	2 10 4	80	2 10 4	80	2 10 4	80	2 10 4	80
	Absturz	Überlastung, Softwarefehler	6 10 4	240	5 10 3	150	4 10 3	120	3 10 3	90
	Fehlfunktion	Virenbefall, Softwarefehler	2 10 5	100	2 10 5	100	2 10 5	100	2 10 5	100
Intranetverbindung zum Medienbüro (über Administration)	Unterbrechung	Stecker gezogen	1 10 4	40	1 10 4	40	1 10 4	40	1 10 4	40
		mech. Belastung (z. B. Baumaßnahme)	1 10 5	50	1 10 5	50	1 10 5	50	1 10 5	50
Leitungsverbindung zu Client-	Unterbrechung	Kabel defekt, Stecker gezogen; Baumaßnahmen	1 3 3	9	1 3 3	9	1 3 3	9	1 3 3	9
Modem (am VR und am Client-	defekt	Gerät / Baugruppe defekt	6 3 3	54	6 3 3	54	6 3 3	54	6 3 3	54
Leitungsverbindung VR-Modem und	Unterbrechung	Kabel defekt, Stecker gezogen; Baumaßnahmen	1 3 3	9	1 3 3	9	1 3 3	9	1 3 3	9
Client-Rechner	defekt	Ausfall einer Baugruppe / Systemkomponente, Kurzschluss	2 3 4	24	2 3 3,5	21	2 3 3,5	21	2 3 3,5	21
	Absturz	Überlastung, Softwarefehler	3 3 1	9	3 3 1	9	3 3 1	9	3 3 1	9
	Fehlfunktion	Virenbefall, Softwarefehler	2 3 2	12	2 3 2	12	2 3 2	12	2 3 2	12
		Summe RPZ		**3.245**		**2.794**		**2.651**		**2.588**
						86%		**82%**		**80%**

*: vor Einführung von QS-Maßnahmen

In Summe über das Gesamtsystem ergeben sich Risikoprioritätszahlen in Höhe von 3.245 im Bestand (ohne Qualitätssicherungsmaßnahmen) sowie 2.794, 2.651 und 2.588 mit kurz-, mittel- und langfristig wirksamen Qualitätssicherungsmaßnahmen. Die erneute FMEA-Bewertung zeigt daher, dass im Vergleich zu der Bestands-FMEA das Risiko von Ausfällen und deren Auswirkungen infolge der eingeführten QS-Maßnahmen kurzfristig um ca. 14 %, mittelfristig um ca. 18 % und langfristig um ca. 20 % reduziert werden kann.

2.1.2 Qualitätsparameter

Im Unterschied zum Nachweis der Wirksamkeit der QS-Maßnahmen auf der Basis von FMEA-Analysen beruht der Nachweis über Qualitätsparameter auf empirischen Daten. Zu diesem Zweck werden die in AP 8-4 definierten Globalparameter zur Beschreibung der Verkehrskenngrößen (vgl. Laufer 2008) einmal für den Bestand (vor Einführung der QS-Maßnahmen) und einmal für einen Zeitraum nach Einführung und Wirksamwerden der kurzfristigen QS-Maßnahmen über alle Messwerte von stationären Erfassungssystemen hinweg ermittelt und beide Messreihen miteinander verglichen. Der Nachweis der Wirksamkeit der QS-Maßnahmen ist erbracht, wenn sich die einzelnen beeinflussbaren Qualitätsparameter gegenüber der Vorher-Untersuchung signifikant verbessern. Es ist jedoch darauf hinzuweisen, dass die Parameter stark von Einzelereignissen abhängen und daher möglicherweise die positiven Effekte des Qualitätsmanagements überlagert werden. Insofern erscheint zunächst die Untersuchung einer möglichst langen Zeitreihe sinnvoll, andererseits muss als limitierender Faktor auch die äußerst umfangreiche zu untersuchende Datenmenge berücksichtigt werden. Als Kompromiss werden die gesamten Messdatensätze jeweils einer Woche ohne besondere Vorkommnisse (wie z.B. langwieriger Ausfall des Verkehrsrechners, Umstellung des Netzproviders, Update des Concert-Rechners etc.) ausgewertet und anschließend alle dokumentierten Ausfallereignisse eingerechnet. Die Vorher-Untersuchung wurde für die KW 40/2008 (29.09.-05.10.2008) mit Einrechnung aller Ausfallereignisse von Juli bis November 2008 und die Nachher-Untersuchung für die KW 6/2009 (02.02.-08.02.2009) unter Ausschluss aller besonderen Vorkommnisse der Monate Januar und Februar 2009 durchgeführt. Im Einzelnen können die in der folgenden Tabelle grau hinterlegten Qualitätsmerkmale bzw. –parameter aus den Messdatenreihen empirisch ermittelt werden:

Tabelle 3: Qualitätsparameter der SES (Laufer 2008)

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
P1.1	Verfügbarkeit VE	Ausfallrate	AUR [%]	Beschreibt die Verfügbarkeit eines einzelnen Sensors in einem Zeitraum
P1.2		Globale Ausfallrate	AUR_{global} [%]	Beschreibt die mittlere Verfügbarkeit der Sensoren im betrachteten Gebiet in einem best. Zeitraum (aus P1.1)
P2.1	Aktualität AK	Meldeverzögerung	Δt_M [min]	Alter der Daten eines Sensors; Differenz aus aktueller Zeit – (letztem) Zeitstempel, gemittelt über Zeitraum
P2.2		Mittlere Meldeverzögerung	$\overline{\Delta t}_M$ [min]	Mittleres Alter der Daten für ein bestimmtes Gebiet (aus P2.1)
P3.1	Vollständigkeit VO	Abdeckungsrate	ABR [%]	Grad der Abdeckung des betrachteten Straßennetzes im Gebiet mit Sensoren (Begriff: „vollständige Abdeckung“ ist zu definieren)
P4.1	Konsistenz KO	Konsistenzrate	KON [%]	Übereinstimmung der Daten mit dem Datenmodell (Anzahl Werte; Einheiten, Wertebereiche,...) je Sensor in einem bestimmten Zeitraum
P4.2		Globale Konsistenzrate	KON_{global} [%]	Wie oben, jedoch im gesamten betrachteten Gebiet in einem Zeitraum (aus P4.1)
P5.1	Korrektheit KR	Korrektheitsrate der Messwerte V, Q, BG, LOS	KOR_{MW} [%]	Korrekt bestimmte Messwerte eines Sensors in einem Zeitraum (Entscheidung anhand von Schwellwerten für die Genauigkeit)
P6.1	Genauigkeit GE	Lagegenauigkeit	G_L [± m]	Geometrische Genauigkeit der Einmessung der Schleifen/Sensorfeld (Stdw. 2D)
P6.2		Genauigkeit der Geschwindigkeit	G_V [± km/h]	Standardabweichung der Geschwindigkeitsbestimmung eines Sensors
P6.3		Genauigkeit der Verkehrsstärke	G_q [± KFZ/Δt]	Standardabweichung der angegebenen Verkehrsstärken eines Sensors
P6.4		Genauigkeit der Belegung	G_B [± %]	Standardabweichung der angegebenen Belegungsgrade eines Sensors

Das Qualitätsmerkmal der Vollständigkeit, welches über den Parameter der Abdeckungsrate als Grad der Abdeckung des Straßennetzes mit Sensoren definiert ist, wird unabhängig von den Messreihen ermittelt und ist sowohl vor als auch nach Einführung der QS-Maßnahmen ein fixer Wert. Ebenso sind mit den eingeführten Maßnahmen die Korrektheit und die Genauigkeit der SES nicht positiv veränderbar. Beide Merkmale beruhen im Wesentlichen auf Herstellertoleranzen und sind teilweise über Referenz-

messungen ermittelbar, welche einmalig nach Inbetriebnahme der SES vorgenommen wurden.

Für stationäre Schleifensensoren auf der Autobahn A81 konnte im Rahmen einer Studienarbeit unabhängig von den Herstellerangaben mit einer maximalen Abweichung von ca. 1 % zur exemplarischen Referenzzählung eine hohe Korrektheit der Verkehrsstärke nachgewiesen werden (Karrer 2008). Wie eine Studie von (Sachse 2002) beispielhaft an zwei TEU-Sensoren nachweist, weichen die ermittelten Verkehrstärken in ca. 95 % der Testintervalle nicht mehr als 1Fz/5Min von der tatsächlichen Verkehrsstärke ab.

Im Einzelnen wurden folgende Qualitätsmerkmale der SES vor und nach Einführung der QS-Maßnahmen ermittelt:

Tabelle 4: Qualitätsparameter der SES mit und ohne QS-Maßnahmen

Merkmal	Parameter	Abk. und Einheit	Vorher (KW40/2008)	Nacher (KW06/2009)
Verfügbarkeit VE	Globale Ausfallrate	AUR_{global} [%]	93,1% (6,9%AUR)	95,8% (4,2%AUR)
Aktualität AK	Mittlere Meldeverzögerung	$\overline{\Delta t}_M$ [min]	1´13	1´15
Vollständigkeit VO	Abdeckungsrate (Gesamt: 64,3km)	ABR [%]	41,9%	41,9%
Konsistenz KO	Globale Konsistenzrate	KON_{global} [%]	99,3%	99,1%
Korrektheit KR	Korrektheitsrate der Messwerte	KOR_V [%]	n.a.*	n.a.*
Genauigkeit GE	Lagegenauigkeit	G_L [± m]	± 5m	± 5m
	Genauigkeit der Geschwindigkeit	G_V [± km/h]	± 3 km/h / 5%	± 3 km/h / 5%
	Genauigkeit der Verkehrsstärke	G_q [± KFZ/Δt]	± 10 KFZ/ 5min	± 10 KFZ/ 5min
	Genauigkeit der Belegung	G_B [± %]	n.a.*	n.a.*

n.a.*: nicht auswertbar, keine Aussage möglich

Es zeigt sich, dass nach Einführung der QS-Maßnahmen insbesondere die Verfügbarkeit der Messdaten von 93,1% auf 95,8% gesteigert werden konnte. Die Ausfallrate der Daten aus stationären Erfassungssystemen konnte entsprechend deutlich von 6,9% auf 4,2% zurückgeführt werden – das entspricht einem Rückgang in Höhe von ca.

39,1%. Zu diesem positiven Ergebnis trägt sicherlich die implementierte Qualitätssicherung einen großen Teil bei. Es muss allerdings hinzugefügt werden, dass die Vorher-Untersuchung in 2008 zeitlich in die erste Betriebsphase des Concert[3]-Systems fällt und insofern hier auch die typischen „Kinderkrankheiten“ zu datenausfallbedingten Störungen beigetragen haben.

Gleichzeitig ist ein leichter Rückgang der Datenkonsistenzrate von 99,3% auf 99,1% zu verzeichnen, was durch die stringenteren Vorgaben zur Plausibilitätskontrolle innerhalb der Concert-Software zum schnelleren Erkennen von Messwert-Implausibilitäten bedingt sein dürfte.

Die Aktualität der Messdaten ist mit einer mittleren Meldeverzögerung von im Schnitt einer Minute und dreizehn Sekunden in der Vorher-Untersuchung zu einer Minute und 15 Sekunden in der Nachher-Untersuchung annähernd gleich geblieben.

[3] Bei dem Concert-System handelt es sich um ein Verkehrsrechnersystem der Firma Siemens (http://www.siemens.com)

2.2 Nachweis der Wirksamkeit der QS-Maßnahmen bei der Generierung von FPD-Trajektorien

In diesem Abschnitt werden im Wesentlichen die bereits in dem Maßnahmenkatalog zur Qualitätssicherung (Laufer 2008) dargestellten Maßnahmen aufgegriffen. Zur Beurteilung der Wirksamkeit der eingeführten Maßnahmen wurden geeignete Qualitätsparameter bzw. -kennzahlen aus den bestehenden konkretisierten Qualitätsmodellen herangezogen.

Zu Beginn der Abschnitte 2.2.2 und 2.2.3 werden daher jeweils nochmals die Qualitätsmodelle zur Beschreibung der Positionsfolgen und der FPD-Trajektorien aufgeführt, wie sie bereits im Bericht zum Evaluierungskonzept (DoiT 2008) enthalten sind. Im Falle der Berechnung der FPD-Trajektorien aus A-Daten, wurde das Qualitätsmodell noch ergänzt, da die Entwicklung der Algorithmen zum Zeitpunkt der Berichterstellung noch nicht abgeschlossen war.

Zur Wahrung der Übersichtlichkeit erfolgt die Darstellung der QS-Maßnahmen in drei Teilabschnitten. Zunächst wird auf den gesamten FPD-Server eingegangen, auf dem alle Prozessierungsschritte zur Berechnung von FPD-Trajektorien implementiert sind. Im den beiden nachfolgenden Kapiteln werden die Berechnung von Positionsfolgen aus Mobilfunkdaten sowie die eigentliche Berechnung der FPD-Trajektorien behandelt.

Aufgrund zeitlicher Einschränkungen war die Durchführung einer FMEA weder vor Entwicklung der Prozesse noch nach deren Fertigstellung möglich. Daher konnten die möglichen Probleme, die zu einer Verschlechterung der Datenqualität führen können, nicht systematisch aufgearbeitet und in formaler Weise dargestellt werden. Die Maßnahmen zur Fehlervermeidung und zur Verbesserung der Datenqualität wurden größtenteils parallel mit den Algorithmen und der Auswertesoftware entwickelt. Daher sind viele der aufgeführten Maßnahmen bereits erfolgreich umgesetzt. Der Nachweis der Wirksamkeit wird daher – soweit möglich – durch den Vergleich geeigneter Qualitätskennzahlen geführt.

2.2.1 FPD-Server

Auf dem FPD-Server laufen alle vom IAGB entwickelten Programme zur Generierung von Trajektorien von Verkehrsteilnehmern aus Mobilfunkdaten. Daher wird die Verfügbarkeit der einzelnen Zwischenprodukte unmittelbar durch die Verfügbarkeit des FPD-Servers beschrieben.

Maßnahmen zur Steigerung der Verfügbarkeit

Es erfolgt hier keine Unterscheidung in **A**- und **Abis**-Daten, da die Maßnahmen identisch sind.

Vor der Inbetriebnahme der ersten online-fähigen Version der gesamten Prozesskette zur Generierung von FPD-Trajektorien aus Abis-Daten wurde die Laufstabilität des IAGB-Systems beim Projektpartner DDG in Bonn getestet. Die Testphase erstreckte sich über etwa 6 Wochen von Ende September bis Anfang November 2008. Zu Beginn der Testphase traten fast täglich Ausfälle des Systems und damit Datenlücken auf. Nachdem unter anderem die unten aufgeführten Maßnahmen zur Verbesserung und Sicherung der Verfügbarkeit implementiert waren, lief das System am Ende der Testphase 2 Wochen lang ohne Unterbrechung. Leider wurde aus Zeitgründen keine exakte Protokollierung der Ausfallzeiten und -Gründe vorgenommen, daher kann die Senkung der Ausfallrate des Systems im Nachhinein nur grob abgeschätzt und nur unzureichend numerisch belegt werden. Fest steht jedoch, dass es innerhalb des Zeitraums von etwas weniger als 5 Monaten (12.11.08 – 31.03.09), während dem in Karlsruhe die Abis-Datenauswertung betrieben wurde, nur zu einer geringen Anzahl von Fehlermeldungen mit zumindest teilweiser Unterbrechung des Datenstroms gekommen ist. Aufgrund der fehlenden Protokollierung der Ausfälle bis Ende Dezember konnte jedoch nur für den Zeitraum 1. Januar bis 31. März 2009 die durchschnittliche Verfügbarkeit des Systems ermittelt werden. Geht man jeweils von ganztägigen Ausfällen aus, so ergeben sich in diesem Zeitraum nach Monaten sortiert die folgenden durchschnittlichen Verfügbarkeiten:

Tabelle 5: Verfügbarkeitsrate des Systems in Karlsruhe zur Abis-Auswertung

	Januar 2009	Februar 2009	März 2009	Gesamt
Systemausfall	8 Tage	13 (2) Tage[4]	1 Tag	22 (11) Tage
Verfügbarkeitsrate	74 %	54 (92) %	97 %	76 (88) %

Zur Senkung der Ausfallrate des FPD-Servers im online-Betrieb und damit zur Steigerung der Verfügbarkeit der Daten wurden die folgenden Maßnahmen ergriffen, die die Laufstabilität des Auswerterechners erhöht haben:

- Abbruch und Neubeginn der Auswertung alle 24 Stunden.
  Damit wird eine komplette Neuinitialisierung erzwungen, die die Anhäufung von Fehlern bis hin zum Absturz der Programme verhindern soll.
- Start der Neuberechnung um 3 Uhr.
  Dies minimiert die Anzahl der während des Neustarts nicht auswertbaren MFD, da zu dieser Zeit die minimale Anzahl an Telefonaten gemessen wird (vgl. Czommer 2008). Die Effizienz des Systems wird hierdurch erhöht.
- Einbau eines Fernzugriffs mit Reset-Möglichkeit.

[4] Es handelt sich hier lediglich um 2 Ausfälle, deren Behebung aufgrund personeller Engpässe jeweils erst nach rund einer Woche durchgeführt werden konnte. Daher sind in Klammern die Werte für eine angenommene Behebung jeweils innerhalb eines Tages dargestellt.

Durch die Möglichkeit der Fernwartung ist eine schnellere Behebung von auftretenden Problemen möglich, was eine geringere Ausfallrate zur Folge hat.

- Skript zur automatischen Überprüfung der Prozesse
  (Umsetzung nur bei der **A**-Daten-Prozessierung).
  Sofern möglich werden ausgefallene Prozesse umgehend automatisch neu gestartet und damit die Verfügbarkeit erhöht (sonst wäre das System erst mit dem planmäßigen Neustart um 3 Uhr wieder funktionstüchtig und es würden Datenlücken von bis zu 24 h entstehen).

Die automatische Überprüfung und ggf. Neustart der Prozesse wurde nur bei der Verarbeitung von **A**-Daten implementiert. Aufgrund der kurzen Online-Phase der letzten Version von nur einer Woche Ende März kann jedoch keine gesicherte numerische Aussage über die Steigerung der Verfügbarkeit durch diese Maßnahmen getroffen werden. Nach dem offiziellen Ende der Projektlaufzeit mussten die installierten Systeme unmittelbar wieder außer Betrieb genommen werden.

Die anderen genannten Maßnahmen wurden bereits zu Beginn der ersten Online-Phase am 12.11.08 auf dem **Abis**-System in Karlsruhe realisiert.

Das System zur Verarbeitung der **A**-Daten wurde aufgrund der bekannten Projektverzögerungen erst am 28.01.09 in Fellbach in Betrieb genommen. Eine Woche vor Projektende erfolgte noch ein Update der Software. Im Onlinebetrieb von insgesamt 6 Wochen wurden an 5 Tage Ausfälle verzeichnet, was einer Verfügbarkeit von knapp 90 % entspricht. Weitere Erfahrungen im dauerhaften online-Betrieb hätten voraussichtlich zu einer weiteren Verbesserung der Verfügbarkeit geführt.

Maßnahmen zur Steigerung der Aktualität

Es wurde hier lediglich auf eine hohe Rechnerperformance geachtet um eine schnellere Datenverarbeitung und damit die Echtzeitfähigkeit der Prozesse zu ermöglichen.

Der Rechner zur A-Daten-Prozessierung war mit 4 Prozessoren ausgestattet und zur Zeit der Außerbetriebnahme noch nicht voll ausgelastet. Die Möglichkeit der parallelen Berechnung zweier Auswerteprozesse konnte jedoch aus Zeitmangel nicht mehr umgesetzt werden.

Die weiteren Qualitätsmerkmale Genauigkeit und Korrektheit dienen der Beschreibung von Daten und sind daher zur Beschreibung des FPD-Server nicht sinnvoll.

2.2.2 Berechnung von Positionsfolgen aus Mobilfunkdaten

Zur Berechnung von Trajektorien aus **Abis**-Daten unter Verwendung der Signalstärken wurde das RX-Matching-Verfahren verwendet. Hierbei werden zunächst Positionsfolgen einzelner Mobilfunkteilnehmer berechnet, die in einem weiteren Schritt mit Hilfe

des Map-Aiding-Verfahrens zu Trajektorien im digitalen Straßennetz zusammengesetzt werden.

Im Unterschied dazu ist bei der Berechnung von Trajektorien aus **A**-Daten nicht explizit die Berechnung von Positionen in Form von Koordinaten möglich. Hier werden direkt die Zellwechsel sowie die Wechsel der *Location Area* als Ortsinformation zur Bestimmung der wahrscheinlichsten Trajektorien von Mobilfunkteilnehmern verwendet.

Eine detaillierte Darstellung des Verfahrensentwicklung zur Berechnung der Positionsfolgen sowie zur Trajektoriengenerierung aus **Abis**-Daten findet sich in (Ramm u. Schwieger 2008), (Schwieger et al. 2007) und in (Czommer 2008). Die Trajektorienberechnung auf Basis der **A**-Daten wird im gemeinsamen Abschlussbericht zum AP4.6 sowie AP 10.3 noch ausführlich dargestellt, vgl. hierzu (Beetz et al. 2009).

In der Tabelle 6 ist das Qualitätsmodell für die berechneten Positionen und Positionsfolgen dargestellt.

Qualitätsbeschreibung der aus Mobilfunkdaten (MFD) ermittelten Positionen (Abis)

Tabelle 6: Qualitätsparameter der Positionsfolgen aus MFD

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
P1.1	Verfügbarkeit VE	Globale Ausfallrate (offline+relativ)	AUR_{global} [%]	Beschreibt die Verfügbarkeit des FPD-Servers in einem best. Zeitraum
P2.1	Aktualität AK	Alter einer Einzelposition (online+relativ)	Δt_{Pos} [s]	Das Alter einer Position wird durch die Differenz *aktuelle Zeit – Zeitstempel der MFD* berechnet, dieser Zeitstempel ist im Datenformat für die Positionen enthalten.
P2.2		Alter der Positionsfolge (online+relativ)	Δt_{Posf} [s]	Das Alter einer Folge von Positionen wird durch das Alter der ältesten Position bestimmt.
P2.3		mittlere Berechnungsdauer (offline+relativ)	$\overline{\Delta t}_{Pos\,Rech}$ [s]	Mittlere Berechnungsdauer für eine Position (unabhängig vom FPD-Alter) zur Beurteilung des Prozesses
P3.1	Vollständigkeit VO	Da es sich bei den berechneten Positionen nur um Koordinaten handelt, gibt es keine sinnvolle Definition einer vollständigen Abdeckung, daher wird die VO nur bei den Trajektorien behandelt		
P4.1	Konsistenz KO	Wird durch Einhaltung des Datenmodells gewährleistet		
P5.1	Korrektheit KR	Korrektheitsrate (offline+absolut)	KR [%]	Schwellwert für „korrekt“ und „nicht mehr korrekt“ erforderlich, dieser hängt von der jeweiligen Anwendung ab und kann daher nicht allgemein berechnet werden

P6.1	Genauigkeit GE	Standardabweichung (2D) (online+absolut)	S_{Pos} [m]	Geometrische Genauigkeit einer einzelnen Position (nur im Rahmen der Testfahrten ermittelbar)
P6.2		mittlere Stdw. einer Positionsfolge (online+absolut)	S_{Posfo} [m]	Geometrische Genauigkeit einer Positionsfolge als Mittelwert über alle Einzelpositionen (nur im Rahmen der Testfahrten ermittelbar)
P6.3		mittlere Stdw. über Zeitraum oder Gebiet (offline+absolut)	$\overline{S}_{Posfo}$ [m]	Durchschnittliche geometrische Genauigkeit der Positionsfolgen über einen Zeitraum oder ein Gebiet (nur im Rahmen der Testfahrten ermittelbar)

Maßnahmen zur Steigerung der Verfügbarkeit

Die technische Verfügbarkeit der Positionsfolgen hängt im Wesentlichen nur von der Verfügbarkeit des FPD-Servers ab.

Maßnahmen zur Steigerung der Aktualität

Aufgrund der Funktionsweise des Mobilfunknetzes kann von einer Genauigkeit der in den Daten enthaltenen Zeitstempel von wenigen Zehntelsekunden ausgegangen werden. D. h. die Aufzeichnung der Mobilfunkdaten zusammen mit den Zeitstempeln erfolgt nahezu in Echtzeit (vgl. Baumstark 2007). Diese minimale Zeitverzögerung spielt angesichts der erforderlichen Zeit zur Berechnung der FPD-Trajektorien keine Rolle und wird daher nicht weiter betrachtet. Es muss jedoch sichergestellt sein, dass die Berechnung und das zur Verfügungstellen der Trajektorien so schnell wie möglich erfolgt, um zeitkritische Anwendungen, wie z. B. die Verkehrslageschätzung bedienen zu können. Verzögerungen zwischen Aufzeichnung und dem zur Verfügung stellen der aktuellen Verkehrslage von mehr als 30 Minuten sind auf Grund des meist sehr dynamischen Verhaltens von Verkehrsstörungen inakzeptabel.

Zur Trajektorienberechnung aus Mobilfunkdaten in zufriedenstellender Genauigkeit müssen jedoch zunächst lange Telefonate zur Verfügung stehen, dasselbe gilt für die Unterscheidung von bewegten und stationären Mobiltelefonen. Daher kommt der Verkürzung der erforderlichen Rechenzeit eine besondere Bedeutung zu. Zur Minimierung der Rechenzeit zur Bestimmung der Positionen wurden daher einige Maßnahmen untersucht, die in der folgenden Tabelle 7 kurz aufgeführt sind.

Tabelle 7: Maßnahmen zur Reduktion der Rechenzeit zur Bestimmung von Positionen

	Maßnahme	umgesetzt?
1	Vergrößerung des Rasters der Signalstärkenkarten	Ja
2	Ausdünnen der Informationsdichte der Mobilfunkdaten	Nein
3	Vergrößerung des Arbeitsspeichers	Ja
4	Laufzeitoptimierte Programmierung (z. B. Vorhalten der Signalstärkekarten im Arbeitsspeicher)	Ja
5	Parallelisierung der Prozesse	Ja

Die Verminderung der Rechenzeit wurde durch die Verkleinerung des Parameter P2.3 $\overline{\Delta t}_{Pos\,Rech}$, der *mittleren Berechnungsdauer* bzw. durch Erhöhung der berechneten Trajektorien pro Zeit nachgewiesen. Es musste hierbei jedoch die ggf. durch die Beschleunigung bedingte Verschlechterung der Genauigkeit und Korrektheit der Trajektorien mit berücksichtigt werden. Dazu konnten die Parameter *Zuordnungskorrektheit Typ A* und *B* sowie im Falle der **Abis**-Daten die *Mittlere Querabweichung* herangezogen werden.

Der Nachweis für die Wirksamkeit der Maßnahme 1 in Tabelle 7 wurde im Rahmen der ersten Testfahrten erbracht. Wie aus internen Untersuchungen hervorgeht, konnte die Berechnung der Trajektorien durch ausdünnen der Signalstärkekarten um den Faktor 12 beschleunigt werden. Wie in (Czommer 2008) beschrieben bzw. in der Abbildung 5 erkennbar, ist dadurch nur ein sehr geringer und daher akzeptabler Genauigkeitsverlust zu erwarten. Wie aus Tabelle 9 hervorgeht, steigt die Standardabweichung der Positionsfolgen im Mittel nur von 651 auf 655 m, daher hat sich als Kompromiss zwischen einer Verschlechterung der Ergebnisse bei zu groben Signalstärkekarten und einer massiven Verlängerung der Rechenzeit bei hoher Kartenauflösung ein 5''-Raster für die Signalstärkekarten als Optimum erwiesen.

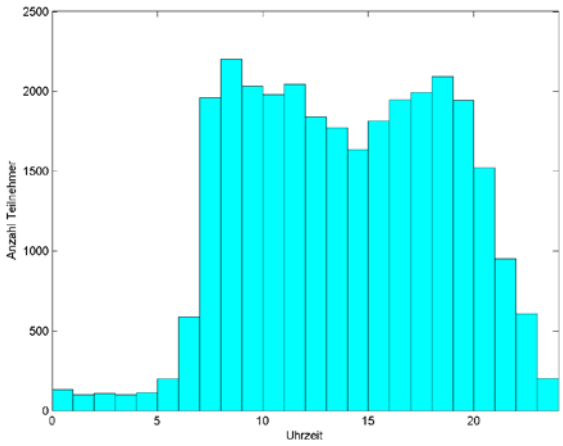
Die Ausdünnung der Mobilfunkdaten (Maßnahme 2 in der Tabelle 7) zur Steigerung der Performance wurde trotz einer Beschleunigung um Faktor 10-20 wieder verworfen, da die gesamte Informationsdichte der anfallenden Mobilfunkdaten zur Identifikation der aktiven Verkehrsteilnehmer erforderlich ist.

Durch Vorhalten der Signalstärkekarten im Arbeitsspeicher (Maßnahmen 3 und 4) entfällt das ständige rechenintensive Einlesen. Dadurch wurde eine signifikante Steigerung der Rechenleistung erzielt.

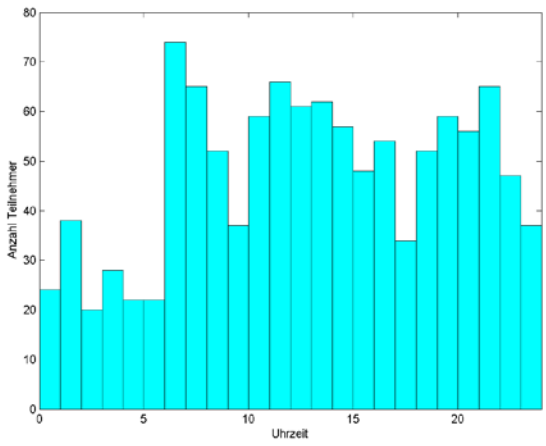
In der online-Variante der **Abis**-Datenauswertung wurden sämtliche berechenbare Teilnehmer sequenziell ausgewertet. Durch die parallele Auswertung (Maßnahme 5) mehrerer Teilnehmer konnte im Wesentlichen durch die Reduktion der Eingabe- und

Ausgabeoperationen die Auswertung um den Faktor 2 beschleunigt werden. Zusätzlich wurden in der neuesten Version der Auswertesoftware , die aus Zeitgründen nur in der offline-Variante realisiert wurde, immer vier parallele Prozesse gestartet, wodurch der Durchsatz – die Anzahl der berechneten Trajektorien pro Stunde – nochmals nahezu vervierfacht werden konnte.

Mit der letzten Version, die in Karlsruhe im Echtzeitbetrieb online installiert war, konnten an einem durchschnittlichen Wochentag innerhalb von 24 Stunden ca. 1 150 Teilnehmer-Trajektorien im Stadtgebiet Karlsruhe aus **Abis**-Daten berechnet werden. Im gesamten Projektgebiet wurden aus **A**-Daten durchschnittlich ca. 30 000 Teilnehmer berechnet. In der folgenden Abbildung 1 ist der Tagesverlauf der berechneten Teilnehmer an einem durchschnittlichen Wochentag dargestellt.



A-Daten Auswertung (gesamtes Projektnetz)



Abis-Daten Auswertung (Karlsruhe)

Abbildung 1: Darstellung der Anzahl berechneter Teilnehmer innerhalb 24 h an einem durchschnittlichen Wochentag;(jeweils Di., 03.03.2009); Vorsicht: verschiedene Maßstäbe in beiden Bildern

Maßnahmen zur Steigerung der Genauigkeit / Korrektheit

Die maximal zur Verfügung stehende Auswertezeit wurde auf 15 Minuten begrenzt. Es werden immer 15 Minuten dauernde Blöcke sortiert und anschließend ausgewertet, daher sind die Trajektorien maximal 30 Minuten alt. Dabei werden – im Hinblick auf die Auswertung – Erfolg versprechende Teilnehmer priorisiert, wodurch insbesondere Genauigkeit und Korrektheit aber indirekt auch die Vollständigkeit (höhere Anzahl von brauchbaren Trajektorien) der generierten Positionsfolgen und damit der Trajektorien gesteigert wird. Leider wurde diese Steigerung aufgrund Zeitmangels nicht numerisch nachgewiesen.

Da die Mobilfunkdaten und deren Erfassung nicht beeinflusst werden können, konzentrieren sich die weiteren Maßnahmen auf geeignete Plausibilitätsprüfungen der zu verarbeitenden Eingangsdaten. Dabei werden fehlerhafte Daten herausgefiltert und von der weiteren Verarbeitung ausgeschlossen. Bei den notwendigen Daten zur Ortung auf Abis-Ebene, handelt es sich um

- den Zeitpunkt der Informationsentstehung,
- die Teilnehmer-ID (Identifizierungsnummer des Teilnehmers),
- die Call-ID (Identifizierungsnummer des Telefonats),
- die Cell-ID (Identifizierungsnummer der Mobilfunkzelle) und
- die Signalstärke

der bedienenden Zelle sowie um

- die Cell-IDs der benachbarten Zellen sowie die
- Signalstärken der benachbarten Zellen (bis zu 6),

die alle zweimal je Sekunde während des Gesprächs zur Verfügung stehen. Auf A-Ebene stehen keine Signalstärken und keine Informationen benachbarter Zellen zur Verfügung.

Mögliche Fehlerquellen in den Mobilfunkdaten mit Auswirkungen auf die Qualitätsparameter Genauigkeit und Korrektheit der Positionen sowie die Implementierten Gegenmaßnahmen sind in der folgenden Tabelle 8 aufgeführt.

Tabelle 8: Fehlerquellen in den Mobilfunkdaten und Gegenmaßnahmen

	Mögliche Fehlerquelle	**Gegenmaßnahme**	**umgesetzt?**
1	Cell-ID kann vom Mobilfunkbetreiber nicht zugeordnet werden	Belegung mit Default-Wert „0“ -->keine Berücksichtigung	Ja
2	Cell-ID ist nicht bekannt bzw. ID ist in keiner Signalstärkenkarte vorhanden	Prüfung mit Hilfe einer Lookup-Tabelle; Zellen, die nicht in den Stammdaten auftreten erhalten Default-Wert „0“ -->keine Berücksichtigung	Ja
3	Falsche Zuordnung einer oder mehrerer Nachbarzellen	Überprüfung mit Hilfe der Koordinaten der Signalstärkekarten	Ja
4	Signalstärke ist nicht ermittelbar	Belegung mit Default-Wert; wird nicht in der Auswertung berücksichtigt	Ja
5	Signalstärke ist zu niedrig für eine Auswertung	Einführung eines variablen Schwellwerts für die Signalstärke; ggf. Elimination zu kleiner Werte	Ja

Erläuterungen:

Zu 2: Da die Daten für die Punktfolgen blockweise ausgewertet werden, kann vor der Auswertung ein Plausibilitätscheck für alle aufgeführten Zellen durchgeführt werden (Blockweise heißt es wird eine Datei mit allen Abis-Zeilen eines Telefonanrufs übergeben). Es werden alle vorkommenden Zellen eines Telefonats an Hand einer Lookup-Tabelle mit den Stammdaten verglichen. Dabei werden Zellen, die in den Stammdaten nicht vorkommen, aus den Abis-Zeilen gelöscht und nehmen somit nicht an der Auswertung teil.

Zu 3: Wenn im laufenden Betrieb Mobilfunkzellen falsch zugeordnet werden, so ist dies nicht unmittelbar feststellbar. Da die Positionsbestimmung mit Hilfe eines Suchgebiets in Form einer Sektormaske über mehrere Zellen geschieht, können grob falsch zugeordnete Zellen mit Hilfe der Signalstärkenkartenkoordinaten überprüft werden. Das Suchgebiet richtet sich immer nach der bedienenden Zelle. Da die Signalstärkenkarte (67 km x 67 km) einer Zelle einen ausreichend großen Überlappungsbereich mit ihren Nachbarzellen besitzt, kann davon ausgegangen werden, dass eine Nachbarzelle falsch zugeordnet wird, wenn diese nicht in das Suchgebiet einer bedienenden Zelle fällt. Dies ist bereits in der Onlineauswertung implementiert. Falsch zugeordnete Nachbarzellen werden somit nicht in die Auswertung mit einbezogen. Wird eine bedienende Zelle falsch zugeordnet, so wird dies nicht berücksichtigt. Dies könnte nur entdeckt werden, wenn z.B. mehrere Nachbarzellen nicht in das Suchgebiet fallen würden. Dies konnte jedoch aus Zeitgründen nicht mehr umgesetzt werden.

Durch die dargestellten Maßnahmen konnten fehlerhafte Eingangsdaten weitestgehend vermieden werden. Da die Umsetzung der Vorfilterung bereits vor Implementierung einer lauffähigen Online-Version der Auswertesoftware abgeschlossen war, ist der nachträgliche Vergleich und ein damit verbundener numerischer Nachweis der Wirksamkeit der Maßnahmen nicht mehr mit vertretbarem Aufwand möglich.

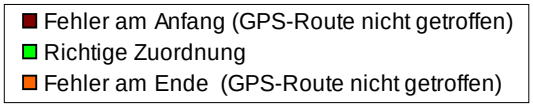
Die Einführung eines Kalman-Filters zur stützenden Prädiktion der Positionsfolgen der **Abis**-Daten-Auswertung konnte die Korrektheit der Ergebnisse leicht steigern, wie bereits im Bericht zum AP4.5 (vgl. Czommer 2008) anhand der gesteigerten Zuordnungskorrektheit A und B nachgewiesen wurde. In den nachfolgenden Abbildungen 2 und 3 wurden verschiedene Auswerteansätze exemplarisch anhand der resultierenden Zuordnungskorrektheiten verglichen. Die Darstellungen beziehen sich auf Testfahrten mit GPS, die im Stadtgebiet Karlsruhe durchgeführt wurden. Es wurden jeweils die Testfahrten auf der B3 zwischen Wolfartsweiler und Weingarten zur Darstellung ausgewählt (vgl. Abbildung 2).

Abbildung 2: Testfahrt mit GPS in Karlsruhe

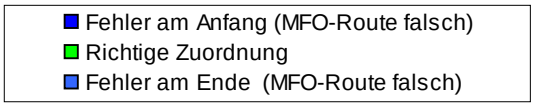
Die Beurteilung der Maßnahme erfolgt hier im Vorgriff auf Kapitel 2.2.3 anhand der Qualität der FPD-Trajektorien, da keine separate Auswertung der GPS-Testfahrten hinsichtlich Korrektheit und Genauigkeit der berechneten Positionsfolgen durchgeführt wurde. Daher sind in den folgenden Abbildungen die Zuordnungskorrektheit Typ A und B als Qualitätsparameter zur Beschreibung der Korrektheit der FPD-Trajektorien dargestellt.

Legende:

Fehler 1. Art: GPS-Route nicht getroffen



Fehler 2. Art: MFO findet falsche Route



Zur Untersuchung der verschiedenen Algorithmen zur Bestimmung der Positionsfolgen wurden (nach dem Unterstrich) die folgenden Abkürzungen verwendet:

- RX reines RX-Matching
  (RX1 mit 1“-Signalstärkenkarte; RX5 mit 5“-Signalstärkenkarte)
- gRX* = RX* + Glättung
- KF_RX* = RX* + Kalmanfilter
- RX_S10R1 = RX-Matching mit 1“-Signalstärkenkarte + Kalmanfilter mit 10-Schritten
- RX_S10R5 = RX-Matching mit 5“-Signalstärkenkarte + Kalmanfilter mit 10-Schritten
- KF_RX5M90 = RX-Matching mit 5“-Signalstärkenkarte, wobei nur Messungen größer 90dB verwendet wurden
- KF55M90 = RX-Matching mit 5“-Signalstärkenkarte, wobei nur Messungen größer 90dB verwendet wurden, + Kalmanfilterung mit 5-Schritten

Wie in den folgenden beiden Abbildungen ersichtlich ist, liefert die Auswertung unter Verwendung eines Kalman-Filters in der Regel die besten Ergebnisse. Das heißt, in den Grafiken sind die grünen Balken, die den jeweils richtig zugeordneten Trajektorienteil darstellen bei den Auswertevarianten mit Kalman-Filter am größten. Hier wurde dieser Zusammenhang nur exemplarisch gezeigt, weitere Beispiele finden sich in (Czommer 2008).

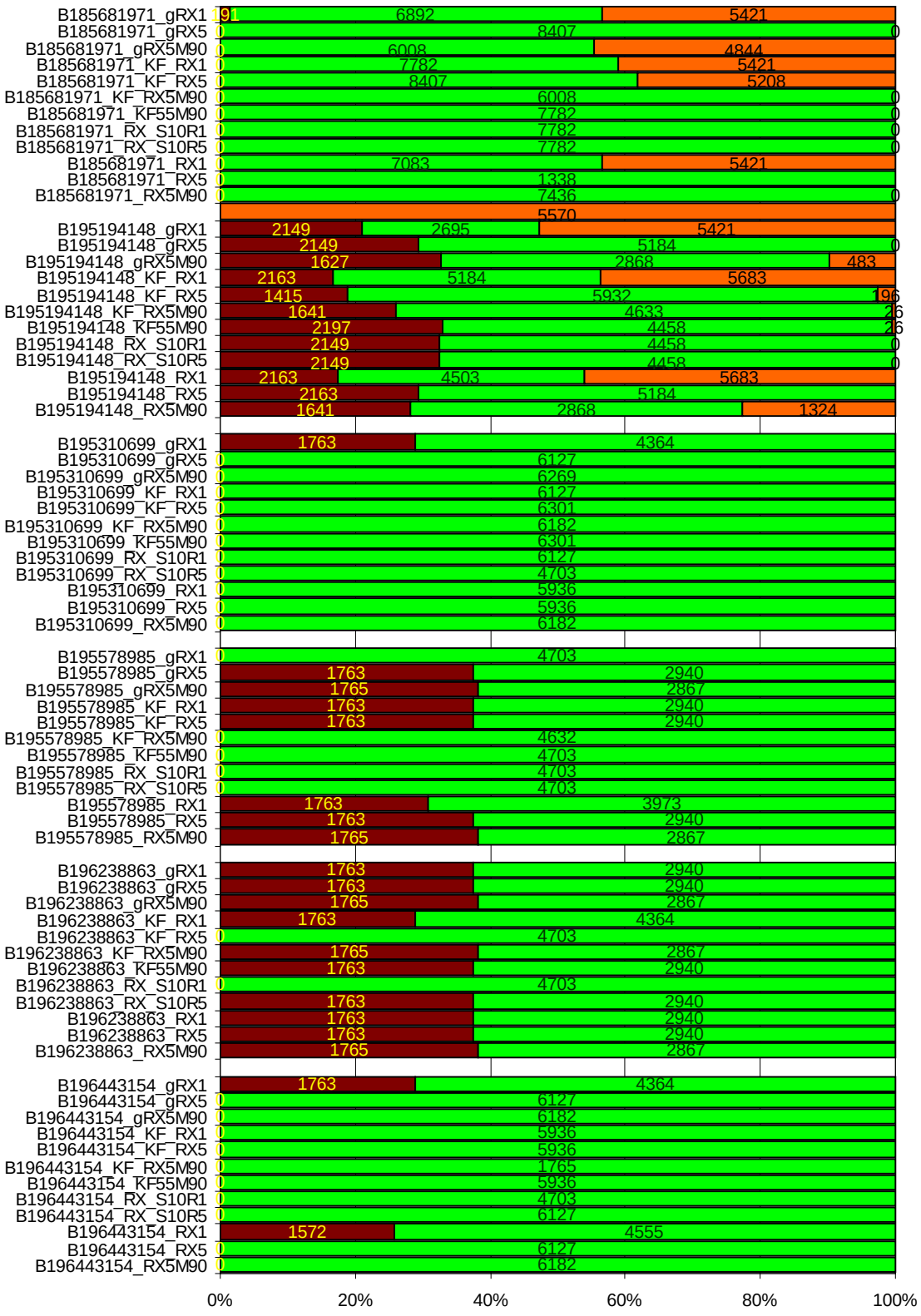


Abbildung 3: Exemplarische Darstellung der Zuordnungskorrektheit Typ A verschiedener Trajektorien auf der B3 im Raum Karlsruhe unter Verwendung verschiedener Auswertestrategien (Czommer 2008)

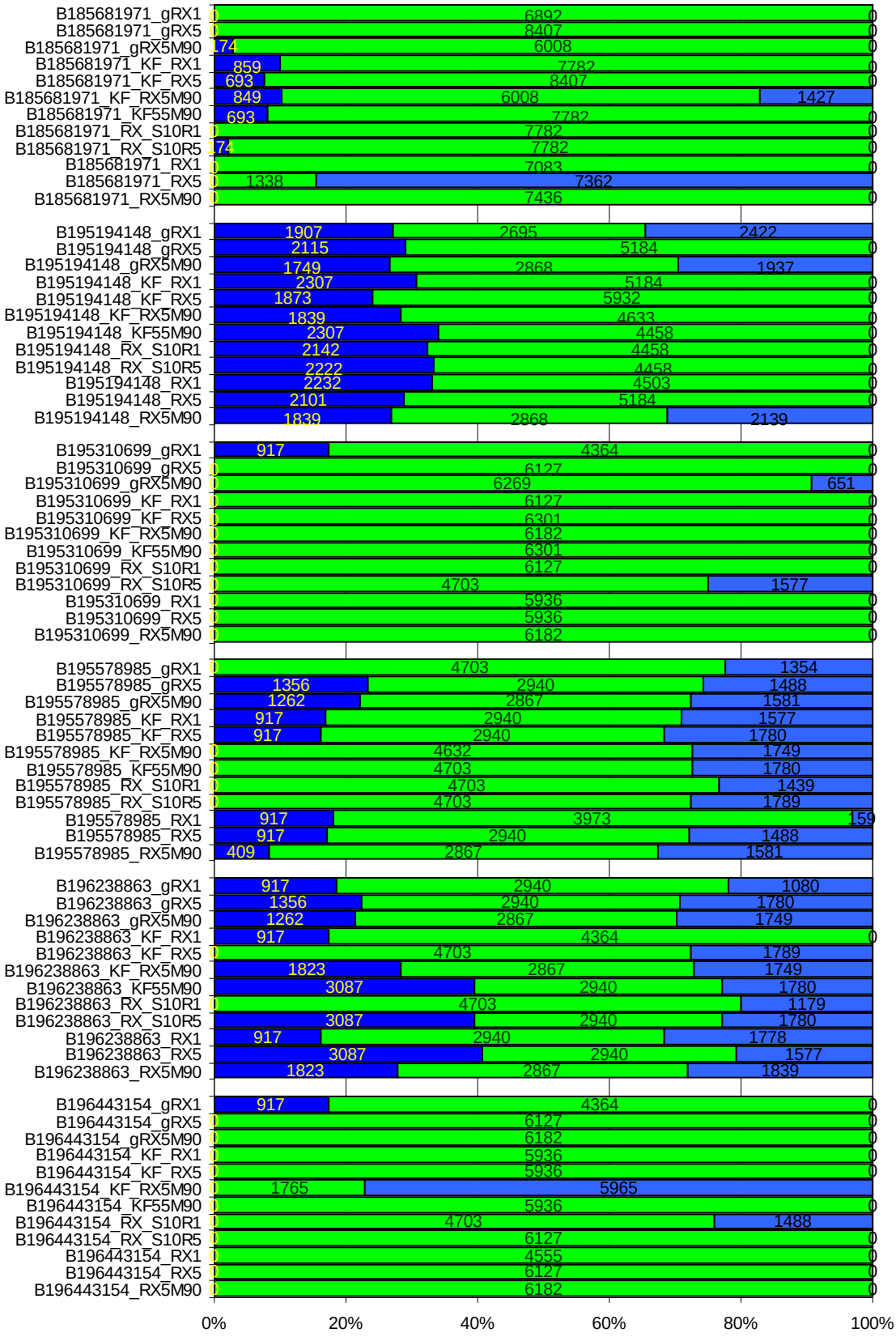


Abbildung 4: Exemplarische Darstellung der Zuordnungskorrektheit Typ B verschiedener Trajektorien auf der B3 im Raum Karlsruhe unter Verwendung verschiedener Auswertestrategien (Czommer 2008)

Auch der Vergleich der nachfolgenden Schaubilder in Abbildung 5 lässt eine deutliche Verschiebung des Genauigkeitsniveaus hin zu kleineren Standardabweichungen erkennen. Im Mittel sinken die Standardabweichungen der Positionen von 650 m auf 570 m (vgl. Tabelle 9).

Tabelle 9: Genauigkeiten für verschiedene Mobilfunkortungsverfahren

Ortung	Standardabweichung [m]		
	Mittel	Max	Min
Signalstärken-Matching 1“	651	1780	206
Signalstärken-Matching 5“	655	1709	280
Kalmanfilter 1“	587	1985	170
Kalmanfilter 5“	560	1883	108

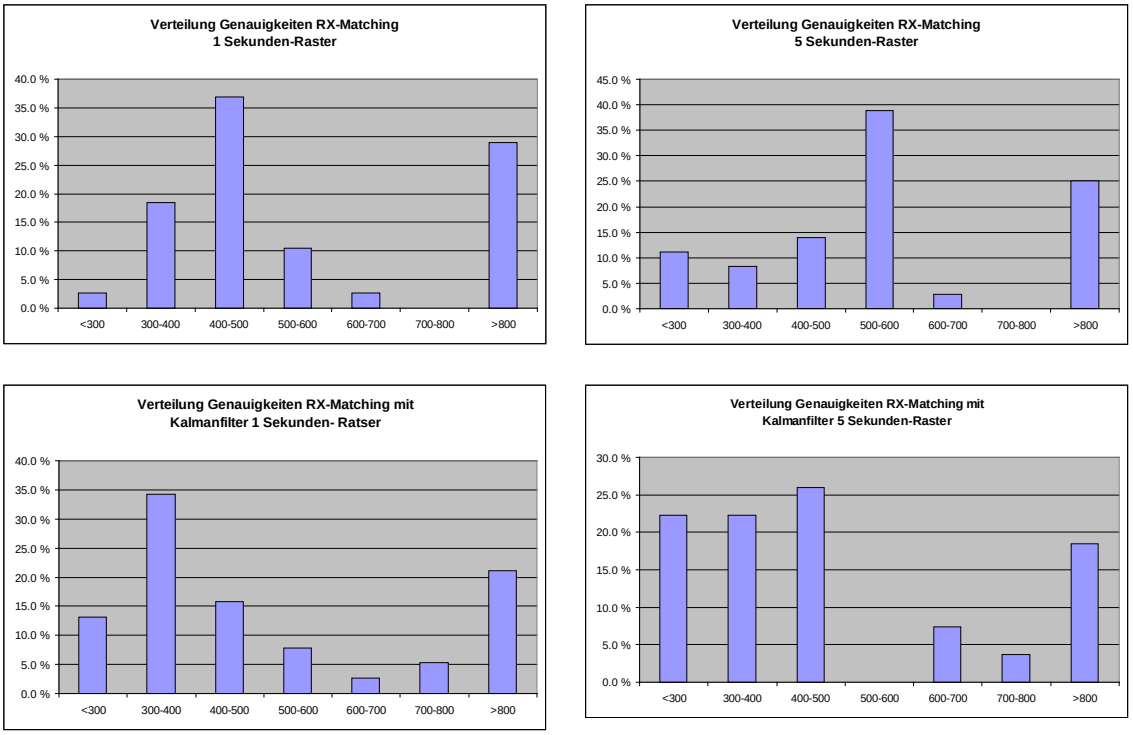


Abbildung 5: Verteilung der Genauigkeiten (Czommer 2008)

Die Verwendung von 5‘‘ Signalstärkekarten hingegen lässt keine signifikante Genauigkeitsänderung erkennen, ist jedoch – wie bereits erwähnt – aufgrund der Datenreduktion um Faktor 25 mit einer erheblichen Beschleunigung des Berechnungsverfahrens verbunden.

In (Baumstark 2007) konnten schlechte Ortungsergebnisse mit schlechten Signalstärken in Verbindung gebracht werden. Durch die Einführung eines Schwellwertes für die

gemessenen Signalstärken (vgl. Tabelle 8, Maßnahme 5) der umliegenden Mobilfunkantennen von -90 dBm konnte die Korrektheit der Positionsfolgen signifikant gesteigert werden. Offensichtlich sind extrem schwache Signale mit einer sehr schlechten Genauigkeit behaftet, die keinen Beitrag zur Positionsberechnung des Mobilfunkgerätes leisten können, sondern das Ergebnis sogar verschlechtern. Exemplarisch sind in den folgenden Abbildung 6 bis Abbildung 9 vier Beispielstrecken angeführt. Das linke Schaubild zeigt die Signalstärke[5] der bedienenden Zelle, das rechte jeweils die zugehörige zweidimensionale Koordinatenabweichung zur GPS-Referenz. In allen Bildern ist eine signifikante Verschlechterung der Genauigkeit und eine Zunahme des Rauschens bei sinkender Signalstärke und umgekehrt zu erkennen.

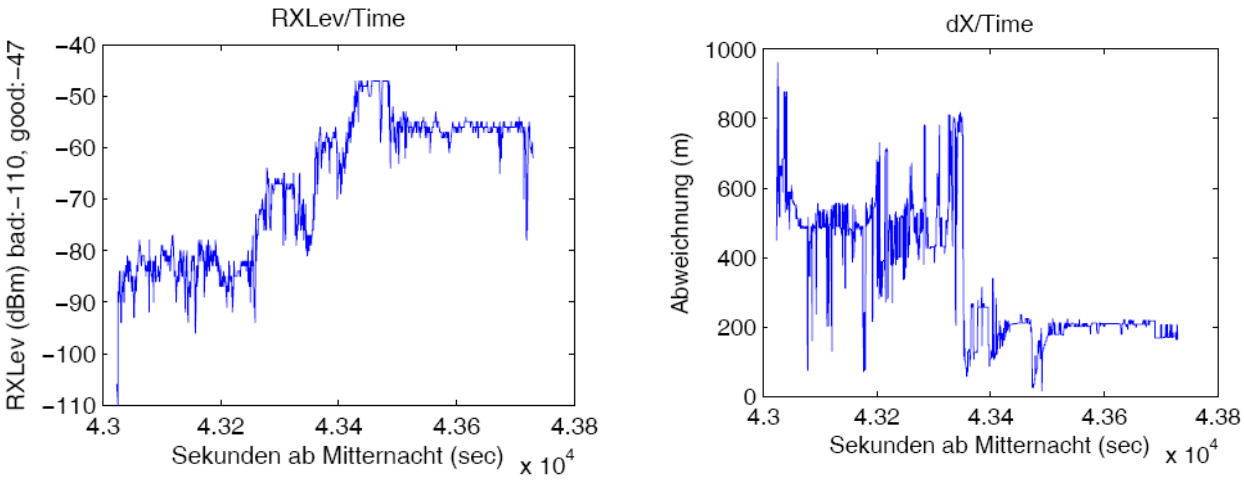


Abbildung 6: Signalstärke der bedienenden Zelle (links) sowie die zugehörige Koordinatenabweichung (rechts), Trajektorie 186274121, 5“-Karte; keine Ausdünnung

[5] Die Signalstärke wird in dbm (Dezibel-Milliwatt) angegeben; es handelt sich dabei um eine logarithmische Größe zur Angabe der Signalstärke mit der Bezugseinheit der Leistung (1 mW)

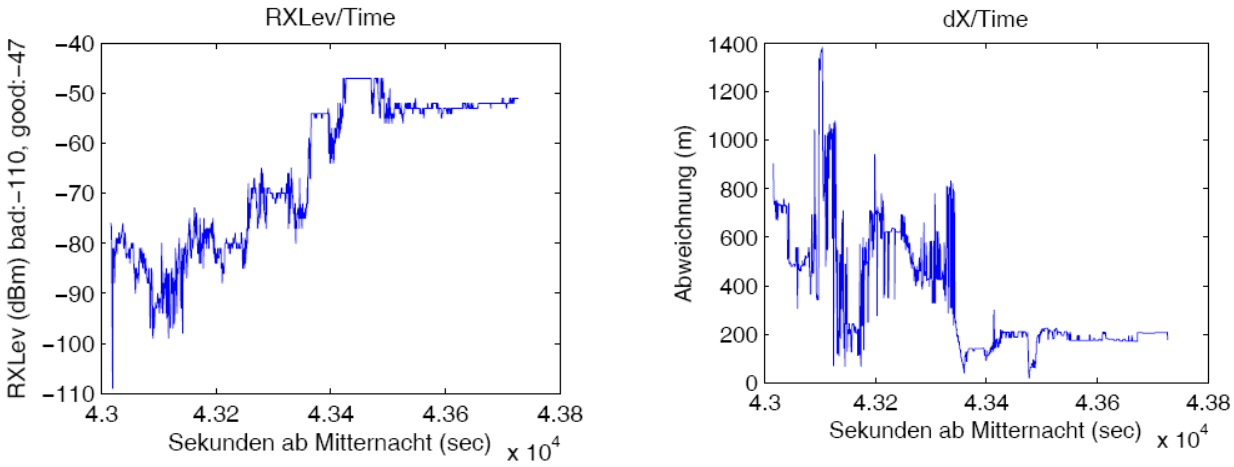


Abbildung 7: Signalstärke der bedienenden Zelle (links) sowie die zugehörige Koordinatenabweichung (rechts), Trajektorie 168668827, 5‘‘-Karte; keine Ausdünnung

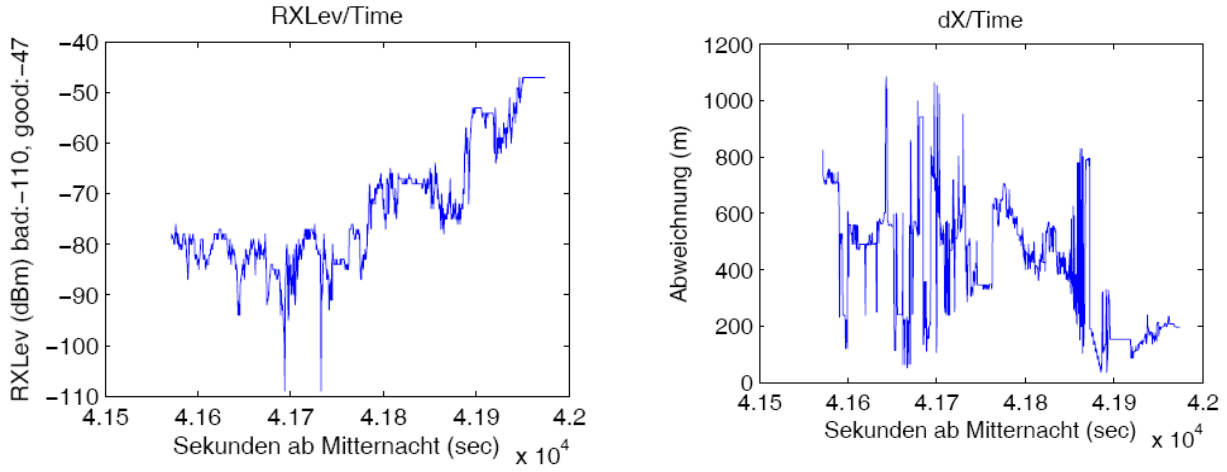


Abbildung 8: Signalstärke der bedienenden Zelle (links) sowie die zugehörige Koordinatenabweichung (rechts), Trajektorie 192940906, 5‘‘-Karte; keine Ausdünnung

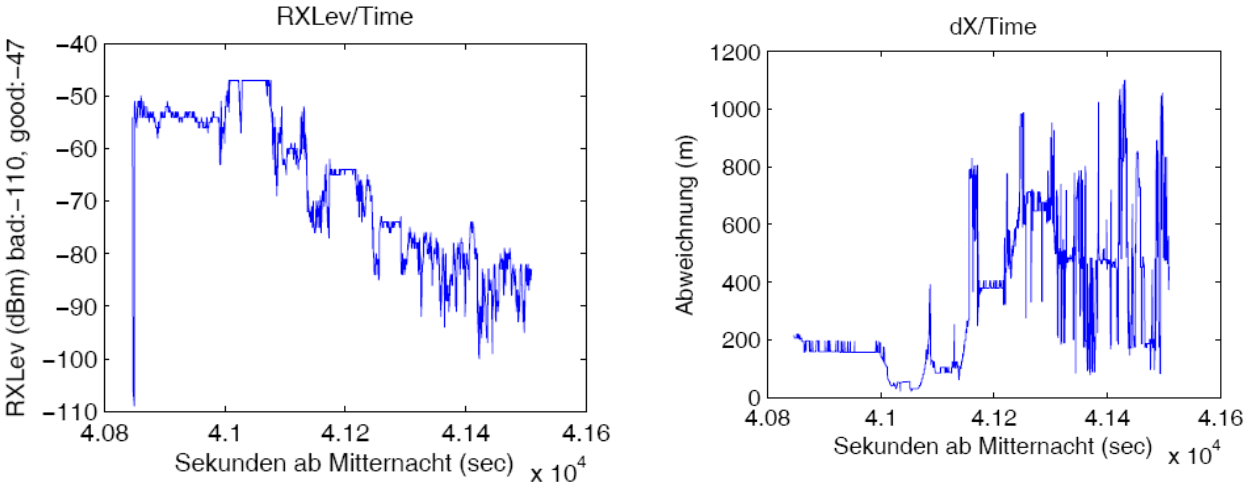


Abbildung 9: Signalstärke der bedienenden Zelle (links) sowie die zugehörige Koordinatenabweichung (rechts); Trajektorie 1953394167, 5‘‘-Karte; keine Ausdünnung

Der Nachweis der Wirksamkeit der Maßnahme wurde ausschließlich visuell anhand der Schaubilder geführt, eine umfassende numerische Untermauerung konnte auch hier aus Zeitgründen nicht erfolgen.

2.2.3 Berechnung von FPD-Trajektorien

Die Berechnung von Trajektorien aus den aus **Abis**-Daten ermittelten Positionsfolgen basiert auf Map-Matching Verfahren. Dabei beinhalten die berechneten Trajektorien sowohl die Route des Verkehrsteilnehmers auf dem Straßennetz als auch ein Weg-Zeit-Diagramm zur Verknüpfung von Ort und Zeit.

Die Trajektorienberechnung aus **A**-Daten hingegen erfolgt unmittelbar aus den Zellinformationen sowie den Location-Updates. Das im Folgenden dargestellte Qualitätsmodell der FPD-Trajektorien unterscheidet sich jedoch nur in wenigen Details.

Das nachfolgende Qualitätsmodell für FPD-Trajektorien wurde unter Berücksichtigung der Weiterverwendung der Trajektorien zur Generierung von nutzerspezifischen Verkehrsinformationen entwickelt (Laufer 2006). Das Modell zur Beschreibung der Qualität der Verkehrslagebeschreibung ist in (Do-iT 2008) dargestellt.

Qualitätsbeschreibung der FPD-Trajektorien

Tabelle 10: Qualitätsparameter der FPD-Trajektorien

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
P1.1	Verfügbarkeit VE	Globale Ausfallrate (offline+relativ)	AUR_{global} [%]	Beschreibt die Verfügbarkeit des FPD-Servers in einem best. Zeitraum
P1.2		Trajektoriendichte (online+relativ)	ρ_{Tra} [1/s]	Summe(Routenlängen)/L/T mit L = Länge des betrachteten Streckenabschnitts und T = Dauer
P2.1	Aktualität AK	Berechnungszeit (online+relativ)	Δt_{ber} [s]	Berechnungszeit für eine FPD-Trajektorie aus der Differenz *FPD_CalculationTime – FPD_Starttime*
P2.2		Mittlere Berechnungszeit (offline+relativ)	$\overline{\Delta t}_{ber}$ [s]	Mittlere Berechnungszeit (Stunde, Tag) aus P2.1
P2.3		Alter der Trajektorie (online+relativ)	Δt_{FPD} [min]	Alter einer Trajektorie, berechnet aus der Differenz zwischen der aktuellen Zeit und dem Zeitstempel der ersten Position *aktuelle Zeit –FPD_Starttime*
P2.4		Mittleres Trajektorienalter (offline+relativ)	$\overline{\Delta t}_{FPD}$ [min]	Mittleres Alter der Daten (aus P2.3)
P3.1 Abis	Vollständigkeit VO	Trajektorienvollständigkeit (online+relativ)	VO_{FPD} [%]	**Abis:** Anteil der verwendeten Positionen einer Positionsfolge eines Teilnehmers, der für die Trajektorie verwendet wird *FPD_Quality (2)*
P3.1 A				**A:** Anteil der verwendeten Zellwechsel eines Teilnehmers der für die Trajektorie verwendet wird *FPD_Quality (2)*
P3.2		Abdeckungsrate (offline+absolut)	ABR_{FPD} [%]	Grad der Abdeckung des betrachteten Straßennetzes mit FPD (Begriff: „vollständige Abdeckung“ ist zu definieren)
P4.1	Konsistenz KO	Wird durch Einhaltung des Datenmodells gewährleistet		
P5.1	Korrektheit KR	Klassifizierungskorrektheit (online+absolut)	KR_{Kl} [%] oder [0...3]	Wahrscheinlichkeit, mit der die Teilnehmerklassen identifiziert werden (Aggregation auf einen QP "aktiver IVT" geplant) *FPD_Identification*
P5.2		Mittlere Klassifizierungskorrektheit (offline+absolut)	$\overline{KR}_{Kl}$ [%] oder [0...3]	Mittlere Klassifizierungskorrektheit der Klassen z. B. über einen Tag berechnet (aus P5.1)
P5.3		Trajektorienlänge (online+relativ)	L_{FPD} [km]	Länge der FPD-Trajektorie (*FPD_RouteLength*)
P5.4		Mittlere Trajektorienlänge (offline+relativ)	$\overline{L}_{FPD}$ [km]	Mittlere Länge der FPD-Trajektorie (aus P5.3)

P5.5		Zuordnungskorrektheit Typ A (offline+absolut)	KR_{ZuA} [%]	Korrekter Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, bezogen auf die Länge der GPS-Route (GPS als Referenz erforderlich)
P5.6		Zuordnungskorrektheit Typ B (offline+absolut)	KR_{ZuB} [%]	Korrekter Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, bezogen auf die Länge der ermittelten FPD-Route (GPS erforderlich)
P6.1 ABis	Genauigkeit GE	Mittlere Querabweichung (online+relativ)	QA_{FPD} [m]	Mittel der orthogonalen Abweichungen der verwendeten Positionen von der wahrscheinlichsten Route *FPD_Quality (1*) **nur Abis-Daten**
P6.1 A		Anzahl der Zellwechsel eines Teilnehmers	Anz_{ZW} [-]	Anzahl der Zellwechsel, die einem Teilnehmer zugeordnet werden konnten *FPD_Quality (1)* **nur A-Daten**
P6.2		Standardabweichung der Geschwindigkeit (offline+absolut)	S_V	Gibt die Genauigkeit der aus den Trajektorien ermittelbaren Geschwindigkeit an. Referenz erforderlich (Messfahrt oder TEU)

Anmerkungen:

Der Parameter 5.1 **Klassifizierungskorrektheit** wird im Bericht zum AP3.6: *Qualitätsbewertung der Identifikation aktiver Verkehrsteilnehmer* verwendet.

Die **Zuordnungskorrektheit Typ A** (P5.5) berechnet sich als Verhältnis zwischen der Länge einer FPD-Route, die korrekt zugeordnet wurde zu der Gesamtlänge der GPS-Trajektorie, die mit Hilfe von Testfahrten ermittelt wurde. Dieser Parameter sagt nichts darüber aus, wie gut die Zuordnung in den Abschnitten ist, in denen die richtige Straße nicht gefunden wurde und wie groß der nicht richtige Teil der Trajektorie bezogen auf die gesamte Trajektorie ist. Daher wird mit der **Zuordnungskorrektheit Typ B** (P5.6) ein weiterer Parameter eingeführt, der als Bezugsgröße die Gesamtlänge der FPD-Trajektorie zu Grunde legt.

Maßnahmen zur Steigerung der Verfügbarkeit

Im Abschnitt 2.2.1 sind die Maßnahmen zur Steigerung der Systemverfügbarkeit bereits dargestellt. Die Maßnahmen sind unabhängig von dem jeweiligen Verarbeitungsschritt in der Prozesskette zur Generierung von Mobilfunktrajektorien.

Maßnahmen zur Steigerung der Aktualität und der Vollständigkeit

Wie bereits beim Ortungsverfahren handelt es sich hier unter Berücksichtigung der anfallenden Datenmengen ebenfalls um ein sehr rechen- und damit zeitintensives Ver-

fahren. Daher waren auch hier einige Maßnahmen notwendig, um Rechenzeit zu sparen. Neben der Erweiterung des Arbeitsspeichers und der Datenhaltung in einer SQL-Datenbank wurden noch weitere Optimierungen durchgeführt. Die wesentlichen Maßnahmen neben einer laufzeitoptimierten Programmierung sind in der folgenden Tabelle 11 aufgeführt:

Tabelle 11: Maßnahmen zur Reduktion der Rechenzeit zur Bestimmung von Trajektorien

	Maßnahme	umgesetzt?
1	Ausdünnung der Positionsfolgen	Ja
2	Kachelung des Lokalisierungsnetzes	Ja
3	Enge Fenster zur Suche von Start- und Zielknoten	Ja

Das Ausdünnen der Positionsfolgen aus den **Abis**-Daten vor der weiteren Auswertung (Maßnahme 1 in der Tabelle 11) führt zu einer signifikanten Beschleunigung der Auswertung. Wie bereits in (Czommer 2008) untersucht wurde, ist damit keine nachweisbare Verschlechterung der Trajektorienkorrektheit vom Typ B verbunden. Die Abbildung 10 zeigt den Einfluss des Abtastintervalls auf die Korrektheit Typ B. Es sind für jedes Beispiel die Abtastintervalle 2 Hz, 0.4 Hz und 0.2 Hz dargestellt. Wie aus der Abbildung ersichtlich ist, steigt oder stagniert der korrekt zugeordnete Teil der Trajektorie in 7 Fällen, in 3 Fällen steigt er sehr stark und fällt nur in 3 Fällen. Die Verbesserung ist vermutlich auf eine zufällig richtige Herausfilterung von Ausreißern zurückzuführen, im umgekehrten Fall wurden vermutlich gerade die Ausreißer gewählt und die richtigen Nachbarwerte eliminiert. Aufgrund des begrenzten Datenumfangs ist eine numerische Angabe des Effekts jedoch nicht möglich.

Als Kompromiss wurde schließlich eine Ausdünnung der Positionsfolge zur Herstellung des digitalen Kartenbezugs mit der Verwendung aller Punkte zur Kalmanfilterung kombiniert.

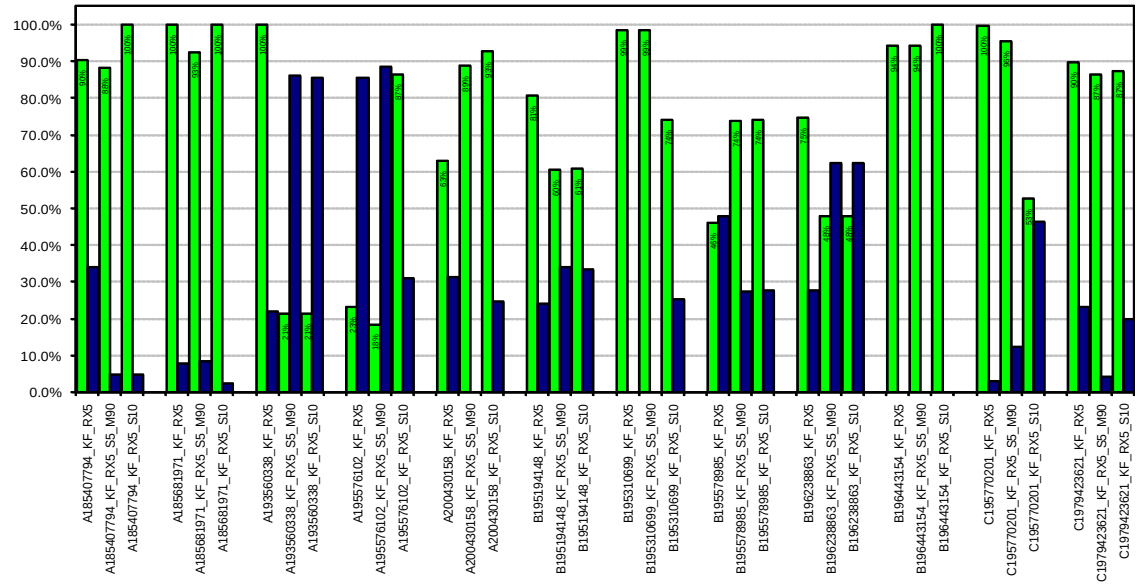


Abbildung 10: Abis-Trajektoriengenerierung; Einfluss des Abtastintervalls auf die Korrektheit Typ B (Czommer 2008)

Die Notwendigkeit der Kachelung (Maßnahme 2 in Tabelle 11) des Lokalisierungsnetzes wurde bereits zu Beginn der Algorithmenentwicklung erkannt und umgesetzt (vgl. Schollmeyer 2008). Dies ermöglicht das sequentielle Einlesen der jeweils zur Auswertung notwendigen kleinen Bereiche des Lokalisierungsnetzes anstelle des speicherintensiven Vorhaltens des gesamten Straßennetzes. Zur Auswertung auf Abis-Ebene wurde das Lokalisierungsnetz auf das die Location Areas umschließende Rechteck begrenzt. Dadurch wurde für den Auswerterechner in Karlsruhe (LACs Ka und Ettlingen) eine durchschnittliche Datenreduktion von 900 MB auf 80 MB bzw. etwa um den Faktor 11 erreicht. Obwohl keine konkreten Zahlen zum Nachweis existieren, kann davon ausgegangen werden, dass die Anzahl berechenbarer Trajektorien pro Stunde durch die Maßnahme ebenfalls um ein Vielfaches erhöht werden konnte.

Im Rahmen der **A-Daten** Prozessierung wurden auf Grund der enormen Datenmengen, die in Echtzeit im gesamten Projektgebiet anfallen, die Kacheln nochmals von 100 m Kantenlänge auf 500 m vergrößert. Dadurch konnte die Rechenzeit auf ein Zehntel verkürzt werden. Da dies jedoch noch immer nicht ausreichend war um mit der im Rahmen des Projektes zur Verfügung stehenden Rechenkapazität die Echtzeitanforderungen zu erfüllen, musste das Projektnetz ausgedünnt werden. Es wurden nur noch Straßen mit einer angegebenen Funktionsklasse von ≤ 4 verwendet.

Um die Anzahl der möglichen Start- und Zielknoten für die **Abis**-Auswertung eines Teilnehmers einzugrenzen, werden zunächst enge Suchfenster verwendet (Maßnahme 3 in Tabelle 11). Finden sich keine möglichen Knoten, so werden die Fenster bis zu einer maximalen Größe schrittweise vergrößert. Wie sich bereits im Rahmen der ersten

Testfahrten nachweisen ließ, wurde damit die Anzahl der möglichen Routen massiv eingegrenzt und damit das Verfahren wesentlich beschleunigt.

Maßnahmen zur Steigerung der Genauigkeit und der Korrektheit

Die Qualität der berechneten Trajektorien hängt elementar von der Qualität der Eingangsdaten ab. Insbesondere eine geeignete Auswahl der Positionsfolgen, für die Routen auf dem Straßennetz ermittelt werden sollen, beeinflusst die Korrektheit und Genauigkeit der Trajektorien entscheidend. Zur Qualitätssicherung werden daher die folgenden Filterkriterien eingesetzt:

A-bis Daten

- Hohe Zuordnungswahrscheinlichkeit der Positionsfolge zu bewegten Verkehrsteilnehmern erforderlich;
- Mindestanzahl an Einzelpositionen eines Teilnehmers: 15;
- Mindestlänge eines Telefonats: 2 km (Luftlinie zwischen erster und letzter Position);
- Mindestdauer des Telefonats: 1 min;
- Maximalgeschwindigkeit um von der ersten zur letzten Position zu gelangen: < 200 km/h.

A-Daten

- Auswertung wird erst nach dem 8. Event (Handover oder Location Area Update) gestartet;
- Löschen von nervösen Rücksprüngen an Zellgrenzen;
- Löschen unbekannter Zellen (z. B. von fremden Netzanbietern);
- Überprüfung auf Plausibilität und ggf. Reduktion des Weg-Zeit-Diagramms durch:
    - Mittelung der Werte bei kurzen Zeitintervallen (≤ 1 min),
    - Elimination von Rückwärtsfahrten, sowie
    - Elimination von Fahrten schneller als das 1.3-fache der maximal vorkommenden Kantengeschwindigkeit.
- Ausgabe nur von Trajektorien, deren Weg-Zeit-Diagramm nach der Plausibilitätsprüfung noch mehr als 2 Werte enthält, und
- Eine Länge von über 5 km aufweist.

Aufgrund der wenigen Zeit, die am Ende des Projektes noch zur Algorithmenentwicklung zur Verfügung stand, wurde die Festlegung der obigen Filter-Parameter bereits

während der Verfahrensentwicklung getroffen und unzureichend dokumentiert. Daher entfällt hier der numerische Nachweis.

Eine Genauigkeitssteigerung der Positionsfolgen und damit auch der Trajektorien aus **Abis**-Daten konnte durch Implementierung eines Kalman-Filters erzielt werden. Dies wurde bereits im Abschnitt 2.2.2 dargestellt.

2.2.4 QS durch Protokollierung der Tagesergebnisse

Die Protokollierung der täglich berechneten FPD-Trajektorien kann einen wichtigen Beitrag zur zukünftigen Sicherung der Datenqualität leisten. Änderungen bzw. Verschlechterungen einzelner Qualitätsparameter werden durch Vergleich verschiedener Tagesergebnisse schnell erkannt und es können entsprechende Maßnahmen eingeleitet werden.

Während der Entwicklungsphase hat die Protokollierung einige wertvolle Hinweise zur Softwareoptimierung geliefert. Im Falle einer zukünftigen Nutzung der FPD als Verkehrsdatenquelle für verschiedene Anwendungen, stellt die Protokollierung der Ergebnisse und insbesondere der ständig ermittelbaren Qualitätskennzahlen im laufenden Prozess das wesentliche Werkzeug zur Qualitätssicherung dar. Da der Schwerpunkt im Projekt Do-iT jedoch auf der Verfahrensentwicklung lag, spielt die Protokollierung im Zusammenhang mit der Qualitätssicherung der Daten nur eine untergeordnete Rolle. Sie wird vielmehr zur Evaluierung des Projekts herangezogen und liefert wertvolle Hinweise zur Eignung der generierbaren Trajektorien aus MFD in den verkehrstechnischen Anwendungen der Projektpartner. Daher wird hier auf eine ausführlichere Darstellung verzichtet und stattdessen auf den Evaluierungsbericht verwiesen (vgl. Do-iT 2009).

Literatur

Baumstark, Eduard (2007): *Qualitätsbeurteilung von Mobilfunkdaten auf Abis-Ebene zur Positionsbestimmung*. Diplomarbeit am IAGB, Universität Stuttgart, unveröffentlicht, 2007.

Beetz, V.; Czommer, R.; Laufer, R.; Schwieger, V. (2009): *Zusammenfassender Bericht zum AP4-6 und AP10-3.* Interner Abschlussbericht im Projekt Do-iT, noch nicht fertiggestellt, 2009.

Czommer, Renate (2008): *Qualitätsbewertung Map-Matching gestützer Positionsbestimmungsverfahren*. Interner Do-iT Bericht zum AP 4-5, ZMS3, 2008.

Do-iT (2009): *Evaluierungsbericht 10.2.* Interner Abschlussbericht des AP10 in Do-iT, 2009, noch nicht fertiggestellt.

DoiT (2008): *Evaluierungskonzept.* Interner Do-iT Bericht zum AP10-1, Bestandteil des ZMS3, 2008.

Karrer, Katrin (2008): *Untersuchung der Qualität von stationären Erfassungssystemen am Beispiel der Autobahn A81.* Studienarbeit am IAGB, Universität Stuttgart, nicht veröffentlicht, 2008.

Laufer, Ralf (2008): *Analyseergebnisse der vorliegenden Informationsstrukturen und -ketten (V1.0)*. Interner Do-iT Bericht zum AP8-4, 2008.

Laufer, Ralf (2007): *Analyseverfahren zur Bewertung der Datenqualität (V1.2)*. Interner Do-iT Bericht zum AP8-3, 2007.

Laufer, Ralf (2006): *Nutzerspezifische Qualitätsanagaben (V1.0)*. Interner Do-iT Bericht zum AP8-1, 2006.

Laufer, R.; Schwartz, S. (2008): *Qualitätssicherungsmaßnahmen (V1.0).* Interner Do-iT Bericht zum AP8-5, Bestandteil des ZMS3, 2008.

Laufer, R., Wiltschko, T. (2006): *Qualitätsmodell für Do-iT (V1.0).* Interner Do-iT Bericht zum AP8-2 sowie zum ZMS1, 2006.

Ramm, K., Schwieger, V. (2008): *Mobile Positioning for Traffic State Acquisition.* Journal of Location Based Services, Taylor & Francis : London, 2008.

Sachse, Thomas (2002): *Die Messwertqualität autarker Detektionssysteme am Beispiel der Traffic Eyes Universal.* Straßenverkehrstechnik, Nr. 9, Seiten 463-470, Kirschbaum-Verlag GmbH : Bonn, 2002.

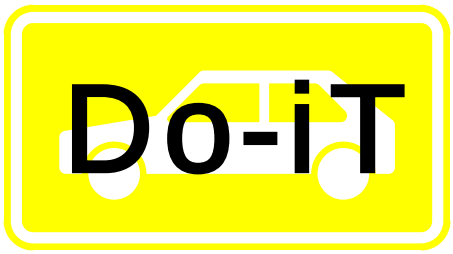
1. Schollmeyer, Ralf (2008): *Identifizierung aktiver Verkehrsteilnehmer.* Interner Do-iT Bericht zum AP3 sowie zum ZMS3, 2008.

Schwieger, V.; Ramm, K.; Czommer, R.; Möhlenbrink, W. (2007): *Mobile Phone Positioning for Traffic State Acquisition.* Journal of Applied Geodesy, volume 1, issue 3, Walter de Gruyter : Berlin, New York, 2007.

Evaluierungsbericht (MS10-2)

Datenoptimierung für integrierte Telematik

Do-iT



Förderkennzeichen: 19 B 5007

Ergebnisse der Evaluierung

August 2009

DDG Gesellschaft für Verkehrsdaten mbH

Landeshauptstadt Stuttgart

Stadt Karlsruhe

Innenministerium Baden-Württemberg

Universität Stuttgart, Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik

Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen

Dokument-Informationen

Version	Datum	Bearbeiter	Bemerkung
0.1	**08.04.09**	**Laufer**	• Erstellung der Vorlage
0.2	**15.07.09**	**Laufer**	• Einfügen der Beiträge der Partner
0.3	**24.07.09**	**Laufer**	• Einfügen des Beitrags des IAGB
1.0	**31.08.09**	**IAGB/Laufer**	• Internes Review, Überarbeitung des Berichts

Inhalt

1 Einleitung 4

2 Evaluierung der Generierung von FPD-Trajektorien 5
2.1 Evaluierung der Positionsbestimmung durch höherwertige Ortungsverfahren 5
2.2 Protokollierung der Tagesauswertungen 6
2.3 Vergleich mit stationären Erfassungssystemen 27
2.3.1 Vergleich der FPD-Trajektorien aus Abis-Daten mit SES 39
2.3.2 Vergleich der FPD-Trajektorien aus A-Daten mit SES 45

3 Evaluierung der Verwendbarkeit von FPD in verkehrstechnischen Anwendungen 60
3.1 FPD zur Generierung der Verkehrslage 60
3.2 Demonstrator der Stadt Karlsruhe 60
3.2.1 Beschreibung des Demonstrators 60
3.2.2 Analyse der durchgeführten Feldversuche 62
3.2.3 Bewertung der Ergebnisse 73
3.3 Demonstrator der Landeshauptstadt Stuttgart 74
3.4 Demonstrator des IM BW, Landesstelle für Straßentechnik 74
3.4.1 Beschreibung des Demonstrators 75
3.4.2 Analyse des durchgeführten Feldversuchs 76
3.4.3 Bewertung der Ergebnisse 77
3.5 Evaluierung der Bestimmung von QZ-Matrizen und Tagesprognosen aus FPD 78
3.5.1 Beschreibung des Demonstrators 78
3.5.2 Analyse der durchgeführten Feldversuche 80
3.5.3 Bewertung der Ergebnisse 85

4 Zusammenfassende Bewertung 87

5 Literatur 88

Anhang A 89

1 Einleitung

Der vorliegende Bericht ist eine Zusammenstellung sämtlicher Evaluierungsergebnisse aller Projektpartner in Do-iT. Dabei werden im Kapitel 2 zunächst die implementierten Methoden zur Generierung von Trajektorien aus Mobilfunkdaten als Grundlage aller weiteren Anwendungen in den Demonstratoren evaluiert. Im anschließenden Kapitel 3 erfolgt dann die Untersuchung der Verwendbarkeit von FPD-Trajektorien in diversen verkehrstechnischen Anwendungen der Projektpartner.

Für die Evaluierung wurden Demonstratoren definiert, deren Laufzeit aufgrund der bekannten Projektverzögerungen und trotz der gewährten Verlängerung allerdings von ursprünglich 8 Monaten auf 3 Monate (Abis-Daten) bzw. 2 Monate (A-Daten) verkürzt werden mussten. Betrachtet man nur die Zeiträume, in denen die zweite, optimierte Version der Verarbeitungssoftware online im Einsatz war, so beschränkt sich der Zeitraum für die Evaluierung der A-Daten sogar auf 10 Tage Ende März (letztes Update am 20.03.09 um ca. 16:00 Uhr), und für die Abis-Daten auf ca. 50 Tage (letztes Update am 10.02.09). Dies ist darin begründet, dass die Aufnahme des Online-Betriebs immer mit einer zwingend notwendigen ersten Testphase verbunden ist, in der die ersten wichtigen Erfahrungen in der Praxis gesammelt und in der zweiten Softwareversion umgesetzt werden. Diese Phase zählt prinzipiell noch nicht zur eigentlichen Evaluierungsphase, die sich in der Regel anschließt. Aufgrund des von Seiten der Projektpartner unverschuldeten Zeitproblems im Projekt, wurde hier jedoch jeweils schon die erste online lauffähige Version der Software zur Generierung von Trajektorien aus Abis- und A-Daten zur Evaluierung zur Verfügung gestellt. Die Trajektorien wurden auf dem FPD-Server gespeichert und waren dort für den Rest der Projektlaufzeit für alle Partner frei verfügbar.

Es fehlen derzeit noch die Beiträge der Projektpartner DDG (Kapitel 3.1) sowie der Stadt Stuttgart (Kapitel 3.3), deren Berichte bis zur Fertigstellung des Berichts am 30.09.09 noch nicht vorlagen.

2 Evaluierung der Generierung von FPD-Trajektorien

Bevor die Evaluierung der Verwendbarkeit des Mobilfunks als Verkehrsdatenquelle erfolgen kann, erfolgt die Evaluierung der aus Mobilfunkdaten erzeugbaren Trajektorien. Dazu wurden zur Schaffung einer Referenz einerseits Testfahrten mit GPS durchgeführt und andererseits die Daten von stationären Verkehrsdatenerfassungssystemen herangezogen. Zusätzlich wurden die berechneten Ergebnisse tageweise protokolliert.

2.1 Evaluierung der Positionsbestimmung durch höherwertige Ortungsverfahren

Zur objektiven Beurteilung der aus realen Mobilfunkdaten bestimmten Trajektorien wurden für Abis-Daten im Stadtgebiet Karlsruhe und mit A-Daten im gesamten Projektgebiet Testfahrten mit GPS durchgeführt. Die GPS-Messungen liefern dabei eine Referenz für Zeit und Position der Mobilstation, die sich aufgrund ihrer höheren Genauigkeit dazu eignet, die Qualität der FPD-Trajektorien absolut zu bewerten. Dazu wurden im Qualitätsmodell zur Beurteilung der FPD-Routen die Parameter **Zuordnungskorrektheit Typ A** und **Zuordnungskorrektheit Typ B** und zur Beurteilung des Weg-Zeit-Diagramms die **Standardabweichung der Geschwindigkeit** und die vorzeichenfreie **Differenz der Durchschnittsgeschwindigkeit** definiert (vgl. Kapitel 3 in [AP4-6 / AP10-3, 2009]).

Zur Evaluierung mit GPS wurden die in der folgenden Tabelle 1 dargestellten Qualitätsparameter herangezogen. Die Nummern der Parameter stammen aus dem für Mobilfunktrajektorien definierten Qualitätsmodell wie im Kapitel 3 in [AP4-6 / AP10-3, 2009] dargestellt ist.

Tabelle 1: Zur Evaluierung mit GPS herangezogene Qualitätsparameter für FPD-Trajektorien (vgl. [AP4-6 / AP10-3, 2009])

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
P3.1 Abis	Vollständigkeit VO	Trajektorienvollständigkeit (online+relativ)	VO_{FPD} [%]	**Abis:** Anteil der verwendeten Positionen einer Positionsfolge eines Teilnehmers, der für die Trajektorie verwendet wird *FPD_Quality (2)*
P5.5	Korrektheit	Zuordnungskorrektheit Typ A (offline+absolut)	KR_{ZuA} [%]	Korrekter Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, bezogen auf die Länge der GPS-Route (GPS als Referenz erforderlich)
P5.6		Zuordnungskorrektheit Typ B (offline+absolut)	KR_{ZuB} [%]	Korrekter Streckenanteil der FPD-Route, der sich mit der GPS-Route deckt, bezogen auf die Länge der ermittelten FPD-Route (GPS erforderlich)

P6.1 ABis	Genauigkeit GE	Mittlere Querabweichung (online+relativ)	QA_{FPD} [m]	Mittel der orthogonalen Abweichungen der verwendeten Positionen von der wahrscheinlichsten Route *FPD_Quality (1*) **nur Abis-Daten**
P6.2		Standardabweichung der Geschwindigkeit (offline+absolut)	S_V	Gibt die Genauigkeit der aus den Trajektorien ermittelbaren Geschwindigkeit an. Referenz erforderlich (Messfahrt oder TEU)
P6.3		Differenz der Durchschnittsgeschwindigkeiten (offline+absolut)	δ [km/h]	Vorzeichenfreie Differenz zwischen Durchschnittsgeschwindigkeit aus GPS und aus MFD

Die Untersuchungen zeigen für Abis-Daten eine starke Abhängigkeit der Korrektheit der Routen von der Länge des zurückgelegten Wegs des Teilnehmers. Erst ab einer Weglänge von über 3 km zeigen die untersuchten Trajektorien in über 80 % der Fälle weniger als 10 % falsche Informationen. In der Praxis wurde jedoch aufgrund der höheren Performance sowie der höheren Abdeckungsrate auf Kosten der Genauigkeit als Kompromiss eine Mindestlänge von 2 km gewählt, der jedoch durch die gleichzeitigen Bedingungen

- mehr als 15 Einzelpositionen
- mindestens 60 s Gesprächsdauer sowie
- berechnete Geschwindigkeit < 200 km/h

verschärft wurde.

Eine realitätsnahe Durchschnittsgeschwindigkeit lässt sich bei längeren Trajektorien ab 8 km gut aus der Weg-Zeit-Information ableiten.

Die Evaluierung der Trajektorien aus A-Daten zeigt gute Ergebnisse bei durchgängigen Telefonaten. Die Zuordnungskorrektheit vom Typ B liegt bei 73 % und fällt bei mehrfach unterbrochenen Telefonaten auf ca. 67 % ab. Bleibt das Mobiltelefon passiv, d. h. es wird kein Telefonat geführt, dann sinkt die Zuordnungskorrektheit stark ab auf ca. 14 %. Eine zuverlässige FPD-Generierung ist dann nicht mehr möglich. Die Durchschnittsgeschwindigkeit kann mit einer Standardabweichung von ca. 20 km/h jedoch auch aus passiven Telefonen mit ähnlicher Genauigkeit bestimmt werden, wie es bei telefonierenden Teilnehmern möglich ist.

Die detaillierte Auswertung der Evaluierung von Trajektorien aus Abis- und A-Daten mit Hilfe von GPS-Testfahrten ist im Bericht [AP4-6 / AP10-3, 2009] zu finden.

2.2 Protokollierung der Tagesauswertungen

Als elementarer Bestandteil der Qualitätssicherung während des online-Betriebs wurde eine einfache statistische Auswertung der protokollierten Tagesauswertungen der FPD-Generierung umgesetzt. Diese dient grundsätzlich der Überwachung der Auswer-

tung über längere Zeiträume und dem schnellen Erkennen von Fehlern und Problemen bei der Auswertung. Aus Zeitgründen wurde diese Protokollierung nur im post processing umgesetzt und diente in erster Linie der Evaluierung der Projektergebnisse. Grundsätzlich ist eine Übertragung auf einen möglichen Dauerbetrieb der FPD-Generierung jedoch problemlos möglich.

Die Protokollierung wurde mit Hilfe von Datenbankabfragen mit MySQL realisiert. Die visuelle Darstellung mit Schaubildern wurde anschließend in Matlab implementiert.

Es werden im Einzelnen die folgenden, in Tabelle 2 dargestellten, Berechnungsergebnisse tages- und stundenweise protokolliert. Bei der mittlere Querabweichung sowie die mittlere Trajektorienvollständigkeit der Abis-Trajektorien handelt es sich um Qualitätsparameter des Qualitätsmodells für FPD-Trajektorien (vgl. [AP4-6 / AP10-3, 2009] oder [AP8-7, 2009]).

Tabelle 2: Protokollierung der Tagesauswertungen

Information	Zweck
Anzahl ausgewerteter Teilnehmer (je h und d)	Beurteilung des Gesamtaufkommens, Hinweise auf Durchdringung mit FPD; Überwachung der Auswertequalität
Gesprächsdauer in [min] (Abis) Dauer der Ereigniskette je Teilnehmer (A)	Beurteilung des Gesamtaufkommens, Hinweis auf die Qualität der FPD; Überwachung der Auswertequalität
Trajektorienlänge in [km]	Beurteilung des Gesamtaufkommens und der Qualität der FPD; Beurteilung der Auwertequalität
Mittlere Querabweichung in [m] (Abis) *Parameter P6.1*	Beurteilung der FPD-Trajektorienqualität; Beurteilung der Auswertequalität
Mittlere Trajektorienvollständigkeit in [%] (Abis) *Parameter P3.1*	Beurteilung der FPD-Trajektorienqualität; Beurteilung der Auswertequalität
Häufigkeit der auftretenden RE_IDs in den FPD-Trajektorien	Beurteilung des Gesamtaufkommens und der Durchdringung des PN mit FPD; Überwachung der Auswertequalität

Aus den protokollierten Daten kann unter anderem der Vollständigkeitsparameter Abdeckungsrate ermittelt werden. Das Verhältnis der in den Trajektorien vorkommenden RE_IDs zu der gesamten Anzahl RE_IDs ist ein erster Hinweis auf die Abdeckungsrate. Um besser und schlechter abdeckbare Gebiete bzw. Streckenabschnitte zu erkennen, muss zusätzlich ein Vergleich mit den SES angestrebt werden. Dieser Vergleich wurde im nachfolgenden Kapitel 2.3 durchgeführt.

© Do-iT 2009

Im Folgenden werden die protokollierten Auswertungen von A- und Abis-Daten im Untersuchungszeitraum vom 23.-31. März 2009 weitgehend vollständig dargestellt und diskutiert.

Anzahl der auswertbaren Teilnehmer

Im jeweils linken Bild ist die Anzahl ausgewerteten Teilnehmer aus A-Daten für das gesamte Projektnetz in Stundenintervallen dargestellt. Das rechte Bild zeigt diejenigen aus Abis-Daten ausgewerteten Teilnehmer in Karlsruhe, wo Abis-Daten verfügbar waren.

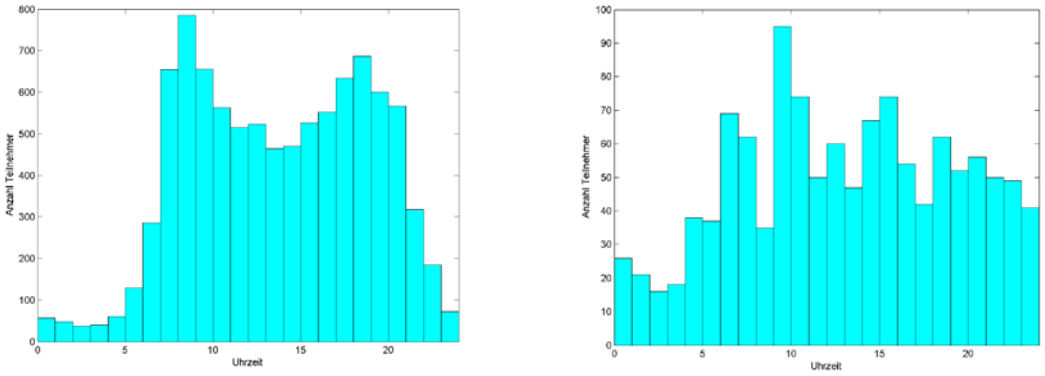


Abbildung 1: Anzahl auswertbarer Teilnehmer am Montag, 23.03.09 (A und Abis)

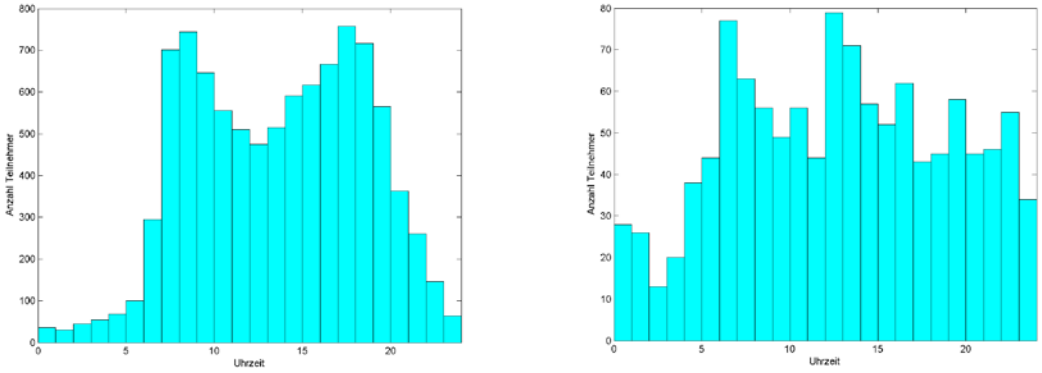


Abbildung 2: Anzahl auswertbarer Teilnehmer am Dienstag, 24.03.09 (A und Abis)

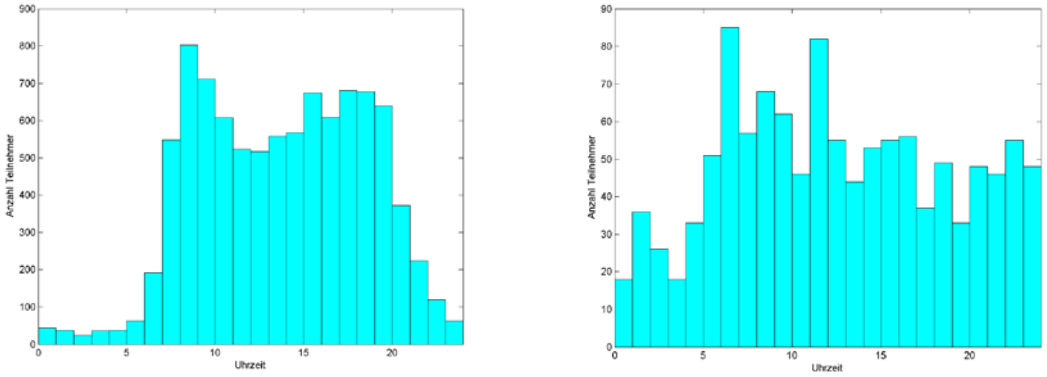


Abbildung 3: Anzahl auswertbarer Teilnehmer am Mittwoch, 25.03.09 (A und Abis)

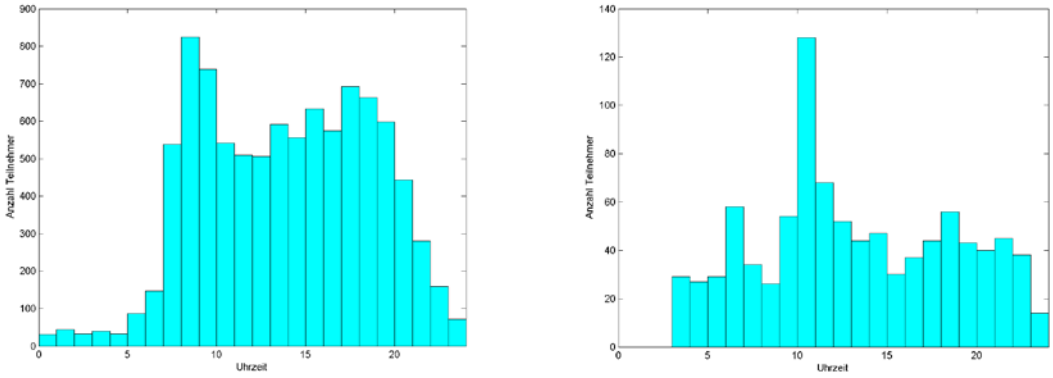


Abbildung 4: Anzahl auswertbarer Teilnehmer am Donnerstag, 26.03.09 (A und Abis)

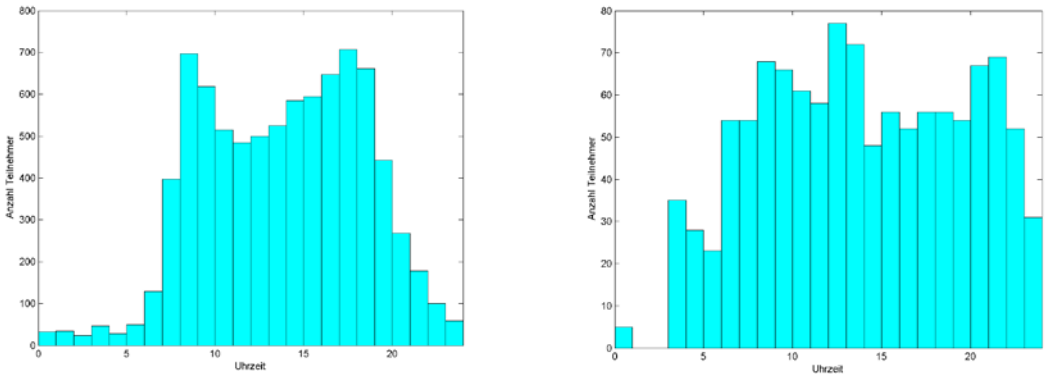


Abbildung 5: Anzahl auswertbarer Teilnehmer am Freitag, 27.03.09 (A und Abis)

Keine A-Daten verfügbar[1]

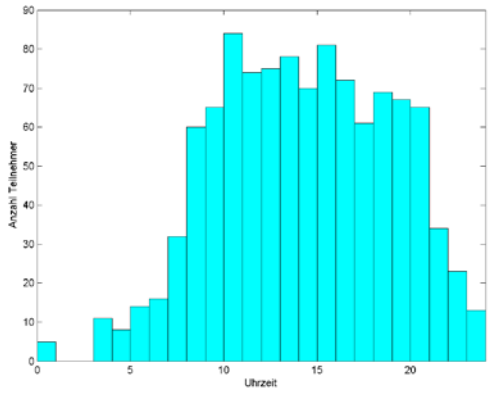


Abbildung 6: Anzahl auswertbarer Teilnehmer am Samstag, 28.03.09 (es waren nur Abis-Daten verfügbar)

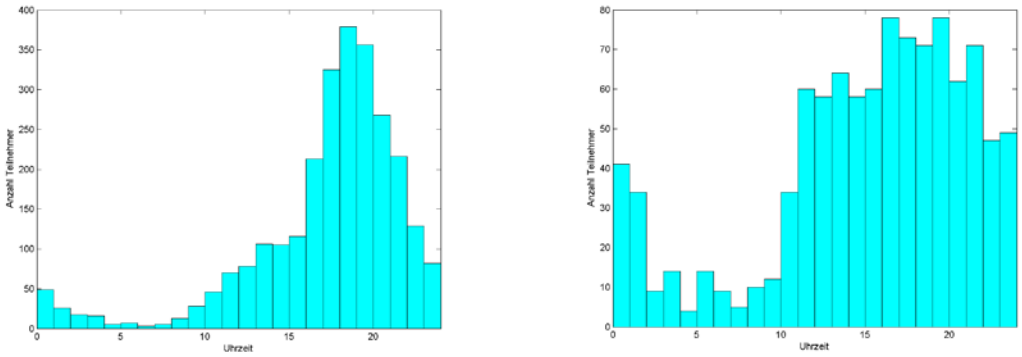


Abbildung 7: Anzahl auswertbarer Teilnehmer am Sonntag, 29.03.09 (A und Abis)

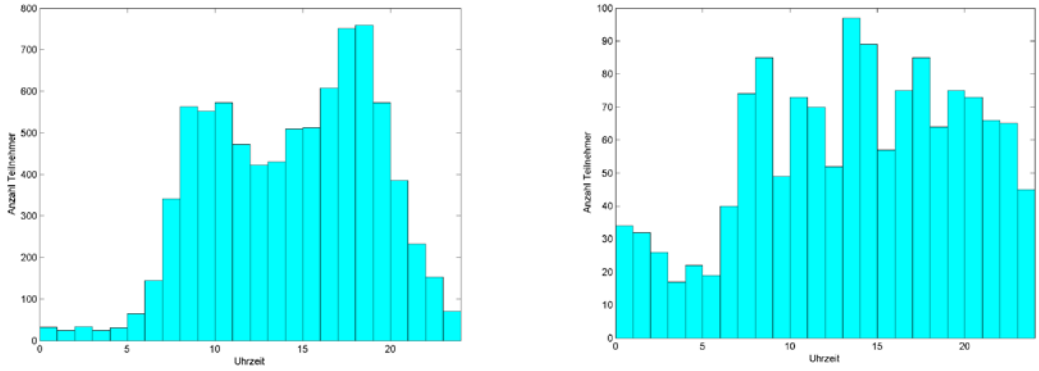


Abbildung 8: Anzahl auswertbarer Teilnehmer am Montag, 30.03.09 (es waren nur A-Daten verfügbar)

[1] Der Grund hierfür konnte leider nicht ermittelt werden.

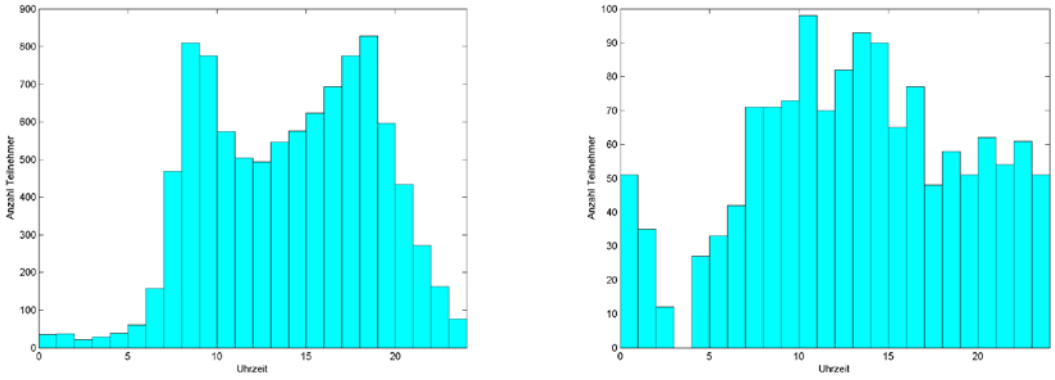


Abbildung 9: Anzahl auswertbarer Teilnehmer am Dienstag, 31.03.09 (A und Abis)

Das Aufkommen an Teilnehmern mit auswertbaren A-Daten (jeweils linke Darstellung in den Abbildung 1 bis Abbildung 9 spiegelt klar das tägliche reale Verkehrsaufkommen wieder. Die Wochentage zeigen die charakteristische Morgenspitze zwischen 7 und 10 Uhr sowie das etwas flachere, dafür breitere Feierabendmaximum zwischen 15 und 20 Uhr, wie sie auch die SES zeigen. Der Sonntag mit dem typischen langsamen Anstieg am Vormittag und dem Maximum zwischen 15 und 20 Uhr wird ebenfalls durch die ausgewerteten Teilnehmer der A-Daten-Auswertung realistisch wiedergegeben. Dies ist bereits ein erstes Indiz für die Tauglichkeit der FPD-Trajektorien aus A-Daten zur Feststellung von Änderungen in der Verkehrsstärke und damit auch zur Ableitung der Verkehrslage.

Aus der Anzahl ausgewerteter Teilnehmer aus Abis-Daten lässt sich offensichtlich nicht direkt die Verkehrsstärkeänderung ablesen. Dies liegt an der Mindestlänge und Mindestdauer, die notwendig sind, damit ein Gespräch überhaupt erst der Auswertung zugeführt wird. Das Telefonierverhalten der Straßenverkehrsteilnehmer ist jedoch unterschiedlich über den Tag hinweg. Morgens werden wenige lange Gespräche geführt, am späteren Nachmittag und Abend hingegen mehr. Nachts hingegen telefonieren relativ wenige Teilnehmer jedoch offensichtlich lange, daher ist hier der nächtliche Einbruch nicht so ausgeprägt wie bei A-Daten (vgl. Abbildung 1 bis Abbildung 4). Donnerstag bis Samstag gab es offensichtlich Probleme mit der Auswertung in den frühen Morgenstunden (vgl. Abbildung 4 bis Abbildung 6). Der Sonntag wiederum zeigt ein realistisches Verhalten mit wenigen Telefonaten in den Morgenstunden und einem sehr starken Anstieg gegen 10 Uhr (vgl. Abbildung 7).

Die absolute Anzahl ausgewerteter Teilnehmer liegt an den Wochentagen im Schnitt bei ca. 8000 bis 9000 pro Tag im gesamten Projektnetz (A-Daten) bzw. bei ca. 1000 bis 1300 pro Tag in Karlsruhe mit Abis-Daten (vgl. Tabelle 3). Sehr auffällig ist hierbei der fehlende Einbruch an den Tagen Samstag und Sonntag, an denen die Verkehrsstärke im Allgemeinen stark abnimmt (es lag hier kein Ferienbeginn vor). Die A-Daten zeigen am Sonntag einen Rückgang auf weniger als ein Drittel des Niveaus der Wochentage, für Samstag liegen leider keine Angaben vor.

Tabelle 3: Anzahl ausgewerteter Teilnehmer aus A und Abis für verschiedene Tage

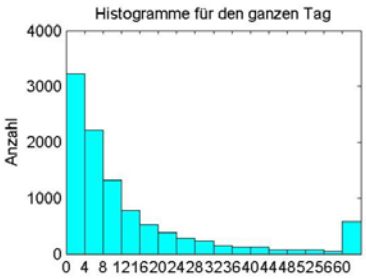
Tag	Mo 23.	Di 24.	Mi 25.	Do 26.	Fr 27.	Sa 28.	So 29.	Mo 30.	Di 31.
A	9434	9516	9322	9333	8320	k. A.	2658	8266	9593
Abis	1195	1161	1161	943	1142	1077	1015	1384	1375

Dauer der Ereigniskette (A) / Gesprächsdauer (Abis)

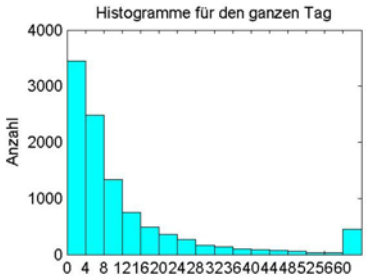
Die Darstellung der protokollierten Dauer der ausgewerteten Teilnehmer - im Falle von A-Daten handelt es sich um Ketten von verknüpfbaren Ereignissen eines Teilnehmers im Falle der Abis-Daten sind dies i. d. R. Gespräche – gibt Aufschluss über die Struktur der ausgewerteten Daten und erleichtert die Detektion von Veränderungen und/oder Problemen.

Grundsätzlich unterscheiden sich die Verteilungen der Längen von Ereignisketten in den 8 Histogrammen in Abbildung 10 nur unwesentlich, dies deutet auf eine gleichmäßige Auswertequalität über den Untersuchungszeitraum hin. Auch der Sonntag zeigt einen ähnlichen Verlauf auch wenn insgesamt nur etwa ein knappes Drittel der Trajektorien an Wochentagen berechnet wurde (dieses Verhältnis konnte bereits in den Abbildung 1 bis Abbildung 9 anhand der Gesamtzahl ausgewerteter Teilnehmer je Tag gezeigt werden). Die Häufung bei mehr als 60 Ereignissen je Teilnehmer bleibt jedoch am Sonntag aus. Dies könnte darauf hindeuten, dass es sich bei diesen aufgezeichneten Langzeitteilnehmern z. B. um GSM-Modems in Navigationsempfängern oder Toll-Collect-Abrechnungseinheiten in Lastkraftwagen und im Lieferverkehr handeln könnte. Da diese geometrischen Informationsketten sehr lange sind, eignen sie sich hervorragen für die FPD-Trajektoriengenerierung aus Mobilfunkdaten. Diese Trajektorien könnten daher auch gut zur Detektion überregionaler Verkehrsereignisse sowie zur Detektion des Routenwahlverhaltens (insbesondere des Warentransportverkehrs) geeignet sein.

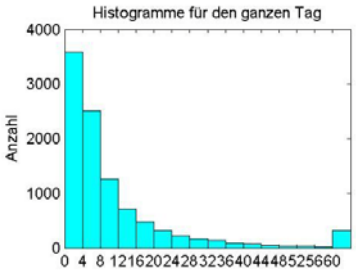
Der Großteil der auswertbaren Ereignisketten liegt jedoch im kurzen bis sehr kurzen Bereich von bis 4 bis 8 Ereignissen. 4 Ereignisse stellen hierbei die Mindestgrenze dar, ab wann ein Teilnehmer ausgewertet wird. Diese kurzen Ereignisketten können aufgrund der relativ schlechten geometrischen Einzelinformationen (LAC-Wechsel oder Zellwechsel erlauben nur die Ortung im Umkreis von bis zu mehreren Kilometern) in Ballungsräumen nur schwer eindeutig lokalisiert und zu FPD-Trajektorien verwandelt werden. Sehr gut funktionieren kann dies jedoch in dünn besiedelten und daher auch mit weitmaschiger Infrastruktur durchzogenen ländlichen Gegenden.



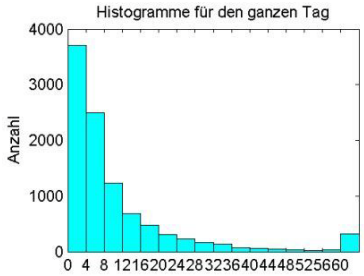
Montag, 23.03.09



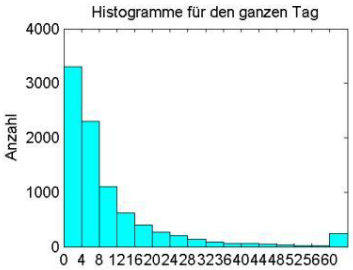
Dienstag, 24.03.09



Mittwoch, 25.03.09



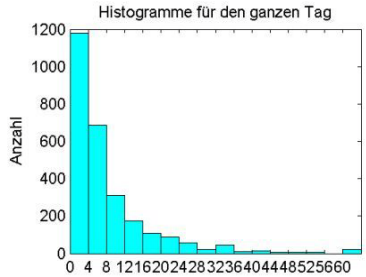
Donnerstag, 26.03.09



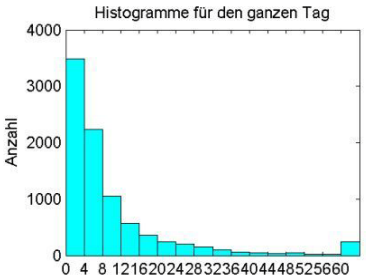
Freitag, 27.03.09

Keine A-Daten

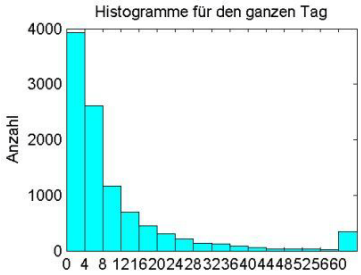
Samstag, 28.03.09



Sonntag, 29.03.09



Montag, 30.03.09



Dienstag, 31.03.09

Abbildung 10: Übersicht über die Anzahl zusammenhängender Ereignisse eines Teilnehmers im Untersuchungszeitraum 23.-31.03.09 aus A-Daten (am 28. konnten keine A-Daten prozessiert werden)

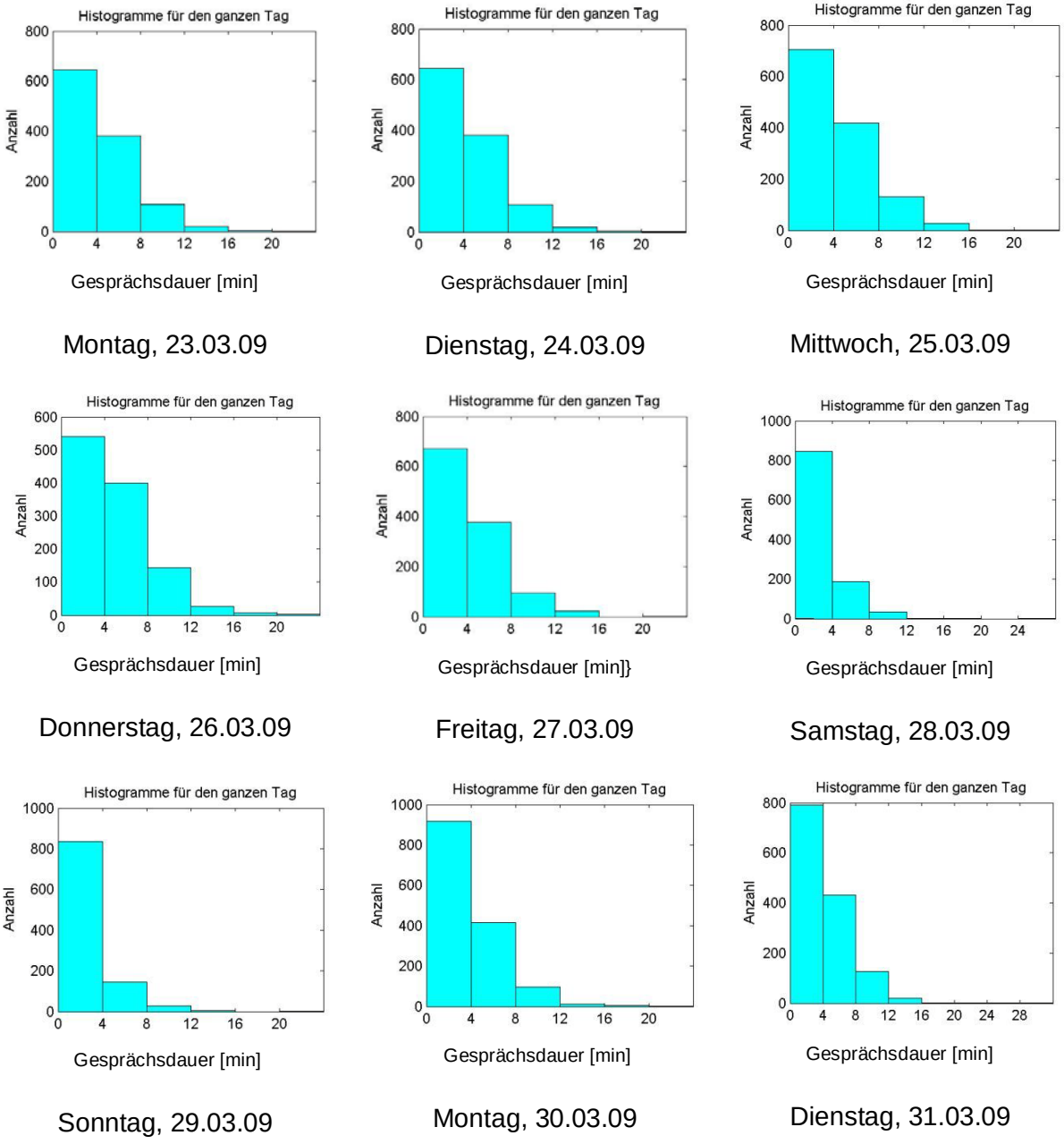


Abbildung 11: Übersicht über die Gesprächsdauer der ausgewerteten Teilnehmer im Untersuchungszeitraum 23.-31.03.09 aus Abis-Daten in Ka

Die Histogramme der Gesprächslängen für die ausgewerteten Abis-Daten zeigen im Untersuchungszeitraum relativ gesehen ein ähnlich unauffälliges Verhalten (vgl. Abbildung 11). Daraus kann auch hier eine Auswertung in vergleichbarer Qualität geschlossen werden. Auffällig ist, dass am Samstag und Sonntag weniger lange mobil im KFZ telefoniert wurde. Die Anzahl Gespräche bis zu 4 Minuten bleibt mit ca. 800 ähnlich wie zu Wochenbeginn. Dies scheint realistisch, da es sich bei längeren, von Mobiltelefonen in Fahrzeugen aus, geführte Gespräche häufig um geschäftliche Gespräche handelt die naturgemäß vorwiegend an Wochentagen stattfinden.

Länge der FPD-Trajektorien (A und Abis)

Die Länge der berechenbaren FPD-Trajektorie ist ein Hinweis auf deren Qualität. Je länger ein Teilnehmer verfolgt werden kann, desto besser lässt sich seine Fahrtroute rekonstruieren. Des Weiteren kann hier unter Umständen eine Veränderung in der Auswertequalität festgestellt werden, daher ist die ständige Protokollierung im Echtzeit-Betrieb sinnvoll.

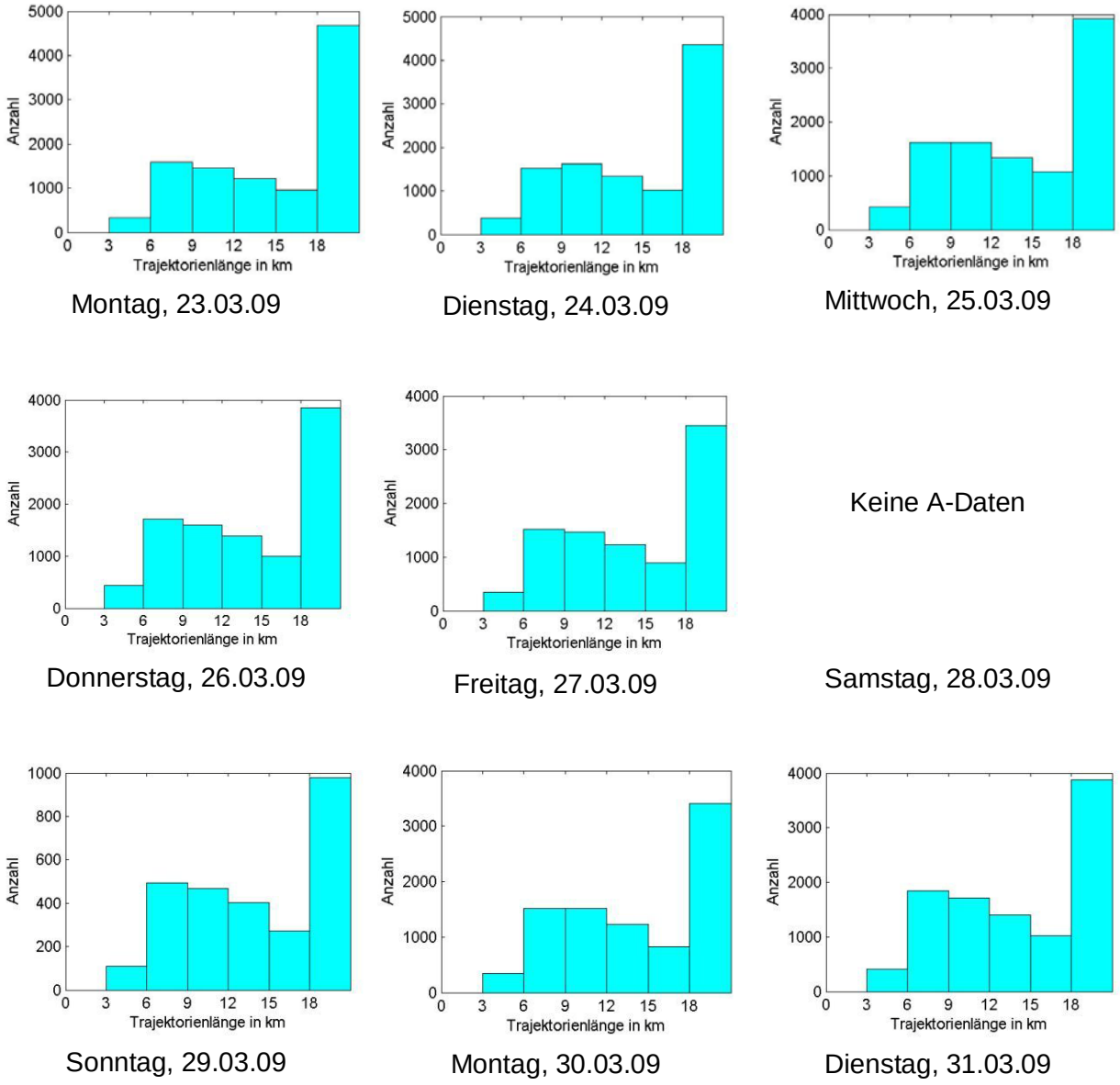


Abbildung 12: Übersicht über die Trajektorienlängen der ausgewerteten Teilnehmer im Untersuchungszeitraum 23.-31.03.09 aus A-Daten (am 28. standen keine A-Daten zur Verfügung)

Auch die in Abbildung 12 dargestellten Längen der aus A-Daten generierten Trajektorien zeigen keine Auffälligkeiten im Hinblick auf die Auswertequalität. Die Mindestlänge

von 3-6 km ergibt sich aus der minimalen Anzahl von 4 Ereignissen, ab der eine Auswertung stattfindet. Neben dem Bereich zwischen 6 und 15 km ist ein Großteil der Trajektorien länger als 18 km. Das in den Histogrammen für die Anzahl der Ereignisse in Abbildung 10 erkennbare ausbleibende Peak für sehr lange Ereignisketten geht hier in der großen Anzahl an Trajektorien über 18 km unter, daher zeigt sich für den Sonntag im Vergleich zu den Wochentagen ein nahezu identisches Häufungsbild. Durch eine bessere Aufschlüsselung der langen Trajektorien könnte die Beurteilung nochmals verfeinert werden, dies war aus Zeitmangel leider nicht mehr möglich.

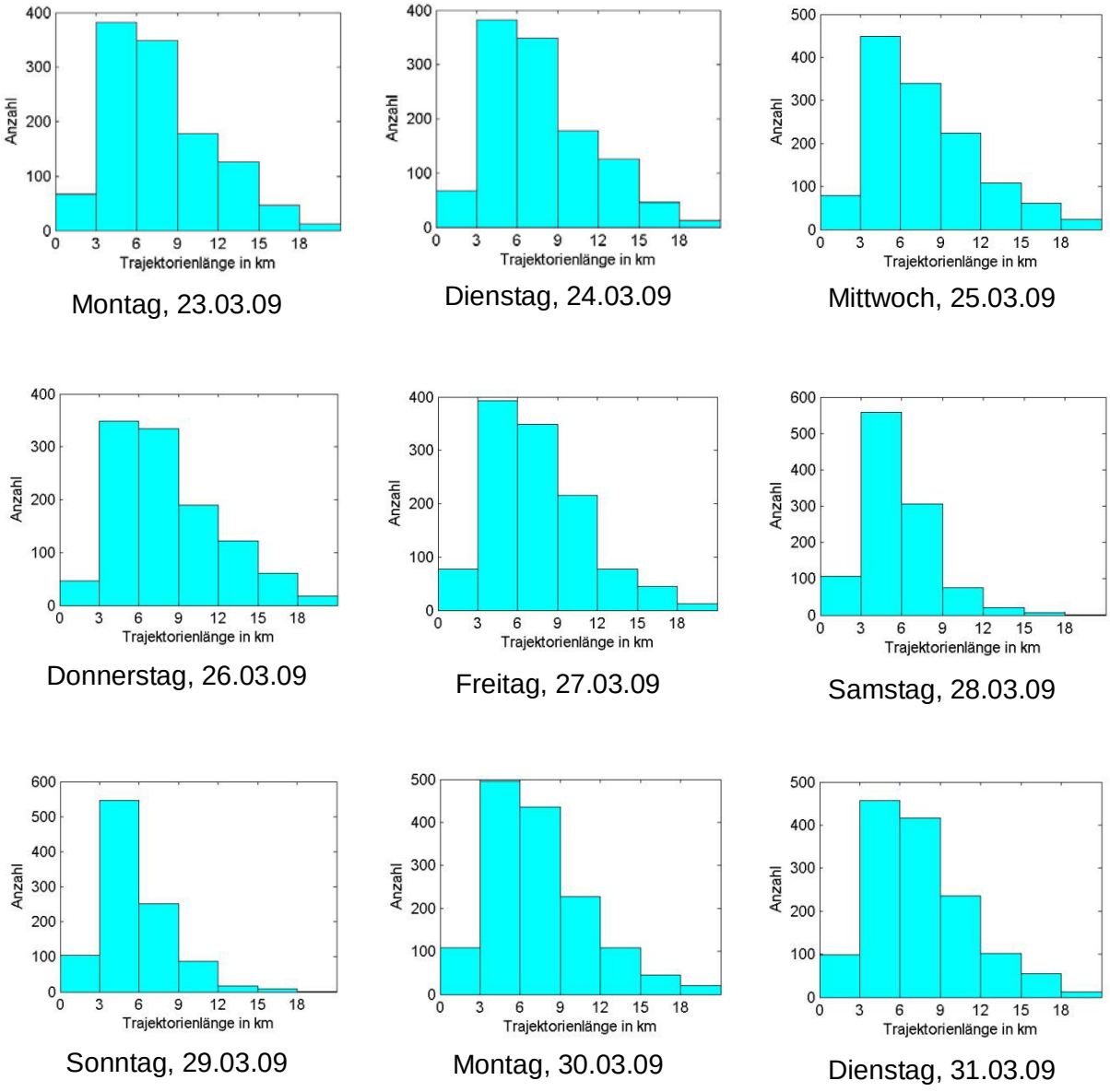


Abbildung 13: Übersicht über die Trajektorienlängen der ausgewerteten Teilnehmer im Untersuchungszeitraum 23.-31.03.09 aus Abis-Daten in Ka

Die Trajektorienlängen der ausgewerteten Abis-Daten im Raum Karlsruhe zeigen eine völlig andere Verteilung als die Trajektorien aus A-Daten (vgl. Abbildung 13 und Abbildung 12). Dies liegt insbesondere an dem relativ kleinen Gebiet, in dem Abis-Daten verfügbar waren. Die maximale Ausdehnung der LACs Ettlingen und Karlsruhe beträgt weniger 20x20 km², daher waren Trajektorien von über 15 km kaum möglich.

Am Wochenende ist eine Verschiebung hin zu kürzeren Trajektorien sichtbar, dies ist vermutlich wieder durch die fehlenden geschäftlichen Gespräche in den Fahrzeugen an den Wochentagen zu erklären.

Mittlere Querabweichung der FPD-Trajektorien (nur Abis)

Zur Beurteilung der Qualität der gefundenen Trajektorie im digitalen Straßennetz wurde der Qualitätsparameter *mittlere Querabweichung* entwickelt. Dieser ist ein metrisches Maß für den mittleren orthogonalen Abstand zu der gefundenen Solltrajektorie auf der digitalen Karte (vgl. [AP10-1 2008]).

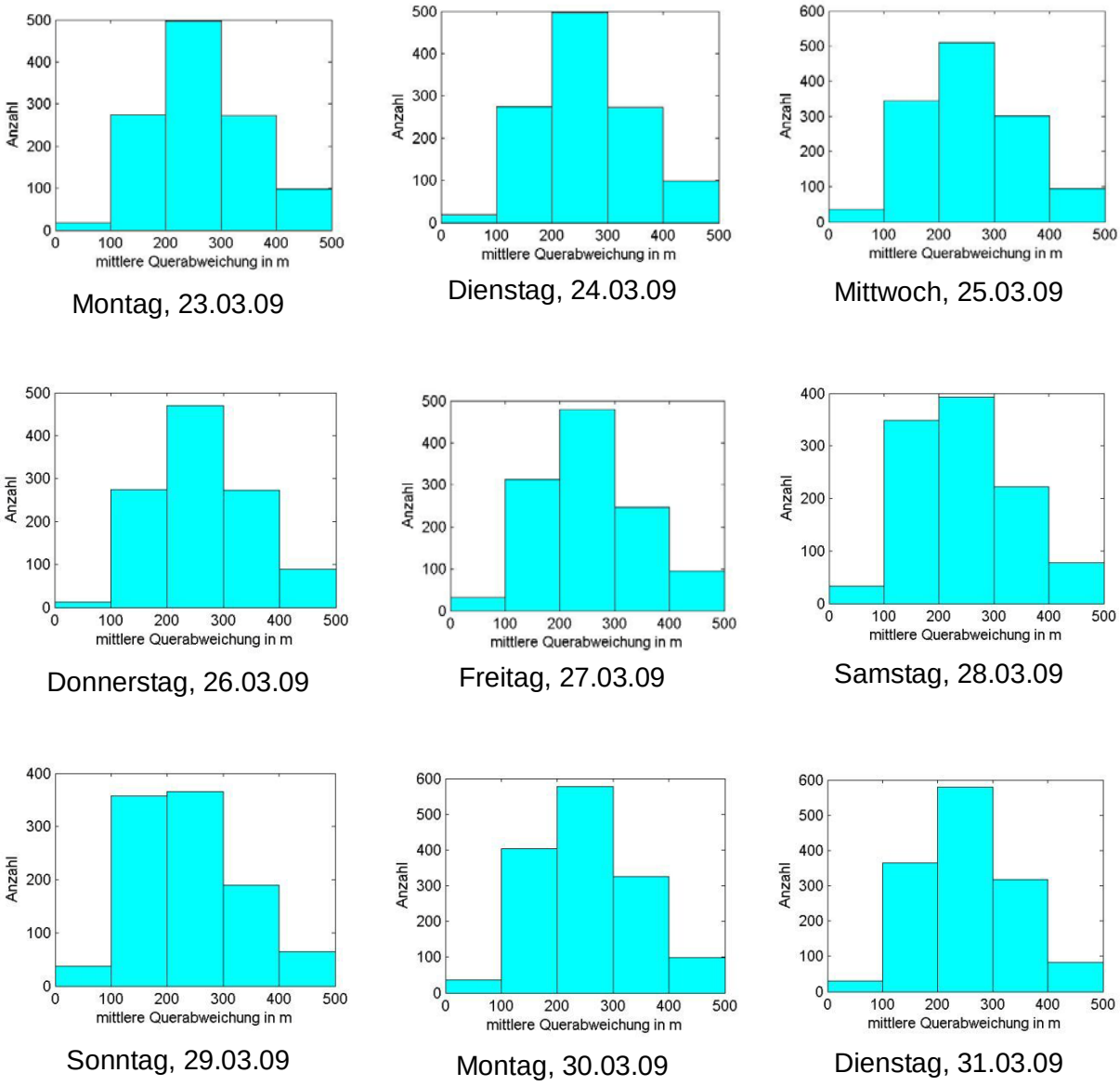


Abbildung 14: Übersicht über die mittlere Querabweichung der ausgewerteten Teilnehmer im Untersuchungszeitraum 23.-31.03.09 (Abis)

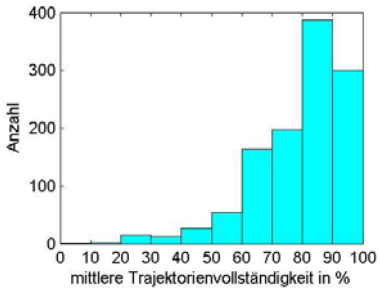
Die Querabweichungen an den Wochentagen zeigen ein sehr homogenes Bild mit einer Häufung bei einer Querabweichung von 200-300 m (vgl. Abbildung 14). Auch hier ist eine konstante Auswertequalität ableitbar. Auffällig ist jeweils die Verschiebung der Spitze hin zu kleineren Querabweichungen am Wochenende (Samstag und Sonntag). Die Ursache dafür konnte bislang jedoch noch nicht ermittelt werden.

Der Parameter Querabweichung in Verbindung mit der Trajektorienlänge erlaubt eine relativ gute Einschätzung der Qualität einzelner berechneter FPD-Trajektorien. Detaillierte Untersuchungen hierzu finden sich in [AP 4-5 2008].

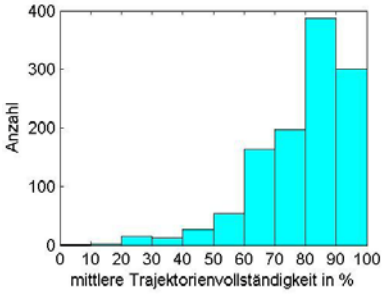
Mittlere Trajektorienvollständigkeit der FPD-Trajektorien (nur Abis)

Die mittlere Trajektorienvollständigkeit gibt an, welcher Anteil der intern berechneten Positionsfolgen eines Teilnehmers im Anschluss für die damit ermittelte Trajektorie verwendet wurde. Bei der Berechnung von Trajektorien aus A-Daten entfällt der Zwischenschritt der Koordinatenberechnung (jeder Funkzelle sind direkt Koordinaten zugeordnet), daher ist dieser Parameter für A-Daten nicht bestimmbar.

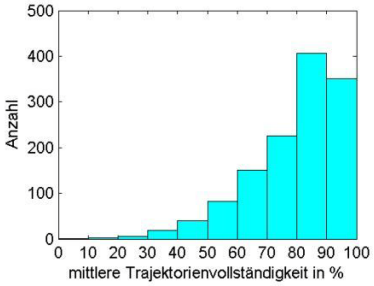
Der Parameter ist nicht direkt zur absoluten Beurteilung der generierten Trajektorien geeignet, sondern dient der Beurteilung einzelner Prozessschritte und kann daher Änderungen in der Auswertequalität sichtbar machen.



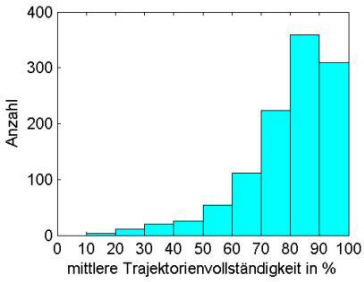
Montag, 23.03.09



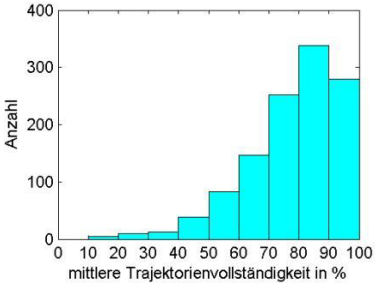
Dienstag, 24.03.09



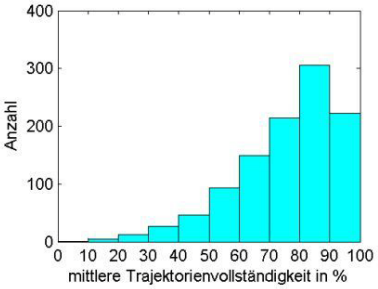
Mittwoch, 25.03.09



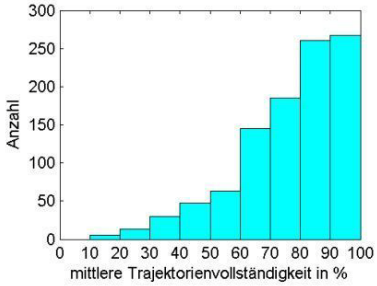
Donnerstag, 26.03.09



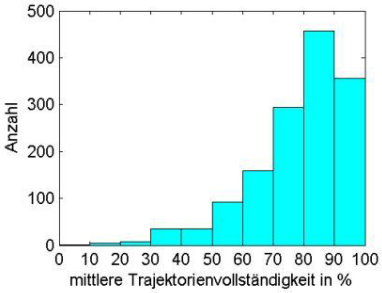
Freitag, 27.03.09



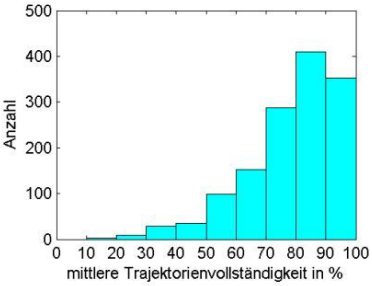
Samstag, 28.03.09



Sonntag, 29.03.09



Montag, 30.03.09



Dienstag, 31.03.09

Abbildung 15: Übersicht über die Trajektorienvollständigkeit der ausgewerteten Teilnehmer im Untersuchungszeitraum 23.-31.03.09 (Abis)

Offensichtlich kann bei einem kleineren Teil der Trajektorien nur ein relativ kleiner Anteil von 50-60 %der primär berechneten Positionsfolge verwendet werden. Bei weit mehr als der Hälfte sind es jedoch mehr als 80 % der berechneten Positionen. Der Ausschluss einzelner Positionen stellt eine Maßnahme zur Qualitätssicherung dar – die geometrische Korrektheit der Trajektorie wird dadurch erhöht.

Der Vergleich der dargestellten 9 Tage des Untersuchungszeitraumes zeigt kaum Auffälligkeiten in der Trajektorienvollständigkeit. Es kann allenfalls eine leichte Verschiebung am Sonntag hin zu einer höheren Vollständigkeit von 90-100 % abgelesen werden. Dies bestätigt in gewisser Weise die etwas höhere Abis-Trajektorienqualität, die bereits in den Darstellungen der Querabweichungen für den Sonntag (Abbildung 14) erkannbar war. Die Ursache ist noch unklar, es sind weitere Untersuchungen zur Klärung dieses Phänomens erforderlich.

Häufigkeit mit FPD getroffener RE (Road Elements)

Die folgenden Abbildungen stellen die Häufung der in den FPD-Trajektorien enthaltenen Straßenstücke (RE) farblich codiert und in unterschiedlichen Dicken für jeweils 24h dar. Die Farbpalette reicht dabei von hellblau (maximal 24 Treffer pro Tag oder durchschnittlich 1 Treffer/h) über dunkelblau bis hin zu violett (mehr als 240 Treffer/d oder durchschnittlich 10 Treffer/h). Daraus lässt sich die Abdeckung des Projektnetzes mit aus Mobilfunk gewonnenen Straßeninformationen entnehmen. Außerdem liefern die Abbildungen wichtige Hinweise auf mögliche Ausfälle einzelner Location Areas (LAC) oder größerer Gebiete sowie auf Änderungen in der Auswertequalität.

Es wurden hier nur exemplarisch die beiden Montage, 23. und 30.03. sowie der Sonntag 29. März ausgewählt. Neben den Übersichtsdarstellungen für das gesamte Projektnetz sind die Großräume Stuttgart und Karlsruhe nochmals detaillierter dargestellt. In Karlsruhe wird dabei noch zwischen A- und Abis-Daten unterschieden. Die Schaubilder aller Untersuchungstage für A- und Abis-Daten finden sich nochmals im Anhang A.

Beim Vergleich der Betrachtung der Abbildung 16 fällt zunächst das stark ausgeprägte Autobahnnetz auf. Dieses wird mit FPD-Trajektorien mit Ausnahme der A6 im Norden sowie der Autobahnkreuze Walldorf und Weinsberg im Nordosten bzw. Nordwesten mit mehr als 240 Trajektorien pro Tag, teilweise sogar mehr als 480 Trajektorien pro Tag sehr gut lückenlos abgedeckt. Dies entspricht durchschnittlich mehr als 10 bis 20 Treffern pro Stunde. Die nachgeordneten Hauptachsen wie zum Beispiel B10/Südtangente in Karlsruhe, der B10/B14 Uferstraße, Weinsteige sowie den Autobahnzufahrten B295 Stammheim und B10 Zuffenhausen im Raum Stuttgart sind ebenfalls gut erkennbar mit durchschnittlich 1 bis 5, teilweise sogar 5 bis 10 Trajektorien pro Stunde. Der Großteil der weiteren Straßen im Projektnetz ist weniger stark befahren und können daher mit weniger als 1, Abschnittsweise auch mit bis zu 120 Trajektorien pro Tag zumindest nachgewiesen werden. Wie weit diese der Einschätzung der Verkehrslage dienen können, muss in den Demonstratoren untersucht werden. Es existieren nur vereinzelte

Straßenabschnitte, die in keiner FPD-Trajektorie enthalten sind (in Abbildung 16 schwarz dargestellt).

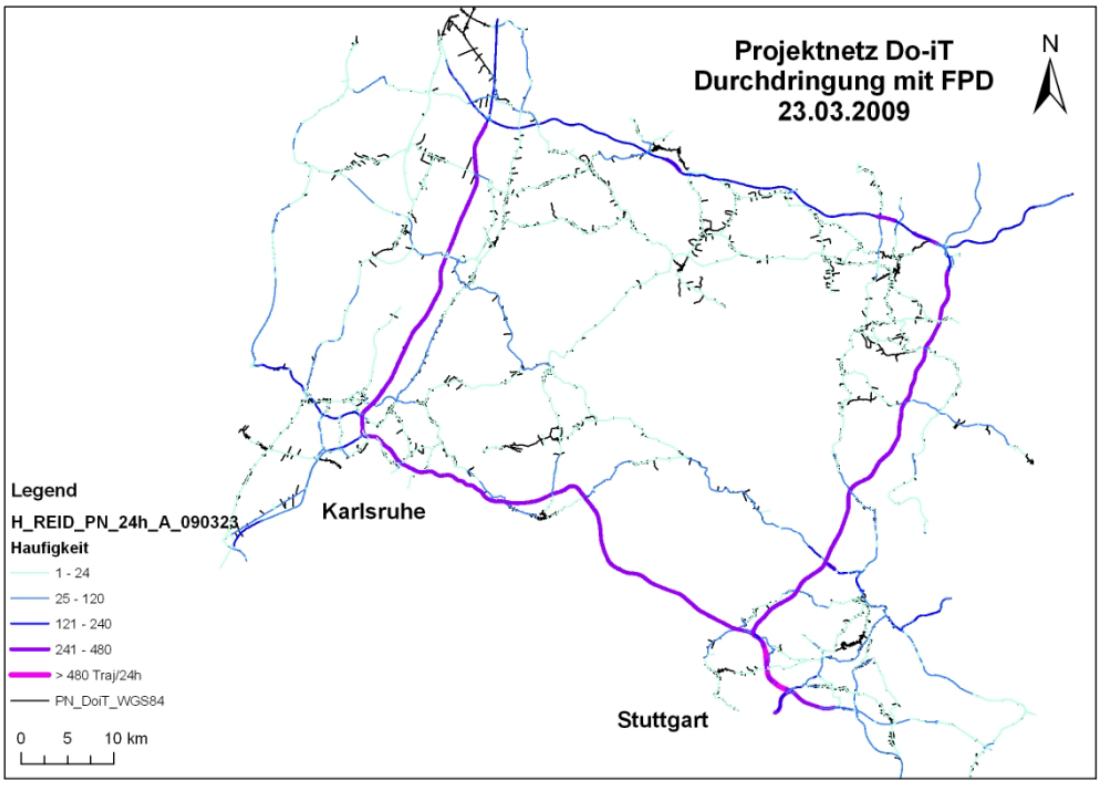


Abbildung 16: Durchdringung des gesamten PN mit FPD am Montag, 23.03.09

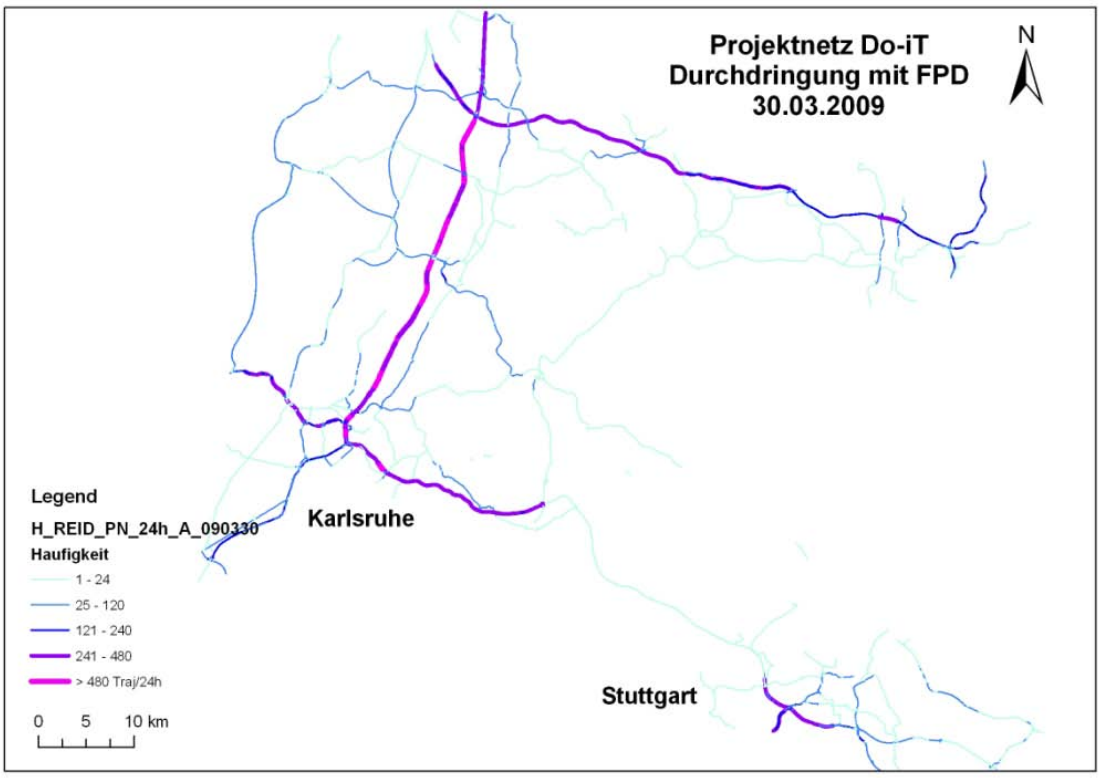


Abbildung 17: Durchdringung des gesamten PN mit FPD am Montag 30.03.09

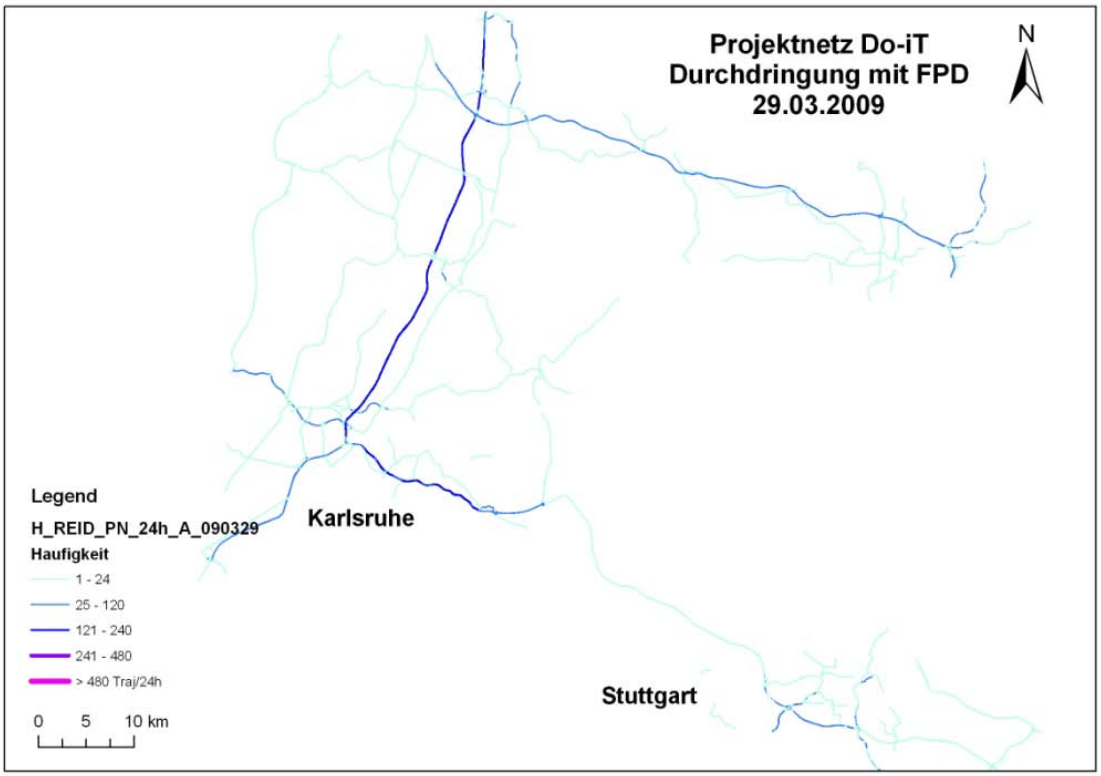


Abbildung 18: Durchdringung des gesamten PN mit FPD am Sonntag 29.03.09

Vergleicht man Abbildung 16 und Abbildung 25, fällt sofort die Lücke auf, die sich als ca. 30 bis 40 km breiter Streifen von Pforzheim im Südwesten bis zur A81 erstreckt. Diese Datenlücke findet sich in allen weiteren Tagen bis zum Projektende. Wie sich herausgestellt hat, gab es am Dienstag, 25. März einen technischen Defekt bei der T-Mobile, auf Grund dessen 20 der 52 überwachten LACs ausgefallen sind. Dieser Defekt wurde auch bis zum Projektende nicht mehr behoben.

Die Abbildung 19 zeigt eine relativ geringe Abdeckung mit FPD aus A-Daten von durchschnittlich ca. 1 - 5 Trajektorien pro Stunde im Raum Stuttgart. Teilweise werden auch 5 – 10 Trajektorien pro Stunde erreicht. Bis auf den Wagenburgtunnel im Zentrum der Abbildung ist nahezu das gesamte Projektnetz in Stuttgart mit FPD-Trajektorien überdeckt. Es bestätigt sich die Vermutung, dass die bewegte Topographie und die Kessellage Stuttgarts trotz der kleineren Mobilfunkzellen relativ zum Umland eine Bestimmung von Mobilfunktrajektorien erschwert.

Der Vergleich mit Abbildung 20 zeigt wiederum den Ausfall der 20 LACs, der bereits angesprochen wurde. Es bleiben nur wenige Bereiche von Stuttgart übrig, in denen weiterhin A-Daten prozessiert werden konnten.

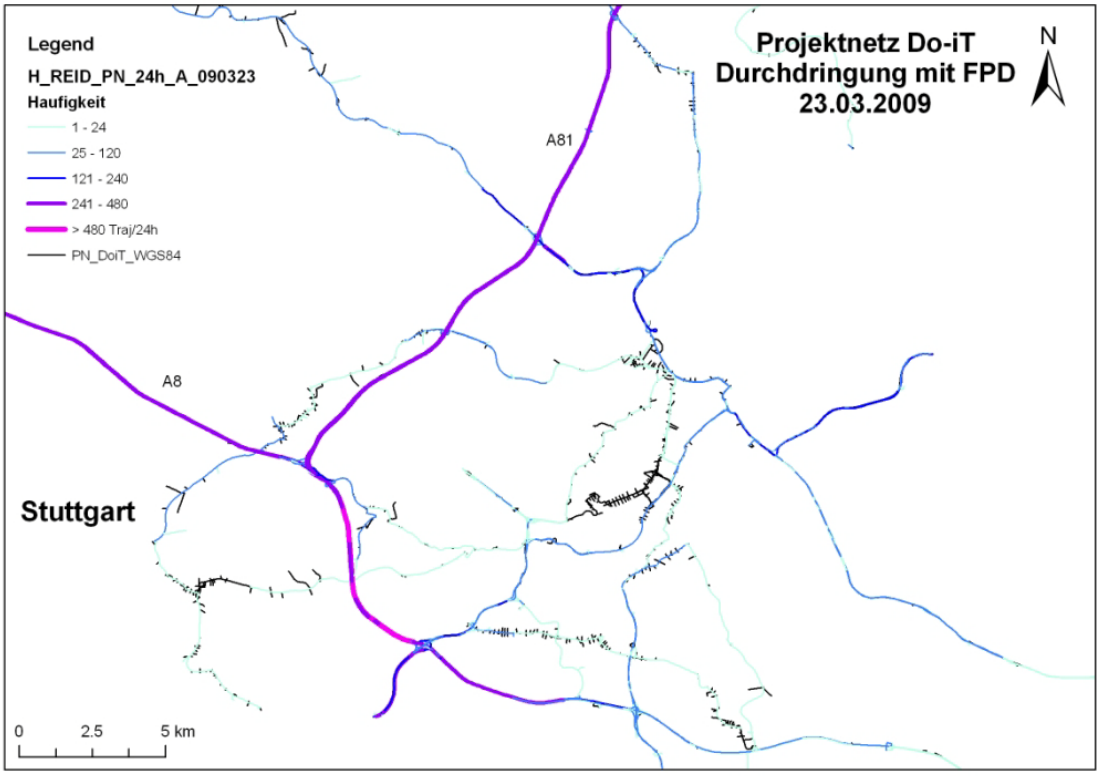


Abbildung 19: Häufigkeit auftretender RE im Raum Stuttgart, Montag, 23.03.09 (A)

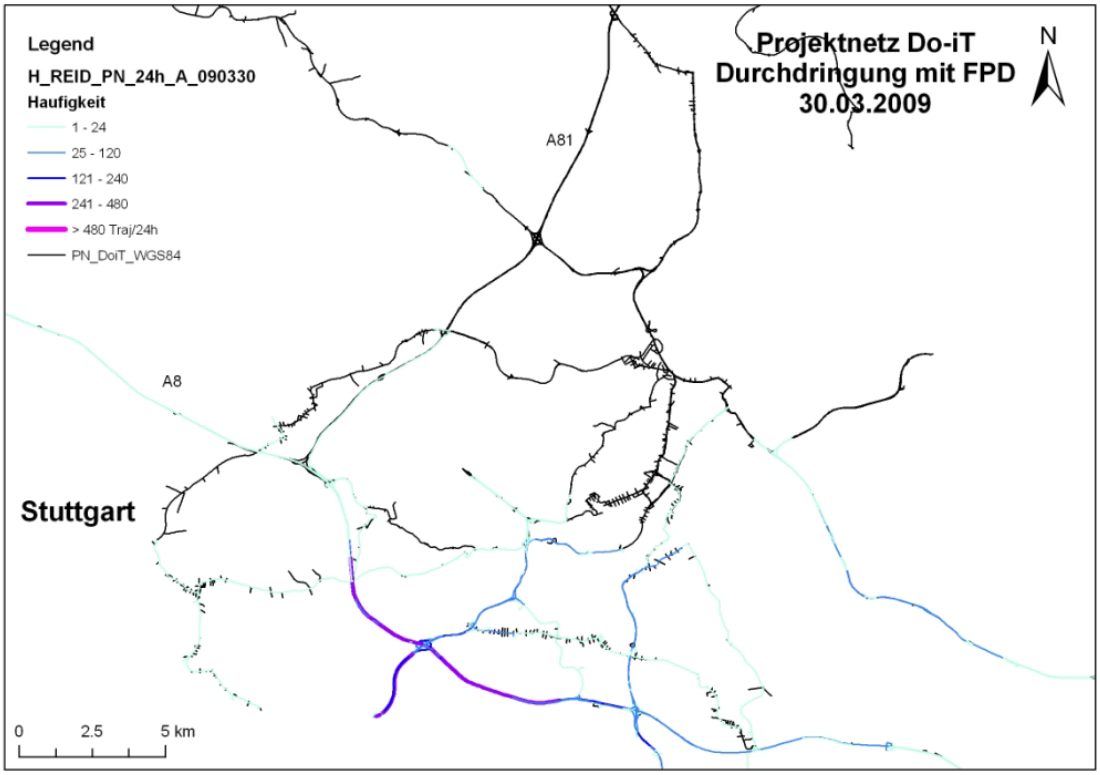


Abbildung 20: Häufigkeit auftretender RE im Raum Stuttgart, Montag, 30.03.09 (A)

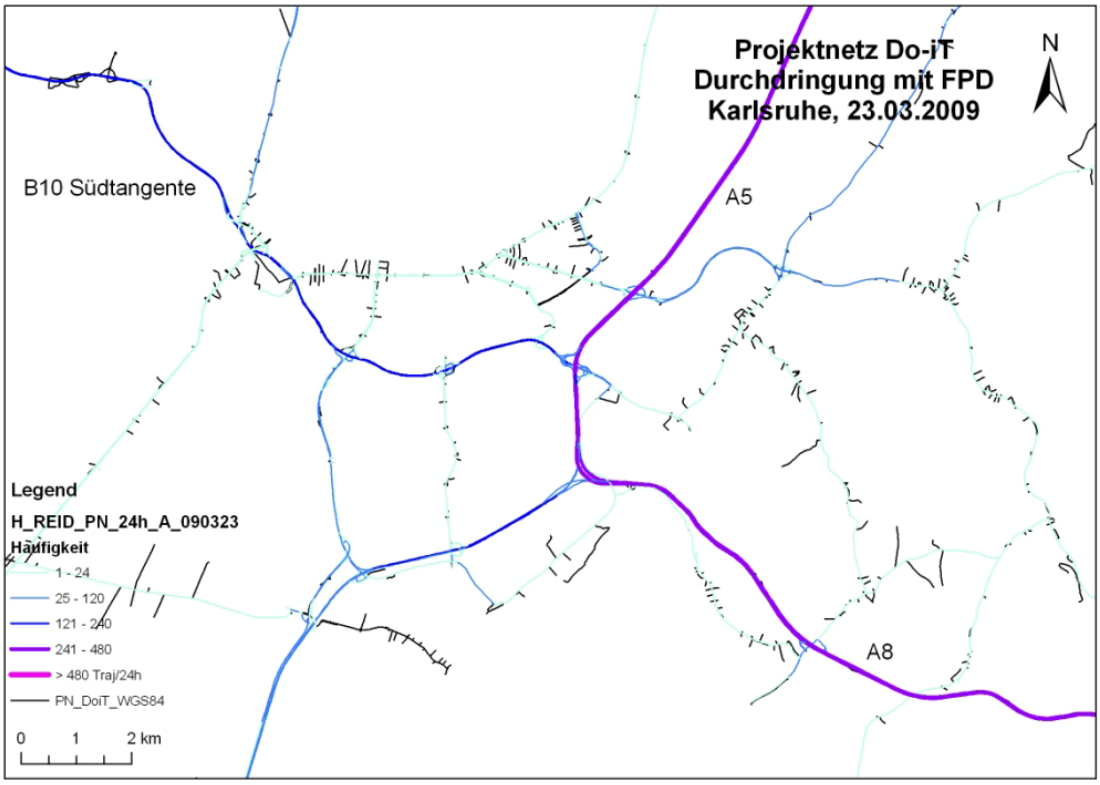


Abbildung 21: Häufigkeit auftretender RE im Raum Karlsruhe, Montag, 23.03.09 (A)

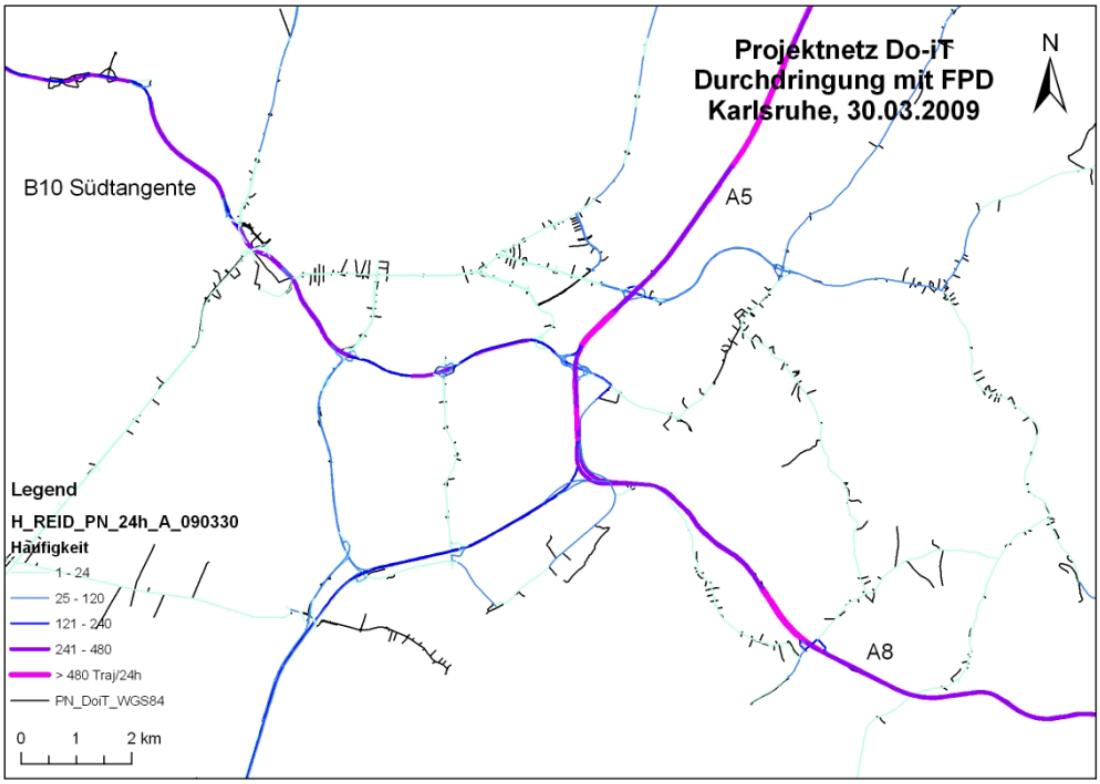


Abbildung 22: Häufigkeit auftretender RE im Raum Karlsruhe, Montag, 30.03.09 (A)

In Karlsruhe werden ebenfalls nahezu alle Straßen aus dem Projektnetz abgedeckt. Allerdings zeigt sich ähnlich wie in Stuttgart eine sehr dünne Überdeckung von weniger als 24 Trajektorien pro Tag für die innerstädtischen Bestandteile des PN. Die Südtangente (B10), die Einfallsstraße (B10/Durlacher Allee) aus Osten sowie die westliche Spange zwischen A5 und Südtangente (L606/Brauerstraße) zeigen hingegen eine stärkere Überdeckung von durchschnittlich 1 bis 10 Trajektorien pro Stunde.

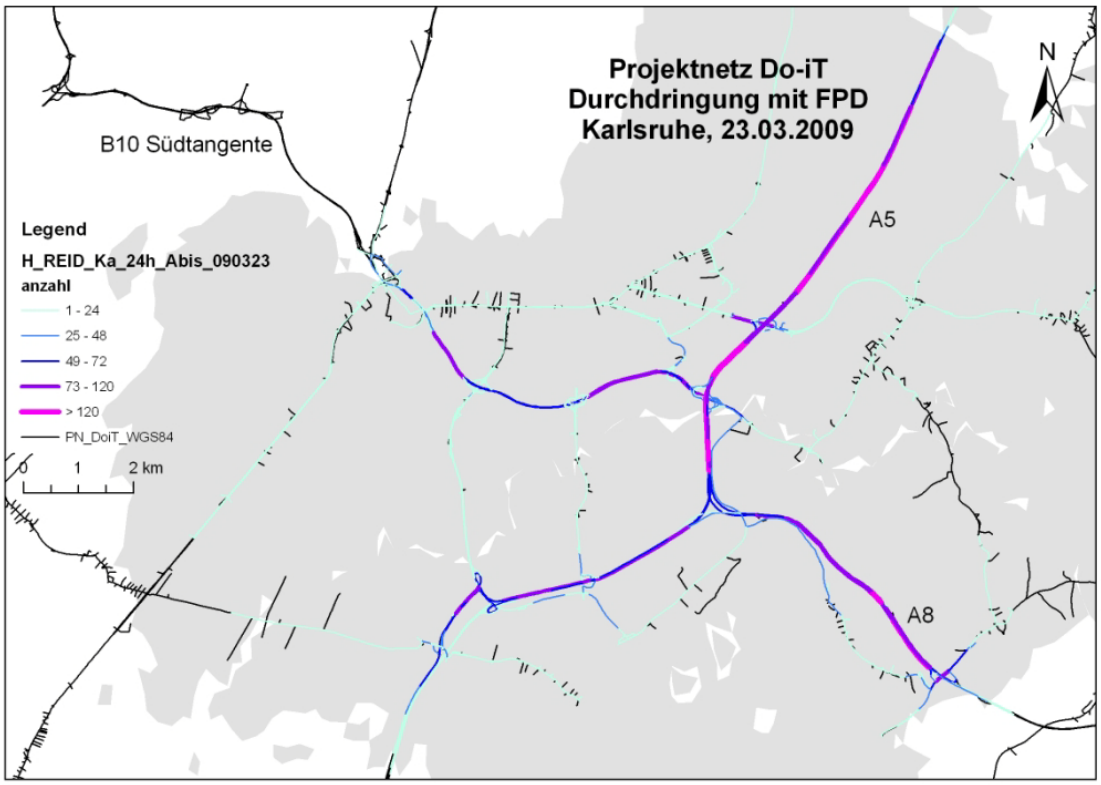


Abbildung 23: Häufigkeit auftretender RE im Raum Karlsruhe, Montag, 23.03.09 (A-bis)

In der Umgebung von Karlsruhe zeigt sich eine durchschnittliche Abdeckung der Autobahn A8 und A5 mit Abis-Trajektorien mit mehr als 3, teilweise auch mehr als 5 Trajektorien pro Stunde. Insbesondere im Norden von Karlsruhe wird die Autobahn relativ häufig getroffen (vgl. Abbildung 24 und Abbildung 23). Am Rand der grauen Fläche, in der Abis-Daten zur Verfügung standen, wird die Anzahl Trajektorien allmählich weniger aufgrund der noch nicht auswertbaren Fahrzeuge, die in das überwachte Abis-Gebiet einfahren. Die Südtangente wurde vollständig mit durchschnittlich mehr als 2 Trajektorien, im Ostteil mit mehr als 3 Trajektorien pro Stunde identifiziert. Die untergeordneten Straßen des Projektnetzes in Karlsruhe wurden bis auf wenige Ausnahmen nur mit weniger als 1 Trajektorie je Stunde getroffen.

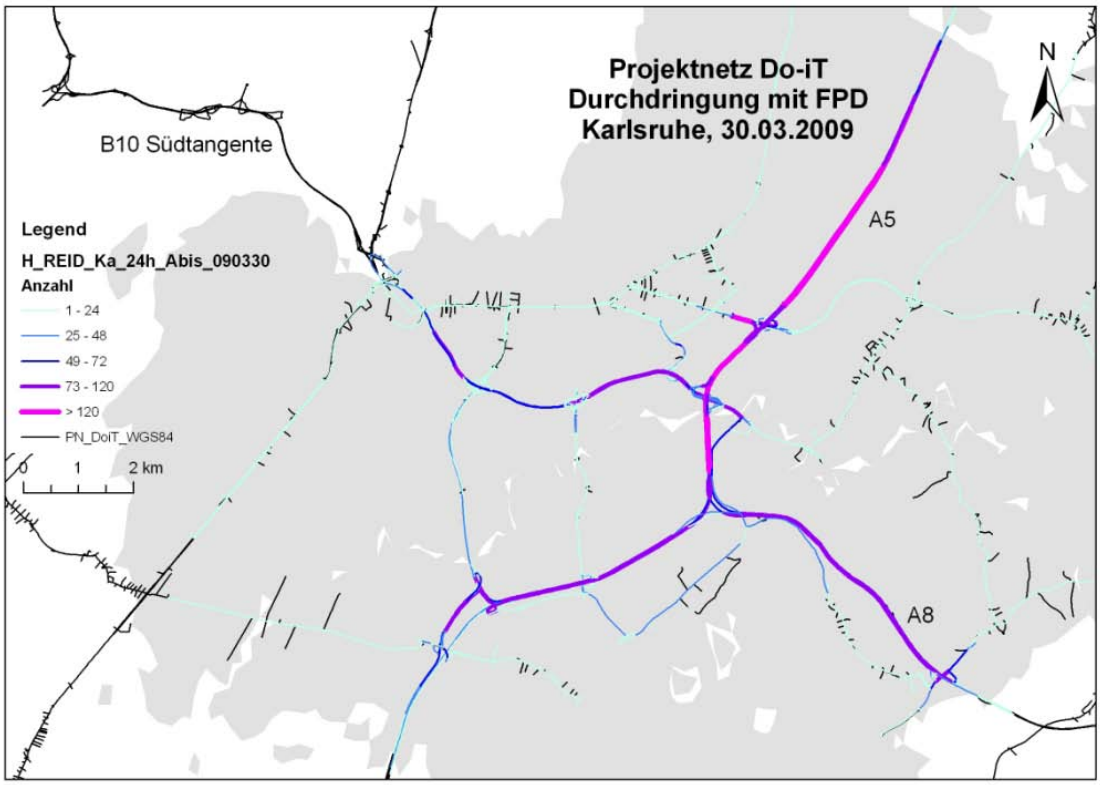


Abbildung 24: Häufigkeit auftretender RE im Raum Karlsruhe, Montag, 30.03.09 (A-bis)

2.3 Vergleich mit stationären Erfassungssystemen

Zur Evaluierung der berechneten FPD-Trajektorien im Hinblick auf die Nutzung als Verkehrsdatenquelle waren Referenzdaten aus dem Straßennetz erforderlich. Es wurden zunächst insgesamt 253 stationäre Sensoren der Projektpartner zur Verkehrserfassung ausgewählt, um die aus Mobilfunkdaten ermittelten Verkehrsstärken zu beurteilen sowie einen ersten Eindruck von der ableitbaren Durchschnittsgeschwindigkeit zu bekommen (eine genauere Untersuchung der aus FPD ableitbaren Geschwindigkeiten sowie der Weg-Zeit-Diagramme wurde mit GPS-Referenz durchgeführt, vgl. Abschnitt 2.1). Im Untersuchungszeitraum von Montag, 23. bis Dienstag, 31. März standen jedoch nur die Messdaten von 173 der ausgewählten stationären Verkehrsdatenerfassungssystemen (SES) ständig zur Verfügung. Einige Sensoren haben in diesem Zeitraum nur teilweise Daten geliefert oder waren überhaupt nicht verfügbar. Daher beziehen sich alle weiteren Beschreibungen und Darstellungen weitgehend auf die Sensoren, deren Daten in diesem Zeitraum Ende März zur Verfügung standen.

Zusätzlich eingeschränkt wird die Evaluierungsmöglichkeit durch den bereits im vorigen Kapitel 2.2 erwähnten Ausfall von 20 LACs im Süden und Osten des Projektgebietes. Daher können für den betroffenen Bereich nur am Montag, den 23.03.09 Vergleiche angestellt werden, im restlichen Untersuchungszeitraum standen dort keine FPD-Trajektorien aus A-Daten zur Verfügung.

Die Qualität der aus Mobilfunkdaten ableitbaren Verkehrsstärke q wird mit Hilfe der beiden Parameter **Durchdringung** und **Rangkorrelation** beurteilt (vgl. Tabelle 3). Die Durchdringung ist der prozentuale Anteil der realen Verkehrsstärke, der mit Hilfe der Mobilfunkortung ermittelt werden kann. Als reale Verkehrsstärke wird dabei die punktuell mit stationären Sensoren ermittelte Verkehrsstärke herangezogen. Die Rangkorrelation nach Spearman ist ein robustes Maß für die Stärke des Zusammenhangs zweier Messreihen respektive Tagesganglinien (vgl. Sachs 2006). Der in der Regel verwendete Korrelationskoeffizient nach Pearson kommt hier nicht zum Einsatz, da dieser zu stark durch einzelne Ausreißer beeinflusst wird. Grundsätzlich ergeben sich jedoch, wie in Tabelle 5 dargestellt, vergleichbare Werte für die Korrelationen, so dass die hier gemachten Grundaussagen von der Wahl weitgehend unbeeinflusst bleiben. Durch die Transformation der Tagesganglinien in Rangreihen wird der Einfluss von Ausreißern minimiert und der Zusammenhang der beiden Messreihen kann besser beurteilt werden.

Tabelle 4: Definition der verwendeten Qualitätsparameter zur Evaluierung der FPD mit SES

Parameter	Berechnung	Beschreibung
Durchdringung (diskret)	$d = \frac{q_{FPD}}{q_{SES}}[\%]$	Anteil des Verkehrs, der mit FPD erfasst werden kann; Verkehrsstärke in KFZ/h aus FPD bezogen auf die Verkehrsstärke aus den SES
Rangkorrelation nach Spearman (diskret)	$r = 1 - \frac{6 \cdot \sum D^2}{n \cdot (n^2 - 1)}$	Robustes Maß zur Beurteilung des Zusammenhangs der Tagesganglinien der Verkehrsstärke aus FPD und aus SES; zunächst werden die Reihen in Rangreihen transformiert, dann werden die Differenzen D der n Rangpaare gebildet

Tabelle 5. Vergleich der Korrelationswerte nach Pearson (klassisch) sowie nach Spearman (Rangkorrelation) für Montag, den 23. März 2009

Differenz Spearman-Pearson	Anzahl	Anteil in %
< 0,1	174	97
0.1< D < 0,4	4	3
> 0,4	0	0

Durchdringung und Korrelation werden lokal an den Sensorstandorten berechnet. Daher ist eine hohe Sensordichte im Projektgebiet wichtig, um sinnvolle flächendeckende Aussagen treffen zu können. Je weiter ein Streckenabschnitt von dem nächsten SES-Standort entfernt ist, desto unsicherer kann die Qualität der FPD-Trajektorien beurteilt werden. Durch die relativ große Anzahl von Sensoren, die annähernd gleichmäßig über die arteriellen Verkehrswege im Projektgebiet verteilt sind, ergibt sich im Mittel jedoch ein gutes Bild vom Potenzial der Mobilfunkortung wie sie in Do-iT umgesetzt wurde.

Im Folgenden werden zunächst die verwendeten Sensoren der Projektpartner sowie deren Lage im Projektnetz dargestellt. Es handelt sich dabei sowohl um Doppelschleifen, wie sie insbesondere auf den Autobahnen betrieben werden, als auch um Einfachschleifen und passive Infrarotsensoren (TEU-Traffic Eye Universal). Letztere werden vorwiegend innerstädtisch sowie auf Durchgangsstraßen eingesetzt.

Die Qualität der verwendeten stationären Erfassungssysteme war ausreichend. Die Korrektheit der Verkehrsstärke q, die mit Doppelschleifen auf der A81 erfasst wird, konnte bereits im Rahmen einer Studienarbeit exemplarisch nachgewiesen werden. Es

wurde anhand von punktuellen Referenzzählungen eine maximale Abweichung von 1 % zur tatsächlichen Verkehrsstärke ermittelt (Karrer 2008). Diese Ergebnisse können aufgrund der ähnlichen Bau- und Funktionsweise aller Sensoren auf den Autobahnen im Projektnetz auf diese übertragen werden. Wie eine Studie von (Sachse 2002) beispielhaft an zwei TEU-Sensoren nachweist, weichen die ermittelten Verkehrsstärken in ca. 95 % der Testintervalle nicht mehr als 1Fz/5Min von der tatsächlichen Verkehrsstärke ab und erfüllen damit die Anforderungen der Bundesanstalt für Straßenwesen (TLS 1997). Die Geschwindigkeitsmessungen der Sensoren im Autobahnnetz sind aufgrund der Bauweise als Doppelschleifen als ausreichend genau anzusehen. Für die stadtnah installierten TEU liefert die Untersuchung von (Sachse 2002) eine erreichbare Genauigkeit von ±10 km/h für die mittlere Fahrzeuggeschwindigkeit. Wie sich jedoch bereits in den Untersuchungen mit GPS-Referenz herausgestellt hat, wird diese Genauigkeit zur Beurteilung der aus FPD ermittelbaren Geschwindigkeiten in Stadtnähe ausreichen, da sich für die FPD Werte im günstigsten Fall (lange Trajektorien, ausserstädtisch) Genauigkeiten von ±12,5 km/h ergeben haben. Innerstädtisch wurden jedoch wesentlich schlechtere Ergebnisse erzielt.

Damit können aus funktionaler Sicht grundsätzlich alle in Do-iT zur Verfügung stehenden Verkehrssensoren zur Evaluierung herangezogen werden. Einige Sensoren insbesondere innerstädtisch müssen jedoch aufgrund der ungünstigen Lage oder nur teilweiser Abdeckung des Fahrbahnquerschnitts im Bereich von Kreuzungen zunächst ausgesondert werden oder eignen sich nur als Referenz für die Verkehrsstärke und nicht für die mittlere Geschwindigkeit.

Zur groben geografischen Orientierung ist auf der folgenden Abbildung 25 zunächst das gesamte Projektnetz dargestellt. Sämtliche Straßenelemente (REIDs) auf denen während des Untersuchungszeitraumes stationäre Sensoren Referenzdaten geliefert haben, sind rot hervorgehoben. Die grau hinterlegte Fläche markiert die beiden LACs KA und Ettlingen, in denen Abis Daten zur Verfügung standen. In diesem Bereich war sowohl die Auswertung von A- als auch Abis-Daten möglich. Im restlichen Projektgebiet standen aus technischen Gründen nur die A-Daten zur Verfügung.

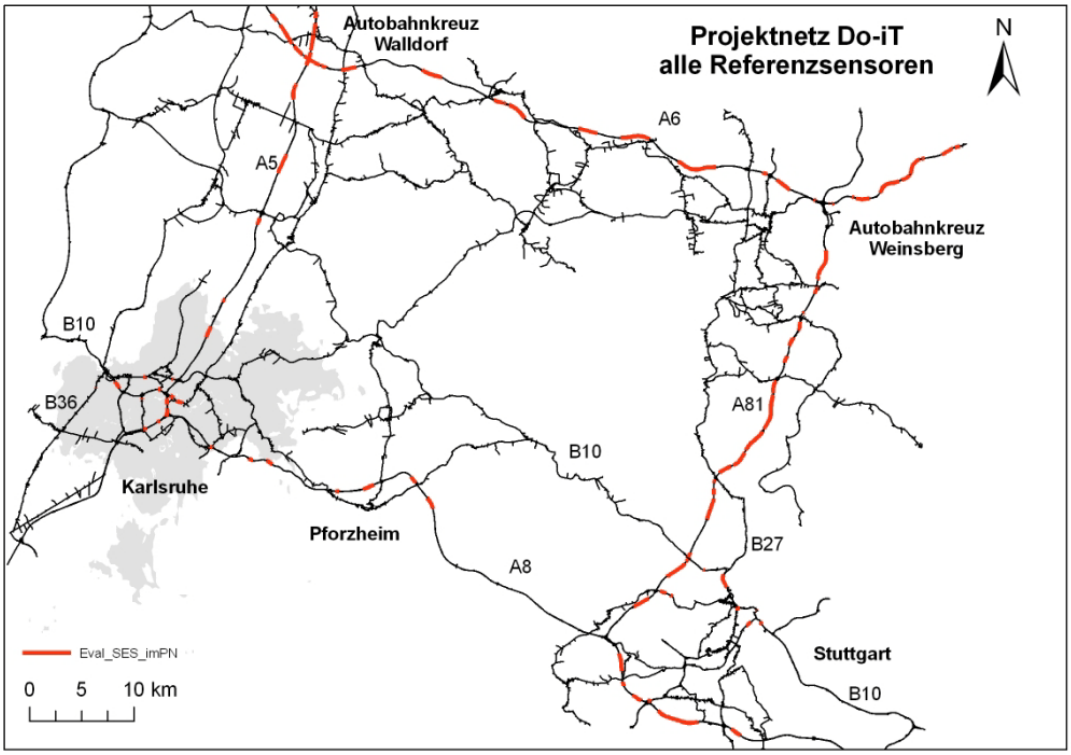


Abbildung 25: Übersicht über das gesamte Projektstraßennetz mit rot hervorgehobenen Road Elements (RE) der digitalen Straßenkarte auf denen geeignete SES installiert sind

In den folgenden Darstellungen der Sensorstandorte wird zur Unterscheidung jeweils die RMQ-ID (ID des Regelmessquerschnitts[2]) der Sensoren mit angegeben. Diese projektinterne Bezeichnung wurde eingeführt, um sämtliche Sensoren der beteiligten Partner eindeutig identifizieren zu können. Daher werden im Weiteren nur noch diese Bezeichnungen verwendet.

Sensoren der DDG (Autobahnnetz)

Die DDG stellte insgesamt 122 Sensoren zur Verfügung, von denen 92 im Untersuchungszeitraum Daten geliefert haben. Es handelt sich vorwiegend um Doppelschleifen auf dem Autobahnnetz, wobei Verkehrsstärke und Durchschnittsgeschwindigkeit i. d. R. aus Messungen auf der Überholspur abgeleitet werden. Wie Langzeiterfahrungen gezeigt haben, stimmen die hochgerechneten Werte jedoch sehr gut mit den realen Verhältnissen überein, weshalb die daraus extrapolierte Verkehrsstärke und der generierte LOS (Level of Service) die tatsächlichen Straßenverhältnisse auf allen Spuren des Messquerschnitts (MQ) sehr gut wiedergibt. Die Verteilung der 92 Sensoren

[2] Als Regelmessquerschnitt wird die Zusammenfassung aller Spuren in einer Fahrtrichtung bezeichnet

auf dem Projektnetz ist in Abbildung 26 dargestellt, die Straßenelemente (Road Elements, RE) auf denen Sensoren installiert sind, wurden rot hervorgehoben. Aus Platzgründen konnten in der Übersicht nicht alle Regelmessquerschnitte mit Ihrer RMQ-ID dargestellt werden. Auffällig ist jedoch die teilweise ungleichmäßige Verteilung mit einer hohen Sensordichte auf der A81 sowie in der Nähe des Autobahnkreuz Weinsberg im Nordosten sowie des Autobahnkreuzes Walldorf im Nordwesten des Projektgebietes. Auf der A5 sowie auf der A8 ist die Sensordichte eher klein. Dies liegt teilweise an dem Ausschluss von Sensoren, die im Untersuchungszeitraum keine Daten geliefert haben. Auf der A8 zwischen PF und KA sowie Auf der A6 zwischen Steinsfurt und Neckarsulm befinden sich jedoch echte Sensorlücken. Beide sind vermutlich auf aktuell noch bestehende und erst kürzlich abgeschlossene Bautätigkeiten zurückzuführen.

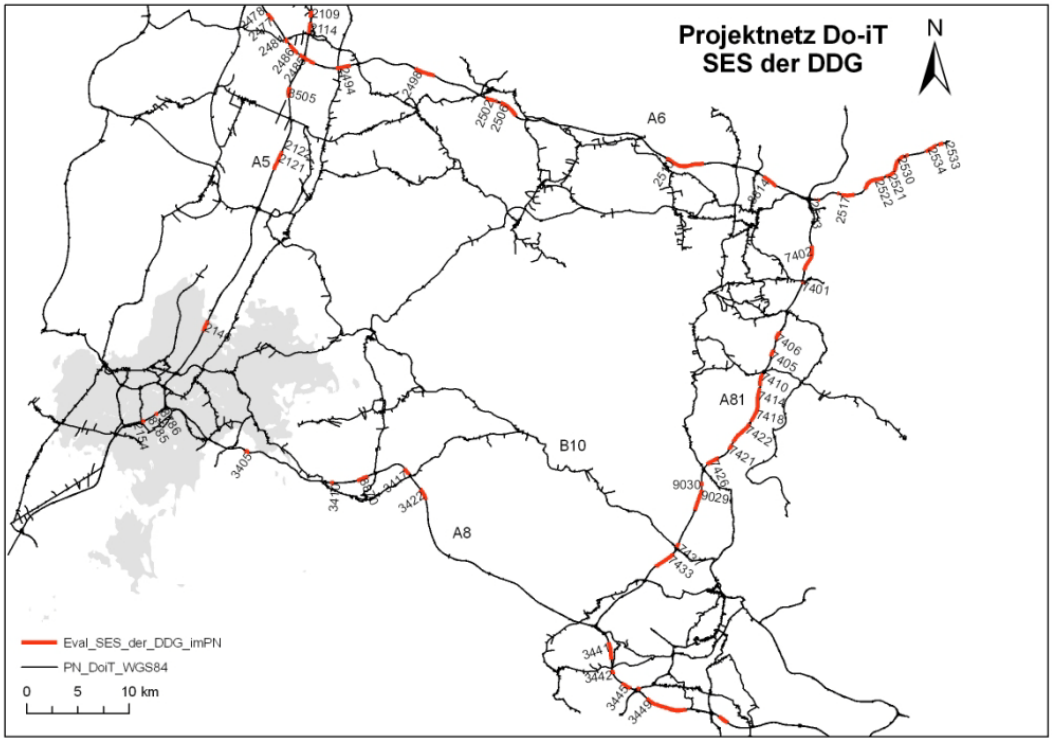


Abbildung 26: Übersicht aller verwendeten SES der DDG mit ihrer RMQ-ID auf dem Projektnetz (aus Platzgründen sind nicht alle RMQ_IDs angegeben)

Nach dem technischen Defekt am 24.03., der zum Ausfall von 20 LACs geführt hat, können die Sensoren auf der A81 nicht mehr zur Evaluierung herangezogen werden. Dasselbe gilt für die Sensoren auf der A8 im Raum Pforzheim.

Sensoren der Landesstelle für Straßentechnik BW (Autobahnnetz)

Wie in Abbildung 27 ersichtlich, wurden Sensoren der Landesstelle für Straßentechnik ebenfalls aus dem gesamten Projektnetz zur Verfügung gestellt. Insgesamt standen 91 RMQ zur Verfügung. Grundsätzlich waren alle Sensoren der Landesstelle als Refe-

renz geeignet, allerdings standen nur von 56 Sensoren im Evaluierungszeitraum Ende März Messdaten zur Verfügung. Hier fielen einerseits einige Sensoren auf der A8 vermutlich aufgrund des bereits erwähnten Autobahnausbaus sowie einige im Bereich des Kreuz Weinsberg aus.

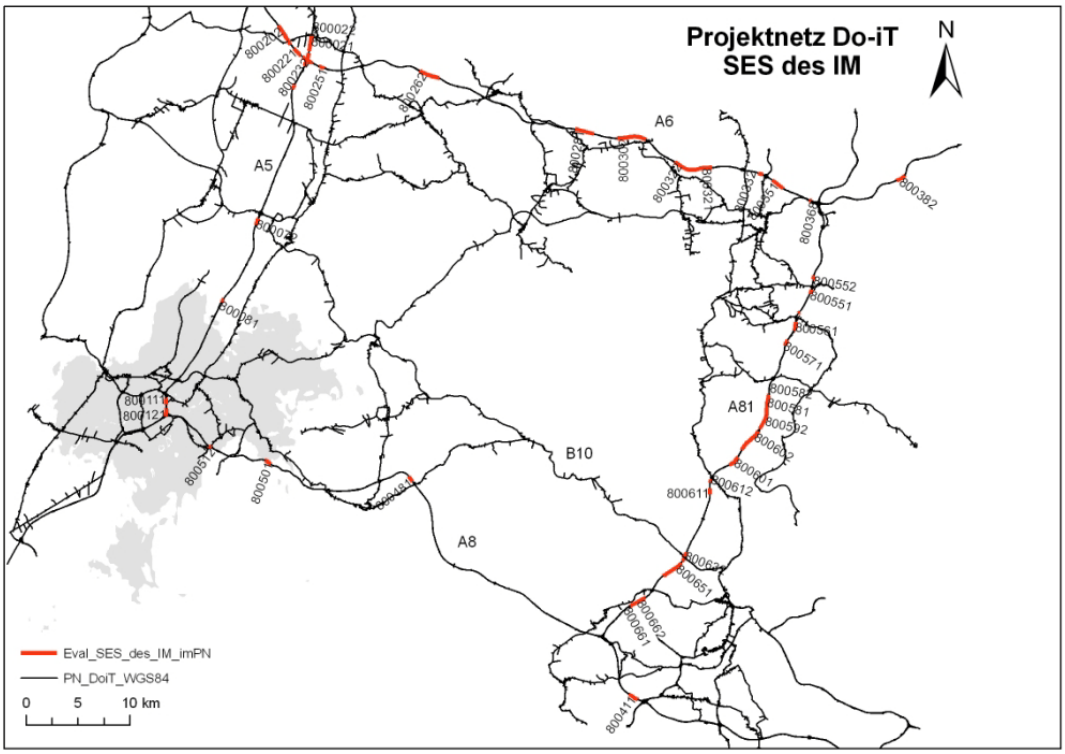


Abbildung 27: Projektnetz mit allen MQ der Landesstelle für Straßentechnik BW in Do-iT (aus Platzgründen sind nicht alle RMQ_ID angegeben).

Da die Messquerschnitte vorwiegend der Steuerung einer Netzbeeinflussungsanlage im Autobahnviereck sowie statistischen Zwecken dienen, wird an den einzelnen Messquerschnitten jeweils richtungsscharf der gesamte Verkehr auf allen Spuren erfasst. Es handelt sich ausschließlich um Doppelschleifen, die neben der Verkehrsstärke auch gute Durchschnittsgeschwindigkeiten liefern.

Das gesamte Autobahnviereck im Projektnetz wird somit mit insgesamt 153 Messquerschnitten sehr dicht abgedeckt, auch wenn einzelne Abschnitte aus verschiedenen Gründen im Untersuchungszeitraum nicht überwacht werden konnten. Damit wird eine sehr gute Referenz zur Evaluierung der aus A-Daten generierten FPD-Trajektorien auf dem Autobahnnetz zur Verfügung gestellt. Die Qualitätsaussagen können aufgrund der vergleichbaren Infrastrukurdichte auf die weniger dicht abgedeckten Autobahnabschnitte übertragen werden und es ist daher möglich, allgemeine Aussagen zur Eignung von FPD zur Bestimmung der Verkehrsstärke auf Autobahnen zu treffen.

Auch hier konnten die Sensoren auf der A81 auf Grund des technischen Defekts bei der T-Mobile ab dem 24.03. nicht mehr zur Evaluierung verwendet werden, da die Berechnung von FPD-Trajektorien nicht mehr möglich war.

Sensoren der Stadt Karlsruhe

In Karlsruhe wurden nur TEU Sensoren ausgewählt, da die Induktionsschleifen meist nur einzelne Fahrspuren abdecken und sich die Sensoren i. d. R. unmittelbar vor oder nach Kreuzungen befinden. Daher sind sie weder als Referenz für die Verkehrsstärke des gesamten Richtungsmessquerschnitts noch für die Durchschnittsgeschwindigkeit geeignet. Abbildung 28 zeigt in rot hervorgehoben die mit einem geeigneten Sensor besetzen Straßenelemente.

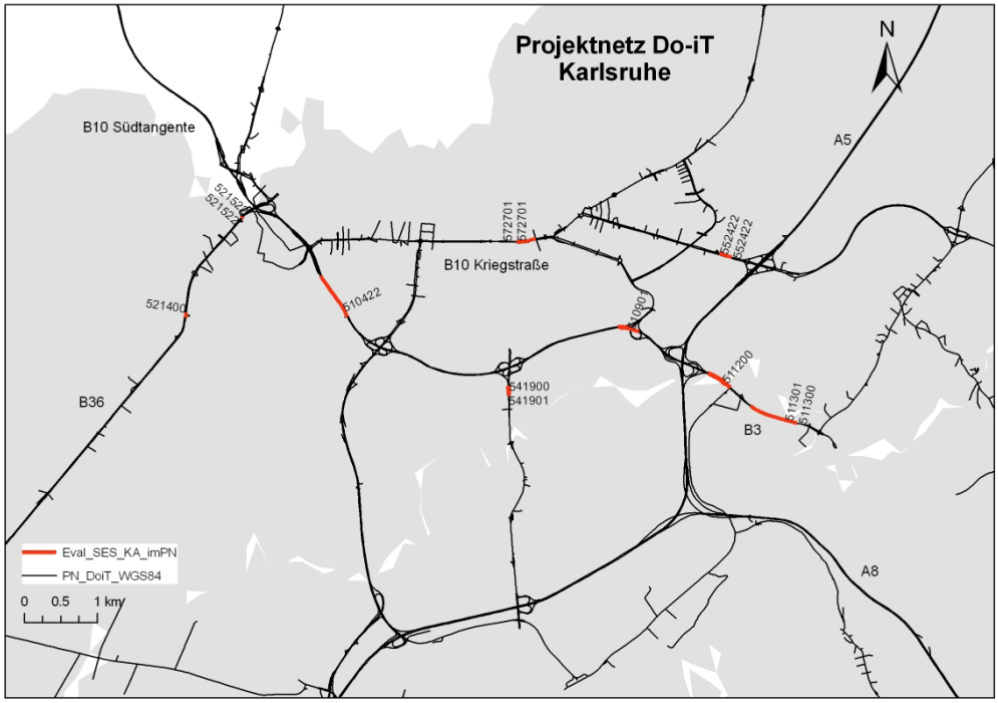


Abbildung 28: Projektnetz mit allen verwendeten TEU sowie ihrer RMQ-ID in Karlsruhe

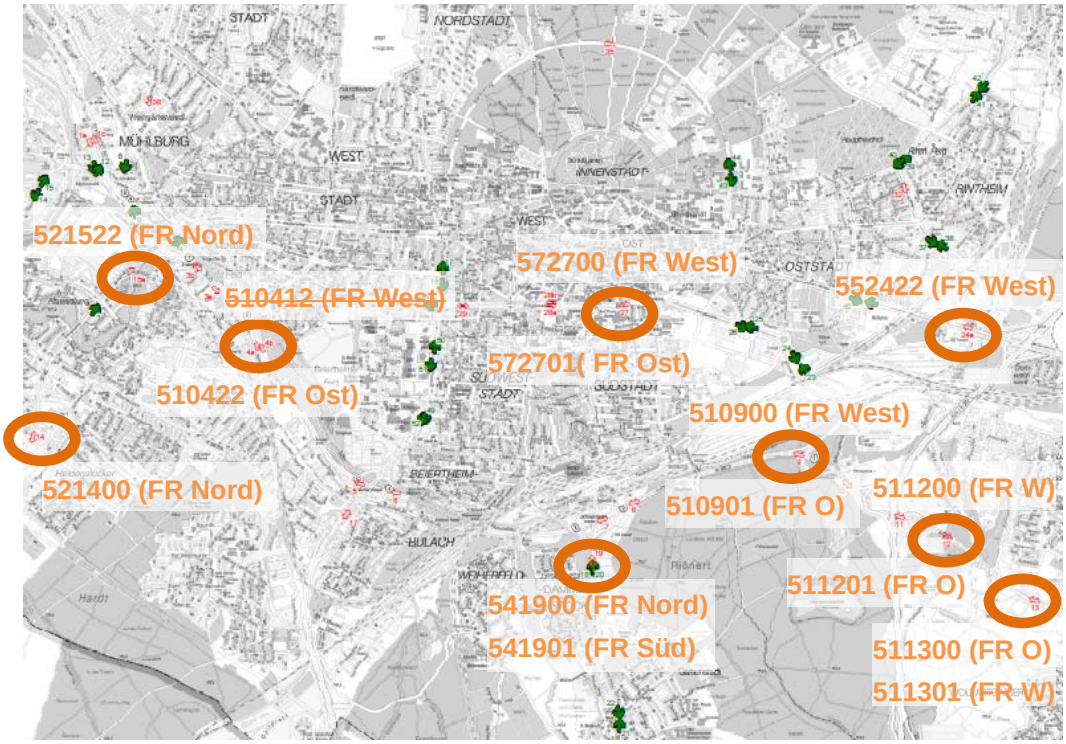


Abbildung 29: Übersicht der ausgewählten TEU-Sensoren in Karlsruhe mit Angabe der Fahrtrichtungen (top. Karte als Hintergrund, Quelle: Stadt Karlsruhe)

Aufgrund der Lage als besonders geeignet eingestuft wurden die in der Abbildung 29 orange hervorgehobenen und mit ihrer projektinternen RMQ-ID dargestellten Infrarotsensoren. Alle Sensoren liegen auf den Hauptein- und ausfallstraßen sowie auf der Südtangente als verkehrstechnisch wichtige Ost-West Verbindung zwischen den Ländern Baden-Württemberg und Rheinland-Pfalz bzw. als Verknüpfung der Autobahn A8 mit der A65. Es handelt sich um 4 Messquerschnitte auf der kreuzungsfreien Südtangente B10/K9657 (RMQ-ID: 510421/422, 510900/901, 511200/01 und 511300/01) sowie um Messstellen auf den Einfallsstraßen B36 aus Südwest stadteinwärts (521400 und 521522), auf der L561 im Süden stadtein- und auswärts (541900/01) sowie der Durlacher Allee aus Osten kommend (552422). Zusätzlich wurde eine Messstelle auf der Kriegsbergstraße in der Stadtmitte hinzugenommen (572700/01). Der Sensor mit der RMQ-ID 510412 auf der Südtangete hat im Untersuchungszeitraum keine Daten geliefert. Damit standen insgesamt 14 RMQ zur Verfügung.

Sensoren der Stadt Stuttgart

In Stuttgart (LHS) standen nur Schleifensensoren zur Verfügung. Allerdings sind – anders als in Karlsruhe – viele Schleifensensoren kreuzungsfern installiert und als Doppelschleifen ausgeführt und bilden daher gut den fließenden Verkehr ab. Insgesamt wurden zunächst 20 Richtungsmessquerschnitte mit insgesamt 46 einzelnen Spursensoren auf 34 einzelnen Road Elementen der digitalen Karte ausgewählt. Von diesen grundsätzlich zur Evaluierung geeigneten Sensoren standen jedoch viele im Untersu-

chungszeitraum nicht zur Verfügung. Daher beschränkt sich die Auswahl auf die in den folgenden Abbildungen dargestellten 11 Richtungsmessquerschnitte.

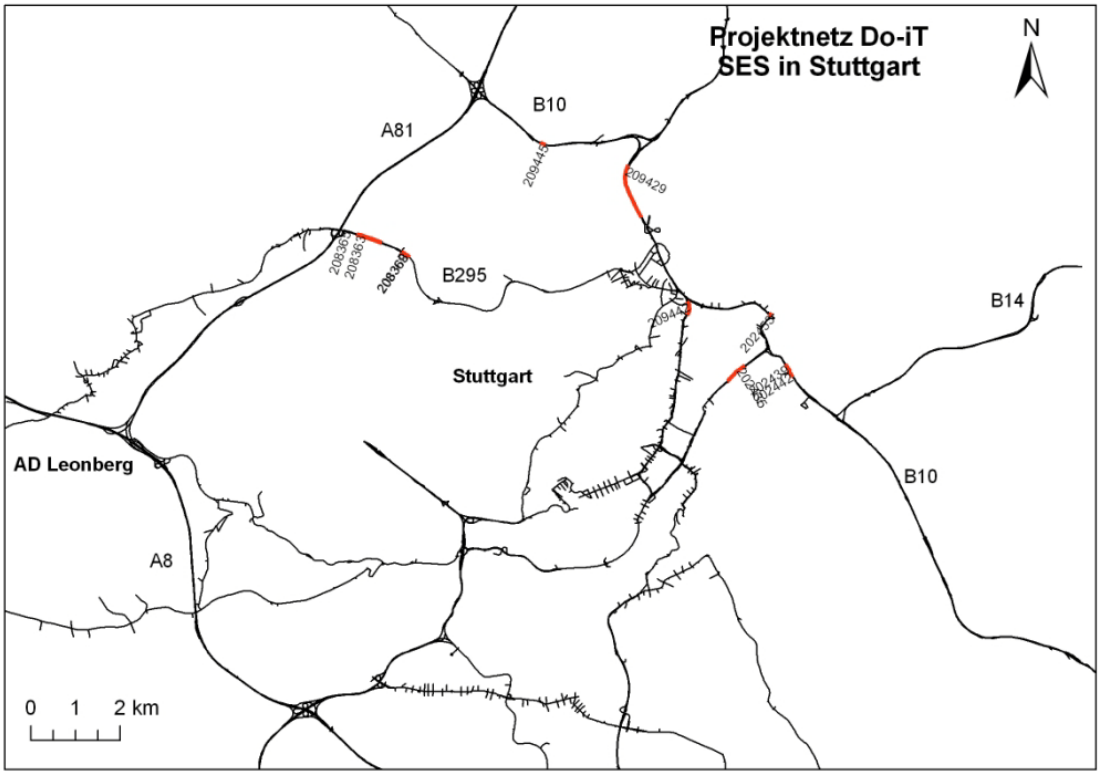


Abbildung 30: Übersicht aller verwendeter RMQ in Stuttgart im Projektnetz

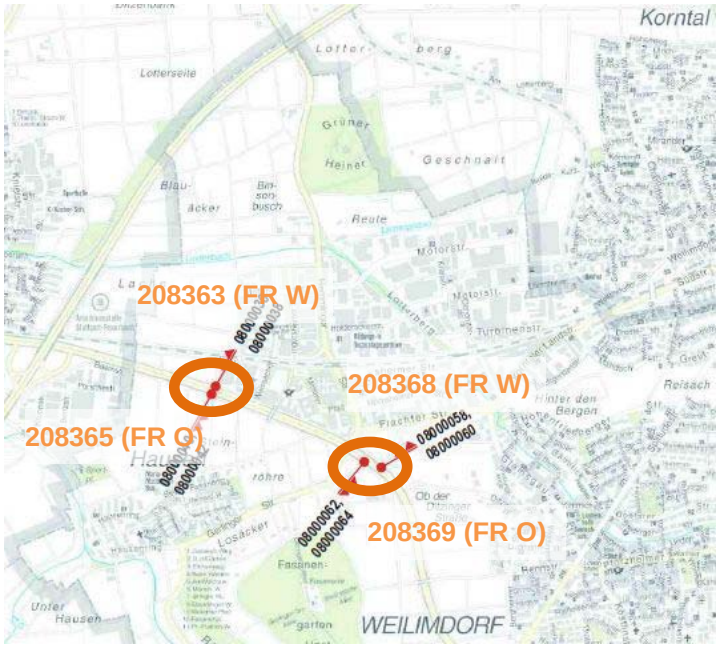


Abbildung 31: Lage der ausgewählten Sensoren der LHS – Zubringer B295 zu A81 (Kartenquelle: Stadt Stuttgart)

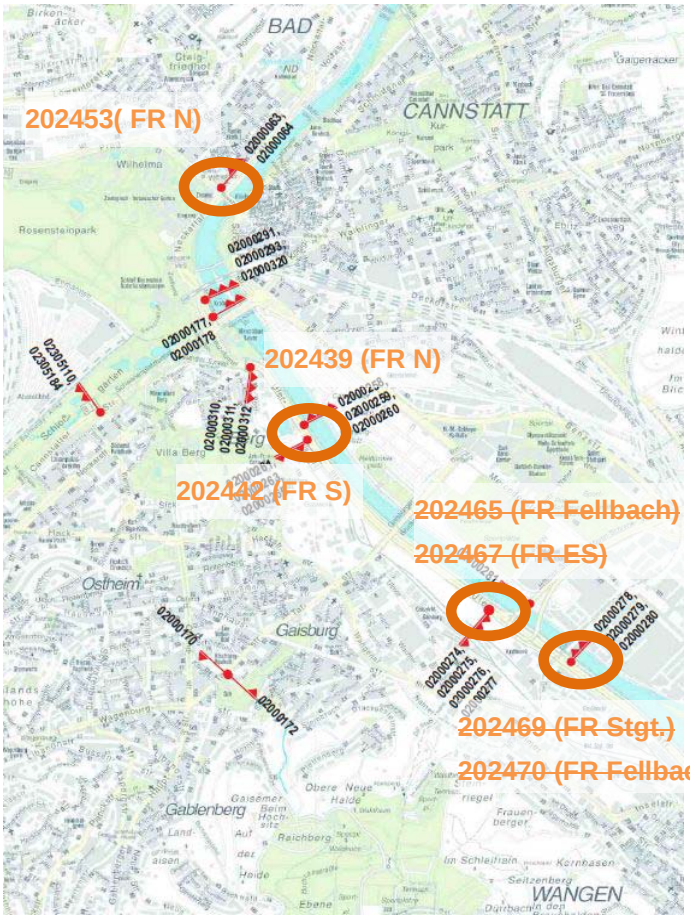


Abbildung 32: Lage der ausgewählten Sensoren der LHS – B10, Uferstraße (Kartenquelle: Stadt Stuttgart)

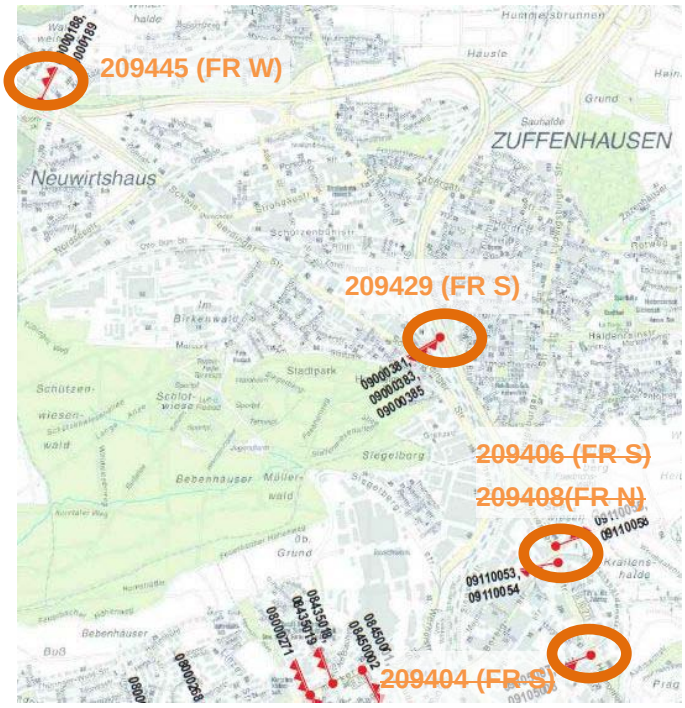


Abbildung 33: Lage der ausgewählten Sensoren der LHS – Bereich Pragsattel (Kartenquelle: Stadt Stuttgart)

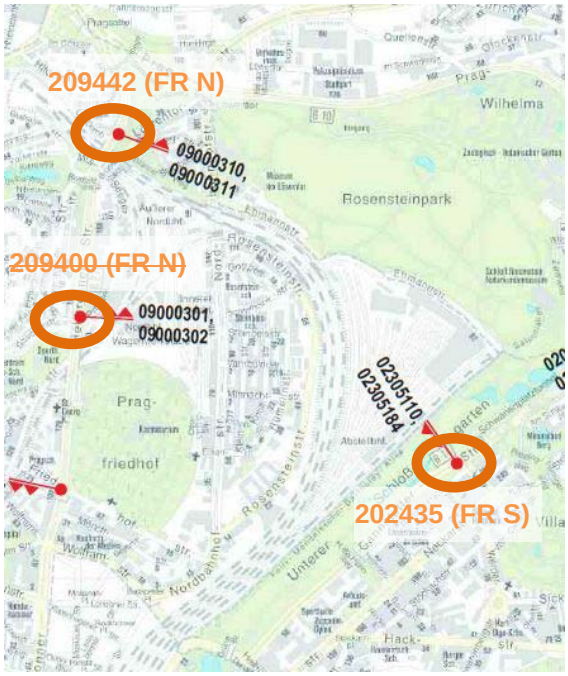


Abbildung 34: Lage der ausgewählten Sensoren der LHS – Bereich Nordbahnhof (Kartenquelle: Stadt Stuttgart)

Es wurden Sensoren auf den innerstädtischen Haupttransversalen wie B10/B14 sowie einige Stuttgarter Ein- und Ausfallstraßen wie die B295 gewählt, da auf diese überregionale Bedeutung haben. Daher wurden zwei Sensoren auf dem Autobahnzubringer B295 bei Korntal (208363/65 und 208368/69), wie in Abbildung 31 dargestellt, sowie die 202439/42 und 202453 auf der B10 entlang des Neckarufers (vgl. Abbildung 32) ausgewählt. Die Sensoren an dem Autobahnzubringer B10/Zuffenhausen wurden ebenfalls genutzt (209445), um die Verteilung des Verkehrs auf die drei alternativen Zufahrten von Stuttgart-Zentrum zur Autobahn A81 nach Norden (über AS Feuerbach, AS Zuffenhausen oder AS Ludwigsburg) erfassen zu können.

Die Erfassung der Verkehrslage auf dem Pragsattel (209429, Abbildung 33) sowie auf dessen unmittelbarer Zufahrt (209442, Abbildung 34) ist auf Grund der enormen täglichen Verkehrslasten und der damit verbundenen Staugefahr aus verkehrstechnischer Sicht von enormer Bedeutung. Daher wurden hier die gezeigten Sensorstandpunkte ausgewählt. Aus demselben Grund wurde schließlich auch noch der Messquerschnitt 202435 auf der B14/Cannstatter Straße stadteinwärts in die Untersuchungen mit einbezogen (vgl. Abbildung 34).

Zur Evaluierung der Grenzen der Mobilfunkortung wurden auch 6 aufeinanderfolgende Messquerschnitte (MQ14-MQ19) auf der B14 im Heslacher Tunnel ausgewählt. Die ungefähre Lage im einspurigen Tunnel ist in der Abbildung 35 ersichtlich. Leider waren Ende März keine Daten der RMQ im Heslacher Tunnel verfügbar, daher sind alle RMQ-IDs gestrichen dargestellt.

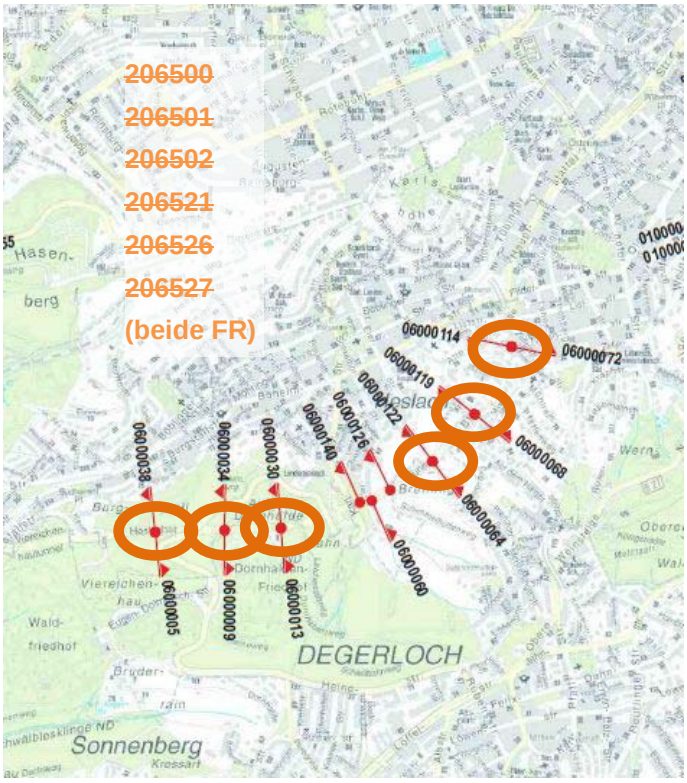


Abbildung 35: Lage der ausgewählten Sensoren der LHS – B14/Heslacher Tunnel (Kartenquelle: Stadt Stuttgart)

Die Messdaten der ausgewählten Stationären Erfassungssysteme der Projektpartner standen weitgehend vollständig ab Februar 2009 täglich zur Verfügung. Als Evaluierungszeitraum wurde der März 2009 festgelegt, da in diesem Zeitraum der endgültige Probelauf der A- und Abis-Datenverarbeitung stattgefunden hat und die letzte Softwareversion zur A-Daten Generierung in Fellbach installiert war.

Da der Raum Stuttgart direkt von dem Ausfall der 20 LACs ab dem 24.03. betroffen war, konnten hier für die restliche Zeit des Untersuchungszeitraumes keine FPD-Trajektorien mehr berechnet werden. Der Vergleich mit den Sensoren der Stadt Stuttgart beschränkt sich daher auf Montag, den 23.03.09.

2.3.1 Vergleich der FPD-Trajektorien aus Abis-Daten mit SES

Zunächst werden die FPD-Trajektorien, die aus Abis-Daten berechnet wurden, mit den SES-Daten verglichen. Das in Abbildung 25 grau dargestellte Gebiet, in dem A-Daten verfügbar waren, überdeckt den Raum Karlsruhe sowie Ettlingen und damit ein Stück der A8 sowie A5. Daher können neben den SES in Karlsruhe noch einige Sensoren der DDG sowie des Innenministeriums zum Vergleich herangezogen werden (vgl. Tabelle 6). Eine Übersicht über alle verfügbaren Sensoren mit ihrer Bezeichnung (RMQ_ID) ist in der Abbildung 36 dargestellt. Die drei Sensoren des Innenministeriums an der Autobahnabfahrt Ka-Durlach (RMQ-ID 800091, 800101 und 800102) haben keine Daten geliefert und werden daher im Weiteren nicht betrachtet. Dasselbe gilt für den Sensor 510421 der Stadt Karlsruhe, der in dem Zeitraum ebenfalls ausgefallen war.

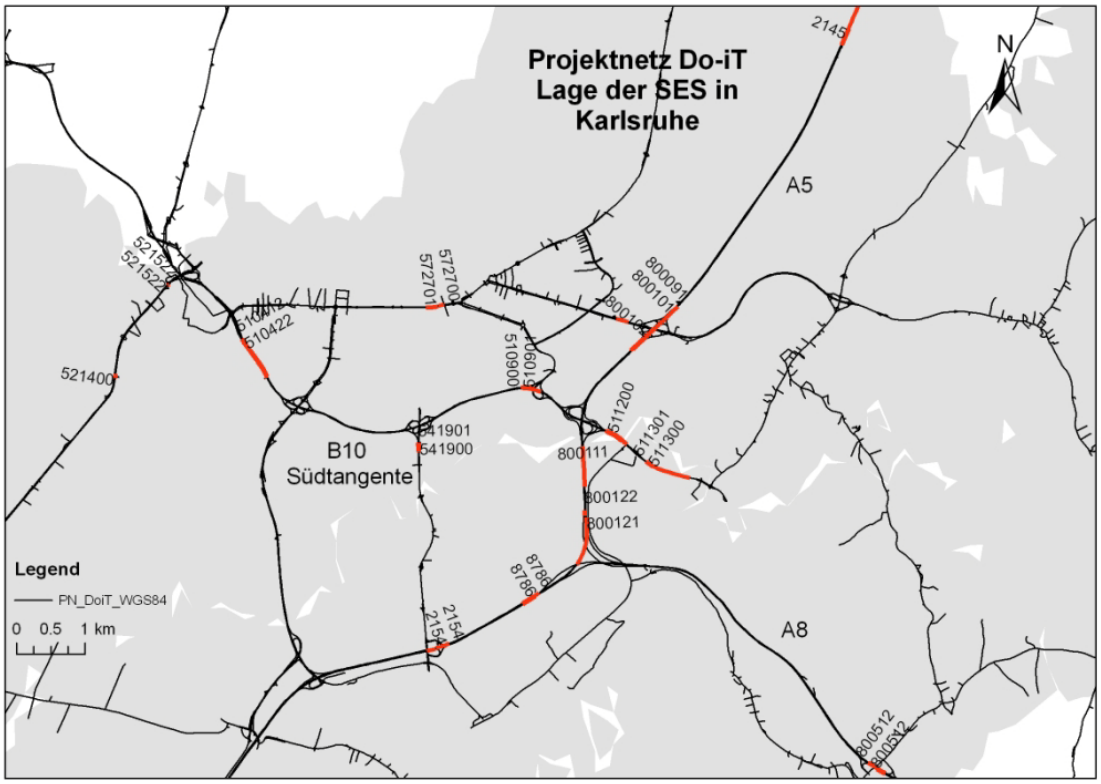


Abbildung 36: Übersicht aller Sensoren im Bereich Karlsruhe, in dem Abis-Daten verfügbar waren

Die Durchdringung mit FPD-Trajektorien kann an den SES-Standorten direkt durch Bildung des Quotienten aus detektierten Trajektorien und tatsächlicher Verkehrsstärke bestimmt werden. Dies erfolgte bei getrennten Fahrbahnen getrennt für beide Fahrtrichtungen und bei einspurigen Straßen als Summe aus beiden RMQ. Die Tabelle zeigt die Situation für die Untersuchungstage Ende März. Die Sensoren des IM haben am Sa. 28. und So. 29. keine Daten geliefert, daher sind die Felder in der Tabelle mit „inf“ belegt.

Tabelle 6: Durchdringung an den Sensoren im Abis-Testgebiet (LAC Ka und Ettlingen)

	RMQ	Mo. 23.03.	Di 24.03.	Mi 25.03.	Do 26.03.	Fr. 27.03.	Sa 28.03.	So 29.03.	Mo 30.03.	Di 31.03.	∅	Streuung	Stdw.
Sensoren der DDG	2145	0.21	0.22	0.33	0.24	0.17	0.15	0.20	0.20	0.24	**0.23**	**0.18**	**0.05**
	2146	0.10	0.13	0.19	0.18	0.13	0.08	0.12	0.13	0.16	**0.14**	**0.11**	**0.03**
	2153	0.19	0.26	0.36	0.19	0.21	0.16	0.15	0.30	0.28	**0.24**	**0.21**	**0.07**
	2154	0.18	0.22	0.20	0.27	0.16	0.10	0.11	0.22	0.23	**0.19**	**0.17**	**0.06**
	8785	0.28	0.33	0.34	0.25	0.27	0.21	0.23	0.32	0.34	**0.29**	**0.13**	**0.05**
	8786	0.24	0.28	0.26	0.36	0.23	0.10	0.17	0.29	0.30	**0.25**	**0.26**	**0.08**
Sensoren der Stadt Karlsruhe	510422	0.15	0.12	0.13	0.13	0.12	0.08	0.13	0.12	0.12	**0.12**	**0.06**	**0.02**
	510900	0.20	0.20	0.17	0.16	0.15	0.15	0.16	0.25	0.20	**0.19**	**0.10**	**0.03**
	510901	0.21	0.18	0.20	0.19	0.18	0.21	0.16	0.20	0.17	**0.19**	**0.05**	**0.02**
	511200	0.35	0.30	0.24	0.26	0.28	0.27	0.48	0.37	0.24	**0.31**	**0.25**	**0.08**
	511201	0.28	0.31	0.33	0.29	0.33	0.39	0.24	0.49	0.31	**0.33**	**0.25**	**0.07**
	511300	0.13	0.08	0.12	0.04	0.10	0.08	0.02	0.12	0.11	**0.09**	**0.11**	**0.04**
	511301	0.08	0.00	0.05	0.01	0.04	0.11	0.02	0.08	0.10	**0.06**	**0.11**	**0.04**
	521400	0.00	0.00	0.00	0.00	0.02	0.02	0.01	0.00	0.00	**0.00**	**0.02**	**0.01**
	521522	0.07	0.05	0.05	0.05	0.07	0.04	0.04	0.06	0.07	**0.05**	**0.03**	**0.01**
	541900	0.17	0.17	0.21	0.14	0.09	0.10	0.19	0.25	0.10	**0.16**	**0.15**	**0.06**
	541901	0.15	0.15	0.13	0.09	0.06	0.10	0.18	0.22	0.17	**0.15**	**0.16**	**0.05**
	552422	0.22	0.24	0.29	0.26	0.19	0.17	0.37	0.26	0.24	**0.26**	**0.20**	**0.06**
	572700	0.05	0.02	0.05	0.03	0.02	0.06	0.09	0.04	0.03	**0.04**	**0.07**	**0.02**
	572701	0.03	0.04	0.02	0.02	0.01	0.06	0.06	0.03	0.05	**0.04**	**0.05**	**0.02**
Sensoren des IM	800111	0.15	0.18	0.24	0.18	0.17	Inf	Inf	0.30	0.16	**0.20**	**0.15**	**0.05**
	800121	0.16	0.22	0.32	0.21	0.19	Inf	Inf	0.33	0.19	**0.24**	**0.17**	**0.07**
	800122	0.17	0.17	0.27	0.28	0.17	Inf	Inf	0.23	0.20	**0.22**	**0.11**	**0.05**
	800511	0.11	0.11	0.16	0.08	0.08	Inf	Inf	0.13	0.13	**0.12**	**0.08**	**0.03**
	800512	0.16	0.17	0.20	0.13	0.11	Inf	Inf	0.25	0.12	**0.17**	**0.13**	**0.05**

Wie der Tabelle 6 entnommen werden kann, schwankt die durchschnittliche Durchdringung zwischen 0,4 und 3,3 ‰ (ohne SES 521400, s. u.). Die Streuung liegt bei 0,2 bis 2,6 ‰ und die Standardabweichung der Durchdringung je Sensor ist mit maximal 0,8 ‰ relativ klein.

Über den Sensor 521400 auf der B36 am Westrand der LAC Ka laufen nahezu keine FPD-Trajektorien. Dies liegt vermutlich an der Lage am Rand des Gebiets, in dem Abis-Daten verfügbar waren. Der benachbarte Sensor 521522 zeigt mit 0,4 ‰ ebenfalls einen sehr geringen Wert. Daher ergibt sich für den Sensor 521400 auch eine unrealistisch kleine Standardabweichung, die nicht weiter beurteilt wird.

Die Sensoren auf der Autobahn zeigen hingegen eine durchweg höhere Durchdringung von 1,9 bis 2,9 ‰ (SES der DDG: 2145/2146, 2153/2154 sowie 8785/8786 und SES des IM: 800111, 800121/122 sowie 800511/512) bei einer Streuung von 1,3 bis 2,6 ‰

und einer Standardabweichung kleiner 0,8 ‰. Die etwas niedrigeren Werte der SES 2146 und 800511/512 erklären sich durch die Lage am Rand des mit Abis-Daten abgedeckten Gebiets. In das Gebiet einfahrende Fahrzeuge können erst nach Zurücklegen einer Mindeststrecke mit aktivem Mobiltelefon prozessiert werden. Auf der kreuzungsfreien vierspurig ausgebauten Südtangente kann ebenfalls 2 bis 3 ‰ des Verkehrs detektiert werden (vgl. SES 510900/901 und 511200/201 in der Tabelle 6). Am benachbarten Sensor Nummer 510422 zeigt sich nur eine Durchdringung von 1,2 ‰, die aber auf die Lage des Sensors zurückzuführen ist. Der Sensor liegt am Rand des Abis-Aufnahmegebiets und es handelt sich um die Fahrrichtung Ost stadteinwärts. Damit steht nur ca. 1 km Fahrtstrecke vom Rand der LAC-Ka bis zum Sensor zur Verfügung.

Über die beiden Sensoren auf der Kriegstraße (Nr. 572700/01) im Stadtzentrum von Karlsruhe werden jeweils nur ca. 0,4 ‰ der tatsächlichen Verkehrsstärke mit FPD erfasst, es ergibt sich eine Streuung von 0,7 bzw. 0,5 ‰ sowie eine Standardabweichung von 0,2 ‰. Dies liegt an dem zu geringen Auflösungsvermögen der Abis-Daten in der eng bebauten Innenstadt. Es werden Nahe der Innenstadt in der Regel nur kurze Strecken zurückgelegt und dabei aufgrund des besondere Aufmerksamkeit erfordernden Innenstadtverkehrs weniger und vor allem kürzere Telefonate als auf kreuzungsfreien Durchgangsstraßen und auf Autobahnen durchgeführt. Die geringe Standardabweichung ergibt sich durch die relativ konstante niedrige Durchdringung an unterschiedlichen Tagen.

Die folgenden drei Abbildungen stellen exemplarisch die Tagesganglinien der detektierten Verkehrsstärke aus Abis-Trajektorien und aus SES-Daten ausgewählter Sensoren für den 23.03.09 dar.

Die beiden oberen Schaubilder der Abbildung 37 zeigen recht ähnliche Verläufe der Verkehrsstärken aus Abis und aus den Sensordaten. Das untere Schaubild hingegen zeigt deutlich, dass das Verkehrsaufkommen im Innenstadtbereich mit vielen Straßen und Kreuzungen auf engem Raum nur schwer mit Mobilfunkdaten erfassbar ist. Es wurden zu wenige Trajektorien berechnet, um eine Hochrechnung der Verkehrslage zu ermöglichen. Dieser Eindruck bestätigt sich auch bei der Betrachtung der Rangkorrelation der beiden Tagesganglinien über den gesamten Untersuchungszeitraum. Hier ergeben sich im Mittel Korrelationen von ca. -0,1 für die beiden Sensoren in der Innenstadt auf der Kriegsbergstraße (vgl. Tabelle 7).

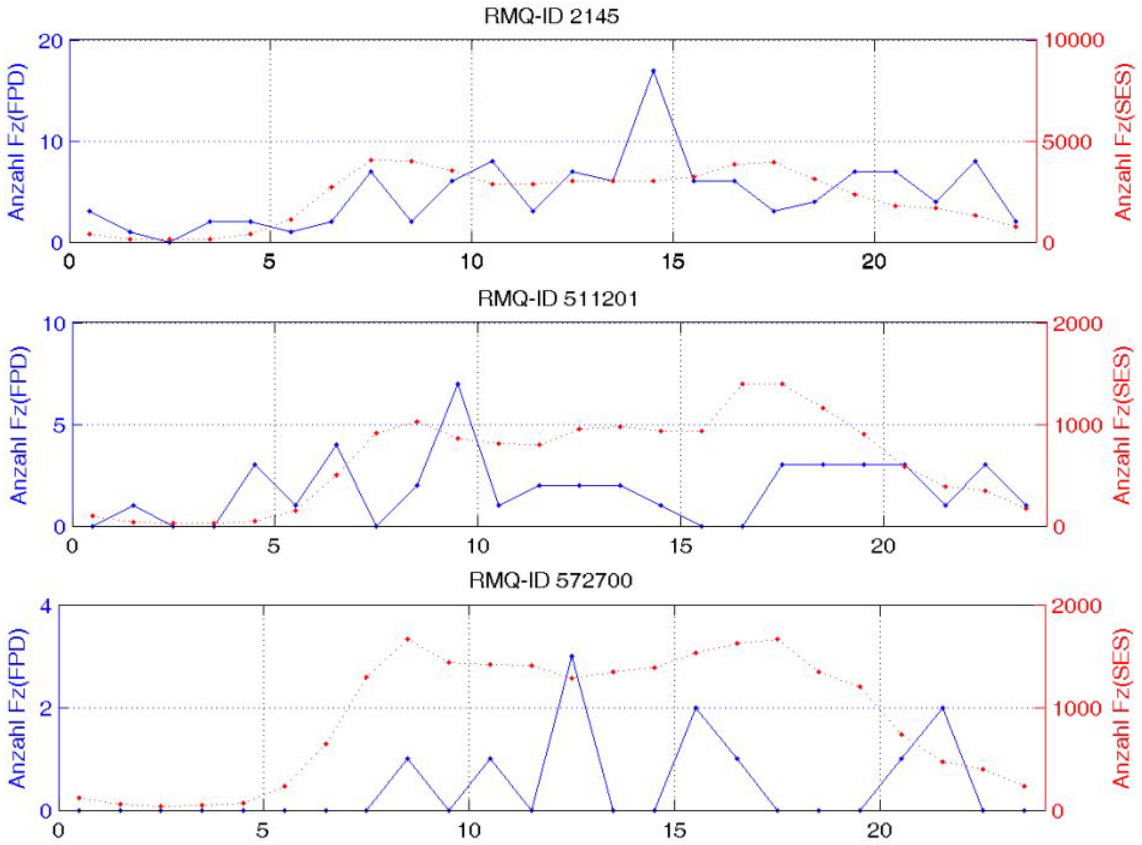


Abbildung 37: Tagesganglinie der realen sowie der aus Abis-Daten ermittelten Verkehrsstärke am Sensor 2145 (Autobahn A5), 511201 (Südtangente) sowie 572700 (Kriegstraße, Innenstadt) am 23.03.09

Der Zusammenhang zwischen den beiden Tagesganglinien wird in der folgenden Tabelle 7 durch den Qualitätsparameter Rangkorrelation quantitativ beschrieben.

Tabelle 7: Rangkorrelationen zwischen den Tagesganglinien aus Abis-Trajektorien und den Sensormessungen

	RMQ	Mo. 23.03	Di 24.03.	Mi 25.03.	Do 26.03.	Fr. 27.03.	Sa 28.03.	So 29.03.	Mo 30.03.	Di 31.03.	∅	Streuung	Stdw.
Sensoren der DDG	**2145**	0.40	0.30	0.41	0.02	0.566	0.82	0.35	0.33	-0.13	**0.31**	**0.94**	**0.28**
	2146	0.06	0.03	-0.16	0.30	0.188	0.38	-0.10	0.22	0.31	**0.13**	**0.55**	**0.19**
	2153	-0.29	-0.10	0.10	0.15	0.326	0.36	0.56	0.13	0.44	**0.17**	**0.84**	**0.27**
	2154	0.09	0.05	-0.46	-0.14	-0.595	0.39	0.17	0.36	0.29	**0.09**	**0.98**	**0.35**
	8785	0.24	-0.06	0.10	0.09	0.278	0.42	0.35	0.32	0.59	**0.26**	**0.65**	**0.19**
	8786	0.53	0.16	-0.53	0.01	-0.48	-0.08	-0.10	0.31	0.36	**0.08**	**1.06**	**0.36**
Sensoren der Stadt Karlsruhe	**510422**	0.44	0.55	0.46	0.44	-0.146	-0.21	-0.29	0.38	0.07	**0.23**	**0.83**	**0.33**
	510900	0.39	0.41	0.52	0.24	0.636	0.32	0.01	0.48	0.08	**0.31**	**0.62**	**0.20**
	510901	0.52	0.47	0.72	0.16	0.235	0.35	0.28	0.37	0.30	**0.40**	**0.56**	**0.17**
	511200	-0.26	0.23	-0.23	0.17	-0.017	0.29	0.29	0.41	-0.36	**0.07**	**0.77**	**0.28**
	511201	0.25	0.35	0.13	0.24	0.035	0.29	-0.29	0.05	0.16	**0.15**	**0.64**	**0.19**
	511300	0.40	NaN	-0.71	NaN	-0.632	NaN	NaN	0.07	-0.41	**-0.16**	**1.11**	**0.48**
	511301	NaN	NaN	NaN	NaN	NaN	0.66	NaN	0.89	-0.21	**0.45**	**1.10**	**0.58**
	521400	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	**0.00**	**0.00**	**NaN**
	521522	-0.19	0.46	0.10	0.25	-0.09	0.26	NaN	-0.22	0.26	**0.13**	**0.69**	**0.25**
	541900	0.08	-0.17	0.33	0.47	0.242	0.83	-0.26	0.09	0.50	**0.23**	**1.09**	**0.34**
	541901	-0.36	-0.33	-0.02	0.00	NaN	0.29	-0.34	-0.03	0.55	**-0.03**	**0.91**	**0.32**
	552422	0.51	0.58	0.18	0.05	0.253	-0.12	0.49	0.38	0.04	**0.26**	**0.70**	**0.24**
	572700	-0.42	NaN	-0.58	-0.71	-0.775	0.81	-0.41	0.45	0.00	**-0.12**	**1.59**	**0.57**
	572701	0.26	0.71	0.00	-0.87	NaN	-0.41	NaN	NaN	NaN	**-0.06**	**1.57**	**0.60**
Sensoren des IM	**800111**	0.27	0.50	0.18	0.04	0.003	NaN	NaN	0.31	0.21	**0.25**	**0.49**	**0.17**
	800121	0.00	-0.01	0.00	-0.16	0.067	NaN	NaN	0.43	0.15	**0.07**	**0.59**	**0.19**
	800122	0.30	0.02	-0.17	-0.25	-0.176	NaN	NaN	0.04	0.53	**0.08**	**0.78**	**0.28**
	800511	0.25	-0.04	0.30	0.01	-0.015	NaN	NaN	0.04	0.01	**0.10**	**0.35**	**0.14**
	800512	0.22	0.05	-0.01	0.04	0.017	NaN	NaN	-0.10	-0.18	**0.00**	**0.40**	**0.13**

Die Tabelle 7 zeigt die berechneten Rangkorrelationen zwischen den Tagesganglinien der Verkehrsstärke aus Abis-Daten und der SES. Die Werte „NaN“ bedeuten es war keine Berechnung der Rangkorrelation möglich. Dies war einerseits der Fall, wenn keine oder fast keine Trajektorien aus FPD-Daten berechnet werden konnten (SES 611300, 611301, 621400, 621622 sowie 672701 der Stadt Karlsruhe). Andererseits war bei den SES des Innenministeriums am Sa. 28.03. sowie So. 29.03. aufgrund fehlender Sensordaten keine Berechnung der Korrelation möglich.

Die Korrelationen erreichen für Abis-Trajektorien maximale Werte von bis zu 0,5. Gleichzeitig liegt die Streuung jedoch meist bei über 0,5, bei 6 von 25 SES sogar über 1,0. Die Standardabweichung erreicht Werte von 0,1 bis 0,6. Im Falle des Sensors 511301 scheint jedoch eine zufällige Anordnung bzw. Häufung der wenigen berechen-

baren Trajektorien zu der relativ hohen Korrelation zu führen (nur ca. 0.06 ‰ bzw. 4 der mittleren Verkehrsstärke von 7200 Fahrzeugen in 24h werden detektiert). Die absoluten mittleren Verkehrsstärken während des Untersuchungszeitraums sind in Tabelle 8 für alle Sensoren gegenübergestellt.

Tabelle 8: Mittlere Verkehrsstärken je 24h aller Sensoren in Ka (SES und Abis)

RMQ-ID	Q_SES	Q_FPD
2145	49963	112
2146	47898	65
2153	33125	78
2154	31018	59
8785	32831	95
8786	26699	67
510422	48105	59
510900	33560	62
510901	40471	77
511200	15204	48
511201	14573	48
511300	7186	6
511301	7206	4

RMQ-ID	Q_SES	Q_FPD
521400	11315	1
521522	19101	10
541900	12191	20
541901	11680	17
552422	23280	60
572700	20765	9
572701	16139	6
800111	51120	103
800121	24871	59
800122	25397	56
800511	29948	36
800512	28858	49

Es können offensichtlich derzeit nur an wenigen Sensoren mehr als 4 Trajektorien pro Stunde oder 96 Trajektorien am Tag berechnet werden (vgl. Tabelle 8). Daraus erklären sich trotz der guten Qualität der Trajektorien die relativ schlechten Durchdringungen verbunden mit hohen Streuungen und die teilweise mangelhafte Korrelation mit der realen Verkehrsstärke insbesondere abseits der Autobahn und der kreuzungsfreien Südtangente. Signifikante Schwankungen der Verkehrsstärke sind daher im untergeordneten Straßennetz kaum feststellbar.

Eine weitere Ursache für die mäßige Korrelation der Abis-Daten auf der Südtangente sowie der Autobahn liegt mit hoher Wahrscheinlichkeit auch in dem tageszeitabhängigen Telefonierverhalten der Verkehrsteilnehmer begründet. Die Auswertbarkeit der Teilnehmer hängt hochgradig von der Anzahl und Länge der Telefonate ab, die in den Fahrzeugen geführt wurden. Die Gegenüberstellung der Tagesganglinien zur Beurteilung der Hochrechenbarkeit der Verkehrstärke geht bislang jedoch von einem konstanten Telefonierverhalten aus. Diese Annahme ist falsch und müsste bei der Berechnung der Korrelation Berücksichtigung finden. Dazu sind jedoch umfassende Untersuchungen des Telefonierverhaltens erforderlich. Nur so ist es möglich, diesen Effekt aus den Tagesganglinien der FPD-Trajektorien zu eliminieren.

2.3.2 Vergleich der FPD-Trajektorien aus A-Daten mit SES

Aufgrund der großen Anzahl von 173 Referenzsensoren im gesamten Projektnetz, werden die Vergleiche für die Ballungszentren Großraum Karlsruhe und Stuttgart sowie für den Rest des Projektnetzes getrennt durchgeführt. Zunächst erfolgt der Vergleich der Trajektorien aus A-Daten mit den stationären Erfassungssystemen in Karlsruhe. Es werden dabei dieselben SES herangezogen, wie zum Vergleich mit den Trajektorien aus Abis im vorigen Kapitel 2.3.1. Wie bereits mehrfach erwähnt, konnten am Samstag, den 28.03. keine Trajektorien aus A-Daten berechnet werden, daher fehlt in der Tabelle diese Spalte. Am Sonntag, den 29.03. standen die Sensoren des Innenministeriums nicht zur Verfügung.

Tabelle 9: Durchdringung an den Sensoren mit A-Daten im Raum Ka

	RMQ	Mo. 23.03	Di 24.03.	Mi 25.03.	Do 26.03.	Fr 27.03.	So 29.03.	Mo 30.03.	Di 31.03.	∅	Streuung	Stdw.
Sensoren der DDG	2145	0.57	0.83	1.20	0.93	0.75	0.30	0.82	0.86	**0.78**	**0.90**	**0.26**
	2146	0.71	1.09	1.20	1.03	0.83	0.44	0.95	1.29	**0.94**	**0.86**	**0.28**
	2153	0.33	0.55	0.66	0.48	0.40	0.23	0.35	0.54	**0.44**	**0.42**	**0.14**
	2154	0.49	0.64	0.67	0.78	0.52	0.32	0.62	0.73	**0.60**	**0.46**	**0.15**
	8785	0.39	0.63	0.61	0.59	0.51	0.30	0.45	0.66	**0.52**	**0.36**	**0.13**
	8786	0.63	0.86	0.83	0.95	0.65	0.42	0.83	0.98	**0.77**	**0.56**	**0.19**
Sensoren der Stadt Karlsruhe	510422	0.44	0.61	0.71	0.67	0.65	0.29	0.57	0.63	**0.57**	**0.42**	**0.14**
	510900	0.43	0.50	0.57	0.53	0.43	0.24	0.54	0.53	**0.47**	**0.33**	**0.11**
	510901	0.35	0.53	0.55	0.48	0.51	0.32	0.51	0.53	**0.47**	**0.23**	**0.09**
	511200	0.70	0.86	0.91	0.99	0.96	0.36	1.01	0.98	**0.85**	**0.65**	**0.22**
	511201	0.29	0.45	0.45	0.35	0.35	0.14	0.37	0.41	**0.35**	**0.31**	**0.10**
	511300	0.05	0.07	0.06	0.09	0.16	0.00	0.05	0.09	**0.07**	**0.16**	**0.05**
	511301	0.03	0.09	0.05	0.05	0.05	0.02	0.10	0.03	**0.05**	**0.08**	**0.03**
	521400	0.07	0.11	0.17	0.19	0.07	0.17	0.10	0.12	**0.13**	**0.12**	**0.05**
	521522	0.05	0.06	0.11	0.11	0.06	0.12	0.07	0.08	**0.08**	**0.07**	**0.03**
	541900	0.01	0.01	0.02	0.01	0.01	0.03	0.03	0.00	**0.01**	**0.03**	**0.01**
	541901	0.01	0.02	0.00	0.01	0.00	0.01	0.02	0.00	**0.01**	**0.02**	**0.01**
	552422	0.09	0.14	0.20	0.14	0.13	0.13	0.19	0.19	**0.15**	**0.10**	**0.04**
	572700	0.02	0.04	0.06	0.06	0.05	0.02	0.05	0.05	**0.04**	**0.04**	**0.02**
	572701	0.04	0.05	0.10	0.05	0.08	0.04	0.08	0.05	**0.06**	**0.06**	**0.02**
Sensoren des IM	800111	0.44	0.66	0.91	0.79	0.68	Inf	0.85	0.62	**0.71**	**0.47**	**0.16**
	800121	0.22	0.42	0.46	0.40	0.40	Inf	0.42	0.40	**0.39**	**0.24**	**0.08**
	800122	0.30	0.40	0.60	0.55	0.38	Inf	0.63	0.50	**0.48**	**0.33**	**0.12**
	800511	0.95	1.25	1.45	1.29	1.06	Inf	1.39	1.12	**1.22**	**0.50**	**0.18**
	800512	0.97	1.36	1.88	1.71	1.35	Inf	1.74	1.32	**1.47**	**0.91**	**0.32**

Die Durchdringung an den Sensorstandpunkten in Karlsruhe erreicht mit A-Daten bis zu 1,5 % bei einer Streuung von bis zu 0,9 %. Die Standardabweichungen liegen im Bereich von maximal 0,3. Innerorts sinken die Werte für die Durchdringung erwartungsgemäß teilweise erheblich bis 0,4 ‰ ab und liegen damit auf dem Niveau der dazugehörigen Streuungen (vgl. Tabelle 9). An den beiden Sensoren mit der RMQ-ID 541900 und 541901 südlich der Stadt vor der Einmündung in die Südtangente, können aus ungeklärter Ursache kaum Trajektorien berechnet werden. Auf der Südtangente (B10) werden ca. 4 bis 9 ‰ der Verkehrsstärke erreicht (SES 10900/901 und 11200/201) jedoch ist die zugehörige Streuung teils erheblich kleiner als die Durchdringung und die Standardabweichung erreicht maximal 2 ‰. Eine Ausnahme bilden die beiden Sensoren 511300/301 östlich der A5, die nur ca. 1 ‰ erreichen und deren Streuung den 1,5 bis 2-fachen Wert der Durchdringung annehmen.

Die beiden Sensoren auf der Kriegstraße im Stadtzentrum (SES 572700/01) zeigen auch für A-Daten erwartungsgemäß eine niedrige Durchdringungen von 0,4 und 0,6 ‰,. Dies liegt ebenso wie bei den Trajektorien aus Abis-Daten an dem begrenzten räumlichen Auflösungsvermögen der A-Daten, mit denen die hohe Infrastrukturdichte in Innenstädten nur schwer abgebildet werden kann.

Im Bereich der Autobahnen (SES der DDG sowie des IM) werden durchweg Durchdringungen von 4 ‰ und mehr erreicht bei Streuungen, die in den meisten Fällen wesentlich kleiner als die Durchdringung sind. Die Standardabweichungen liegen bei 1 bis 3 ‰. Hier funktioniert die Mobilfunkortung auf A-Daten-Ebene trotz der Stadtnähe erwartungsgemäß am besten.

Insgesamt zeigt sich ein überwiegend homogenes Bild der Durchdringungen. In wieweit diese Durchdringungen auf die Extrapolierbarkeit der gesamten Verkehrsstärke hindeutet, wird mit Hilfe der Korrelationskoeffizienten näher untersucht. Die folgenden drei Schaubilder in Abbildung 38 stellen exemplarisch die Tagesganglinien der detektierten Verkehrsstärken aus A-Trajektorien und aus SES-Daten ausgewählter Sensoren für den 23.03.09 dar.

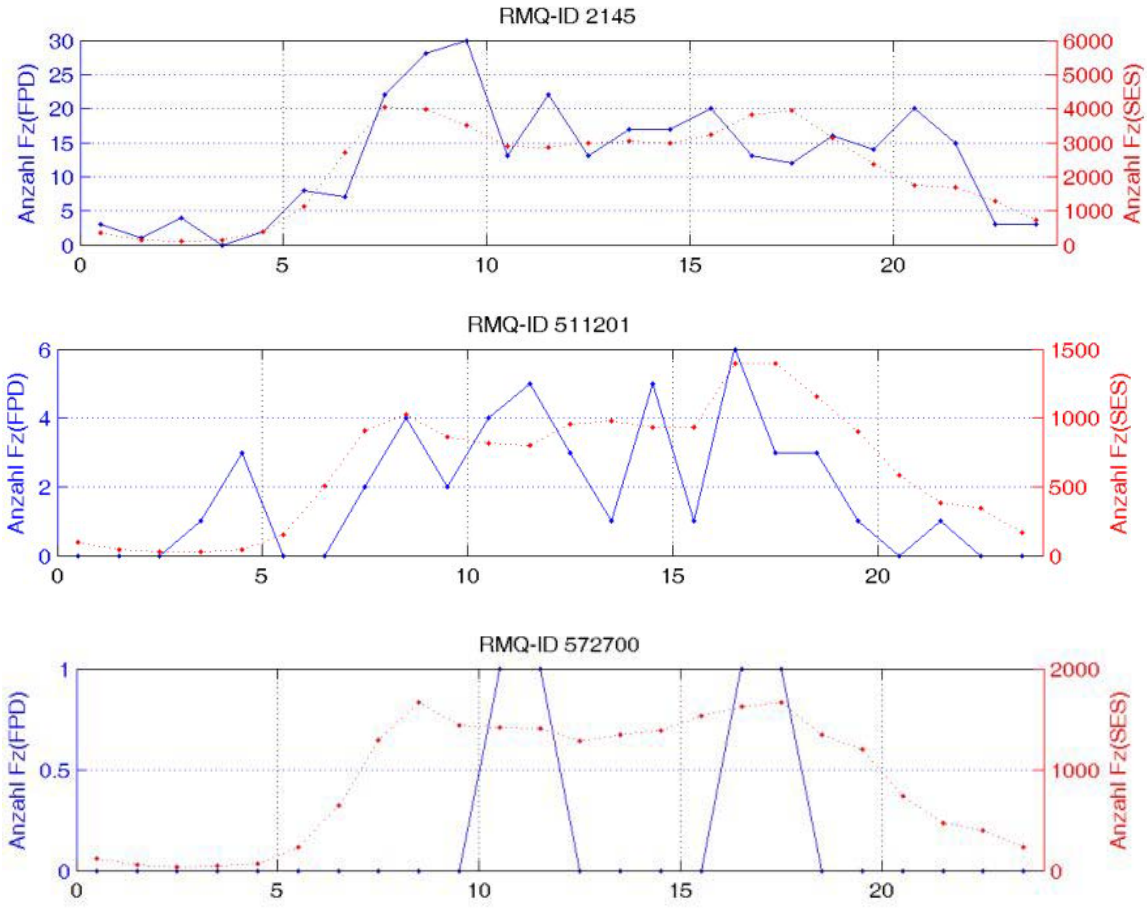


Abbildung 38: Tagesganglinie der realen sowie der aus A-Daten ermittelten Verkehrsstärke am Sensor 2145 (Autobahn A5), 511201 (Südtangente) sowie 572700 (Kriegstraße, Innenstadt) am 23.03.09

In der oberen Grafik der Abbildung 38 ist der tendenzielle Verlauf der Verkehrsstärke über den ganzen Tag klar aus den FPD-Daten erkennbar. In der mittleren Grafik ist die Ähnlichkeit grundsätzlich noch erkennbar, in der unteren Grafik kann aufgrund der sehr geringen Gesamtzahl an Trajektorien (es wurden innerhalb 24 h lediglich 4 Trajektorien zugeordnet) nicht mehr auf den tatsächlichen Verlauf der Verkehrsstärke geschlossen werden.

Als Maß für die Ähnlichkeit der beiden Kurven, und damit für die Möglichkeit, die Verkehrsstärke aus den FPD-Trajektorien abzuleiten, werden im Folgenden die Rangkorrelationen für den gesamten Untersuchungszeitraum von 8 Tagen für alle Sensoren im Bereich Karlsruhe/Ettlingen sowie die Mittelwerte berechnet.

Tabelle 10: Rangkorrelationen zwischen den Tagesganglinien aus A-Trajektorien und den Sensormessungen im Raum Karlsruhe

	RMQ	Mo. 23.03	Di 24.03.	Mi 25.03.	Do 26.03.	Fr 27.03.	So 29.03.	Mo 30.03.	Di 31.03.	∅	Streuung	Stdw.
Sensoren der DDG	2145	0.74	0.89	0.70	0.73	0.83	0.38	0.83	0.82	0.74	0.52	0.16
	2146	0.86	0.79	0.86	0.83	0.83	0.70	0.85	0.87	0.82	0.17	0.06
	2153	0.48	0.52	0.55	0.64	0.61	0.83	0.59	0.80	0.63	0.34	0.13
	2154	0.68	0.85	0.76	0.79	0.82	0.48	0.79	0.86	0.75	0.38	0.12
	8785	0.62	0.64	0.86	0.70	0.62	0.79	0.57	0.77	0.70	0.29	0.10
	8786	0.75	0.87	0.81	0.74	0.82	0.52	0.79	0.84	0.77	0.35	0.11
Sensoren der Stadt Karlsruhe	510422	0.74	0.82	0.85	0.75	0.80	0.42	0.77	0.80	0.74	0.43	0.14
	510900	0.66	0.83	0.76	0.74	0.59	0.30	0.64	0.62	0.64	0.53	0.16
	510901	0.76	0.81	0.74	0.81	0.83	0.45	0.78	0.85	0.75	0.40	0.13
	511200	0.78	0.66	0.74	0.66	0.70	-0.20	0.75	0.49	0.57	0.98	0.32
	511201	0.30	0.69	0.70	0.47	0.58	0.60	0.56	0.32	0.53	0.40	0.15
	511300	NaN	NaN	NaN	-0.13	0.22	NaN	-0.87	0.66	-0.03	1.52	0.64
	511301	NaN	-0.74	NaN	0.87	NaN	NaN	NaN	NaN	0.06	1.60	1.13
	521400	0.51	0.66	0.76	0.13	0.25	0.04	0.08	-0.47	0.24	1.22	0.39
	521522	0.62	0.66	0.61	0.13	0.65	0.04	0.15	-0.07	0.35	0.73	0.31
	541900	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-	-	-
	541901	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-	-	-
	552422	-0.31	0.56	0.36	0.37	0.04	0.08	0.05	0.31	0.18	0.87	0.27
	572700	NaN	-0.39	0.23	0.57	0.00	NaN	0.63	0.41	0.24	1.03	0.39
	572701	0.00	0.26	-0.07	0.63	0.10	NaN	0.36	-0.08	0.17	0.71	0.26
Sensoren des IM	800111	0.85	0.92	0.50	0.73	0.57	NaN	0.72	0.78	0.72	0.42	0.15
	800121	0.35	0.45	0.41	0.54	0.36	NaN	0.47	0.75	0.47	0.41	0.14
	800122	0.51	0.68	0.44	0.74	0.66	NaN	0.54	0.82	0.63	0.37	0.13
	800511	0.86	0.88	0.56	0.94	0.78	NaN	0.66	0.82	0.78	0.38	0.13
	800512	0.91	0.90	0.62	0.85	0.64	NaN	0.75	0.88	0.79	0.28	0.12

Die Rangkorrelationen reichen auf der Südtangente sowie der Autobahn von 0,5 bis 0,8, wobei die Streuungen bis auf eine Ausnahme (SES 511200) kaum mehr als die Hälfte des Wertes beträgt. Auch die Standardabweichung zeigt mit Werten zwischen 0,1 und 0,2 (mit der eben genannten Ausnahme) sehr homogene Werte. Im untergeordneten Straßennetz liegen die Werte weit darunter. In zwei Fällen ist die Korrelation aufgrund der sehr geringen Anzahl von 1 berechneten Trajektorie/24 h nicht ermittelbar.

Wie zu erwarten war, bewegen sich die Korrelationen innerstädtisch abseits der kreuzungsfreien Südtangente im Bereich von nahe 0 bis 0,35 bei einer Streuung von dem doppelten bis vierfachen des Wertes und einer Standardabweichung von 0,3 bis 0,6 (vgl. Tabelle 10).

Für die Tagesganglinie in Abbildung 38 unten konnte am Montag, den 23. auf Grund der geringen Trajektorienanzahl keine Korrelation berechnet werden, für die weiteren Tage ergab sich an diesem Sensor eine sehr niedrige Korrelation von 0,24. Für den SES-Standort 511201 (auf der Südtangente/B10, östlich der Anschlussstelle Ka-Degerloch), exemplarisch in der Abbildung 38 mittig für den Montag, 23.03. dargestellt, lässt sich der Tabelle eine Korrelation von 0,30 entnehmen. Im Mittel lag diese an diesem Standort bei 0,53. Damit scheint eine Korrelation von größer 0,5 bereits ein Indiz zu sein, dass stärkere Änderungen der Verkehrsstärke (z. B. durch Stockungen oder durch Ausweichverkehr) mit FPD-Trajektorien detektierbar sind. Eventuell kann hier bereits auch schon eine Hochrechnung auf die reale Verkehrsstärke erfolgen, diese ist jedoch aufgrund der meist geringen absoluten Trajektorienanzahl abseits der stark befahrenen Autobahnen mit einer großen Ungenauigkeit belegt, da einzelne Trajektorien einen sehr großen Einfluss auf die Hochrechnung haben.

Für den ausgewählten Sensorstandort der DDG auf der Autobahn A5 nördlich von Karlsruhe (RMQ-ID 2145) ergab sich für den 23.03. eine Korrelation von 0,74, dies stellt auch den Mittelwert über den Messzeitraum dar. Die FPD-Trajektorien spiegeln den Verlauf der Verkehrsstärke bereits gut wieder, auch wenn die Abschwächung der Verkehrsstärke um 9 Uhr zu spät registriert wird und aus ungeklärter Ursache ein Peak um 21 Uhr auftaucht.

Es folgen nochmals 3 Beispiele, die eine bessere quantitative Beurteilung des Korrelationskoeffizienten zulassen. Es handelt sich um Beispiele bei denen die Verkehrsstärke gut mit den FPD-Trajektorien abgebildet werden kann. Die Korrelationen betragen für die vorgestellten Sensoren:

- 0,85 (RMQ 510422, Südtangente am Dienstag, 24.03.)
- 0,76 (RMQ 510900, Südtangente am Dienstag, 24.03.)
- 0,92 (RMQ 800111, A5 AS Ka-Degerloch am Mittwoch, 25.03.)

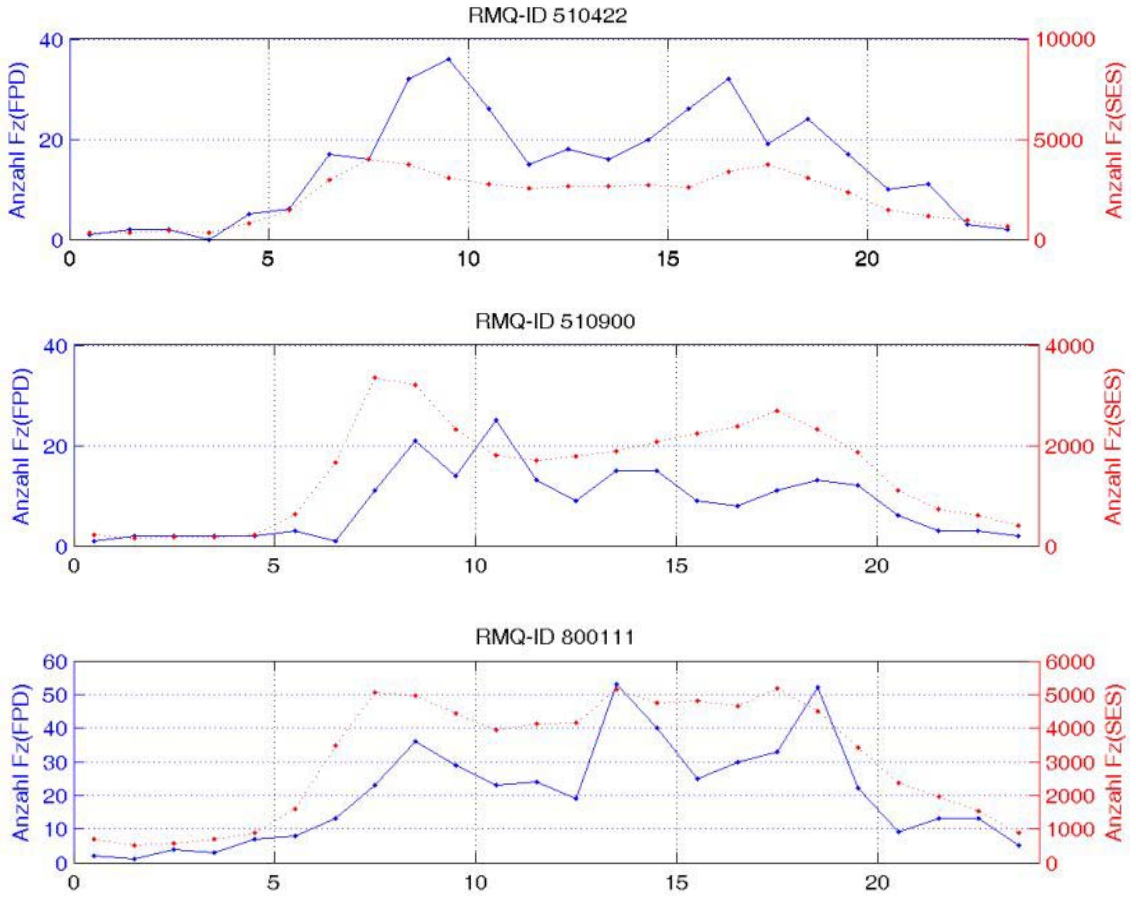


Abbildung 39: Darstellung ausgewählter Beispiele von Tagesganglinien: RMQ 510422, 510900 (Mittwoch, 25.03.09) und 800111 (Dienstag, 24.03.09)

Die Abbildung 39 zeigt abgesehen von einzelnen Ausreißern in den Tendenzen weitgehend übereinstimmende Verläufe der Verkehrsstärken aus FPD und aus SES. Die Korrelationen zeigen dies ebenfalls und sind mit 0,76 bis 0,92 hoch. Als erste Größenordnung zur Beurteilung der Tauglichkeit der FPD zur Erkennung von Änderungen der Verkehrsstärke durch die Korrelation wird daher auf Grund visueller Analyse die Grenze von 0,75 empirisch festgelegt. Es folgen 3 weitere Beispiele, die diesen Vorschlag untermauern. Es handelt sich um Beispiele mit Korrelationen im Bereich von 0,5 bis 0,7, die nur bedingt die wahren Tendenzen wiederspiegeln können. Die Grafiken in Abbildung 40 weisen von Oben nach Unten die folgenden Korrelationen auf:

- 0,64 (RMQ 8785, Autobahn A5 am Dienstag, 24.03.)
- 0,56 (RMQ 552422, Durlacher Allee am Dienstag, 24.03.)
- 0,45 (RMQ 800121, A5 AD Karlsruhe am Dienstag, 24.03.)

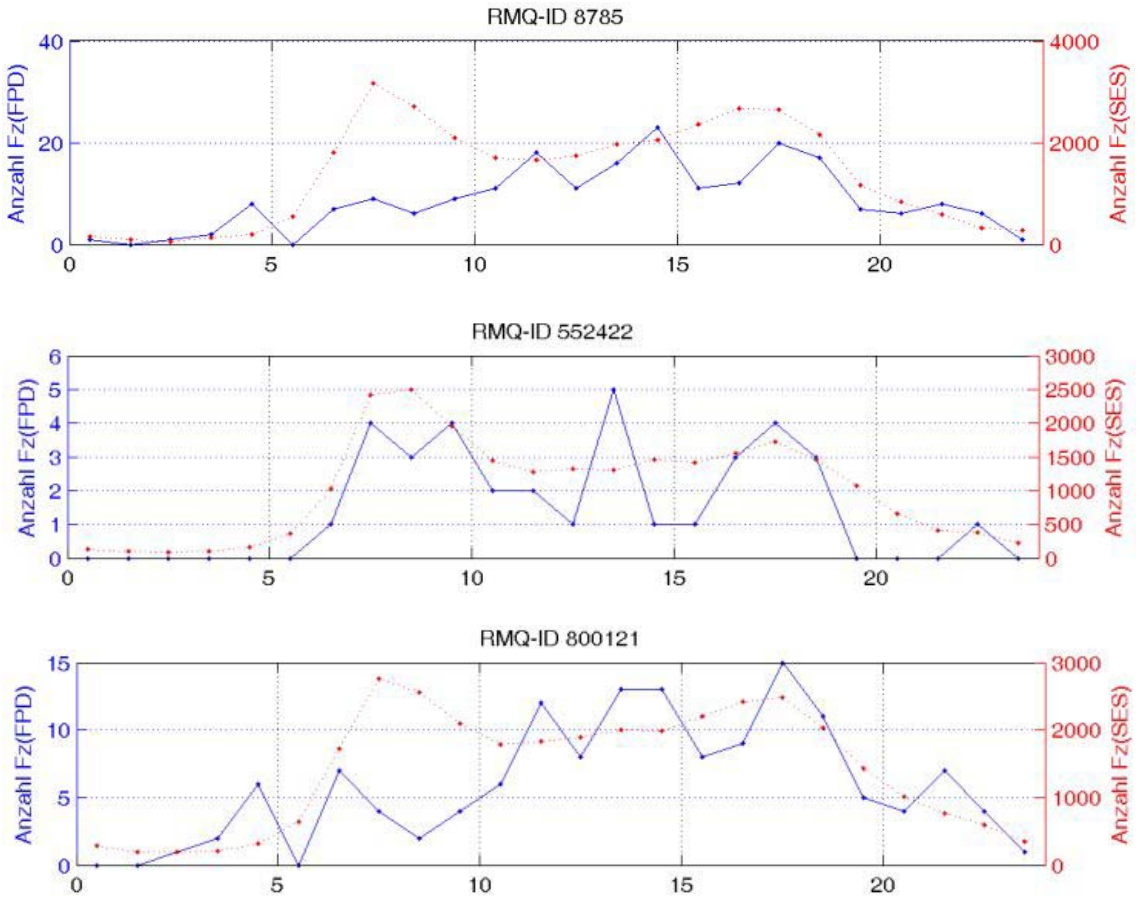


Abbildung 40: Darstellung ausgewählter Beispiele von Tagesganglinien: RMQ 8786, 552422 und 800121 (alle vom Dienstag, 24.03.09)

Korrelationen unter 0,5 deuten auf eine schlechte Eignung der FPD-Trajektorien am SES-Standort zur Beurteilung der Verkehrsstärke hin.

Die folgende Tabelle zeigt die absoluten Verkehrsstärken aller Sensoren im Raum Karlsruhe.

Die absolute Anzahl an berechneten FPD-Trajektorien aus A-Daten bestätigt nochmals die Erkenntnisse aus der Betrachtung der Korrelationen. Können nur wenige Trajektorien in 24 Stunden berechnet werden, so sind auch die Korrelationen gering oder gar nicht berechenbar. Dies gilt hier für die Straßenabschnitte auf denen die SES mit der RMQ-ID 611300/301, 521400, 521522, 541900/01 sowie 572700/01, für die weniger als durchschnittlich 1 Trajektorie/h berechnet werden konnte (vgl. Tabelle 11).

Tabelle 11: Mittlere Verkehrsstärken je 24h aller Sensoren in Ka (SES und FPD aus A)

RMQ-ID	Q_SES	Q_FPD
2145	50283	392
2146	48709	464
2153	33512	147
2154	31806	194
8785	33576	177
8786	27277	214
510422	50016	294
510900	34454	167
510901	41897	202
511200	15696	139
511201	15002	55
511300	7308	6
511301	7324	4

RMQ-ID	Q_SES	Q_FPD
521400	11579	14
521522	19547	16
541900	12399	2
541901	11881	1
552422	23496	36
572700	21021	9
572701	16370	10
800111	57510	414
800121	27980	113
800122	28572	140
800511	33692	421
800512	32465	488

Im Folgenden werden die Durchdringung sowie die Korrelationskoeffizienten der Sensoren im Bereich Stuttgart (vgl. Abbildung 30 ff) dargestellt. Aus Gründen der Übersichtlichkeit wird dabei auf die Darstellung von Beispielen zu den gegenübergestellten Tagesganglinien verzichtet.

Tabelle 12: Durchdringung mit FPD aus A-Daten in Stuttgart

RMQ	Mo. 23.03	Di 24.03.[3]	Mi 25.03.bis Di. 31.03.	∅ Mo. 23.03. bis Di. 24.03.	Streuung
202435	0.24	0.18	-	0.21	0.06
202439	0.20	0.11	-	0.16	0.09
202442	0.27	0.14	-	0.21	0.13
202453	0.20	0.11	-	0.16	0.09
208363	0.19	0.08	-	0.13	0.11
208365	0.30	0.28	-	0.29	0.02
208368	0.26	0.08	-	0.17	0.18
208369	0.30	0.28	-	0.29	0.02
209429	0.63	0.42	-	0.53	0.21
209442	0.04	0.03	-	0.04	0.01
209445	0.74	0.50	-	0.62	0.24

[3] Ab ca. 11 Uhr morgens Ausfall der LACs, daher wurden die Daten extrapoliert

Es können hier nur zwei Tage (23.03. und 24.03.) zum Vergleich herangezogen werden, da in den restlichen Tagen keine A-Daten mehr verfügbar waren. Ab Dienstag ca. 11 Uhr morgens sind alle LACs ausgefallen, die den Bereich Stuttgart abgedeckt haben.

Die Werte für die Durchdringungen liegen im Bereich von 0,4 bis 6,2 ‰ und damit im selben Bereich, der auch in Karlsruhe innerstädtisch erreicht wurde (vgl. Tabelle 12 und Tabelle 9).

Tabelle 13: Rangkorrelationen der FPD aus A mit den Sensordaten in Stuttgart (links) sowie ermittelte Verkehrsstärken aus SES und aus FPD (rechts) für Montag, den 23.03.09

RMQ	Mo. 23.03	Di 24.03.bis Di. 31.03.[4]
202435	0.56	-
202439	0.65	-
202442	0.50	-
202453	0.37	-
208363	0.16	-
208365	0.26	-
208368	0.30	-
208369	0.01	-
209429	0.55	-
209442	0.78	-
209445	0.83	-

RMQ	Q_SES [KFZ/24h]	Q_FPD [KFZ/24h]
202435	27217	64
202439	36543	74
202442	48768	132
202453	28980	59
208363	22184	41
208365	14902	44
208368	14630	38
208369	14846	44
209429	28242	179
209442	14138	6
209445	28209	209

Die Rangkorrelation erreicht im Stadtgebiet Stuttgart bis zu 0,83 bei einer Durchdringung von ca. 7 ‰ für den Sensor auf der B10 kurz vor der Autobahnauffahrt Zuffenhausen (vgl. RMQ-ID 209445 in Tabelle 13). Im Bereich der B295/ Autobahnauffahrt Feuerbach hingegen werden nur geringe Korrelationen von 0,16 und 0,30 erreicht (RMQ-ID 208363 und 208368). Es wurden hier lediglich ca. 40 Trajektorien innerhalb 24 h berechnet, die Durchdringung lag damit bei weniger als 2 ‰.

Die Korrelation des SES 209442 von 0,78 scheint zunächst hoch, allerdings zeigt die Durchdringung von 0,04 % bzw. 6 berechneten Trajektorien pro 24 h, das hier die Verkehrsstärke nicht durch FPD-Trajektorien ermittelbar ist. Offensichtlich kann die Eignung der FPD-Trajektorien zur Ermittlung der Verkehrsstärke nur in Kombination der Parameter **Durchdringung**, **Rangkorrelation** und der **absoluten Anzahl berechneter FPD-Trajektorien** beurteilt werden.

[4] Aufgrund des Ausfalls der Stuttgart abdeckenden LACs keine Angabe möglich

Für das restliche Projektnetz werden nur die Mittelwerte der Durchdringung und Korrelation für den gesamten Untersuchungszeitraum angegeben. Im Falle der SES, die in dem Bereich der ausgefallenen LACs lagen, wurden die Werte vom Montag, den 23.03. verwendet.

Tabelle 14: Durchschnittliche Durchdringung mit FPD aus A-Daten an den SES sowie Korrelation mit den SES der DDG im Autobahnviereck.

RMQ	D [%]	Streuung von D	Stdw. von D	r	Streuung von r	Stdw. von r	Q_SES [KFZ/d]	Q_FPD [KFZ/d]
2109	0.68	0.64	0.21	0.70	0.26	0.08	42276	293
2110	0.70	0.60	0.19	0.77	0.31	0.10	37294	262
2113	0.65	0.63	0.21	0.74	0.24	0.09	34767	229
2114	0.87	0.75	0.27	0.71	0.42	0.15	29572	258
2121	0.87	0.89	0.28	0.78	0.45	0.14	51529	451
2122	1.10	0.96	0.34	0.83	0.16	0.05	51731	576
2473	0.03	0.04	0.02	0.03	0.92	0.37	20862	7
2477	0.76	0.58	0.20	0.75	0.32	0.09	38381	299
2478	0.34	0.25	0.09	0.70	0.25	0.08	38655	134
2481	1.49	1.20	0.39	0.75	0.36	0.11	40068	608
2482	0.45	0.39	0.14	0.73	0.46	0.15	38264	176
2485	1.17	0.91	0.34	0.73	0.35	0.12	24489	294
2486	1.49	1.24	0.42	0.67	0.20	0.07	20127	308
2489	0.96	0.74	0.27	0.81	0.25	0.08	30436	299
2490	0.39	0.37	0.12	0.66	0.19	0.07	25716	103
2493	0.77	0.60	0.20	0.80	0.35	0.11	32082	252
2494	1.37	1.04	0.37	0.81	0.21	0.07	22151	310
2497	1.41	0.86	0.30	0.78	0.23	0.08	20044	289
2498	1.59	1.49	0.51	0.83	0.33	0.11	37476	606
2501	1.61	0.88	0.28	0.78	0.36	0.11	15603	256
2502	4.02	2.92	0.95	0.81	0.28	0.08	14157	583
2505	2.06	1.11	0.36	0.80	0.28	0.10	15720	331
2506	2.10	1.69	0.57	0.82	0.27	0.08	14751	316
2509	0.68	0.41	0.13	0.70	0.43	0.15	25466	177
2510	1.26	1.11	0.35	0.78	0.41	0.13	36066	464
2513	0.51	0.82	0.26	0.56	0.44	0.15	14487	74
2514	0.92	0.82	0.27	0.63	0.37	0.12	14681	139
2517	0.09	0.46	0.16	0.34	0.75	0.31	28712	26
2518	0.43	0.45	0.15	0.63	0.36	0.11	32039	139
2521	0.10	0.50	0.17	0.33	0.75	0.31	26386	25
2522	0.15	0.55	0.19	0.11	0.84	0.29	28095	43
2525*	0.47	-	-	0.57	-	-	25551	121
2526	0.31	1.15	0.41	0.13	0.92	0.29	26943	87
2529*	1.07	-	-	0.57	-	-	11205	120
2530*	1.32	-	-	0.64	-	-	12519	165
2533*	2.58	-	-	0.70	-	-	4419	114
2534*	0.86	-	-	0.66	-	-	19833	171
3405	1.11	0.89	0.28	0.86	0.18	0.06	40119	456
3406	6.66	8.22	2.68	0.72	0.50	0.15	5169	334
3409	0.92	0.96	0.28	0.82	0.20	0.07	36624	341
3410	0.72	0.66	0.20	0.76	0.45	0.14	44186	323
3417	0.35	1.62	0.60	0.37	0.96	0.32	36676	131
3418	0.17	0.91	0.33	0.15	1.25	0.61	39272	68

(*) Sensoren im Bereich der ab Dienstag, 24.03. ausgefallenen LACs (vgl. Abbildung 16 und Abbildung 17)

RMQ	D [%]	Streuung von D	Stdw. von D	r	Streuung von r	Stdw. von r	Q_SES [KFZ/d]	Q_FPD [KFZ/d]
3421	0.21	0.98	0.36	0.27	0.83	0.39	32260	67
3422	0.19	1.04	0.37	0.29	1.12	0.53	35754	68
3441	0.16	0.77	0.29	0.55	1.16	0.52	52155	82
3442	0.62	1.10	0.32	0.76	0.26	0.10	55124	349
3445	0.68	1.67	0.64	0.53	0.96	0.31	34545	229
3446	0.25	0.28	0.08	0.59	0.41	0.14	55094	143
3449	0.19	0.17	0.05	0.66	0.62	0.20	59273	115
3450	0.51	0.42	0.13	0.70	0.25	0.08	56751	294
3453	0.22	0.19	0.06	0.67	0.71	0.22	53612	120
3454	0.59	0.51	0.16	0.67	0.24	0.09	48925	295
3457	0.11	0.11	0.03	0.52	0.73	0.26	35511	39
3458	0.14	0.13	0.04	0.32	0.85	0.28	39535	56
7401*	0.68	-	-	0.64	-	-	37719	257
7402*	0.65	-	-	0.81	-	-	43010	280
7405*	1.09	-	-	0.70	-	-	35565	389
7406*	0.81	-	-	0.80	-	-	39573	321
7409*	1.05	-	-	0.69	-	-	38929	407
7410*	0.94	-	-	0.60	-	-	39631	374
7413*	1.01	-	-	0.68	-	-	43875	441
7414*	0.77	-	-	0.83	-	-	47814	366
7417*	2.00	-	-	0.71	-	-	39858	796
7418*	1.51	-	-	0.87	-	-	47316	716
7421*	1.18	-	-	0.74	-	-	33430	394
7422*	1.77	-	-	0.79	-	-	44766	794
7425*	1.06	-	-	0.81	-	-	35907	380
7426*	0.95	-	-	0.81	-	-	39959	381
7429*	0.79	-	-	0.72	-	-	46515	367
7430*	0.80	-	-	0.77	-	-	47972	383
7433*	4.18	-	-	0.74	-	-	34253	1432
7434*	1.87	-	-	0.85	-	-	33507	628
7438*	0.76	-	-	0.80	-	-	55380	422
8505	0.93	0.81	0.27	0.82	0.27	0.09	46734	443
8506	1.06	0.96	0.33	0.80	0.22	0.07	48659	516
8785	0.52	0.36	0.13	0.70	0.29	0.10	33576	177
8786	0.77	0.56	0.19	0.77	0.35	0.11	27277	214
8809	1.47	1.24	0.41	0.76	0.43	0.13	40394	605
8810	0.47	0.40	0.14	0.72	0.42	0.14	37659	180
8813	0.59	0.44	0.14	0.72	0.41	0.13	28079	167
8814	1.66	1.25	0.42	0.74	0.24	0.08	27213	458
8869	0.48	0.91	0.29	0.77	0.26	0.09	37623	181
8870	0.76	0.75	0.24	0.72	0.77	0.25	39698	305
9029*	0.59	-	-	0.75	-	-	60915	362
9030*	1.05	-	-	0.69	-	-	33979	357

(*) Sensoren im Bereich der ab Dienstag, 24.03. ausgefallenen LACs (vgl. Abbildung 16 und Abbildung 17)

Tabelle 15: Durchschnittliche Durchdringung mit FPD aus A-Daten an den SES sowie Korrelation mit den SES der Landesstelle für Straßentechnik (IM) im Autobahnviereck.

RMQ	D [%]	Streuung von D	Stdw. von D	r	Streuung von r	Stdw. von r	Q_SES [KFZ/d]	Q_FPD [KFZ/d]
800021	0.91	0.74	0.24	0.73	0.40	0.13	31258	291
800022	1.02	0.88	0.30	0.72	0.50	0.16	29189	308
800031	1.01	0.82	0.24	0.68	0.39	0.12	22526	233
800032	1.05	0.87	0.30	0.68	0.52	0.17	21856	238
800041	1.01	0.81	0.24	0.69	0.40	0.12	22600	236
800042	1.04	0.86	0.29	0.68	0.53	0.18	21943	235
800061	1.18	0.83	0.27	0.77	0.35	0.13	42818	519
800062	1.16	0.83	0.29	0.78	0.40	0.14	42674	506
800071	1.01	0.73	0.23	0.76	0.36	0.13	42194	439
800072	1.14	0.81	0.27	0.76	0.38	0.14	41909	488
800081	0.95	0.63	0.21	0.74	0.34	0.11	41061	401
800082	1.11	0.71	0.25	0.78	0.29	0.11	40969	464
800201	0.93	0.61	0.21	0.74	0.41	0.13	32845	311
800202	0.55	0.42	0.15	0.69	0.45	0.15	31509	179
800211	0.00	0.00	0.00	0	0.00	-	32756	0
800212	0.54	0.41	0.14	0.70	0.39	0.14	31596	176
800221	0.00	0.00	0.00	0	0.00	-	32777	0
800222	0.56	0.40	0.15	0.69	0.41	0.14	31454	180
800232	0.00	0.00	0.00	0	0.00	-	18526	0
800241	1.74	1.23	0.43	0.75	0.44	0.15	16736	299
800251	1.30	1.44	0.66	0.64	0.66	0.29	15977	313
800261	0.96	0.54	0.19	0.71	0.39	0.14	29725	295
800262	0.00	0.00	0.00	0	0.00	-	30434	0
800281	0.72	0.34	0.12	0.70	0.44	0.15	27222	203
800282	0.93	0.63	0.20	0.75	0.40	0.16	27411	258
800302	0.00	0.00	0.00	0	0.00	-	24181	0
800321	0.59	0.30	0.11	0.64	0.37	0.15	28568	173
800322	0.00	0.00	0.00	0	0.00	-	28002	0
800331	1.13	0.37	0.16	0.76	0.21	0.09	15212	298
800332	5.79	13.69	7.86	0.55	0.57	0.29	7268	292
800351	0.00	0.00	0.00	0	0.00	-	14217	0
800501	1.26	0.56	0.20	0.81	0.32	0.12	33241	431
800551*	0.68	-	-	0.66	-	-	38752	262
800552*	0.62	-	-	0.84	-	-	44653	278
800561*	0.78	-	-	0.64	-	-	44095	344
800562*	0.57	-	-	0.81	-	-	49971	286
800571*	0.81	-	-	0.64	-	-	41062	332
800572*	0.69	-	-	0.86	-	-	46149	320
800581*	0.94	-	-	0.71	-	-	45862	433
800582*	0.71	-	-	0.85	-	-	51633	365
800601*	0.79	-	-	0.84	-	-	47572	375
800602*	0.00	-	-	0	-	-	52479	0
800611*	1.08	-	-	0.79	-	-	33272	359

RMQ	D [%]	Streuung von D	Stdw. von D	r	Streuung von r	Stdw. von r	Q_SES [KFZ/d]	Q_FPD [KFZ/d]
800612*	0.84	-	-	0.82	-	-	43401	364
800631*	0.00	-	-	0	-	-	38345	0
800632*	0.00	-	-	0	-	-	40687	0
800651*	0.00	-	-	0	-	-	52963	0
800652*	0.54	-	-	0.85	-	-	57155	308
800661*	0.68	-	-	0.68	-	-	49626	338
800662*	0.58	-	-	0.83	-	-	53795	313

(*) Sensoren im Bereich der ab Dienstag, 24.03. ausgefallenen LACs (vgl. Abbildung 16 und Abbildung 17)

Im Autobahnnetz liegt die Durchdringung an 103 der in den beiden Tabelle 14 und Tabelle 15 dargestellten 136 SES über 5 ‰, was ca. 76 % entspricht, bei diesen 103 SES liegt die Streuung der Werte in über 80% (92 von 103) unter dem Wert für die Durchdringung und die Standardabweichung ist bei ca. 60 % (60 von 103) kleiner als 3 ‰, bei ca. 86 % (89 von 103) kleiner als 5 ‰.

Lässt man die 11 Sensoren weg, über die keine einzige FPD-Trajektorie verläuft, dann handelt es sich um ca. 82 % aller Sensoren auf dem Autobahnnetz, für die eine Durchdingung größer als 5 ‰ ermittelt werden konnte. Die REID, die nicht von FPD-Trajektorien berührt werden, sind im zugrundeliegenden Lokalisierungsnetz nicht vorhanden. Es handelt sich daher nicht um einen Verfahrensfehler der FPD-Generierung. An benachbarten SES werden hohe Durchdringungen mit FPD-Trajektorien registriert (Bsp: RMQ 800262: Durchdringung 0 %; benachbarter SES 2498: 1.59 % mit einer Korrelation von 0.83). Diese müssten größtenteils auch über das benachbarte SES verlaufen, dessen REID nicht in den Trajektorien vorkommt. Daher werden diese SES im Weiteren nicht beachtet.

Die Korrelation liegt bei etwa der Hälfte der 125 SES über 0.75 bei einer zugehörigen Streuung von jeweils <= 0.4 sowie einer Standardabweichung von < 0,16. Für diese SES lässt sich eine mittlere Durchdringung von > 6 ‰ (mit einer Ausnahme) feststellen, bei 20 dieser SES liegt die Durchdringung sogar bei mehr als 10 ‰. Die absolute Anzahl an Trajektorien bewegt sich bei diesen SES-Standorten von knapp 200 bis zu über 600 Trajektorien pro Tag, bzw. durchschnittlich 8 bis 25 Trajektorien pro Stunde. An diesen Sensorstandorten sind die FPD-Trajektorien zur Erfassung und Beurteilung der Verkehrsstärke offensichtlich gut geeignet. In Abbildung 41 werden Tagesganglinien von drei beliebig ausgewählten derartigen SES-Standorten dargestellt, die eine gute Eignung der FPD in diversen Anwendungen vermuten lassen. Ungewöhnliche Schwankungen in der Verkehrsstärke, die auf Stauungen hindeuten, scheinen feststellbar zu sein. Die nachfolgenden Kapitel werden Aufschluss über die Verwendbarkeit von FPD in verkehrstechnischen Anwendungen geben. Grundsätzlich scheinen die FPD-Trajektorien aus A-Daten jedoch geeignet zu sein.

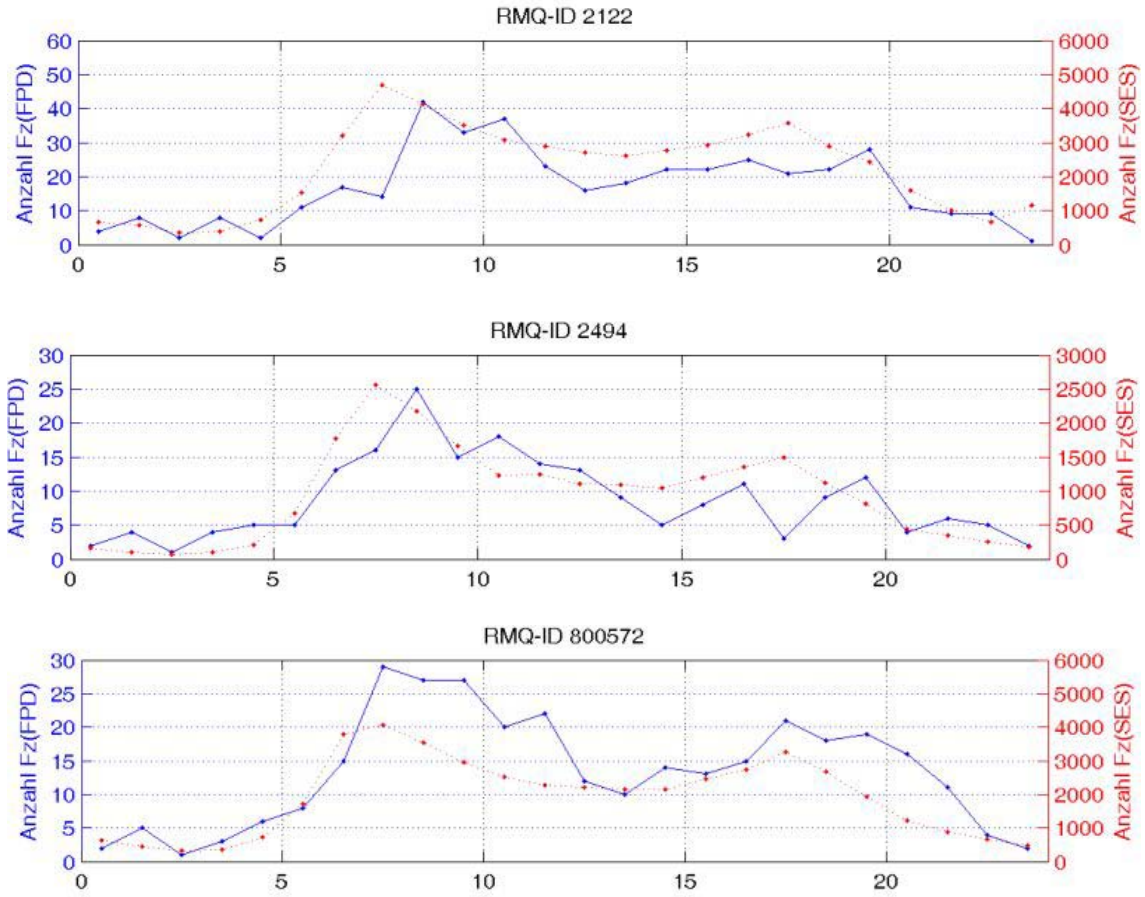


Abbildung 41: Tagesganglinien an drei Beispielen mit Durchdringung > 5 ‰, Rangkorrelation > 0.75 sowie einer durchschnittlichen Trajektorienanzahl von > 252 Traj./24h (RMQ: 2122, 2494

Die Tendenzen der Verkehrsstärke werden meist gut erkannt, aufgrund des bereits erwähnten starken Einflusses einzelner Trajektorien, sind die Tagesganglinien aus FPD wesentlich unruhiger.

Wie die Beispiele untermauern, wurden mit der Durchdringung, der Rangkorrelation sowie der absoluten Anzahl der gefundenen Trajektorien Kriterien gefunden, anhand deren eine Beurteilung einzelner Straßenabschnitte mit Hilfe einer Referenz erfolgen kann. Es wurden erste Vorschläge für die kombinierte quantitative Mindestanforderung an die Parameter gemacht, die auf eine gute Verwendbarkeit der FPD in dem untersuchten Straßenabschnitt hindeuten. Eine genauere Beurteilung und ggf. Festlegung der Grenzen kann jedoch nur im Hinblick auf die gewünschte Anwendung erfolgen. Diese sind für eine Verkehrslageschätzung in Echtzeit wesentlich höher, als beispielsweise für die Erstellung von Quelle-Ziel-Matrizen. Die folgenden Abschnitte sollen die Verwendbarkeit der FPD-Trajektorien aus A- und aus Abis-Daten näher untersuchen und beurteilen.

3 Evaluierung der Verwendbarkeit von FPD in verkehrstechnischen Anwendungen

Im Folgenden werden die Ergebnisse der Evaluierung in den Demonstratoren dargestellt. Dabei stand die Untersuchung der Praxistauglichkeit der aus FPD flächenhaft generierbaren Verkehrsstärke sowie der Weg-Zeit-Diagramme im Fokus.

Aufgrund der bekannten zeitlichen Verzögerungen bei der Bereitstellung der Mobilfunkdaten, die sich über das ganze Projekt fortgesetzt hat, standen hier zur Untersuchung der FPD aus Abis-Daten nur wenige Monate zur Verfügung. Im Falle der A-Daten musste die Evaluierungsphase sogar auf wenige Wochen verkürzt werden.

3.1 FPD zur Generierung der Verkehrslage

Seitens der DDG lag bis zum Abschluss des Berichts Ende September 2009 kein Beitrag vor.

3.2 Demonstrator der Stadt Karlsruhe

Die Pilotanwendung von FPD im städtischen Netz Karlsruhe soll in erster Linie Aufschlüsse über die Qualität und Verwertbarkeit der Verkehrsinformationen in Form von Verkehrslage- bzw. –zustandsberichten liefern. Es soll evaluiert werden, in wie weit die Informationen aus den anfallenden Mobilfunkdaten geeignet sind, auf innerstädtischer Ebene als Informationsquelle für die Verkehrsteilnehmer dienen zu können. Die Qualität der Resultate wird quantitativ anhand von Qualitätsparametern wie z:B. der Verfügbarkeit, der Aktualität und der Korrektheit ermittelt und vor dem Hintergrund unterschiedlicher Streckencharakteristiken (z.B. angebaut/nicht angebaut, kreuzungsfrei/mit LSA) diskutiert.

Die Basisinformation, nämlich die aktuelle Information über die Verkehrszustände in örtlich möglichst genau bezeichneten Streckenabschnitten, soll die Zustände gestaut, stockend (zähflüssig) und frei fließend stabil unterscheiden.

3.2.1 Beschreibung des Demonstrators

Der Demonstrator Karlsruhe zeigt, dass eine vollautomatisierte Übernahme von Floating-Phone-Daten in ein städtisches Verkehrsmanagementsystem möglich ist und insofern diese Daten grundsätzlich zur Auslösung bestimmter verkehrsbeeinflussender Maßnahmen herangezogen werden können.

Die Datenintegration erfolgt in diesem Fall in das bei der Stadt Karlsruhe bereits vorhandene System „Sitraffic Concert“ der Firma Siemens. Hierin sind sämtliche Online-Verkehrsdaten aus stationären Erfassungssystemen SES (Traffic Eyes und Schleifen-

daten) implementiert. Auf der Grundlage der Rohdaten (Verkehrsmenge und mittlere Geschwindigkeiten) findet hier die Berechnung und Visualisierung der Verkehrszustände statt.

Die Datenverarbeitung erfolgt nach folgendem Schema:

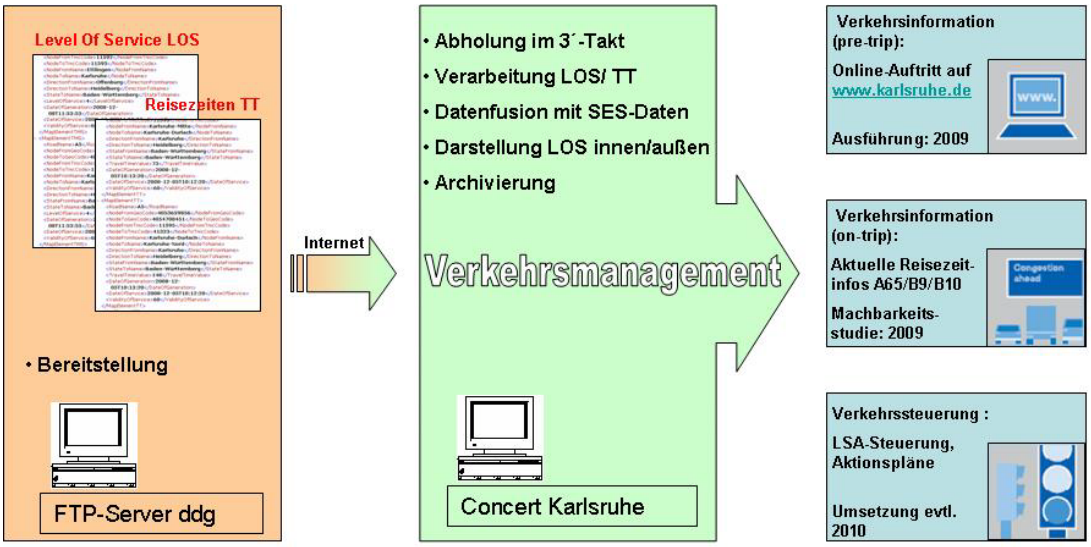


Abbildung 42: Prozesskette Demonstrator Karlsruhe

Auf dem FTP-Server der ddg werden die aktuell berechneten Verkehrszustands- und Reisezeitdateien im xml-Format bereitgestellt. Die beiden Dateien holt der Concert-Rechner aktiv über einen DSL-Anschluss 3-minütlich ab, wobei die Daten bei Vorhandensein von performancebedingt noch nicht vollständig ausgewerteten Datensätzen im Concert wieder verworfen und in der Folgeminute erneut abgeholt werden.

Nach dem Entpacken der beiden Dateien werden die Datensätze nach Referenzobjekten durchsucht - in diesem Fall die TMC-Codes der zu übernehmenden Streckenabschnitte. Da die Do-iT-Produktdaten grundsätzlich eine logische Sicht auf das Straßennetz erlauben und somit Streckenabschnitten, die im Projektnetz mehreren TMC-Kanten zugeordnet sind (z.B. bei Überlagerungsabschnitten B10 / B36 / Südtangente Karlsruhe) auch mehrfach identisch mit den verfügbaren Daten belegt werden, ist eine Übernahme auf TMC-Referenzbasis ohne Informationsverluste in Kauf nehmen zu müssen gewährleistet.

Die für jeden eindeutig identifizierten Streckenabschnitt enthaltene LoS- bzw. Reisezeitinformation wird anschließend einem bestimmten Datenziel im Objektmanager des Concert zugewiesen, womit die Datenkonvertierung abgeschlossen und ab diesem Schritt die vollen Funktionalitäten des Concert-Systems wie Darstellung der Verkehrslage oder Datenarchivierung zum Tragen kommen. Ohne Einschränkung ist ebenfalls eine weitergehende Datennutzung, beispielsweise in Form von Verkehrsinformation

(pre-trip und on-trip) oder zur Auslösung verkehrssteuernder Maßnahmen durch vorgefertigte Aktionspläne möglich.

Der folgende Bildschirmabzug aus dem Concert-Rechner zeigt im oberen Teil die aus den stationären Erfassungssystemen abgeleiteten Verkehrszustände grafisch dargestellt, sowie im unteren Teil die übernommenen Floating-Phone-Daten in tabellarischer Form mit Verkehrslageinformationen und aktuellen Reisezeiten.

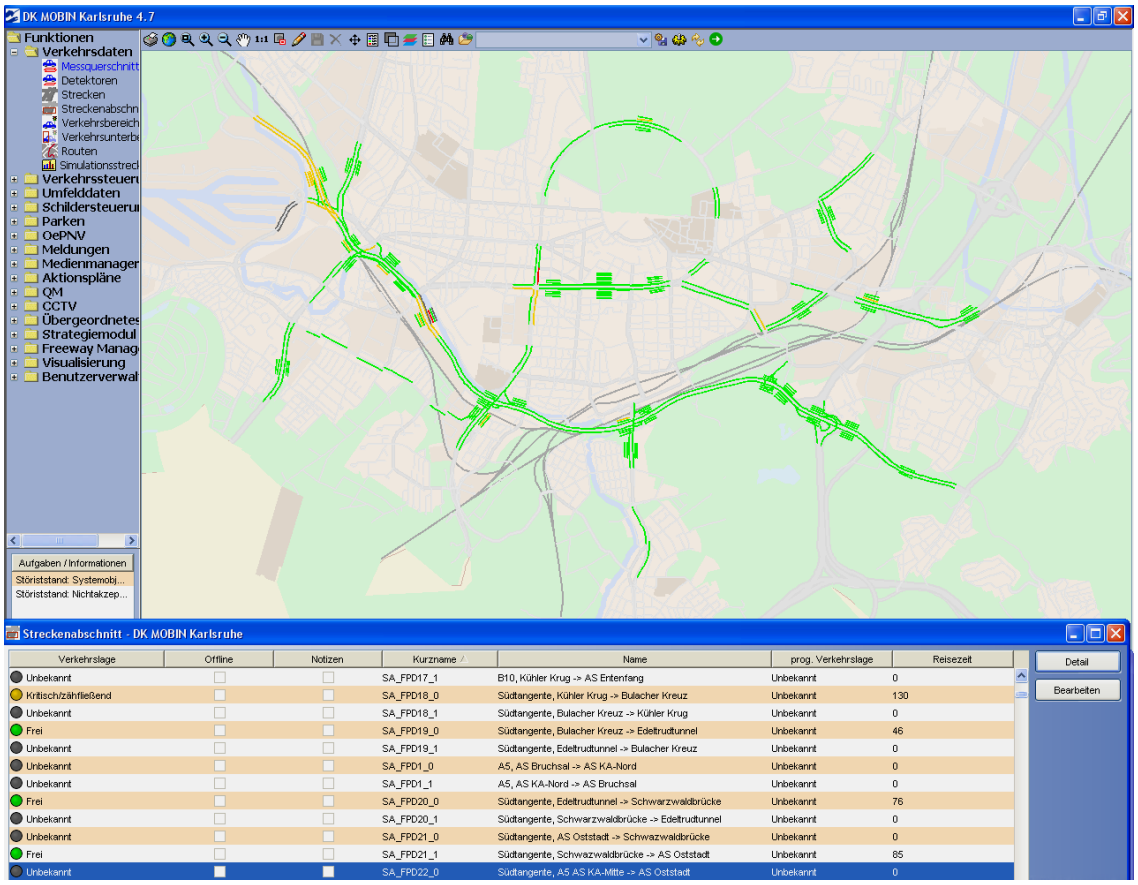


Abbildung 43: Screenshot Demonstrator Karlsruhe

Die Anwendung wurde im Januar und Februar 2009 von der Firma Siemens im Auftrag der Stadt Karlsruhe programmiert und am 24.02.2009 in Betrieb genommen.

3.2.2 Analyse der durchgeführten Feldversuche

Sämtliche Feldversuche wurden im Hinblick auf die Verwertbarkeit von Mobilfunkdaten in städtischen Netzen durchgeführt, wobei das Hauptaugenmerk auf den Möglichkeiten zur Verkehrslageerfassung liegt.

Im Rahmen des Projektes Do-iT wurden sowohl von der ddg als auch vom IAGB Verkehrszustände aus Mobilfunkdaten errechnet. Je nach Kommunikationsebene des Mobilfunknetzes werden Daten unterschiedlichen Detaillierungsgrades verwendet. Wäh-

rend die ddg eine projektnetzweite Erfassung auf A-Ebene realisiert hat, stehen für Karlsruhe die detaillierteren Informationen auf Abis-Ebene des IAGB zur Verfügung. Die Evaluierung selbst wurde ausschließlich für die IAGB-Daten auf **Abis**-Ebene durchgeführt, da aus den A-Daten der ddg erst kurz vor Projektende eine innerstädtische Verkehrslage in Karlsruhe extrahiert werden konnte. Alle Ergebnisse der Feldversuche stehen folglich unter dem Vorbehalt, dass durch Integration der A-Datensätze unter Umständen eine präzisere Verkehrslageerfassung möglich ist, als dies hier in den Feldversuchen für die Abis-Ebene nachgewiesen werden kann.

Die Feldversuche bestehen aus einer systematischen Auswertung der übernommenen Verkehrslagedatensätze sowie entsprechendem Vergleich mit einem geeigneten Referenzsystem. Als Referenzsystem dienen im Stadtgebiet Karlsruhe die Daten aus stationären Erfassungssystemen, deren ermittelte Primär- und abgeleiteten Sekundärdaten als hinreichend genau angenommen werden können. Es sollen dabei die im Qualitätsmodell aufgestellten Qualitätsparameter ermittelt werden:

Tabelle 16: Qualitätsparameter der Verkehrslage

Nr.	Merkmal	Parameter	Abk. und Einheit	Definition
P1.1	Verfügbarkeit VE	Zeitliche Verfügbarkeit	V [%]	Verfügbarkeit der Verkehrslageinformation
P1.2		Raum-Zeitliche Verfügbarkeit	V_{rz} [%]	s. Qualitätsmodell
P2.1	Aktualität AK	Anmeldeverzögerung	MW und Stdw. der ANV [s]	Zeit zwischen dem Entstehen der Störung und der Meldung der Störung
P2.2		Abmeldeverzögerung	MW und Stdw. der ABV [s]	Zeit zwischen dem Verschwinden der Störung und dem Abmelden der Störung
P3.1	Vollständigkeit VO	Detektionsrate	Anteil [%]	Anzahl der gemeldeten Verkehrsbehinderungen im Verhältnis zu der Anzahl tatsächlicher Verkehrsbehinderungen
P4.1	Konsistenz KO	Wird durch Einhaltung des Datenmodells gewährleistet		
P5.1	Korrektheit KR	Falschalarmrate	Anteil [%]	Anzahl der gemeldeten aber nicht tatsächlichen Verkehrsbehinderungen im Verhältnis zu der Anzahl gemeldeter Verkehrsbehinderungen
P5.2		Korrekter Positionierungsanteil	Anteil [%]	Anzahl der gemeldeten tatsächlichen und wohlpositionierten Verkehrsbehinderungen im Verhältnis zu der Anzahl gemeldeter tatsächlicher Verkehrsbehinderungen. Eine Meldung einer Verkehrsbehinderung heißt *wohlpositioniert* wenn der Verkehr vor und hinter dem gemeldeten Anschlussstellenabschnitt frei fließt

P6.1	Genauigkeit GE	Lagegenauigkeit	MW + Stdw. [m]	Genauigkeit der für tatsächliche und wohlpositionierte Verkehrsbehinderungen **innerhalb** der betroffenen Anschlussstellenabschnitte gemeldeten Positionen der Behinderungsgrenzen
P6.2		Längengenauigkeit	MW + Stdw. [m]	Beschreibt die quantitative Abweichung von der Längenangabe eines Ereignis
P6.3		Verzögerungszeitgenauigkeit	MW + Stdw. [s]	Beschreibt die quantitative Abweichung von der gemeldeten Verzögerung durch ein Ereignis

Zur Identifikation geeigneter Vergleichsquerschnitte auf dem FPD-Projektnetz wurde zunächst die Verfügbarkeit der Verkehrszustandsdaten aus FPD streckenabschnittsbezogen für insgesamt 14 Tage im März 2009, d.h. nach Umsetzung der in AP 8 eingeführten Qualitätssicherungsmaßnahmen, ermittelt. Die folgende Abbildung zeigt die minimale, die durchschnittliche und die maximale tägliche Verfügbarkeit des Level of Service aus FPD vor dem Hintergrund der beiden auf Abis-Ebene ausgewerteten Location Areas Ettlingen und Karlsruhe (graue Fläche) für diverse Streckenabschnitte:

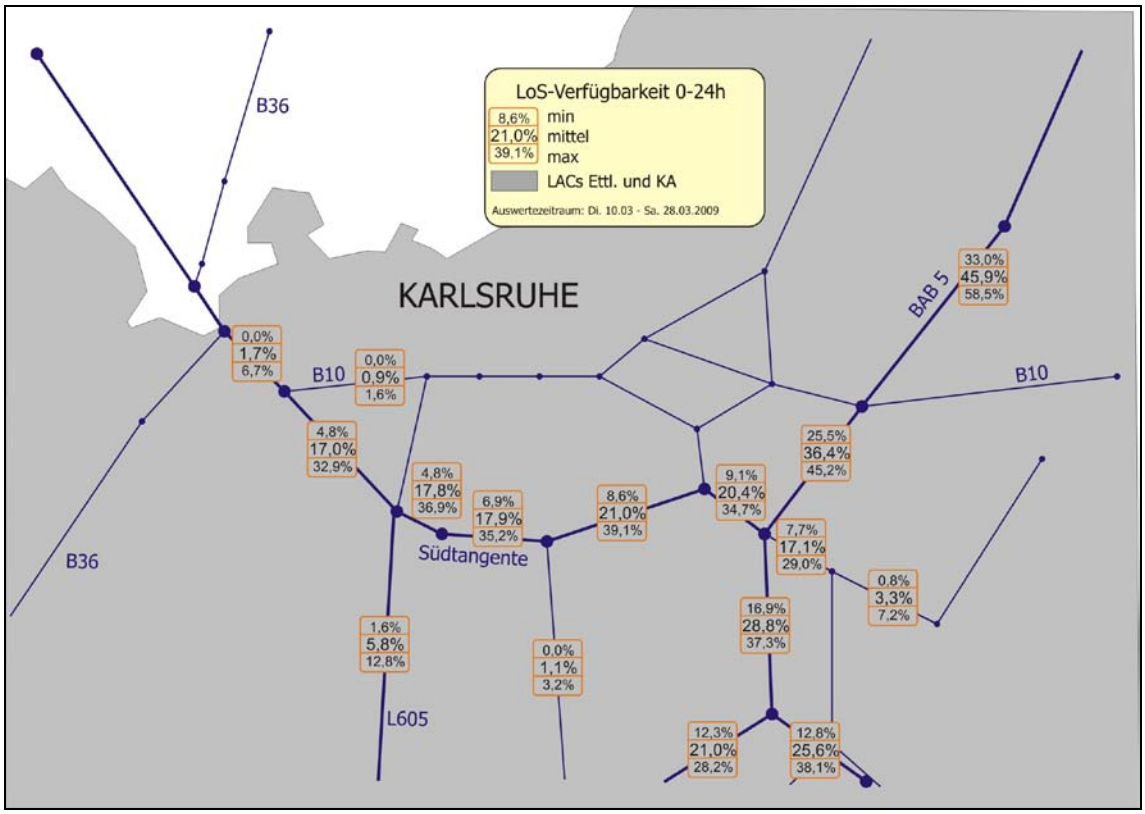


Abbildung 44: LoS-Verfügbarkeit auf ausgewählten Streckenabschnitten

Die höchsten mittleren Verfügbarkeiten an FPD-LoS aus Abis-Daten werden mit bis zu 45,9% auf den Autobahnen erreicht. An einzelnen Tagen ist hier auch in bis zu 60% der Zeit ein aktueller Verkehrszustand verfügbar. Auf dem innerstädtischen Verkehr-

snetz ist die Südtangente zwischen dem Anschlussstelle 6 „Kühler Krug“ im Westen und dem Anschluss der B3 Umgehung Wolfartsweier im Osten mit durchschnittlich ca. 17%-21% verfügbar, wobei die täglichen Verfügbarkeiten relativ stark zwischen nur knapp 5% und fast 40% streuen. Dahinter folgen mit knapp 6% Verfügbarkeit die L605 zwischen der BAB 5 und der Südtangente, die Herrenalber Straße mit 1,1% Verfügbarkeit und mit 0,9 % die B10 zwischen Kühler Krug und Brauerstraße. Im Mittel werden folgende täglichen Verfügbarkeiten erreicht:

Tabelle 17: Verfügbarkeit Abis-Daten im städtischen Netz Karlsruhe

Verfügbarkeit	**0-24h**	**6-22h**
Autobahnen	39,2%	45,8%
Südtangente	18,5%	21,8%
Innerstädtische Hauptverkehrsstraßen	2,6%	3,2%
Raum-Zeitliche Verfügbarkeit (Abis-Netz)	12,8%	15,1%

Über das gesamte städtische Projektnetz inklusive Autobahnen hinweg besteht eine Datenverfügbarkeit von 12,8%, im verkehrsrelevanten Zeitraum von 6:00 bis 22:00 Uhr von 15,1%. Bereits hieran wird deutlich, dass diese Verfügbarkeitswerte für eine zuverlässige flächendeckende innerstädtische LoS-Generierung ausschließlich auf Abis-Ebene wohl nicht ausreichend sein werden. Für eine detaillierte Evaluierung der Daten eignen sich aufgrund dieser Vorüberlegungen in Karlsruhe ausschließlich die folgenden Streckenabschnitte der Südtangente:

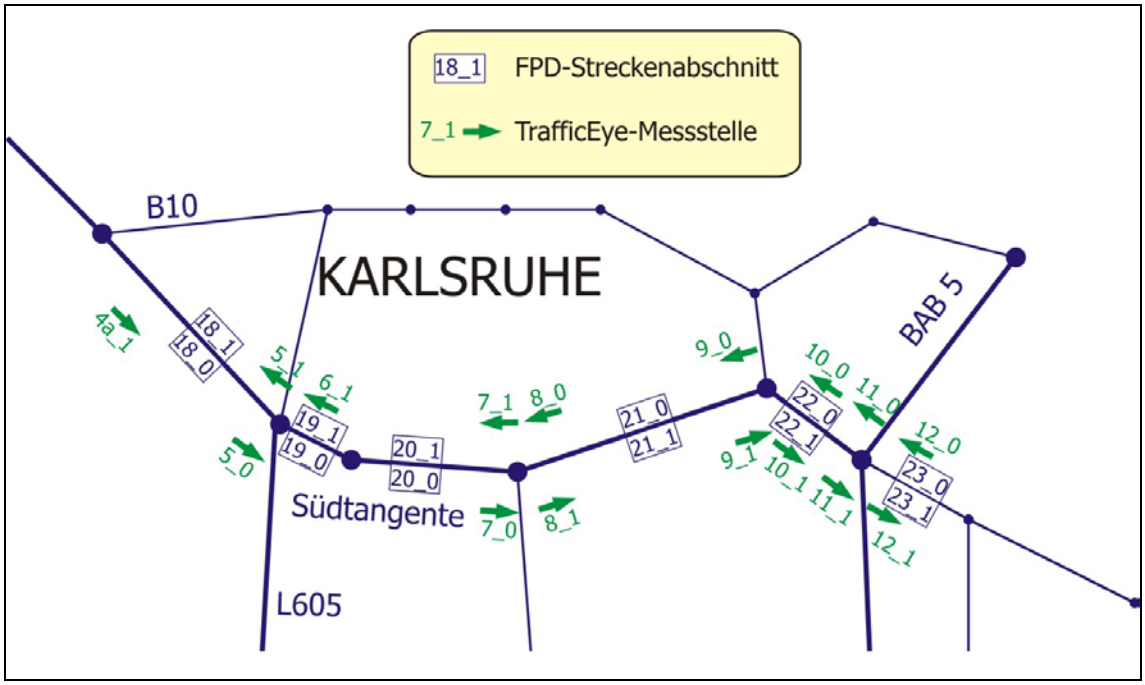
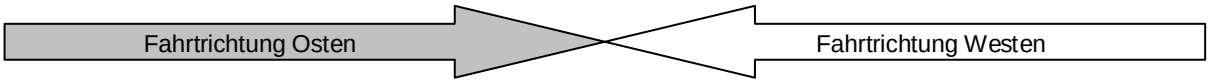


Abbildung 45: Evaluierungsabschnitte mit Referenz-Messstellen

Es werden die in nachstehender Tabelle aufgeführten FPD-Streckenabschnitte (FPD-SA) mit den Ergebnissen der entsprechenden Traffic-Eye Messquerschnitte (TEU-MQ) als Referenzsystem evaluiert:



FPD-SA	18_0	18_0	20_0	21_1	21_1	22_1	22_1	23_1	18_1	19_1	20_1	21_0	21_0	22_0	22_0	23_0
TEU-MQ	4a_1	5_0	7_0	8_1	9_1	10_1	11_1	12_1	5_1	6_1	7_1	8_0	9_0	10_0	11_0	12_0

Um einen ersten Überblick über die Qualität der aus FPD abgeleiteten Verkehrslage zu erhalten, wurde zunächst die Verkehrslage jedes FPD-Streckenabschnitts der Verkehrslage aus dem entsprechenden stationären Detektor vergleichend gegenübergestellt. Als Vergleichszeitraum wurden neun Werktage identifiziert, an denen sowohl die Floating-Phone Daten als auch die Datensätze der Traffic Eyes vollständig erfasst wurden und auf dem betrachteten Streckenabschnitt der Südtangente mindestens eine Verkehrsstörung zu verzeichnen war. Im Einzelnen waren dies der Dienstag, 10.03.2009 bis Freitag, 13.03.2009 und Montag, 24.03.2009 bis Freitag, 27.03.2009. Die entstehenden 16 täglichen Vergleichstabellen FPD-TEU sehen folgendermaßen aus:

Tabelle 18: Beispiel einer Vergleichstabelle Verkehrslage FPD / Traffic Eye (Ausschnitt)

	SA_FPD18_0 / 10.03.2009 (Di.)	SA_TEU4a_1 / 10.03.2009 (Di.)
Typ:	[Typ1] VerkehrslageAktuell	[Typ1] VerkehrslageAktuell
Status:	99.33% OK	100.00% OK
17:01	3	0
17:04	3	0
17:07	3	0
17:10	3	0
17:14	3	0
17:17	3	0
17:20	3	0
17:23	3	1
17:27	3	1
17:30	3	1
17:33	3	1
17:36	3	1
17:39	3	1
17:43	1	1
17:46	1	1
17:49	1	1
17:52	1	1
17:56	1	1
17:59	2	1
18:02	2	1
18:05	2	1
18:09	2	0
18:12	2	0
18:15	2	0
18:18	2	0
18:21	2	0
18:25	3	0
18:28	3	0
18:31	3	0

	SA_FPD18_0 / 10.03.2009 (Di.)	SA_TEU4a_1 / 10.03.2009 (Di.)
Typ:	[Typ1] VerkehrslageAktuell	[Typ1] VerkehrslageAktuell
Status:	99.33% OK	100.00% OK
18:34	3	0
18:38	3	0
18:41	3	0
18:44	3	0
18:48	3	0
18:51	3	0
18:54	3	0
18:57	3	0
19:04	0	0
19:07	0	0
19:10	0	0
19:13	0	0
19:17	0	0
19:20	0	0
19:23	3	0
19:27	3	0
19:30	3	0

0	frei
1	zähfließend
2	gestaut
3	unbekannt

Im Zeitbereich zwischen 17:00 Uhr und 19:30 Uhr weist hier das Traffic Eye 4a_1 in der rechten Spalte eine Verkehrsstörung zwischen 17:23 Uhr und 18:05 Uhr aus. Davor und danach wird freier Verkehr registriert. Dagegen ist die aus FPD generierte Verkehrslage für diesen Streckenabschnitt zunächst bis 17:39 Uhr unbekannt, ab 17:43 Uhr wird analog zu den Referenzdaten des Traffic Eyes ebenfalls eine zähfließende Verkehrslage registriert und ab 17:59 Uhr sogar ein Stauereignis, welches bis 18:21 Uhr anhält. Im Anschluss ist die Verkehrslage wieder unbekannt und im Zeitbereich zwischen 19:04 Uhr und 19:20 Uhr wird in Analogie zu den Werten des Traffic Eyes die Verkehrslage „frei“ angezeigt.

Vergleicht man die entsprechenden Datensätze intervallweise und negiert sämtliche Daten mit unbekannter Verkehrslage, so ergeben sich in diesem Beispiel insgesamt 19 Vergleichsdatensätze, von denen elf übereinstimmen und acht differieren (davon drei Intervalle mit einer Stufe Unterschied und fünf Intervalle mit zwei Stufen Unterschied). In Summe über alle Vergleichquerschnitte konnten auf diese Weise insgesamt knapp 11.000 Vergleichsdatensätze FPD-TEU generiert und folgende Übereinstimmungsquoten ermittelt werden:

Tabelle 19: Übereinstimmungsquoten der Verkehrslage aus FPD- und Traffic Eyes

Datensätze	**Anzahl**	**%**
Übereinstimmung	9.525	87,6%
1 Stufe Abweichung	1.020	9,4%
2 Stufen Abweichung	326	3,0%
Gesamt	10.871	100,0%

Allerdings wird bereits an dem oben aufgeführten Beispiel deutlich, dass einem rein intervallweisen Datenvergleich keine hinreichende Aussagekraft über die Qualität der aus Floating-Phone Daten abgeleiteten Verkehrslage zukommt. Zum einen muss die Prozessierungsdauer der Abis-Daten zur Generierung der Trajektorien berücksichtigt werden, die zwangsläufig zu einer gegenüber den sehr aktuellen Traffic Eyes verzögerten Verkehrslageschätzung führt. In obigem Beispiel führt die Abmeldeverzögerung der Verkehrsstörung dazu, dass die fünf Datensätze von 18:09 Uhr bis 18:21 als stark abweichend gewertet werden, obwohl die Abweichung selbst systematisch bedingt sein dürfte. Daneben bestehen weitere grundsätzliche methodische Unterschiede in Floating-Phone- und stationären Erfassungsdaten. Während FPD letztlich auf Reisezeitmessungen von A nach B beruhen und insofern eine rein streckenbezogene Größe repräsentieren, wird bei stationären Messstellen die Verkehrslage aus lokal gemessenen Verkehrsmengen und mittleren Geschwindigkeiten abgeleitet und diese Information zu Visualisierungszwecken auf einen gesamten Streckenabschnitt extrapoliert. Die genaue Lage der Traffic Eyes auf dem betrachteten Teilabschnitt der Südtangente ist daher ausschlaggebend dafür, ob und in welchem Versatz die beiden Systeme im Falle einer Verkehrsstörung ansprechen. Weil die Traffic Eyes zum großen Teil in unmittelbarer Nähe der Anschlussstellen lokalisiert sind (um auch Verkehrslageinformationen auf den Ausfädelstreifen zu erhalten), sind auch unterschiedliche Ansprechenszeiten im Falle einer sich in Fahrtrichtung rückwärts aufbauenden Stauwelle zu erwarten. Auf dem Streckenabschnitt zwischen Schwarzwaldkreuz und Ottostraße (FPD-SA 21_0) wäre demnach zunächst ein Ansprechen des Traffic Eye 8_0 direkt an der ersten Anschlussstelle realistisch, dann ein Ansprechen der FPD (sobald die Verkehrsstörung die Reisezeit des gesamten Streckenabschnitts signifikant beeinflusst) und als letztes das Traffic Eye 9_0 auf Höhe der zweiten Anschlussstelle. Falls der Verkehr eines Streckenabschnitts nicht in seiner gesamten Länge gestört ist, können zwar die FPD auf eine Störung reagieren, die Stationärdetektoren jedoch weiterhin freien Verkehrsfluss anzeigen.

Insofern können die Floating-Phone Daten nur sinnvoll evaluiert werden, wenn die einzelnen Datensätze nicht nur untereinander, sondern auch mit Lagebezug aufgetragen werden und somit Störungsereignisse im Verkehrsablauf nicht nur zeitlich, sondern auch räumlich lokalisiert werden können. In der folgenden Tabelle sind daher sämtliche von den Traffic Eyes und den FPD registrierten Verkehrsbehinderungen auf der Südtangente für die betreffenden neun Tage aufgeführt und systematisch ausgewertet. Die Übersicht ist folgendermaßen zu lesen: Für jeden Tag sind die hintereinander liegenden Streckenabschnitte und Traffic Eyes nacheinander aufgeführt, wobei die grau hinterlegten Bereiche die Fahrtrichtung Ost und die weißen Bereiche die Fahrtrichtung West kennzeichnen. Jede Zeile beinhaltet die Vergleichsdaten eines FPD-Streckenabschnitts mit dem entsprechenden stationären Referenzdetektor. Die Fahrtrichtung kann anhand des Pfeilsymbols in der Spalte „Nr.+Richtung“ abgelesen werden. In der Spalte „Verkehrsstörungen TEU“ sind sämtliche an dem betreffenden Traffic Eye registrierten Verkehrsstörungen mit Uhrzeit aufgetragen, die bei Vorliegen einer zeitlich übereinstimmenden Verkehrsbehinderung an dem oder den benachbarten

Messstellen zu einer Verkehrsstörung über mehrere Stationärdetektoren bzw. Streckenabschnitte zusammengefasst werden. Durch Vergleich mit den am entsprechenden Floating-Phone-Abschnitt registrierten Verkehrsstörungen werden die Ereignisse entweder als detektiert (grün hinterlegt) oder nicht detektiert (rot hinterlegt, wenn FPD „frei“ anzeigt und unhinterlegt, wenn FPD „unbekannt“ anzeigt) markiert. Falls eine Verkehrsbehinderung durch FPD detektiert wurde, werden in der Spalte „Verzögerung“ die An- und Abmeldeverzögerungen als zeitlicher Versatz zwischen dem Entstehen der Störung und der Meldung bzw. dem Verschwinden der Störung und der Abmeldung in Minuten angegeben. In der letzten Spalte „Falschalarm FPD“ werden die Zeiträume aufgeführt und ggf. über mehrere Streckenabschnitte zusammengefasst, während denen zwar über Floating-Phone Störungen gemeldet werden, diese aber an keiner stationären Referenzstelle bestätigt werden. Gelbe Bereiche kennzeichnen dabei eine zähfließende Meldung, rote Bereiche einen gemeldeten Stau. Grün sind in dieser Spalte Zeitbereiche gekennzeichnet, die richtigerweise eine Störung innerhalb des Streckenabschnitts anzeigen, die Störung aber nicht bis zu dem Vergleichsquerschnitt heranreicht. Nicht farblich hinterlegte Bereiche deuten auf eine fehlpositionierte Störungsmeldung hin.

So wurden beispielsweise am 10.03.2009 am Traffic Eye 4a_1 insgesamt zwei Verkehrsstörungen registriert. Die erste begann um 07:33 Uhr und endete um 09:10 Uhr. Die zweite Störung ist die in obiger Vergleichstabelle bereits dokumentierte Zähflüssigkeit von 17:23 Uhr bis 18:05. Während der ersten Störung wurde kein passendes Pendant in den FPD registriert, die aus Floating Phone ermittelte Verkehrslage hat in diesem Zeitraum „frei“ angezeigt. Die Störung selbst ist deshalb rot hinterlegt (Störung durch FPD nicht detektiert, Status: frei). Die zweite Störung wurde als erkannt definiert und grün hinterlegt, da die FPD-Verkehrslage im Zeitraum von 17:43 Uhr bis 18:21 Uhr ebenfalls eine Störung registriert hat. Gleichzeitig wird durch die tabellarische Darstellung deutlich, dass sich die zweite Verkehrsstörung über einen längeren Streckenabschnitt hinzog und von mehreren Traffic Eyes in gewissen Zeitversätzen gemeldet wurde. Die erste Meldung ist am Traffic Eye 10_1 um 16:41 Uhr registriert, das nächste stromaufwärts gelegene Traffic Eye 9_1 hat um 16:48 Uhr angesprochen, das nächste um 17:17 Uhr und die folgenden drei um jeweils 17:23 Uhr. Insgesamt sind diese sechs von Traffic Eyes registrierten Störungen deshalb als eine einzige Verkehrsbehinderung anzusehen und zusammengefasst als Block dargestellt, um auf dieser Grundlage im Anschluss Aussagen über die Qualitätsparameter Detektionsrate, Falschalarmrate oder korrekte Positionierungen ableiten zu können. Die Verkehrsbehinderung wurde auf ihrer gesamten Länge mit Ausnahme des Abschnitts 20_0 von den FPD an den Streckenabschnitten 18_0, 21_1 und 22_1 um 17:43 Uhr bzw. um 17:59 Uhr detektiert. Die Anmeldeverzögerungen der FPD betrugen zwischen 13 und 78 Minuten, die Abmeldeverzögerungen zwischen -10 und 19 Minuten. Weil die Stauwurzel innerhalb des Streckenabschnitts 22_1 zwischen den Traffic Eyes 10_1 und 11_1 positioniert war, ist die Störungsmeldung des FPD-Streckenabschnitts 22_1 im Vergleich zum Traffic Eye 11_1 nicht als Falschalarm zu werten, wie dies im reinen FPD-TEU-Vergleich der Fall gewesen wäre. Der entsprechende Zeitraum ist daher in

der Spalte „Falschalam FPD“ grün markiert („Störung innerhalb des Anschlussstellen-abschnitts“). Innerhalb der gelb hinterlegten Zeiträume in der Spalte „Falschalarm“ (z.B. von 9:45 Uhr bis 10:04 Uhr) wurde von den FPD eine zähfließende Verkehrslage gemeldet, ohne dass diese von mindestens einem Traffic Eye bestätigt wurde, während der rot hinterlegten Zeitbereiche wurde ein Stau detektiert.

Tabelle 20: Raum-zeitliche Verkehrslagetabellen FPD und Traffic Eyes

Datum	Nr. + Richtung		Verfügbarkeit	Verkehrsstörungen			Verkehrsstörungen	Verzögerung FPD		Falschalarm				
	FPD	TEU	FPD	TEU			FPD	An-	Ab-	FPD				
10.03.	18_0 ↓	4a_1	16%	07:33 - 09:10	17:23 - 18:05		17:43 - 18:21	20	16	09:45 - 10:04				
	18_0 ↓	5_0	16%		17:23 - 18:05		17:43 - 18:21	13	19	09:45 - 10:04				
	20_0 ↓	7_0	23%		17:23 - 18:05					00:00 - 00:22				
	21_1 ↓	8_1	29%		17:17 - 18:05		17:59 - 18:21	42	16					
	21_1 ↓	9_1	29%		16:48 - 18:34		17:59 - 18:21	71	-13					
	22_1 ↓	10_1	29%		16:41 - 18:31		17:59 - 18:21	78	-10		10:34 - 10:53	13:05 - 13:12	14:13-14:32	15:40-16:03
	22_1 ↓	11_1	29%							17:59 - 18:21	10:34 - 10:53	13:05- 13:12	14:13-14:32	15:40-16:03
	23_1 ↓	12_1	21%								10:34 - 10:53	13:05 - 13:28		15:40-16:03
	18_1 ↑	5_1	16%	08:31 - 08:57	16:35 - 18:25					00:00 - 00:09		09:42 - 10:04		
	19_1 ↑	6_1	13%	08:34 - 08:50	16:48 - 18:09					07:43 - 08:08	21:59 - 22:18			
	20_1 ↑	7_1	17%			20:38 - 21:07								
	21_0 ↑	8_0	19%			20:45 - 20:51								
	21_0 ↑	9_0	19%	07:52 - 09:26										
	22_0 ↑	10_0	19%	07:49 - 09:13										
	22_0 ↑	11_0	19%											
	23_0 ↑	12_0	17%							08:38 - 09:00				
11.03.	18_0 ↓	4a_1	12%	14:42 - 15:11										
	18_0 ↓	5_0	12%	14:55 - 15:37										
	20_0 ↓	7_0	13%	15:05 - 15:24						17:59 - 18:15				
	21_1 ↓	8_1	15%	15:05 - 15:27			14:58 - 15:01	-7	-26					
	21_1 ↓	9_1	15%							14:58 - 15:01				
	22_1 ↓	10_1	15%							14:45 - 15:01				
	22_1 ↓	11_1	15%							14:45 - 15:01				
	23_1 ↓	12_1	11%							14:45 - 15:01	18:18 - 18:41			
	18_1 ↑	5_1	11%	08:05 - 08:18	14:00 - 14:52		08:15 - 08:38	10	20					
	19_1 ↑	6_1	13%		14:03 - 14:39									
	20_1 ↑	7_1	12%											
	21_0 ↑	8_0	12%											
	21_0 ↑	9_0	12%											
	22_0 ↑	10_0	13%							21:13 - 21:36				
	22_0 ↑	11_0	13%							21:13 - 21:36				
	23_0 ↑	12_0	13%							21:13 - 21:36				
12.03.	18_0 ↓	4a_1	16%	16:16 - 17:20										
	18_0 ↓	5_0	16%	16:22 - 17:17										
	20_0 ↓	7_0	7%											
	21_1 ↓	8_1	10%							22:40 - 22:59				
	21_1 ↓	9_1	10%	17:10 - 17:36						22:40 - 22:59				
	22_1 ↓	10_1	10%	16:51 - 17:33						20:12 - 20:35				
	22_1 ↓	11_1	10%							20:12 - 20:35				
	23_1 ↓	12_1	8%							20:12 - 20:35				
	18_1 ↑	5_1	15%	16:19 - 19:17			18:05 - 18:25	45	-52	09:03 - 09:23				
	19_1 ↑	6_1	8%	16:19 - 19:07						14:35 - 14:55				
	20_1 ↑	7_1	10%	18:21 - 18:51						14:35 - 14:55	21:59 - 22:05			
	21_0 ↑	8_0	8%											
	21_0 ↑	9_0	8%											
	22_0 ↑	10_0	8%							20:20 - 20:43				
	22_0 ↑	11_0	8%							20:15 - 20:38				
	23_0 ↑	12_0	8%							20:15 - 20:38	02:38 - 03:17			

Datum	Nr. + Richtung FPD	TEU	Verfügbarkeit FPD	Verkehrsstörungen TEU	Verkehrsstörungen FPD	Verzögerung FPD An-	Verzögerung FPD Ab-	Falschalarm FPD
13.03.	18_0 ↓	4a_1	14%					13:48 - 14:08
	18_0 ↓	5_0	14%	09:27 - 10:51				13:48 - 14:08
	20_0 ↓	7_0	10%					15:18 - 15:31 / 17:08 - 17:28
	21_1 ↓	8_1	14%					
	21_1 ↓	9_1	14%					
	22_1 ↓	10_1	13%					
	22_1 ↓	11_1	13%					
	23_1 ↓	12_1	9%					18:29 - 18:51
	18_1 ↑	5_1	5%					
	19_1 ↑	6_1	7%					15:09 - 15:31 / 23:06 - 23:28
	20_1 ↑	7_1	7%					15:09 - 15:31 / 23:06 - 23:28
	21_0 ↑	8_0	7%					
	21_0 ↑	9_0	7%					
	22_0 ↑	10_0	9%					
	22_0 ↑	11_0	9%					
	23_0 ↑	12_0	7%					
23.03.	18_0 ↓	4a_1	32%					
	18_0 ↓	5_0	32%					
	20_0 ↓	7_0	34%					11:06 - 11:25
	21_1 ↓	8_1	39%					
	21_1 ↓	9_1	39%					
	22_1 ↓	10_1	34%					
	22_1 ↓	11_1	34%					08:30 - 08:50
	23_1 ↓	12_1	30%					17:13 - 18:02
	18_1 ↑	5_1	20%					
	19_1 ↑	6_1	20%					
	20_1 ↑	7_1	22%					
	21_0 ↑	8_0	24%					
	21_0 ↑	9_0	24%					
	22_0 ↑	10_0	21%	07:49 - 08:37				13:37 - 14:00 / 15:11 - 15:30
	22_0 ↑	11_0	21%					13:37 - 14:00 / 15:11 - 15:30
	23_0 ↑	12_0	19%					13:37 - 14:00 / 15:11 - 15:30 / 06:30 - 07:09
24.03.	18_0 ↓	4a_1	18%	07:48 - 08:14				
	18_0 ↓	5_0	18%					
	20_0 ↓	7_0	27%					09:22 - 09:28 / 14:43 - 15:25
	21_1 ↓	8_1	32%					09:22 - 09:28 / 08:05 - 08:24
	21_1 ↓	9_1	32%					09:22 - 09:28 / 08:05 - 08:24
	22_1 ↓	10_1	30%					13:09 - 13:19 / 16:23 - 16:42 / 17:17 - 17:40
	22_1 ↓	11_1	30%					13:09 - 13:19 / 16:23 - 16:42 / 17:17 - 17:40
	23_1 ↓	12_1	23%					13:09 - 13:19 / 16:23 - 16:42 / 17:17 - 17:40
	18_1 ↑	5_1	27%					
	19_1 ↑	6_1	30%					01:32 - 02:08
	20_1 ↑	7_1	32%					01:32 - 01:52 / 08:11 - 08:34
	21_0 ↑	8_0	32%					08:11 - 08:34
	21_0 ↑	9_0	32%	07:52 - 08:43	08:11 - 08:34	19	-9	
	22_0 ↑	10_0	31%	08:01 - 08:56	08:11 - 08:34	10	-22	
	22_0 ↑	11_0	31%	08:05 - 08:24	08:11 - 08:34	6	10	
	23_0 ↑	12_0	28%					08:11 - 08:34 / 10:50 - 11:12 / 16:07 - 16:26
25.03.	18_0 ↓	4a_1	13%	13:42 - 14:11 / 15:52 - 17:16				11:34 - 11:40
	18_0 ↓	5_0	13%	13:33 - 14:05 / 15:48 - 16:27 / 23:21 - 24:00				11:34 - 11:40
	20_0 ↓	7_0	23%					
	21_1 ↓	8_1	29%					08:46 - 09:05 / 16:11 - 16:31 / 17:32 - 17:51
	21_1 ↓	9_1	29%					08:46 - 09:05 / 16:11 - 16:31 / 17:32 - 17:51
	22_1 ↓	10_1	27%	16:53 - 17:19				08:46 - 09:05 / 10:00 - 10:20 / 16:11 - 16:31 / 17:32 - 17:51
	22_1 ↓	11_1	27%					08:46 - 09:05 / 10:00 - 10:20 / 16:11 - 16:31 / 17:32 - 17:51
	23_1 ↓	12_1	18%					08:46 - 09:05 / 10:00 - 10:20 / 15:13 - 15:32 / 16:11 - 16:31
	18_1 ↑	5_1	21%	16:05 - 17:48	16:24 - 16:44	19	-64	
	19_1 ↑	6_1	19%	16:05 - 17:16				

Datum	Nr. + Richtung FPD		TEU	Verfügbarkeit FPD	Verkehrsstörungen TEU			Verkehrsstörungen FPD	Verzögerung FPD An-	Verzögerung FPD Ab-	Falschalarm FPD				
	20_1	↑	7_1	19%							08:01 - 08:27	20:26 - 20:42			
	21_0	↑	8_0	16%							06:24 - 06:46	09:05 - 09:25			
	21_0	↑	9_0	16%	07:58 - 09:18			09:05 - 09:25	67	7	06:24 - 06:46				
	22_0	↑	10_0	16%	08:01 - 09:12			08:40 - 09:02	39	-10			05:25 - 06:04		
	22_0	↑	11_0	16%									05:25 - 06:04	08:40 - 09:02	
	23_0	↑	12_0	16%							00:11 - 00:31	08:01 - 08:27	05:25 - 06:04	08:40 - 09:02	
26.03.	18_0	↓	4a_1	13%	07:35 - 08:01	16:34 - 17:59								15:04 - 15:20	
	18_0	↓	5_0	13%										15:04 - 15:20	
	20_0	↓	7_0	14%							12:52 - 13:11	15:04 - 15:20			
	21_1	↓	8_1	20%											
	21_1	↓	9_1	20%	17:07 - 17:49										
	22_1	↓	10_1	22%	17:10 - 19:00						10:33 - 10:46			21:58 - 22:17	
	22_1	↓	11_1	22%							10:33 - 10:46			21:58 - 22:17	
	23_1	↓	12_1	16%							10:26 - 10:46	00:00 - 00:17	12:52 - 13:11	21:58 - 22:17	
	18_1	↑	5_1	7%	16:15 - 16:57						19:07 - 19:26				
	19_1	↑	6_1	10%	16:08 - 16:41						11:37 - 11:54	16:57 - 17:16			
	20_1	↑	7_1	8%							11:37 - 11:54				
	21_0	↑	8_0	8%							06:39 - 07:02				
	21_0	↑	9_0	8%							06:39 - 07:02				
	22_0	↑	10_0	9%	08:01 - 08:26										
	22_0	↑	11_0	9%											
	23_0	↑	12_0	9%							03:05 - 03:25				
27.03.	18_0	↓	4a_1	5%	07:48 - 08:30						19:13 - 19:36			21:42 - 22:01	
	18_0	↓	5_0	5%							19:13 - 19:36			21:42 - 22:01	
	20_0	↓	7_0	9%							18:11 - 18:31	20:31 - 20:50			
	21_1	↓	8_1	13%							18:11 - 18:31	19:13 - 19:36	21:26 - 21:51		
	21_1	↓	9_1	13%							18:11 - 18:31	19:13 - 19:36	21:26 - 21:51		
	22_1	↓	10_1	13%								19:13 - 19:36			
	22_1	↓	11_1	13%								19:13 - 19:36			
	23_1	↓	12_1	9%											
	18_1	↑	5_1	15%											
	19_1	↑	6_1	12%											
	20_1	↑	7_1	10%											
	21_0	↑	8_0	9%											
	21_0	↑	9_0	9%											
	22_0	↑	10_0	13%							12:07 - 12:26				
	22_0	↑	11_0	13%							12:07 - 12:26				
	23_0	↑	12_0	12%							12:07 - 12:26				
	Gesamt			**18,5%**	**Störungen: 28**				**30,9**	**-8,4**		**53**	**20**		
					7	**17**	**4**								

08:05 - 08:18	Störung durch FPD detektiert	12:07 - 12:26	Falschalarm FPD, Status: zähfließend
07:35 - 08:01	Störung durch FPD nicht detektiert (Status: unbekannt)	19:13 - 19:36	Falschalarm FPD, Status: Stau
17:07 - 17:49	Störung durch FPD nicht detektiert (Status: frei)	08:11 - 08:34	Fehlpositionierte Störung
		09:05 - 09:25	Störung innerhalb des Anschlussstellenabschnitts

Insgesamt wurden im neuntägigen Auswertezeitraum 28 verschiedene Verkehrsbehinderungen unterschiedlicher Ausdehnung durch die Referenzmessstellen auf der Südtangente detektiert. Davon konnten sieben durch die Floating-Phone Daten auf Abis-Ebene ebenfalls erkannt werden. In 17 Fällen waren keine FP-Daten verfügbar und in

vier Fällen wurden die Störungen nicht detektiert und fälschlicherweise ein freier Verkehrsfluss angezeigt. Es wird aber auch deutlich, dass in allen Fällen, in denen eine Verkehrsbehinderung über mindestens drei Traffic Eyes hinweg detektiert wurde, die Störung auch durch FPD erkannt werden konnte. Zeitlich und lokal sehr begrenzte Verkehrsstörungen werden dagegen in den meisten Fällen nicht detektiert.

Mit 53 Zähfließend-Meldungen und 20 Staumeldungen, die nicht von den stationären Messstellen bestätigt werden konnten, ist die Menge an Falschalarmen durch FPD doch auffallend hoch. Die Gründe hierfür mögen vielfältig sein. Einerseits kann es sich bei den hier als Falschalarm definierten Störungen um tatsächliche Störungen außerhalb der stationären Detektorerfassungsbereiche handeln oder schlicht um Einzelfahrzeuge, die ein nicht repräsentatives Bewegungsmuster aufwiesen. Andererseits besteht die Möglichkeit, dass zeitlich sehr kurze Störungen durch die intervallbedingte Mittelwertbildung und die Hysteresebereiche der Traffic Eyes durchaus auch einmal von den SES nicht detektiert bzw. nicht angezeigt werden, wenn im unmittelbar darauffolgenden Intervall die Verkehrsverhältnisse wieder als stabil erkannt werden. Und abschließend ist natürlich auch eine im Vergleich zu den im Traffic Eye hinterlegten Werten restriktivere Verkehrslageparametrierung möglich.

Dessen ungeachtet ergeben sich die folgenden Qualitätsparameter der Verkehrslage für FPD auf Abis-Ebene:

Tabelle 21: Qualitätsparameter der Verkehrslage (Abis-Daten)

Nr.	Merkmal	Parameter	Wert
P1.1	Verfügbarkeit VE	Zeitliche Verfügbarkeit	s. Abb. 3-3 bzw. Tab. 2
P1.2		Raum-Zeitliche Verfügbarkeit	12,8%
P2.1	Aktualität AK	Anmeldeverzögerung	30,9 Min. +/- 26,6 Min.
P2.2		Abmeldeverzögerung	-8,4 Min+/- 26,1 Min.
P3.1	Vollständigkeit VO	Detektionsrate	25%
P5.1	Korrektheit KR	Falschalarmrate	91,3 %
P5.2		Korrekter Positionierungsanteil	42,9 %

3.2.3 Bewertung der Ergebnisse

Es konnte der Nachweis erbracht werden, dass eine vollautomatisierte Übernahme von Floating-Phone-Daten in ein städtisches Verkehrsmanagementsystem möglich ist und insofern diese Daten grundsätzlich zur Auslösung bestimmter verkehrsbeeinflussender Maßnahmen herangezogen werden können.

Die Ergebnisse der Evaluierung basieren ausschließlich auf Abis-Daten des Mobilfunknetzes, für die bereits in vorangegangenen Arbeitsschritten auf die Verfügbarkeitsprob-

lematik hingewiesen wurde. Ein durchschnittlicher Erfassungsgrad von 0,4 bis 1,0% aller Fahrzeuge führt zu einer zeitlichen Verfügbarkeit aktueller Verkehrslageinformationen von projektnetzweit ca. 12,8%. Als ausschließliche Datenbasis wird diese Verfügbarkeit nicht für eine zuverlässige Verkehrslageschätzung ausreichend sein können.

Die Datenqualität konnte ausschließlich auf der Karlsruher Südtangente evaluiert werden, einer vierstreifigen, teilweise angebauten Hauptverkehrsstraße mit einer Verkehrsbelastung von ca. 80.000 Kfz/Tag und einer Verkehrslageverfügbarkeit von knapp 20%. Andere innerörtliche Hauptverkehrsstraßen wiesen zu wenige Verkehrslageinformationen auf, als dass hier eine sinnvolle Evaluierung der Datenqualität möglich gewesen wäre.

Es hat sich erwiesen, dass die Verkehrslageinformationen aus FPD in ca. 88% aller Fälle mit den jeweiligen Referenzdaten aus stationären Erfassungssystemen übereingestimmt haben. Die Detektion von Verkehrsbehinderungen kann als grundsätzlich nicht so unzuverlässig eingestuft werden, wie die ermittelte Detektionsrate von 25% dies suggeriert, da alle größeren Verkehrsstörungen detektiert werden konnten und nur die zeitlich und lokal sehr begrenzten Behinderungen in den meisten Fällen nicht detektiert wurden. Weil in über 80% der nicht detektierten Behinderungen die Verkehrslage nach FPD „unbekannt“ war, beruht letztlich die niedrige Detektionsrate größtenteils auf den niedrigen Verfügbarkeitswerten. Das gleiche gilt auch für die gegenüber den Referenzsystemen ermittelten teilweise recht hohen An- bzw. Abmeldeverzögerungen.

Wenn es daher gelänge, die Verfügbarkeit der Abis-Daten noch weiter zu erhöhen, könnte diese Datenquelle tatsächlich einen wichtigen Beitrag zu einer Verkehrslageschätzung auf hochbelasteten innerstädtischen Hauptverkehrsstraßen leisten. Enormes zusätzliches Potenzial bietet auch die Integration der aus A-Daten gewonnenen Verkehrslage, da auf dieser Ebene von einer wesentlich höheren, wenn auch weniger detaillierten Erfassung ausgegangen werden kann.

3.3 Demonstrator der Landeshauptstadt Stuttgart

Seitens der Stadt Stuttgart lag bis zum Abschluss des Berichts Ende September 2009 kein Beitrag vor.

3.4 Demonstrator des IM BW, Landesstelle für Straßentechnik

Im Autobahnviereck A 5/A 6/A 81/A 8 AD Leonberg - AD Karlsruhe - AK Walldorf - AK Weinsberg wird Verkehrsteilnehmern im Falle eines Staus auf der ausgewiesenen Route über Heilbronn mit dWiSta eine Umleitung auf die Alternativroute über Karlsruhe empfohlen. Die Steuerung der Netzbeeinflussungsanlage (NBA) Leonberg - Walldorf ist auf dem Produkt Polydrom aufgebaut, welches die Schaltempfehlungen für das Autobahnviereck generiert. Es werden Meldungen des Verkehrswarndienstes (VWD) und

aus dem Baustelleninformationssystem des Landes Baden-Württemberg (BIS) verwendet. Die NBA wird derzeit über Störungsmeldungen gesteuert werden.

3.4.1 Beschreibung des Demonstrators

Ziel der Untersuchung war es, die Verwendbarkeit von Floating Phone Daten zur Steuerung einer Netzbeeinflussungsanlage zu ermitteln. Dabei waren die Nutzbarkeit und Wirtschaftlichkeit von FPD zur Verwendung in Netzbeeinflussungsanlagen zu bewerten und zu dokumentieren. Die FP-Daten (D1-Netz) wurden von dem Projektpartner Gesellschaft für Verkehrsdaten mbH (DDG) aufbereitet und als Reisezeiten (TIM) bzw.Level of Service (LOS) offline bereitgestellt.

Informationen über Reisezeiten bzw. Verkehrslage sollen es den Verkehrsteilnehmern ermöglichen, das Streckennetz (BAB) optimal zu nutzen. Daher waren Reisezeiten und Verkehrslage auf der BAB miteinander zu vergleichen. Die Qualität der Informationen war zu untersuchen.

Der Demonstrator U-Strecken konnte bis Projektende nicht in Angriff genommen werden, da Trajektorien aus A-Daten im Projekt erst unmittelbar vor Projektende zur Verfügung standen und die Zeit für Auftragsvergabe und -bearbeitung nicht mehr ausreichend war.

Es wurden folgende Steuerungsphilosophien untersucht und Qualitätskriterien für

- Steuerung auf Basis von berechneten Staulängen über VWD und einzelnen Messstellen (aktuell in Betrieb)
- Steuerung über LOS (gebildet aus FPD / Reisegeschwindigkeitsindex, Distanz zwischen den zugehörigen TMC-Locations wird als Staulänge interpretiert),
- Steuerung über TIM (Reisezeiten gebildet aus FPD),

gebildet und miteinander verglichen.

Für die Untersuchung wurden funktionale Steuerungsmodelle erstellt. Beim bestehenden Steuerungsmodell werden die vom VWD gemeldeten Staulängen anhand von Regeln auf die Steuerung übertragen.

Die anschließende Betrachtung bestand aus Offline-Simulationen mit den oben genannten Steuerungsphilosophien und realen Verkehrsdaten. Hierbei waren die jeweiligen Schaltempfehlungen und Schaltzeitpunkte der einzelnen Varianten zu ermitteln, zu vergleichen und zu bewerten. Anschließend wurden die Ergebnisse mit den tatsächlichen Schaltungen der NBA verglichen. Ein Qualitätskriterium war die Korrektheit der Stauprognosen unter Berücksichtigung der verlagerten Verkehrsströme, ein weiteres Qualitätskriterium der optimale Schaltzeitpunkt und die dadurch verkürzte reale Reisezeit im betrachteten Autobahnviereck. Das zur Auswertung entwickelte Evaluationstool kann mit den Online-Daten arbeiten und somit zukünftig direkt zur Steuerung der dWiSta-Anzeigen im Autobahnviereck verwendet werden.

Die tatsächlichen Kosten für die jeweilige Datenerhebung bzw. -beschaffung konnte nicht bestimmt werden, da im Projekt die Daten kostenlos und anschließend noch keine weiteren belastbaren Kosten ermittelt werden konnten. Durch den Vergleich der Routenempfehlungen inkl. der zugehörigen Schaltzeitpunkte wurde die Nutzenveränderung (Reisekosten/Zeitkosten) ermittelt, dargestellt und bewertet.

3.4.2 Analyse des durchgeführten Feldversuchs

Vom 03.02.09 bis zum 04.03.09 wurden während 30 Tagen 24h online Daten von der LST und der DDG aufgenommen. Die Daten von Messquerschnitten und Meldungen des Verkehrswarndienstes (VWD) von der Verkehrsrechnerzentrale (VRZ) der LST und von der DDG LOS-Werte und Reisezeiten pro TMC Location Strecke. Die LOS-Daten wurden zu Staulängen transformiert und die Reisezeiten über die Strecke der Routen aufsummiert. Polydrom kalibrierte alle 3 min die aktuelle Verkehrslage mit den Messquerschnitten. Falls die Online-Daten Schwellenwerte erreichten, wurden 3 Simulationen ausgelöst.

1. LST: Algorithmus für die Schaltung von dWiSta (aktuelles Programm der VRZ)
2. DDG-LOS: wie 1. aber mit Staus aufgrund der LOS-Werte der DDG
3. DDG-TIM: Algorithmus für die Schaltung von dWiSta aufgrund von Reisezeitdifferenzen der Routen.

Ein Tag konnte wetterbedingt nicht gewertet werden. An den verbleibenden 29 Tagen wurden 36 Stauereignisse erkannt. Ein Stauereignis galt dann als erkannt, wenn bei einer der Online-Quellen ein Schwellwert erreicht wurde und endete, wenn alle Online-Quellen unter den Schwellenwerten lagen.

In der „Kategorie grob“ wird angegeben, wer das jeweilige Stauereignis erkannt hat:

Anzahl	in %	Anteil ca.	Kategorie grob	Kommentar
18	50	5 von 10	Nur LST	LST schaltet, DDG meldet nichts
11	31	3 von 10	Nur DDG	DDG schaltet, LST meldet nichts (LOS + TIM)
7	19	2 von 10	LST + DDG	teilweise Überschneidung
36	100		Total	

Die Übereinstimmung zwischen den LST-Daten und den durch FPD ermittelten Werte ist gering (ca. 20%). Über FPD wurden sehr wenige nachgewiesene Staus (Meldungen

des VWD) nicht erkannt (50% der Stauereignisse), d.h. dass die Erkennung von Ereignissen mit FPD noch nicht zuverlässig genug ist, um allein mit ihr Verkehrsinformationen zu generieren und Verkehrsmanagement zu betreiben.

Eine Aggregation der Daten (LST, DDG-LOS, DDG-TIM) erscheint sinnvoll, da ca. 30% der Ereignisse nur über FPD erkannt wurden. Das erste Erkennen eines Ereignisses konnte nicht auf eine Erfassungsart beschränkt werden. Dies wurde von allen Arten gleichmäßig erkannt. Das LOS-Kriterium hat jedoch immer als erstes ein Ereignisende gemeldet.

Die Einsparungen gemäß EWS während einer empfohlenen Umleitung liegen bei 4±1% der Kosten. Dieser Wert gilt nur für Umleitungen von der Route über Heilbronn (103km) auf die Route über Karlsruhe (91km).

3.4.3 Bewertung der Ergebnisse

Momentan ist die Qualität der Online-FP-Daten bezüglich LOS und Reisezeit noch nicht genügend, um die Tauglichkeit von FPD für eine kollektive Netzsteuerung auf der Masche A 5/A 6/A 81/A 8 zwischen Leonberg und Walldorf zu demonstrieren.

Die Berechnungen wurden offline durchgeführt, wobei die Input-Daten und -Formate identisch zu einer Online-Berechnung sind. Da das Netz relativ einfach ist, liegen die Rechenzeiten für die Kalibrierung, Umlegung und Simulation der drei Steuerungsphilosophien unter 3 Sekunden pro 3'-Intervall für die Steuerung der dWiSta auf der Netzmasche Leonberg-Walldorf.

Weiterer Untersuchungsbedarf, ob LOS und TIM anhand von FP-Daten auf A-Ebene weiter verbessert werden können erscheint lohnend.

Wie bereits bekannt ist die ausgeschilderte Route vom AD Leonberg zum AK Walldorf mit einer Länge von 103 km über Heilbronn, 12 km länger als die Alternativroute über Karlsruhe (91 km), die heute bereits von 70% des Durchgangsverkehrs genutzt wird (siehe Anl. 5 „Routenwahl im baden-württembergischen Autobahnviereck - Ergebnisse einer Kennzeichenerfassung“ des Meilensteinberichtes „Zentraler Meilenstein I“ vom April 2006). Eine Änderung der Ausschilderung (über Karlsruhe) und Ausweisung der heutigen Normalroute als Alternativroute würde dem heutigen Fahrverhalten besser entsprechen (kürzere Fahrstrecke und Reisezeit). Die Änderung dieser Beschilderung ist nach dem Ausbau der A 8 zwischen Leonberg und Karlsruhe geplant.

3.5 Evaluierung der Bestimmung von QZ-Matrizen und Tagesprognosen aus FPD

3.5.1 Beschreibung des Demonstrators

Zur Schätzung von QZ-Matrizen aus den Weg-Zeit-Informationen, die durch die Trajektorien zur Verfügung stehen, und den Daten der Querschnittsmessungen wird ein dynamisches Hochrechnungsverfahren verwendet (AP6). Die Verarbeitungskette der Daten ist in Abbildung 46 dargestellt.

Für jeden Untersuchungstag werden die Trajektorien aus dem LAC-basierten Trajektoriengenerierungsverfahren mit den lokalen Verkehrsstärken aus stationären Detektoren hochgerechnet. Ergebnis ist für jeden Tag eine QZ-Matrix. Die Matrizen werden dann nach den Vorgaben der Clusterung von den stationären Detektoren zu QZ-Matrizen von typischen Tagen (z.B. „Montag-Donnerstag Ferien“) zusammengefasst.

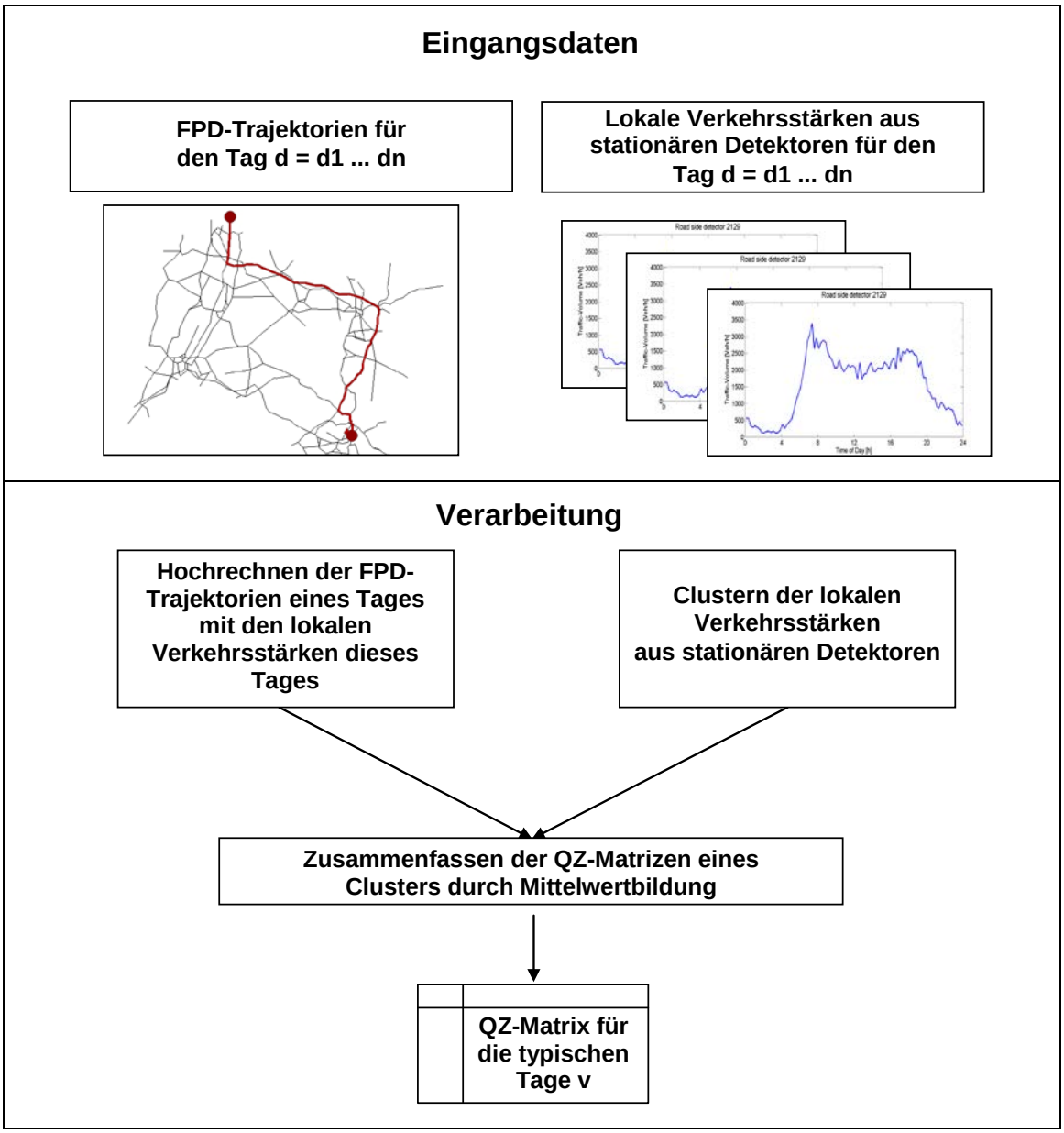


Abbildung 46: Erzeugung typischer QZ-Matrizen.

Für die hochgerechneten QZ-Matrizen existiert kein Referenzwert, d.h. es gibt keine Information über die „richtige“ QZ-Matrix im Projektnetz. Diese kann auch nicht mit alternativen Verfahren für ausgewählte Tage erzeugt werden. Die Qualitätsbewertung erfolgt daher über eine Beschreibung der Vollständigkeit der Eingangsdaten für die Erzeugung von QZ-Matrizen:

- Abdeckung des Netzes durch die eingegangenen FPD-Trajektorien.
- Abdeckung des Netzes durch die eingegangenen lokalen Detektoren.
- Verbleibende Abweichung zwischen hochgerechneten FPD-Belastungen und lokalen Verkehrsstärken.

Die Tagesprognose (AP7) basiert auf der Zuordnung des zu prognostizierenden Tages in die vorhandenen Clustergruppen. Aus dem zugehörigen Cluster ergibt sich die QZ-

Matrix, die die Nachfrage des betrachteten Tages beschreibt. Die Auswirkungen auf das Verkehrsnetz werden unter Berücksichtigung aktueller Gegebenheiten, wie zum Beispiel Baustellen, durch eine Verkehrsumlegung berechnet. Ergebnis sind u.a. Streckenbelastungen sowie Fahrzeiten auf Strecken und Routen. Die Bewertung der Qualität erfolgt mit zwei Kenngrößen:

- Vergleich der Streckenbelastungen auf ausgewählten Querschnitten.
- Vergleich der LOS-Darstellung auf allen Strecken.

Sowohl für die Hochrechnung der QZ-Matrizen als auch für die Tagesprognose gibt es keine standardisierten Bewertungsverfahren. Vielmehr werden im Einzelfall Evaluierungskriterien und zu erreichende Qualitätsmäße festgelegt. Daher werden in Do-iT für die genannten Kriterien keine Zielvorgaben gemacht.

3.5.2 Analyse der durchgeführten Feldversuche

Hochrechnung: Abdeckung des Netzes durch die eingegangen Trajektorien

Zu Beschreibung der Vollständigkeit der Trajektorien aus dem LAC-basierten Trajektoriengenerierungsverfahren wird geprüft, in wie vielen 15min-Intervallen es an für diesen Tag verfügbaren RMQs keine Trajektorienbelastung gibt. Dafür werden die RMQs auf den Autobahnen A5, A6, A8 und A81 zwischen den Autobahnknoten Leonberg, Karlsruhe, Walldorf und Weinsberg herangezogen. Bei einer durchschnittlichen Anzahl von ca. 47 verwendeten RMQ ergeben sich ca. 4.500 RMQ-Intervalle, die geprüft werden.

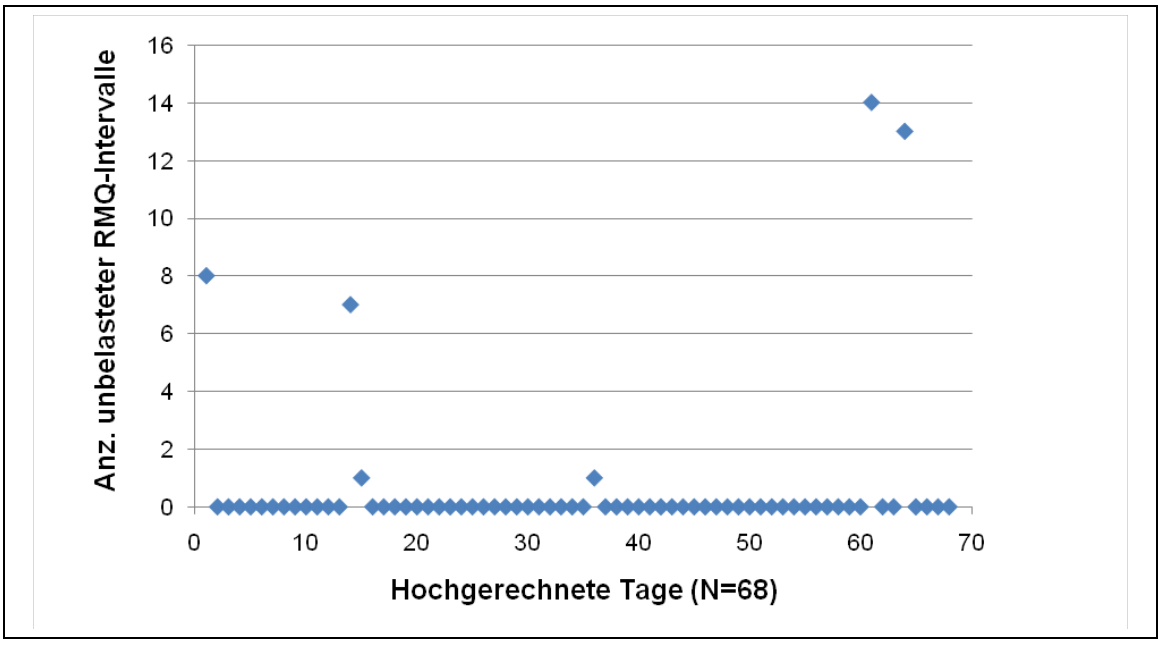


Abbildung 47: Abdeckung durch Trajektorien an den verwendeten RMQs.

Die folgende Abbildung 47 zeigt als Ergebnis die Anzahl RMQ-Intervalle ohne Trajektorienbelastung. Es zeigt sich, dass über 90 % der Tage (62 von 68 Tagen) eine vollständige Abdeckung mit Trajektorien haben. Die verbleibenden 6 Tage haben nur vereinzelte RMQ-Intervalle ohne Belastung während der Nacht. Diese hohe Abdeckung kann dadurch erreicht werden, dass nur Tage hochgerechnet, bei denen mindestens 1.350 A-Daten-Rohdateien (theoretisches Maximum: 1.440) vorhanden sind. Die hohe Abdeckung zeigt aber trotzdem auch die Qualität des LAC-basierten Trajektoriengenerierungsverfahrens, das zu jeder Tageszeit Trajektorien auf allen Strecken erzeugt.

Hochrechnung: Abdeckung des Netzes durch die eingegangenen lokalen Detektoren

Die Verfügbarkeit der RMQ schwankt über die hochgerechneten Tage. Dies liegt an einer umfangreichen Vorfilterung, die zum einen die Qualität der Messwerte sicherstellen soll, und zum anderen einen RMQ nur dann für die Hochrechnung auswählt, wenn für den kompletten Tag aufbereitete Daten vorhanden sind. Die folgende Abbildung 48 zeigt auf der linken Achse (blaue Quadrate) die Anzahl der verwendeten RMQ, die zwischen 38 und 53 liegt. Diese Werte sind aber nur bedingt für die Abdeckung für die Hochrechnung aussagekräftig, da dafür auch die Lage der verfügbaren RMQ wichtig ist. Deshalb werden in einer weiteren Analyse die vier betrachteten Autobahnen A5, A6, A8 und A81 in jeweils zwei Abschnitte pro Richtung unterteilt und geprüft, wie viele dieser insgesamt 16 Abschnitte durch mindestens einen RMQ abgedeckt sind. Das Ergebnis ist ebenfalls in Abbildung 48 abgedruckt (rechte Achse, rote Kreuze). Hier liegen die Werte zwischen 81 und 100 %. Im ungünstigsten Fall sind damit drei Abschnitte ohne verfügbaren RMQ, eine Hochrechnung der Trajektorien, die nur über diese Abschnitte fahren erfolgt daher mit dem durchschnittlichen Hochrechnungsfaktor.

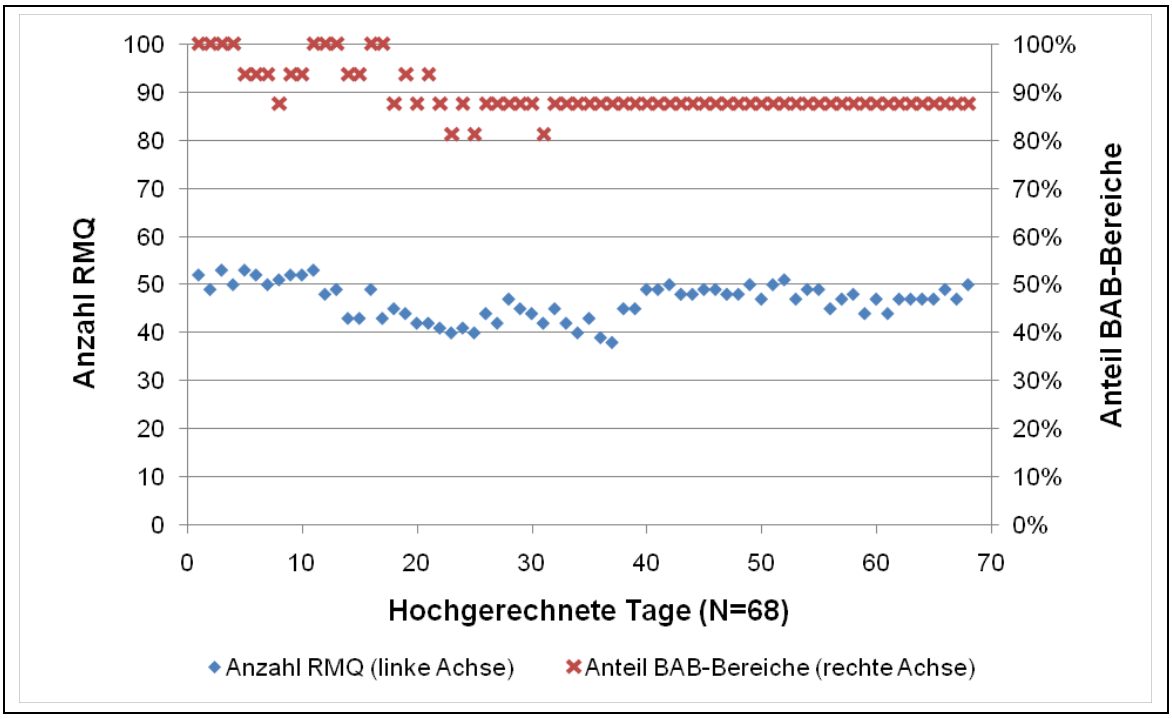


Abbildung 48: Abdeckung durch Trajektorien in BAB-Netzmasche.

Hochrechnung: Verbleibende Abweichung zwischen hochgerechneten FPD-Belastungen und lokalen Verkehrsstärken

Die durchgeführten Hochrechnungen haben gezeigt, dass sich die hochgerechneten FPD-Belastungen mit steigender Anzahl von Iterationen immer mehr den gemessenen Verkehrsstärken annähern. Das Hochrechnungsverfahren ist also in der Lage, die Werte der RMQ sehr gut zu erreichen. Darüberhinaus hat sich gezeigt, dass eine zu genaue Erreichung der RMQ-Werte gar nicht erstrebenswert ist, da diese teilweise hohen Schwankungen unterliegen (siehe dazu den Bericht zur LAC-basierten Trajektoriengenerierung aus A-Daten und zu den Arbeitspaketen 5-7, Anlage zum Schlussbericht).

Die Hochrechnung, deren Prozess iterativ ist, wird dann beendet, wenn

1. der durchschnittliche GEH[5] über alle RMQ-Zeitintervalle kleiner oder gleich 0,25 ist, oder wenn
2. die Änderung des durchschnittlichen GEH in den letzten fünf Zeitintervallen kleiner als 2 %

ist. Dementsprechend liegt der durchschnittliche GEH bei ca. 0,25, wenn das 1. Kriterium zur Beendigung der Hochrechnungen geführt hat. Aber auch, wenn das zweite Abbruchkriterium maßgebend war, liegt der durchschnittliche GEH bis auf eine Ausnahme in diesem Bereich (vgl. Abbildung 49). Alle erreichten Werte liegen in einem so

[5] Zum GEH-Wert siehe den Bericht zur LAC-basierten Trajektoriengenerierung aus A-Daten und zu den Arbeitspaketen 5-7 (VuV)

geringen Bereich, der in der Regel bei anderen verkehrsplanerischen Anwendungen nicht erreicht wird und vermutlich auch über der Genauigkeit der Zählstellen liegt.

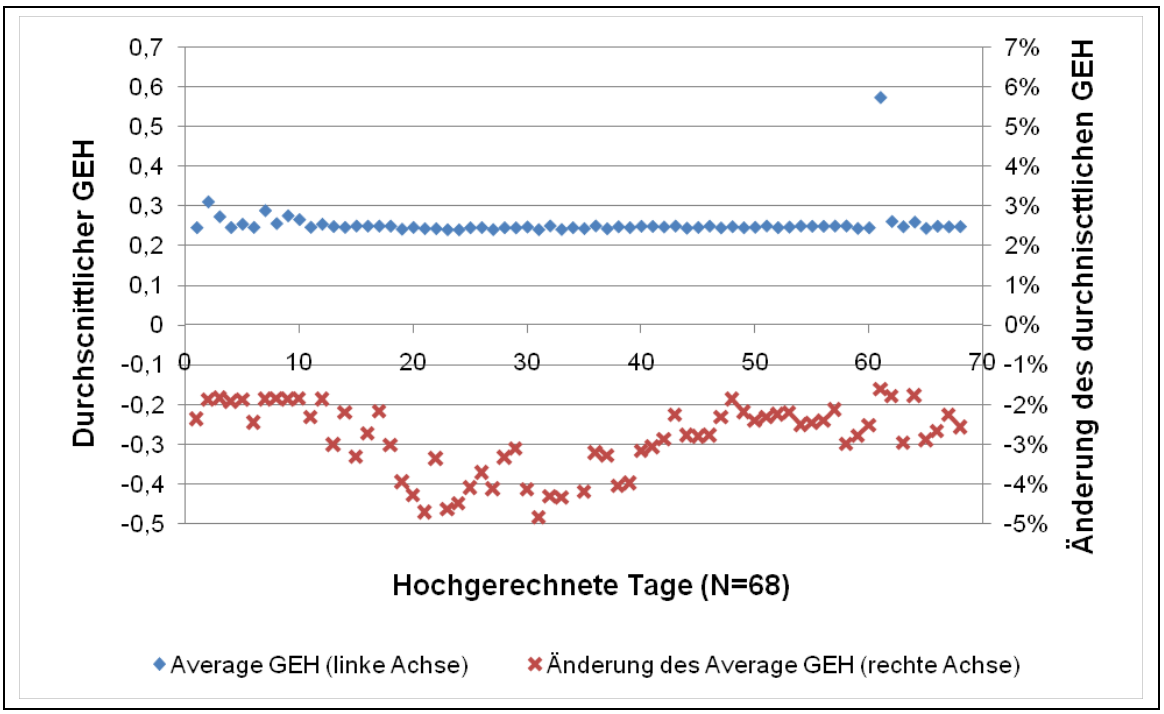


Abbildung 49: Abdeckung durch Trajektorien.

Tagesprognose: Vergleich der Streckenbelastungen

Auf ausgewählten Kontrollstrecken im Autobahnnetz (zwei je BAB-Achse und Richtung) werden die stündlichen modellierten Verkehrsbelastungen aus der Tagesprognose mit den in der Realität gemessenen Verkehrsbelastungen verglichen. Im Ergebnis erhält man somit für jede Stunde 16 GEH-Werte, zwei je Strecke (Hin-und Rückrichtung). Je Stunde werden Mittelwert sowie Maximalwert bestimmt und anschließend über 24 Stunden gemittelt. Abbildung 50 zeigt den mittleren und maximalen GEH für die 68 untersuchten Tage:

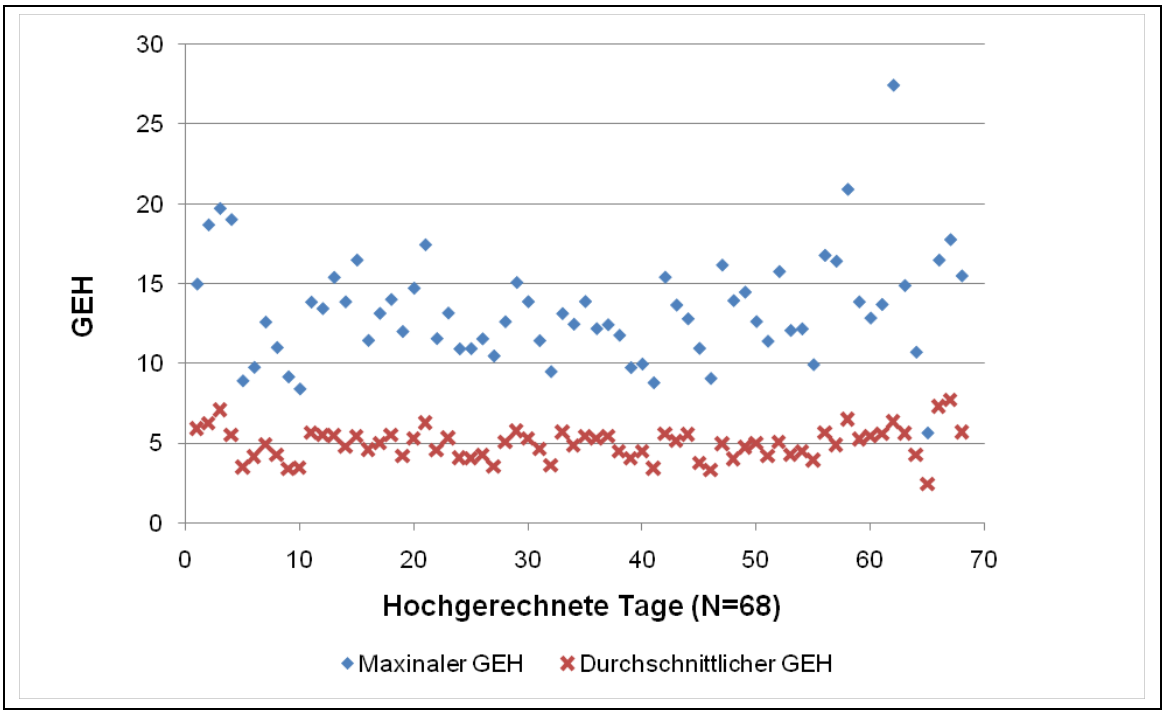


Abbildung 50: Mittlerer und maximaler GEH zwischen gemessenen Streckenbelastungen und den Belastungen aus der Tagesprognose.

Ein GEH-Wert von 5 wird im Allgemeinen für Verkehrsplanungsmodelle als gut bewertet, dieser sollte in 85% der Fälle nicht überschritten werden. Bei der Tagesprognose liegt der mittlere GEH bei ca. 5 %, somit kann die oben genannte Vorgabe bei der Tagesprognose erwartungsgemäß nicht erreicht werden

Da für die Tagesprognose die repräsentative QZ-Matrix des jeweiligen Clusters (z.B. Mo-Do Ferien) verwendet wird und dies im Gegensatz zu klassischen Verkehrsplanungsmodellen der QZ-Matrix bereits eine Prognose darstellt, sind relativ hohe GEH-Werte nicht zu vermeiden. Zu dieser systematischen Abweichung ergeben sich weitere zufällige Abweichungen durch Unfälle etc.

Tagesprognose: Vergleich der LOS-Darstellung

Darstellungen der Verkehrslage erfolgen in der Regel nicht über Verkehrsbelastungen, sondern über die Angabe einer Qualitätsstufen des Verkehrsablaufs (Level of Service, LOS). Dabei haben sich Darstellungen in drei Stufen (rot-gelb-grün für frei-stockend-Stau) durchgesetzt. Die folgende Tabelle 22 zeigt die für die folgende Analyse verwendeten Auslastungsgrade zur Einstufung der Streckenbelastungen in die drei Qualitätsstufen:

Tabelle 22: Ableitung der Qualitätsstufen.

Qualitätsstufe	Auslastung
Frei	< 80%
Stockend	80% bis 99%
Gestaut	> 99 %

In der folgenden Abbildung 51Abbildung 51 ist dargestellt, welcher Anteil der prognostizierten Qualitätsstufen tatsächlich eingetreten ist. Dafür werden alle BAB-Strecken im Autobahnviereck analysiert, indem für die nicht durch Richtungsmessquerschnitte versorgten Strecken die Belastungen aus der für diesen Tag hochgerechneten Matrix herangezogen werden. Dabei ist zu beachten, dass durch die Klassenbildung auch sehr nah beieinander liegende Auslastungen (z.B. 79,8 % und 80,1 %) bereits als falsche Prognose erkannt werden können.

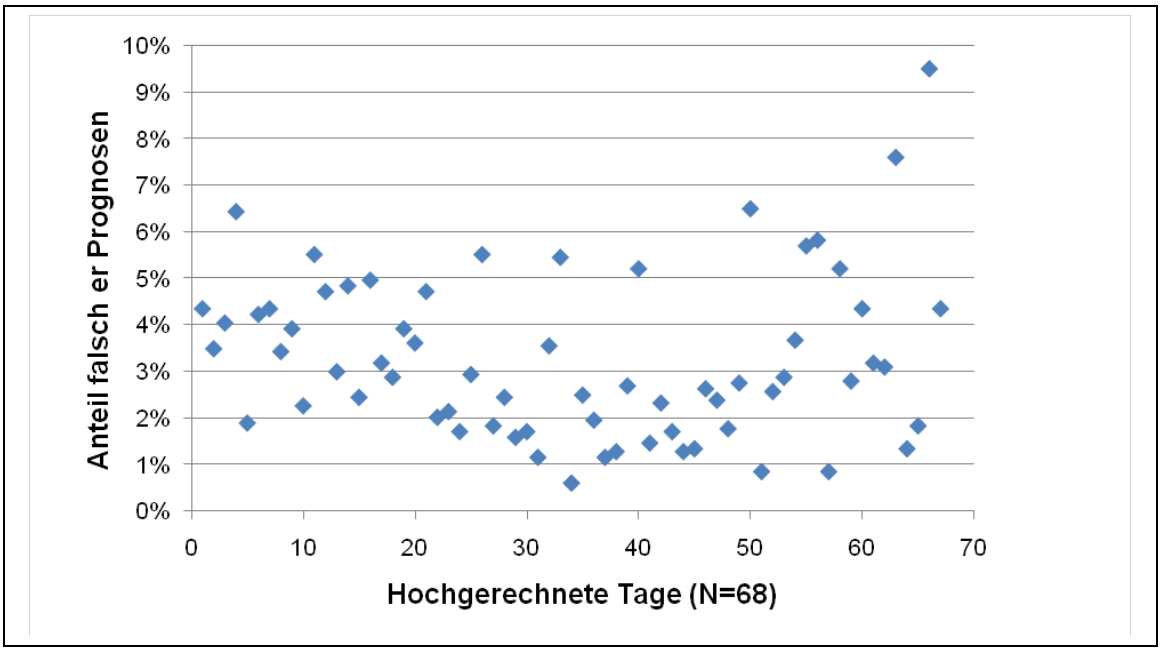


Abbildung 51: Anteil der korrekt prognostizierten Qualitätsstufen.

Der durchschnittliche Wert über alle untersuchten Tage liegt bei ca. 3 %. Das bedeutet, dass durchschnittlich je Strecke seltener als einmal am Tag eine Qualitätsstufe falsch vorhergesagt wird.

3.5.3 Bewertung der Ergebnisse

Für die *Generierung von QZ-Matrizen* können sowohl die Trajektorien aus dem LAC-basierten Trajektoriengenerierungsverfahrens als auch das Hochrechnungsverfahren als sehr gut geeignet bezeichnet werden. Verbesserungspotential wird dagegen bei der

Verfügbarkeit und Korrektheit der RMQ gesehen. Trotz dieser Einschränkung können die erreichten Ergebnisse als positiv und brauchbar bezeichnet werden.

Die Durchführung der *Tagesprognose* hat gezeigt, dass im Durchschnitt ca. 97 % aller prognostizierten Qualitätsstufen korrekt sind und somit durchschnittlicher weniger als ein Zeitintervall pro Stunde nicht korrekt prognostiziert wird.

In diesem Abschnitt 3.5 wurden vor allem Analysen für die gesamte Analyseperiode gezeigt. Detaillierte Beschreibungen der Verfahren sowie Analysen einzelner Tage sind im Bericht zur LAC-basierten Trajektoriengenerierung aus A-Daten und zu den Arbeitspaketen 5-7 (VuV) zu finden.

4 Zusammenfassende Bewertung

Die Evaluierung der Trajektorien aus Mobilfunkdaten mit Hilfe stationärer Erfassungssysteme hat gezeigt, dass diese Datenquelle grundsätzlich für verkehrstechnische Anwendungen geeignet ist. Es ergeben sich aufgrund des begrenzten Auflösungsvermögens insbesondere innerstädtisch einige Einschränkungen in den möglichen Anwendungen insbesondere hinsichtlich der Verfügbarkeit und Korrektheit der Trajektorien.

Grundsätzlich ist die Integration der Daten zur Ergänzung und Unterstützung der innerstädtischen Datenquellen zur Auslösung bestimmter verkehrsbeeinflussender Maßnahmen jedoch möglich. Dies wurde mit Abis-Daten im Demonstrator der Stadt Karlsruhe nachgewiesen (vgl. Kapitel 3.2). Wie die Evaluierung der Verkehrslage aus MFD auf der Südtangente gezeigt hat, wird die noch zu niedrige Detektionsrate im Wesentlichen durch die geringe Verfügbarkeit verursacht. Eine Steigerung der Verfügbarkeit z. B. durch eine Beschleunigung der Berechnung und eine Kombination mit A-Daten könnte die Mobilfunkdaten auch in Stadtnähe zu einer gut zur Verkehrslageschätzung geeigneten und kontinuierlichen Datenquelle machen.

Die Verwendung von A-Trajektorien zur Steuerung der Netzbeeinflussungsanlagen im Autobahnviereck und zur Angabe von Reisezeitinformationen konnte im Untersuchungszeitraum bis Anfang März 2009 nicht demonstriert werden (vgl. Kapitel 3.4). Allerdings wurde erst am 20. März die erste online lauffähige, jedoch aufgrund des enormen Zeitdrucks rudimentäre Version zur A-Daten-Prozessierung durch eine überarbeitete Version abgelöst. Mit dieser konnte die Qualität der Trajektorien wesentlich gesteigert werden, so dass die Verwendbarkeit im genannten Demonstrator nun nochmals geprüft werden müsste. Die Vergleiche der Tagesganglinien im Autobahnnetz mit den SES-Daten zeigen durchweg hohe Durchdringungen und Korrelationen, daher ist die Eignung der A-Daten zur Steuerung von Netzbeeinflussungsanlagen zu erwarten (vgl. Kapitel 2.3.2).

Wie in Kapitel 3.5 nachgewiesen wurde, sind Mobilfunkdaten grundsätzlich gut zur Aufstellung von Quelle-Ziel-Matrizen geeignet. Damit bieten sie eine sehr günstige und flächendeckende Alternative zu den teuren und daher zeitlich und räumlich nur sehr diskret einsetzbaren Kennzeichenerfassungssystemen.

5 Literatur

AP4-6 (2008)	Do-iT: *Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren*. Interner Do-iT Bericht zum AP 4-6, Bestandteil des ZMS3, 2008.
AP10-1 (2008)	Do-iT: *Evaluierungskonzept.* Interner Do-iT Bericht zum AP10-1, Bestandteil des ZMS3, 2008.
AP4-6 / AP10-3 (2009)	Do-iT: *Zusammenfassender Bericht: Qualitätsbewertung Map-Matching gestützter Positionsbestimmungsverfahren auf Abis- und A-Ebene sowie Evaluierung der Positionsbestimmung durch höherwertige Ortungsverfahren.* Interner Do-iT Abschlussbericht, Bestandteil des ZMS4, 2009.
AP8-7 (2009)	Do-iT: *Nachweis der Wirksamkeit der Qualitätssicherungsmaßnahmen*. Interner Do-iT Bericht zum AP 8-7, Bestandteil des ZMS4, 2009.
Karrer (2008)	Karrer, K.: *Untersuchung der Qualität von stationären Erfassungssystemen am Beispiel der Autobahn A81.* Studienarbeit am IAGB, Universität Stuttgart, nicht veröffentlicht, 2008.
Sachs (2006)	L. Sachs, J. Hedderich: *Angewandte Statistik.* Springer : Berlin, Heidelberg, 2006.
Sachse (2002)	Sachse, T.: *Die Messwertqualität autarker Detektionssysteme am Beispiel der Traffic Eyes Universal*. Straßenverkehrstechnik, Nr. 9, Seiten 463-470, Kirschbaum-Verlag GmbH : Bonn, 2002.
TLS (1997)	Bundesanstalt für Straßenwesen BaSt: Technische Lieferbedingungen für Streckenstationen TLS, 1997.

Anhang A

Häufigkeit der auftretenden REID in den FPD-Trajektorien im gesamten Projektnetz sowie in Karlsruhe (A und Abis) und Stuttgart (A):

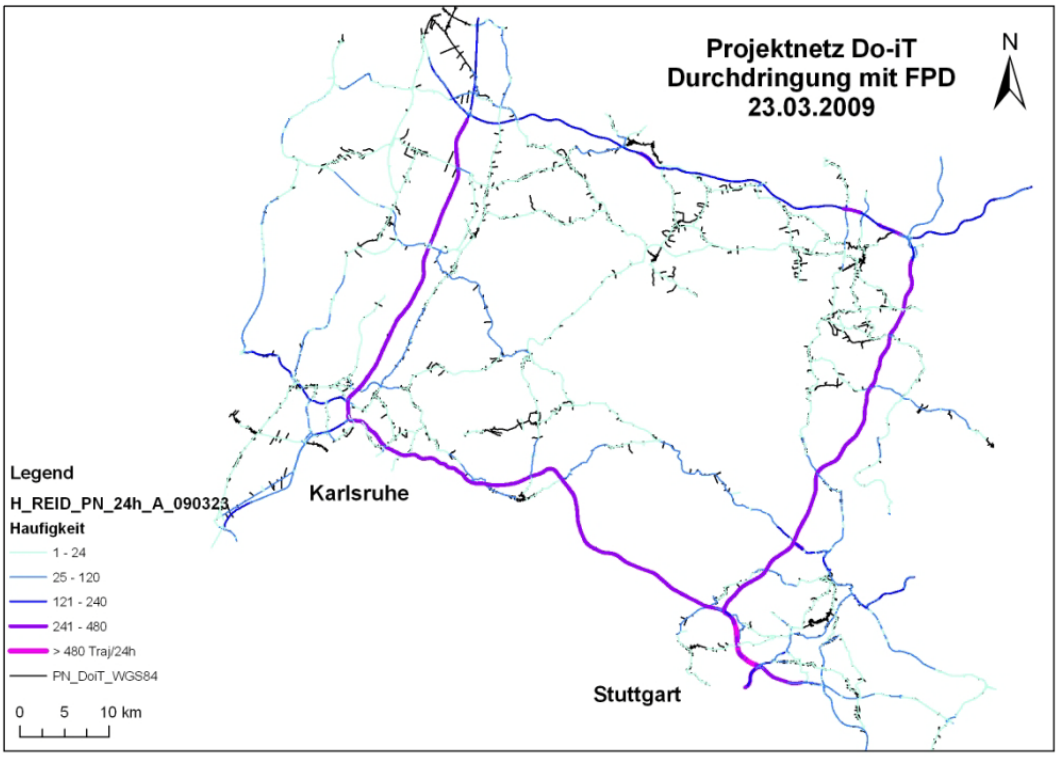


Abbildung 52: Häufigkeit auftretender RE im PN am Montag, 23.03.09 (A)

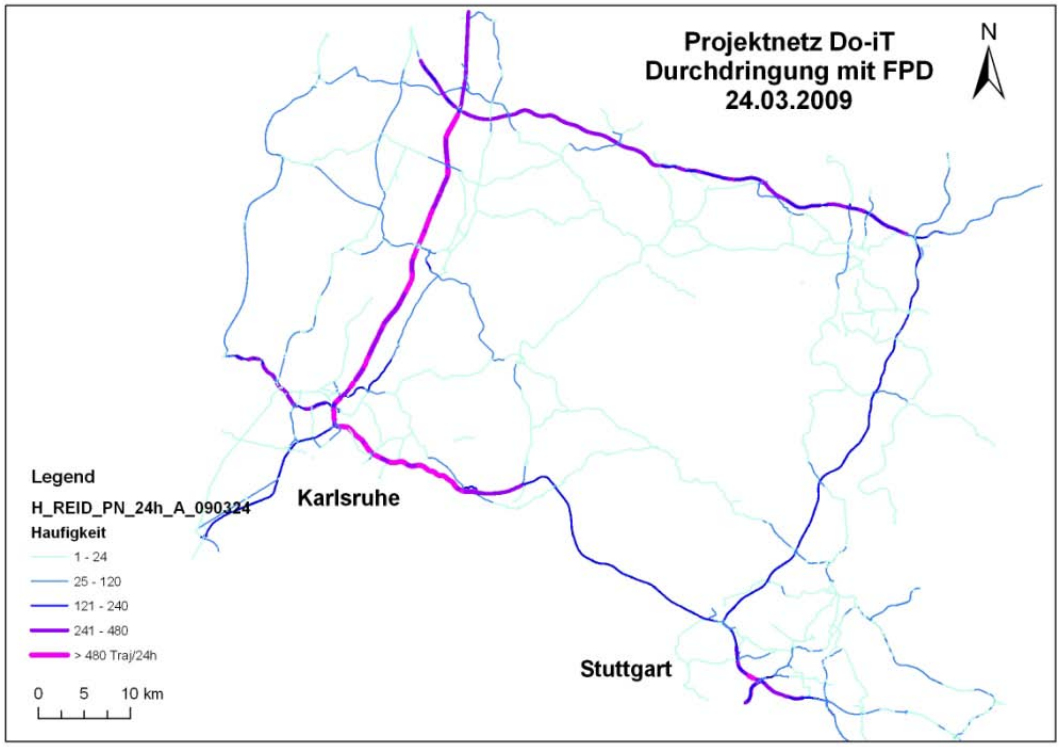


Abbildung 53: Häufigkeit auftretender RE im PN am Dienstag, 24.03.09 (A)

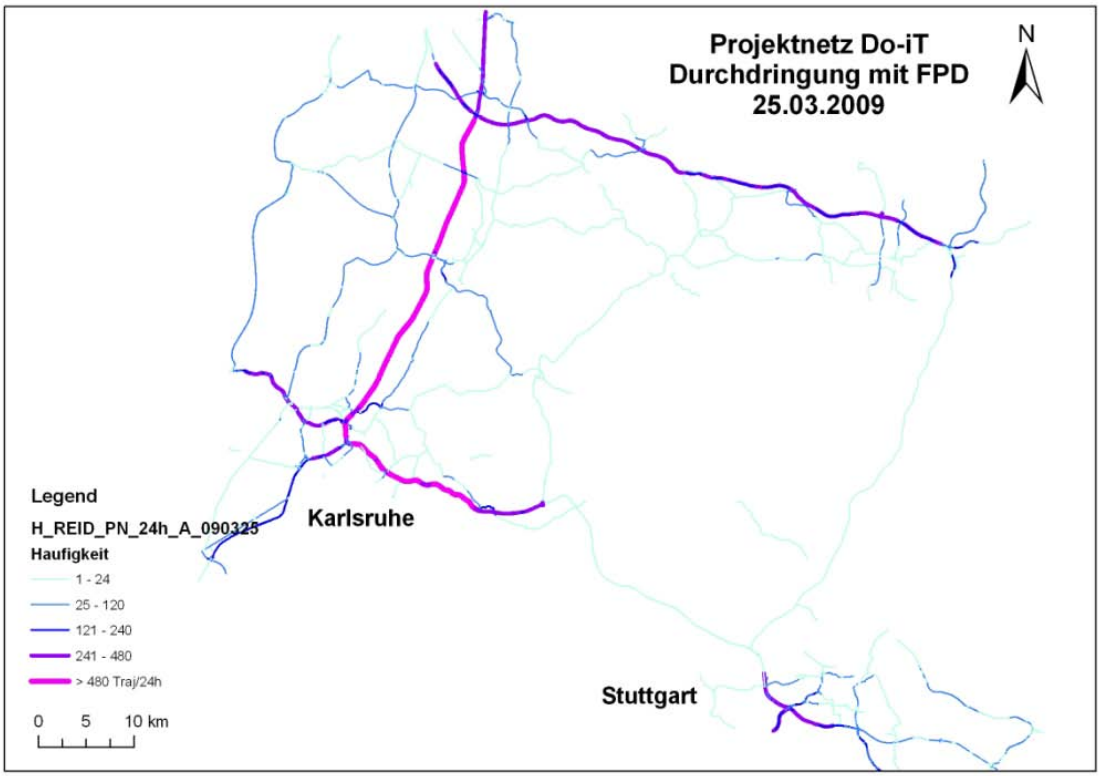


Abbildung 54: Häufigkeit auftretender RE im PN am Mittwoch, 25.03.09 (A)

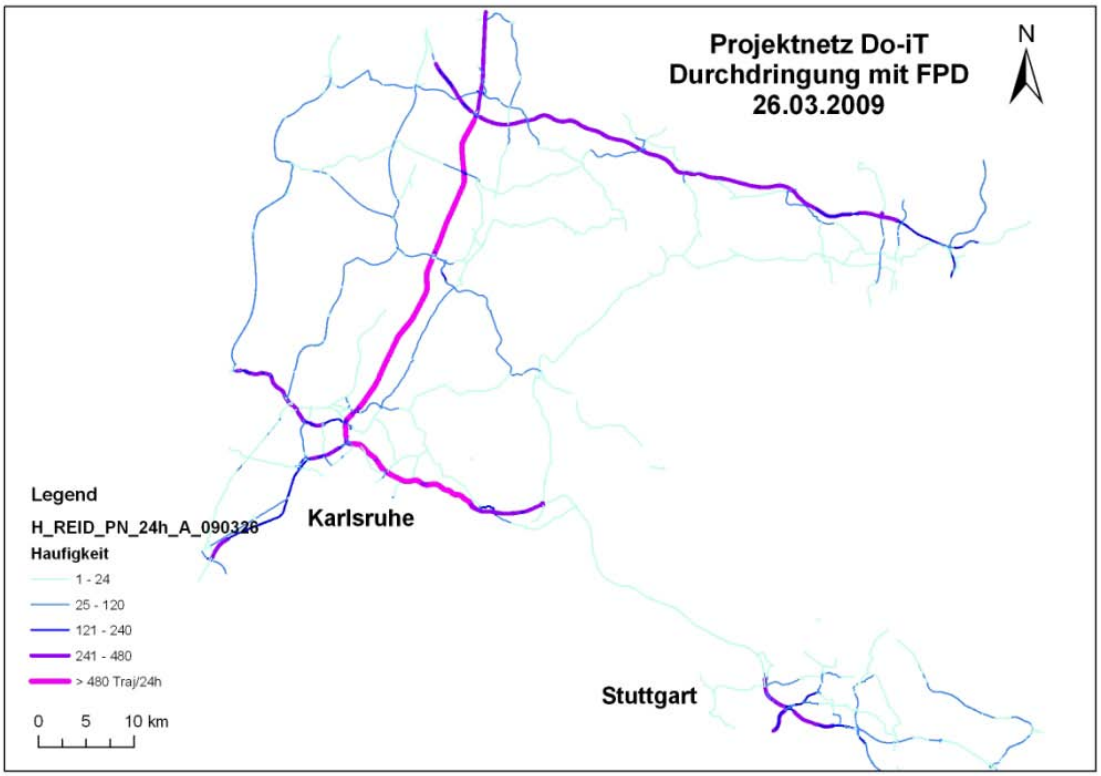


Abbildung 55: Häufigkeit auftretender RE im PN am Donnerstag, 26.03.09 (A)

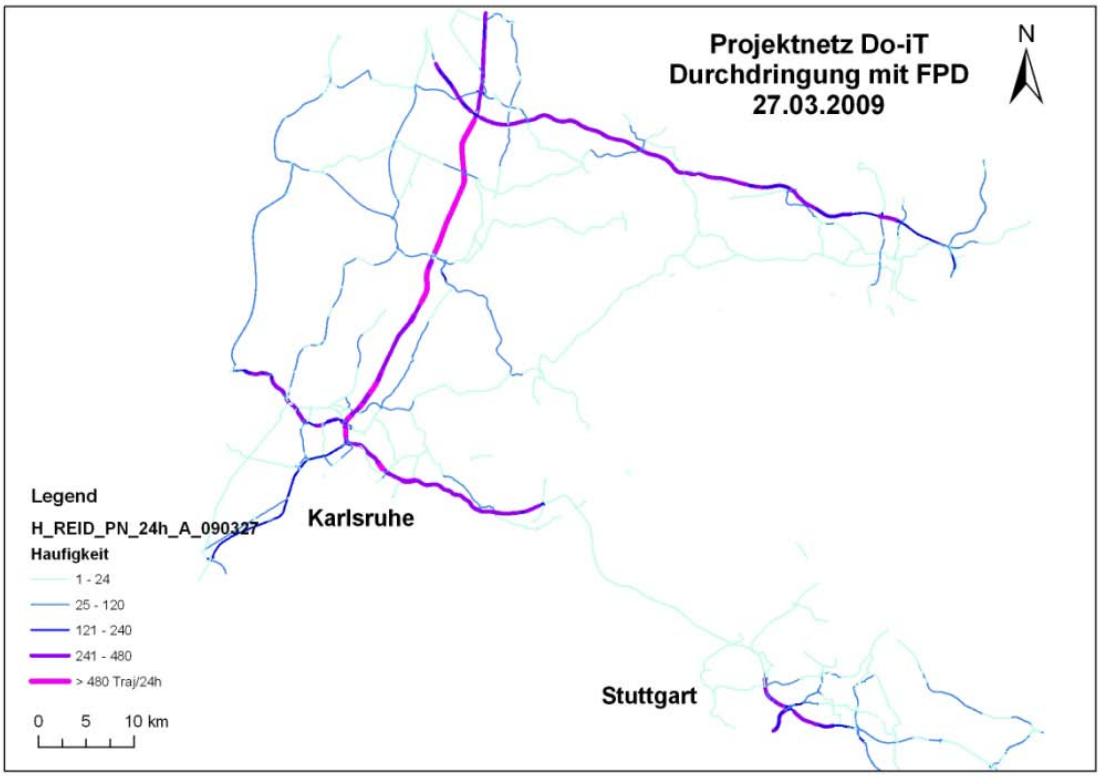


Abbildung 56: Häufigkeit auftretender RE im PN am Freitag, 27.03.09 (A)

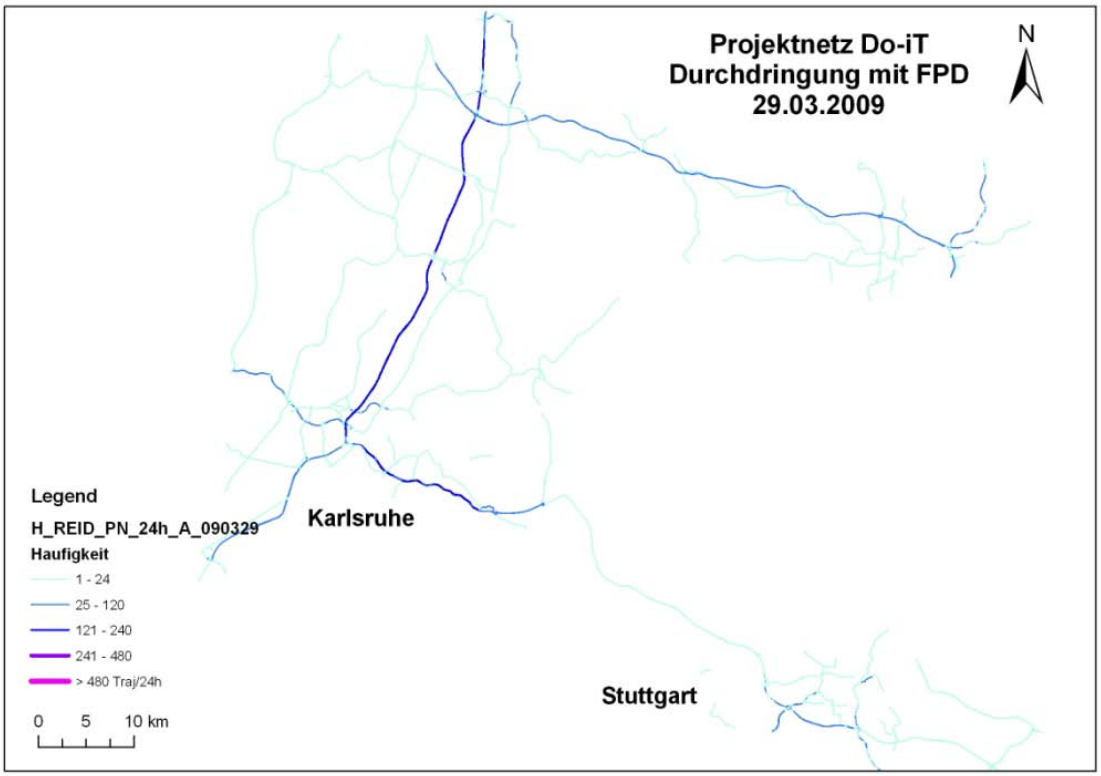


Abbildung 57: Häufigkeit auftretender RE im PN am Sonntag, 29.03.09 (A)

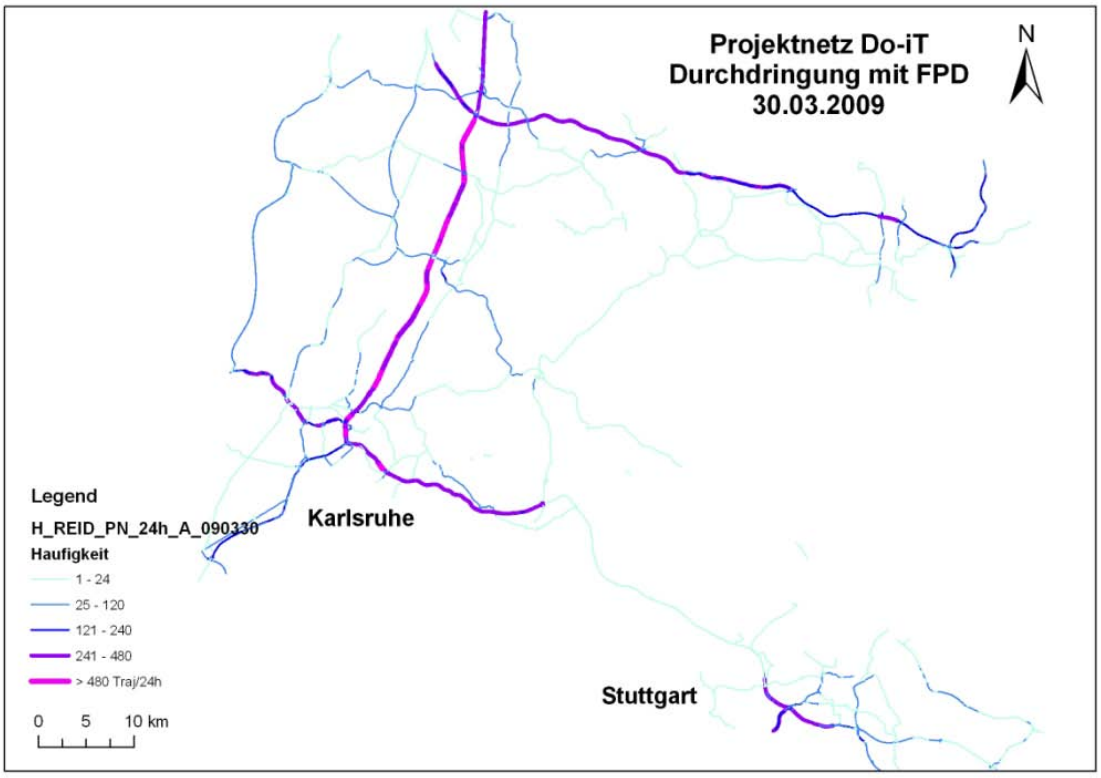


Abbildung 58: Häufigkeit auftretender RE im PN am Montag, 30.03.09 (A)

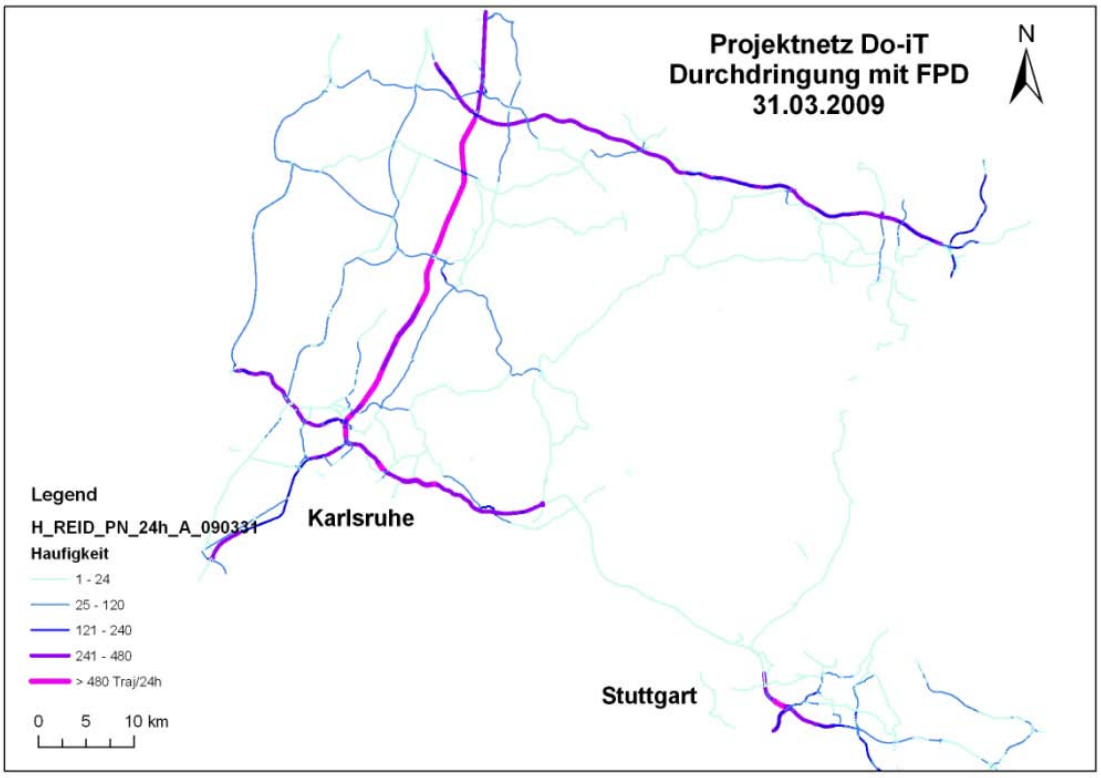


Abbildung 59: Häufigkeit auftretender RE im PN am Dienstag, 31.03.09 (A)

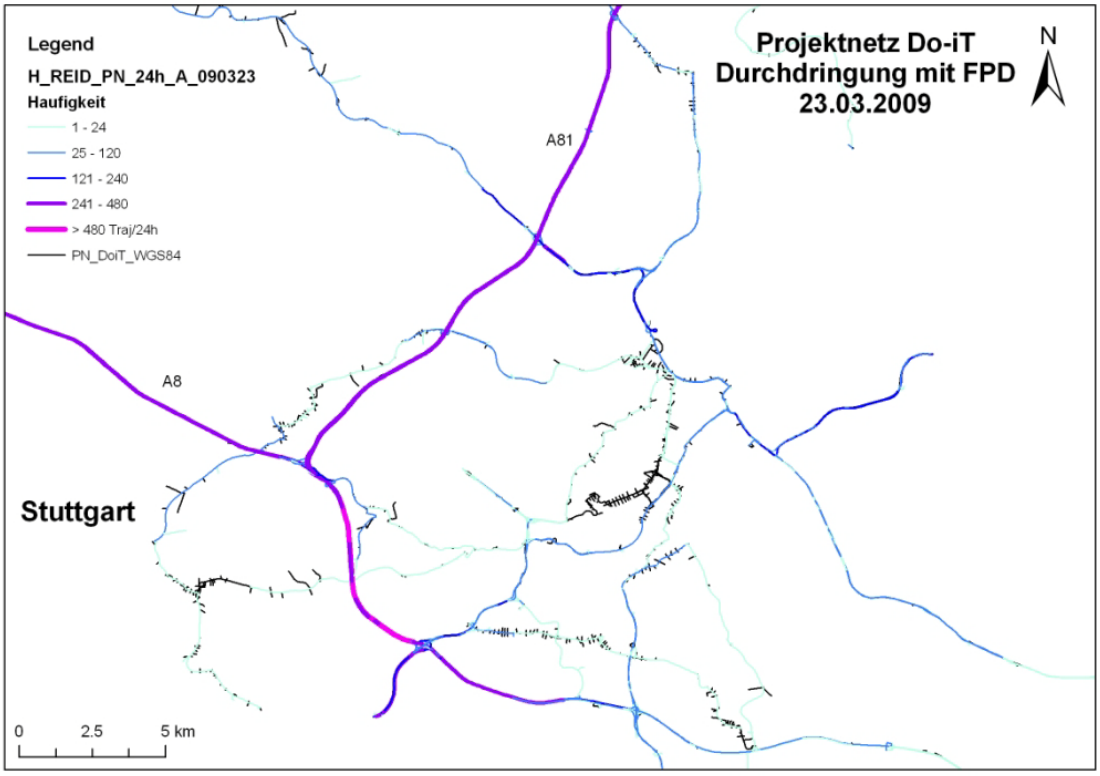


Abbildung 60: Häufigkeit auftretender RE im Raum Stuttgart am Montag, 23.03.09 (A)

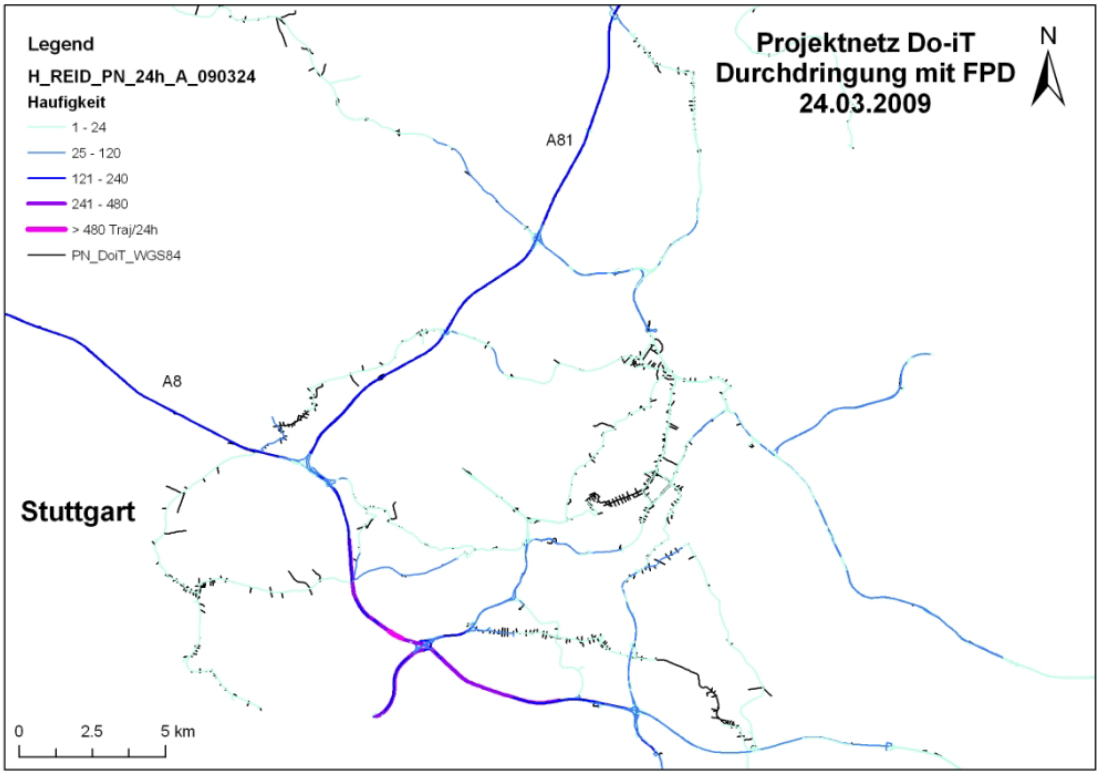


Abbildung 61: Häufigkeit auftretender RE im Raum Stuttgart am Dienstag, 24.03.09 (A)

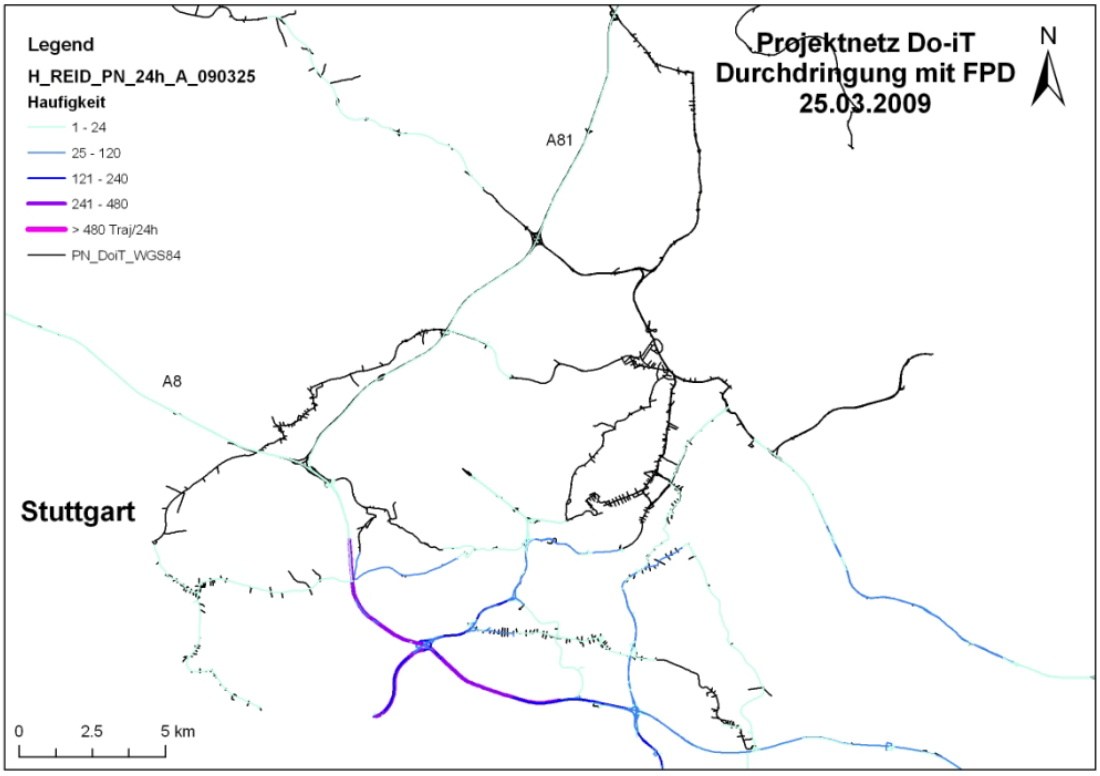


Abbildung 62: Häufigkeit auftretender RE im Raum Stuttgart am Mittwoch, 25.03.09 (A)

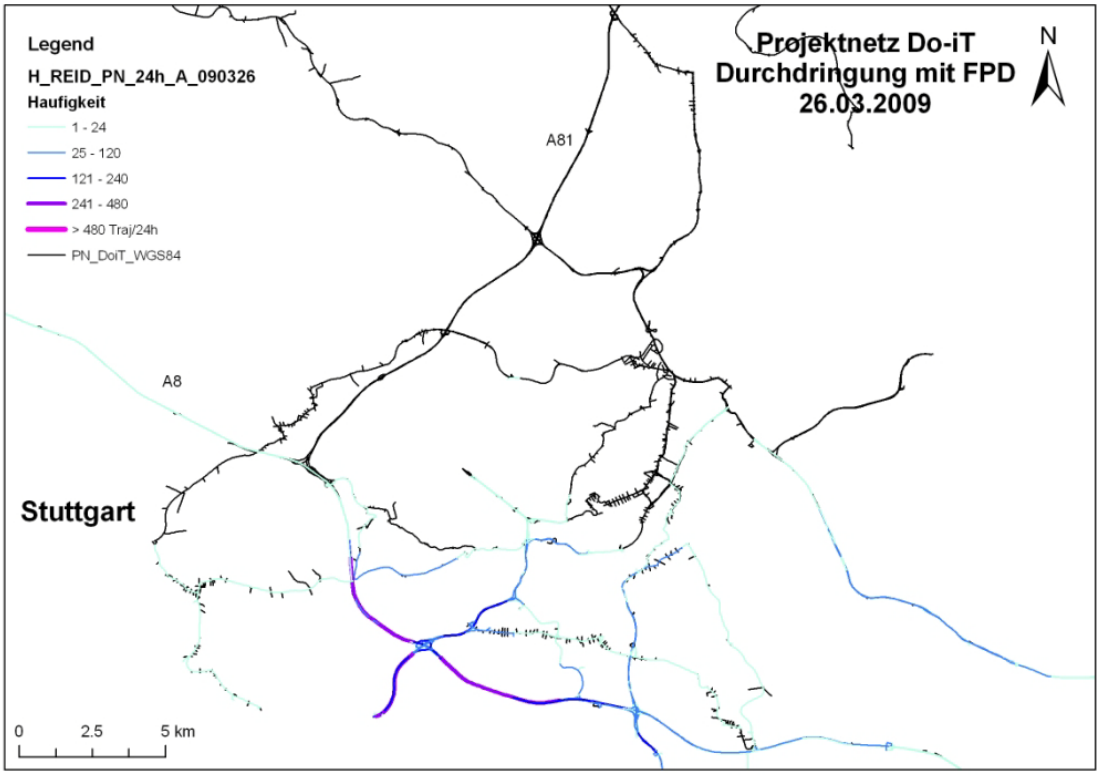


Abbildung 63: Häufigkeit auftretender RE im Raum Stuttgart, Donnerstag, 26.03.09 (A)

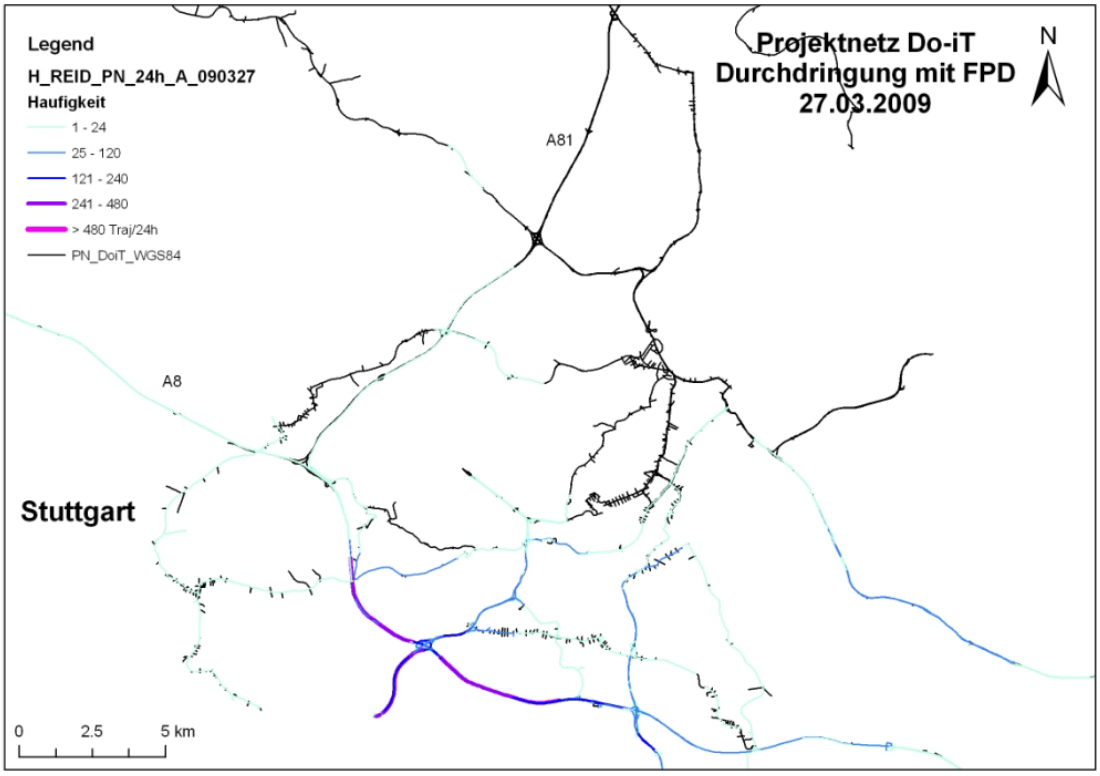


Abbildung 64: Häufigkeit auftretender RE im Raum Stuttgart am Freitag, 27.03.09 (A)

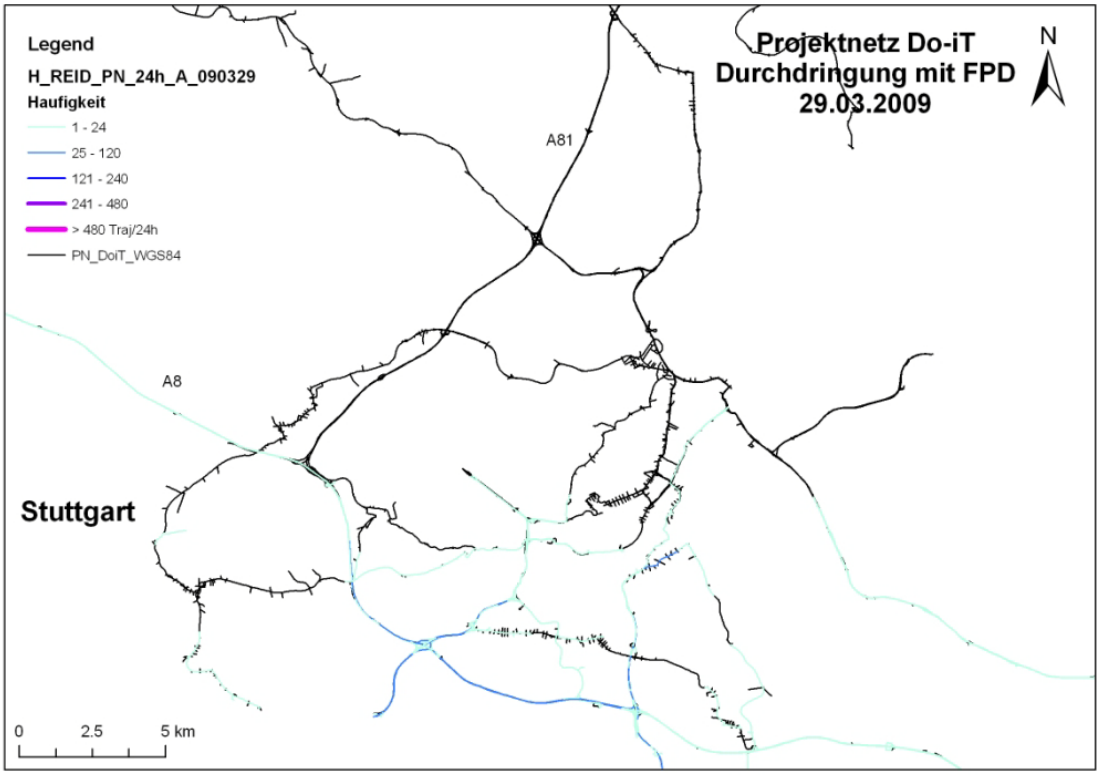


Abbildung 65: Häufigkeit auftretender RE im Raum Stuttgart am Sonntag, 29.03.09 (A)

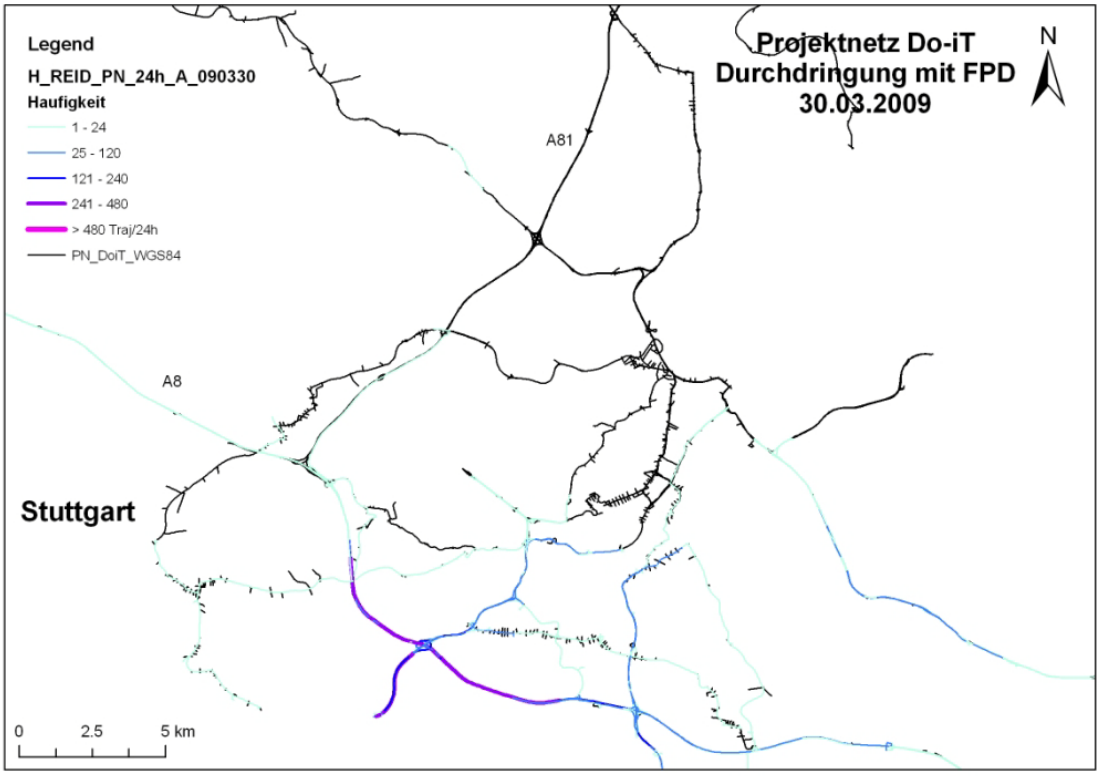


Abbildung 66: Häufigkeit auftretender RE im Raum Stuttgart am Montag, 30.03.09 (A)

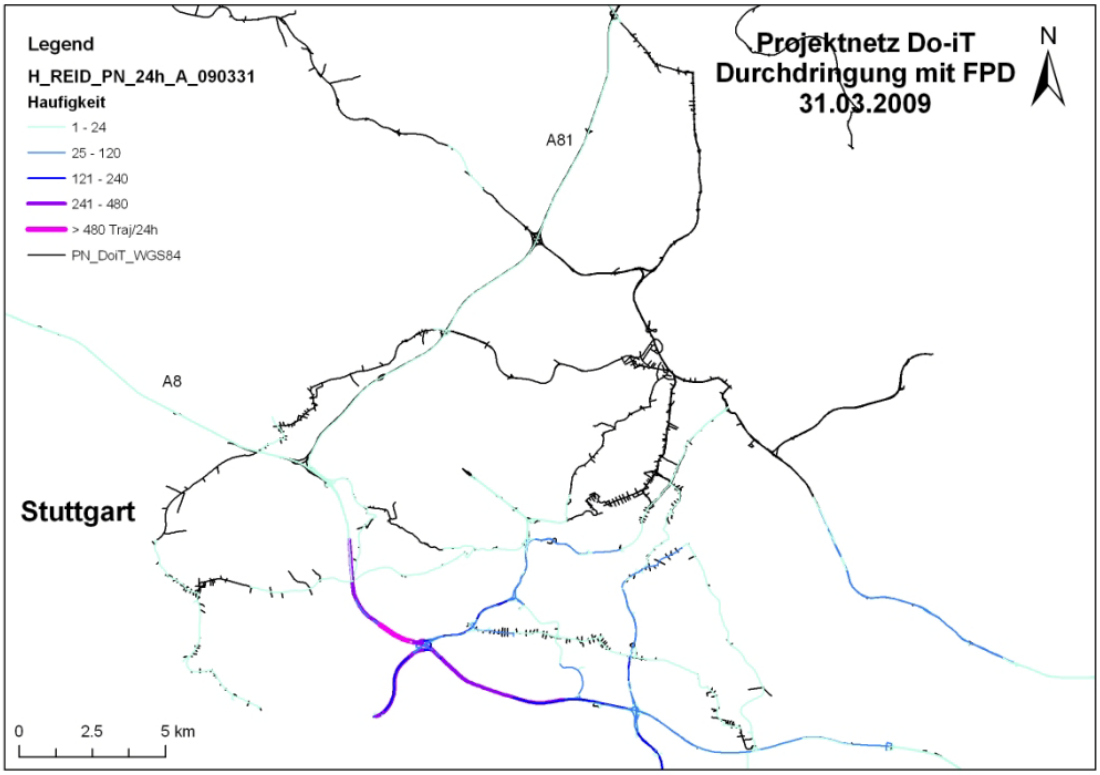


Abbildung 67: Häufigkeit auftretender RE im Raum Stuttgart am Dienstag, 31.03.09 (A)

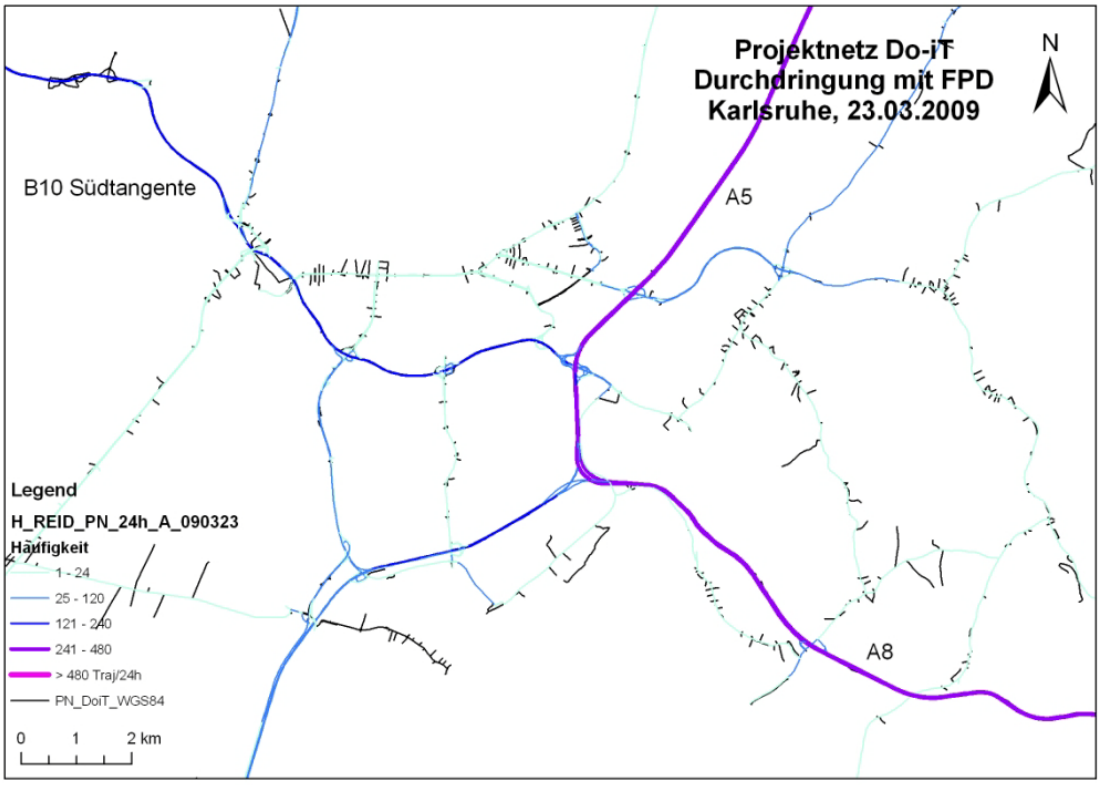


Abbildung 68: Häufigkeit auftretender RE im Raum Karlsruhe, Montag, 23.03.09 (A)

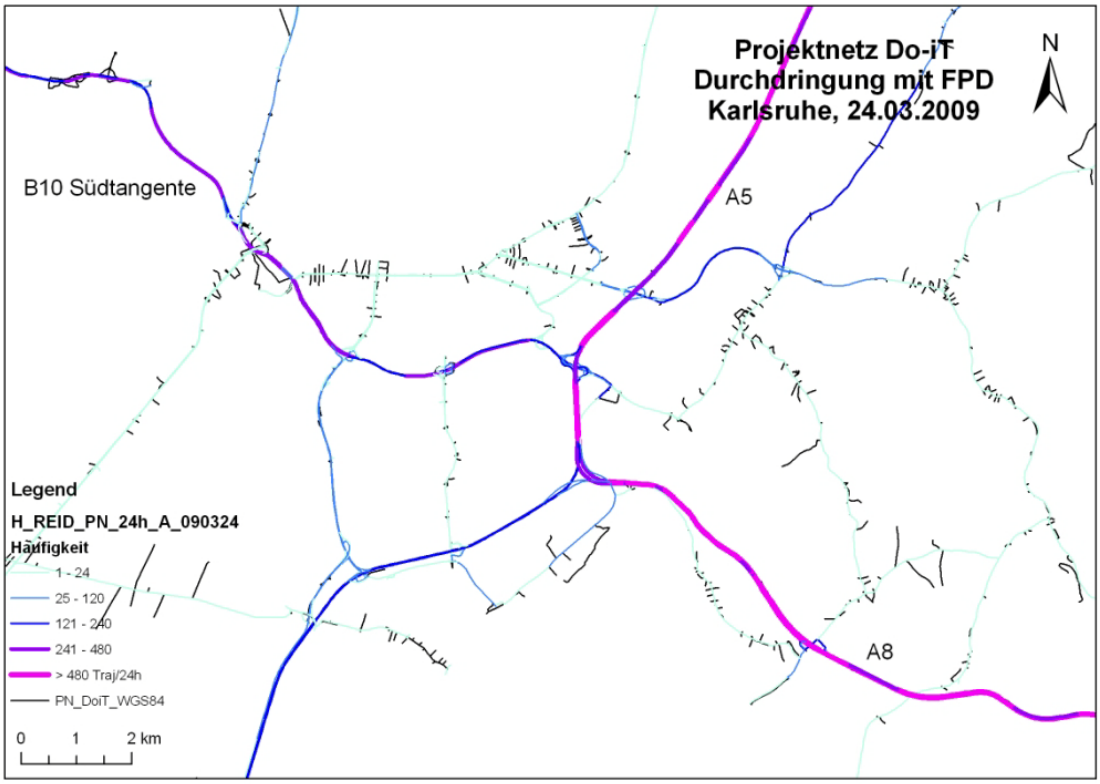


Abbildung 69: Häufigkeit auftretender RE im Raum Karlsruhe, Dienstag, 24.03.09 (A)

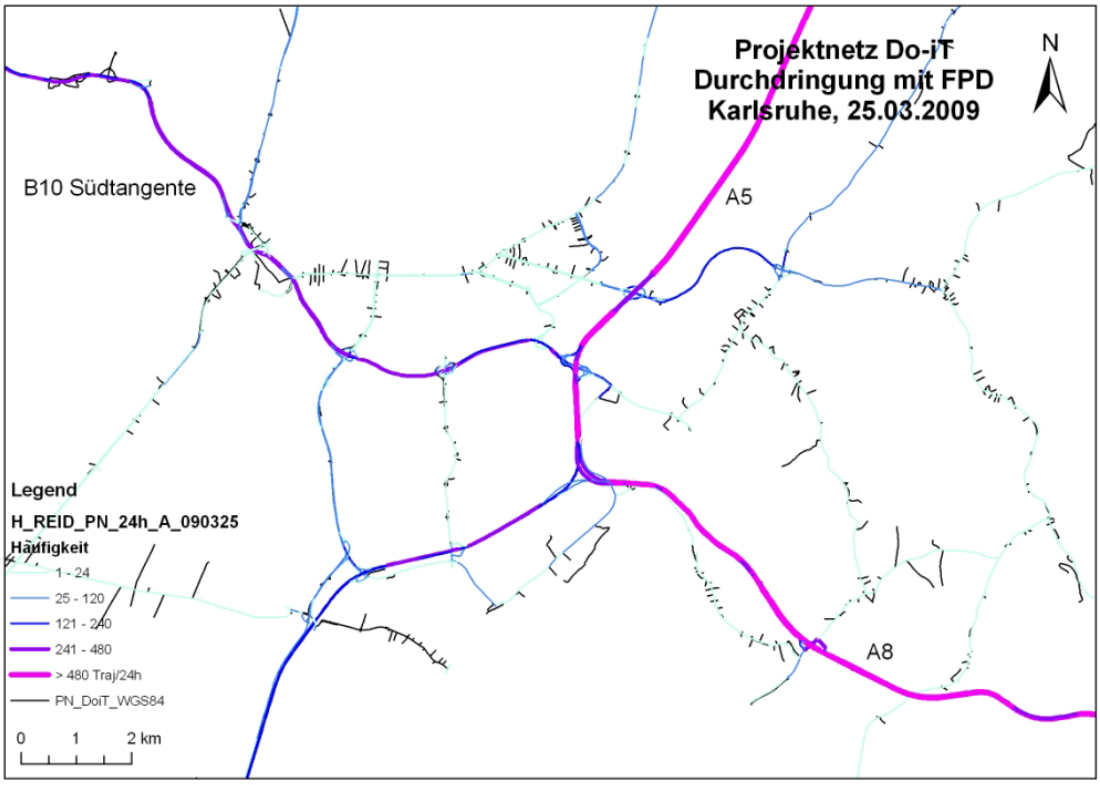


Abbildung 70: Häufigkeit auftretender RE im Raum Karlsruhe, Mittwoch, 25.03.09 (A)

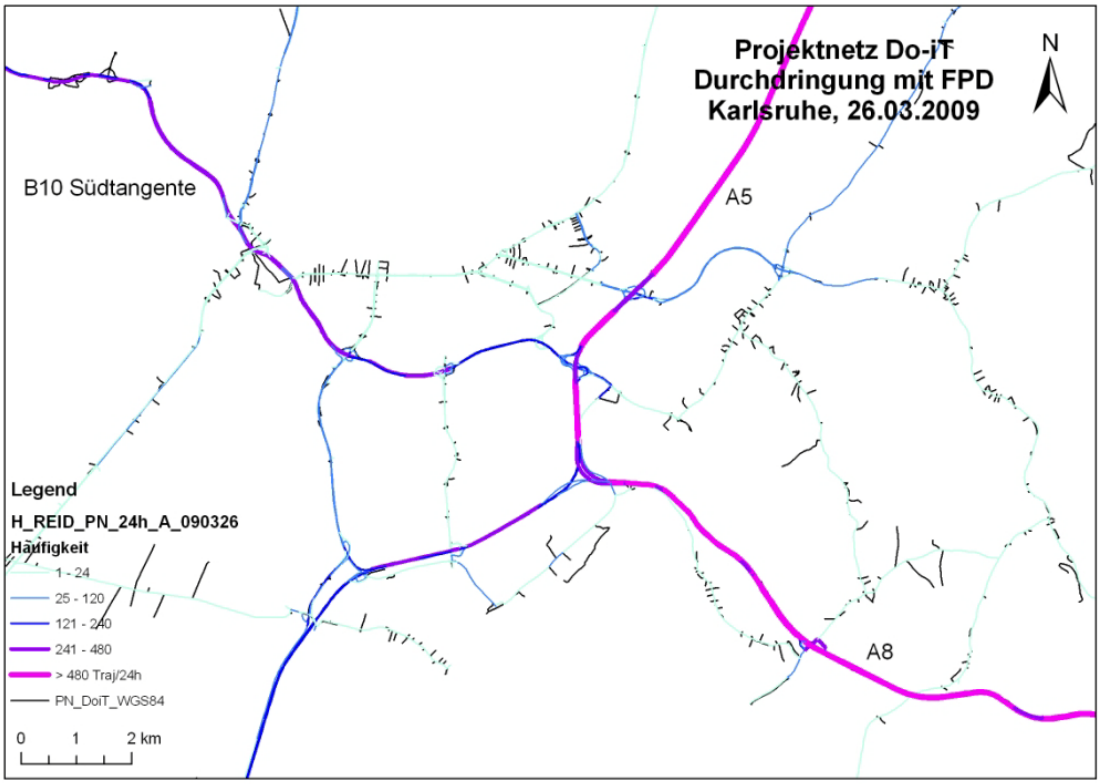


Abbildung 71: Häufigkeit auftretender RE im Raum Karlsruhe, Donnerstag, 26.03.09 (A)

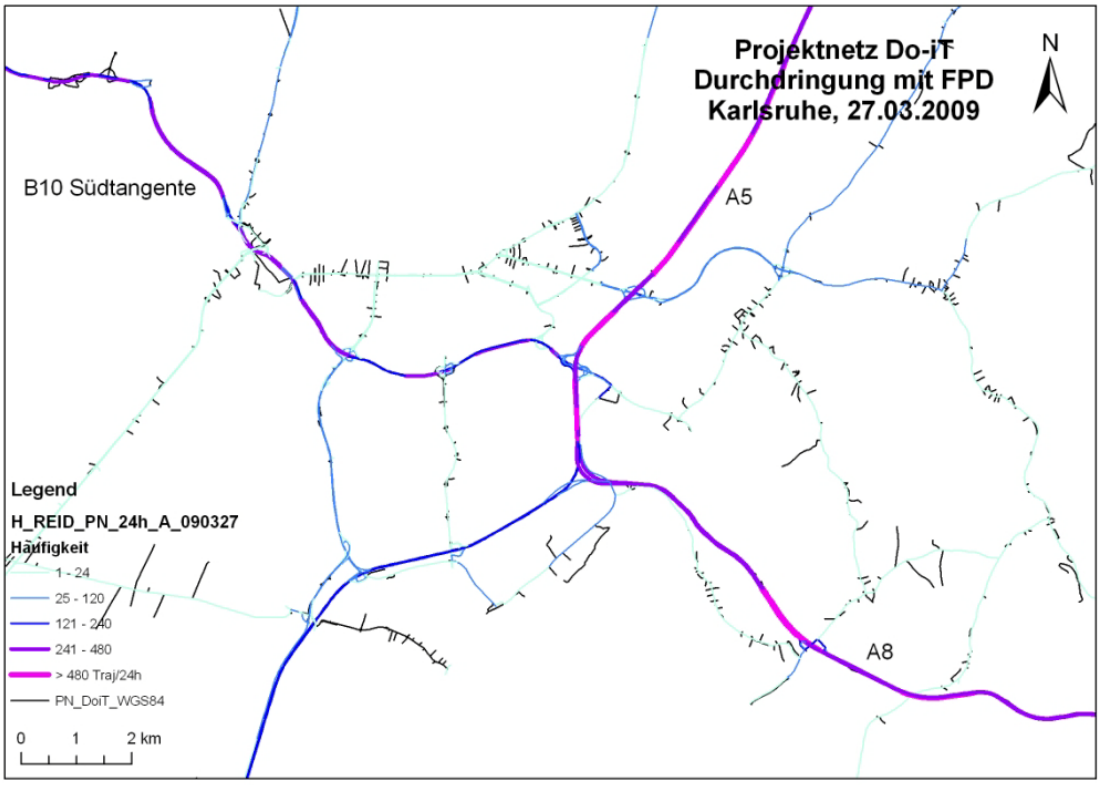


Abbildung 72: Häufigkeit auftretender RE im Raum Karlsruhe, Freitag, 27.03.09 (A)

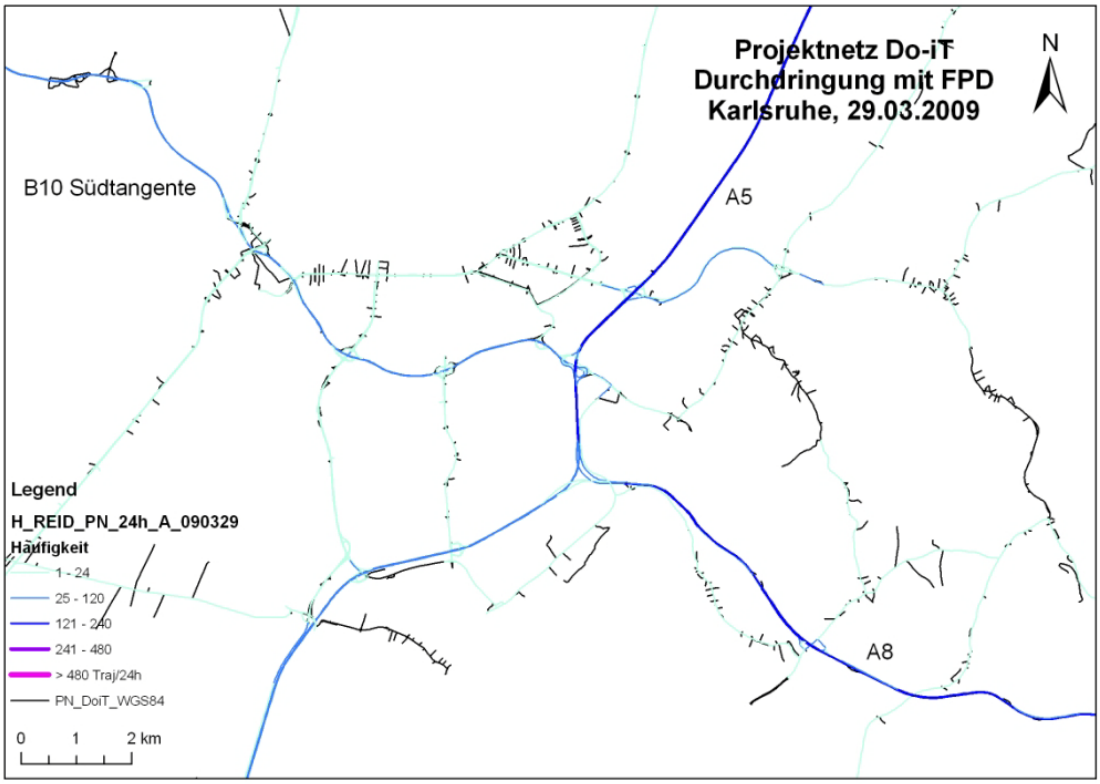


Abbildung 73: Häufigkeit auftretender RE im Raum Karlsruhe, Sonntag, 29.03.09 (A)

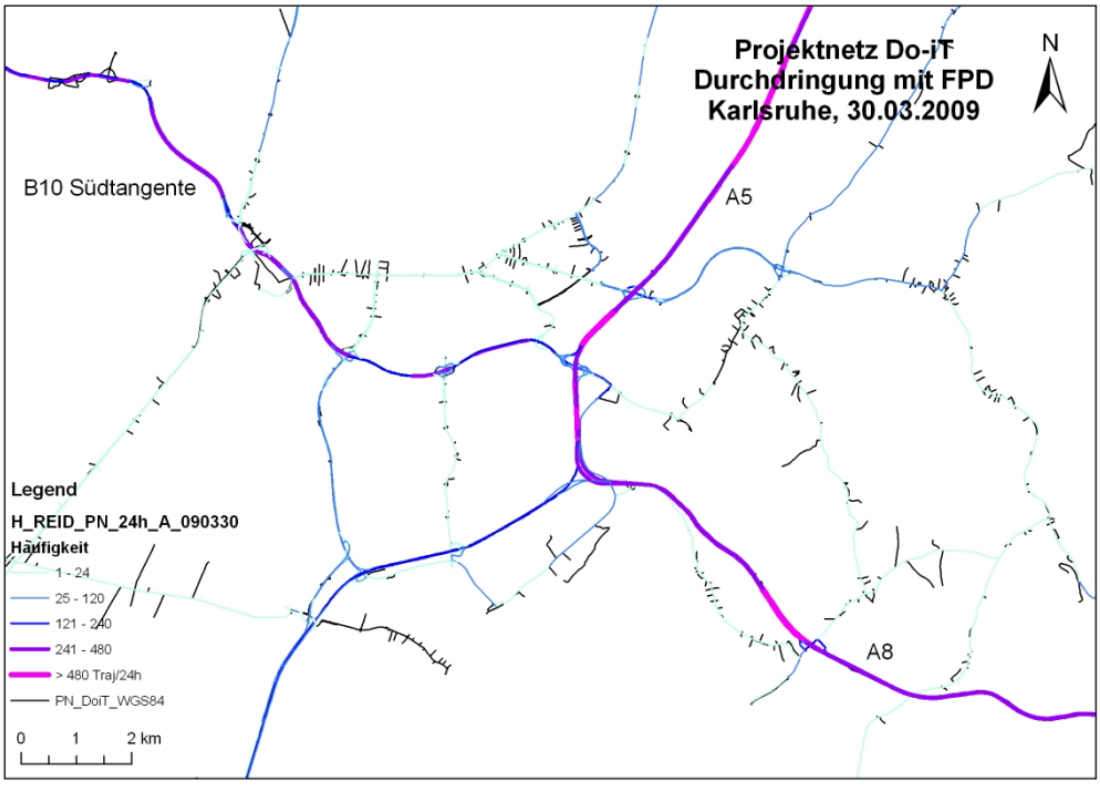


Abbildung 74: Häufigkeit auftretender RE im Raum Karlsruhe, Montag, 30.03.09 (A)

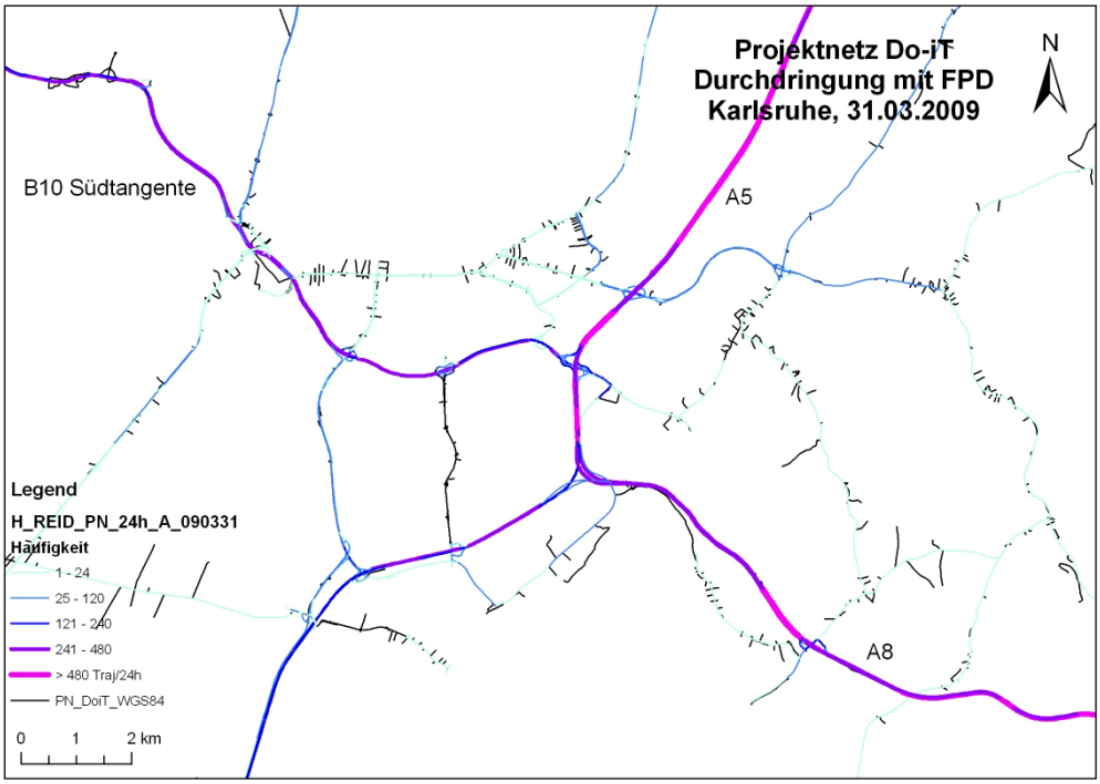


Abbildung 75: Häufigkeit auftretender RE im Raum Karlsruhe, Dienstag, 31.03.09 (A)