



Lehrstuhl für Verkehrstechnik
Technische Universität München
Univ.-Prof. Dr.-Ing. Fritz Busch

Schriftenreihe Heft 13

Friedrich Maier

Abschnittsweise Regressionsanalyse
zur Schätzung von Verkehrskenngrößen

Anwendung auf lokale Daten, Abschnittsdaten
und Informationen aus Positionsmeldungen

München 2010

Die Deutsche Bibliothek – CIP Einheitsaufnahme

Maier, Friedrich:

Abschnittsweise Regressionsanalyse zur Schätzung von
Verkehrskenngrößen - Anwendung auf lokale Daten,
Abschnittsdaten und Informationen aus
Positionsmeldungen

Hrsg.: Fritz Busch, München, 2010

Schriftenreihe des Lehrstuhls für Verkehrstechnik
der Technischen Universität München, Heft 13

Zugl.: München, Techn. Univ., Diss., 2010

ISBN 978-3-937631-13-4

Copyright ©

Lehrstuhl für Verkehrstechnik der Technischen Universität München 2010

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder
vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der
Übersetzung, vorbehalten.

Druck: TypeSet GmbH, Ismaning

ISBN 978-3-937631-13-4

ISSN 1612-9431

Lehrstuhl für Verkehrstechnik · Institut für Verkehrswesen

Technische Universität München · 80333 München

Telefon: 089 / 289 – 22438 · Telefax: 089 / 289 – 22333 · E-Mail: info@vt.bv.tum.de

www.vt.bv.tum.de

Bei der vorliegenden Veröffentlichung handelt es sich um die Dissertation von Herrn Dr.-Ing. Friedrich Maier.

Vorsitzender: Univ.-Prof. Dr.-Ing. Stephan Freudenstein, TU München

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. Fritz Busch, TU München
2. Univ.-Prof. Dr.-Ing. Markus Friedrich,
Universität Stuttgart

München, im Mai 2010

Danke

Diese Dissertation fertigte ich während meiner Zeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Verkehrstechnik, Technische Universität München, an. Bei denjenigen, die mir die Möglichkeit dazu gaben und die Rahmenbedingungen schufen, unter denen dies erfolgte, bedanke ich mich herzlich:

Bei meinem Doktorvater Herrn Professor Fritz Busch für meine Anstellung und die Möglichkeit zur Promotion, für die anregende Betreuung der Arbeit und für die Freiheiten bei ihrer Ausgestaltung.

Bei Herrn Professor Markus Friedrich für die Übernahme des Koreferats und für das Interesse an verkehrswissenschaftlichen Themen, das er mir im Rahmen der Entstehung meiner Diplomarbeit vermittelte.

Bei der Siemens AG für die Förderung des Forschungsvorhabens sowie bei der Landeshauptstadt München und bei der Taxizentrale IsarFunk für die Bereitstellung der Daten.

Bei allen Kolleginnen und Kollegen für den offenen und umfangreichen fachlichen Austausch nicht nur im Rahmen dieser Arbeit und für die wunderbare, freundschaftliche Atmosphäre am Lehrstuhl.

Bei Mircea Palici für die Zuverlässigkeit, mit der er als studentische Hilfskraft über zwei Jahre hinweg am Vorhaben mitwirkte.

Bei meinen Eltern für die jahrzehntelange Förderung und Unterstützung.

Friedrich Maier

Als Kurzbezeichnung für das Verfahren zur Schätzung von Verkehrskenngrößen gemäß der vorliegenden Arbeit wird **SERAFIM** vorgeschlagen.

SERAFIM is proposed as short term for the traffic estimation method described in this work.

SERAFIM – **SE**gmented **R**egression **A**pproach for traf**F**ic esti**M**ation

Kurzfassung

Die zunehmende Vielfalt von Datenquellen und -arten zur Beschreibung der Verkehrssituation erschwert zusehends die Integration sämtlicher verfügbarer Informationen in bestehende Verfahren zur Verkehrslageschätzung. Dadurch können diese Verfahren entscheidende Nachteile hinsichtlich der Qualität ihrer Schätzergebnisse erleiden. Die vorliegende Arbeit beschreibt eine neue Methode, die die Berücksichtigung aller Datenarten erlaubt und für alle eingehenden, auch gering verfügbaren Kenngrößen kontinuierlich Schätzwerte liefern kann. Sie verfolgt einen statistischen Ansatz und setzt folglich kein Verkehrsflussmodell ein.

Durch abschnittsweise lineare Regressionsanalyse werden aus historischen Daten Näherungsfunktionen zur Schätzung von Verkehrskenngrößen aus anderen Kenngrößen ermittelt, sofern die Erhebungsorte von Ein- und Ausgangskenngröße ein Kriterium der räumlichen Nähe erfüllen. Dieser Ansatz erlaubt die Beschreibung auch von Zusammenhängen, die nicht über den gesamten Wertebereich linear sind, sowie die nach Wertebereichen differenzierte qualitative Beurteilung der Zusammenhänge. Bei der Zusammenführung zahlreicher Schätzwerte für eine Kenngröße werden die Einzelschätzungen aus den stärksten Zusammenhängen verwendet. Die zusammengeführten Schätzwerte vervollständigen den aktuellen Datenbestand und dienen der Plausibilisierung von Messwerten. Gehen fahrzeuggenerierte Daten in das Verfahren ein, kann die Verkehrslage netzweit und linkfein geschätzt werden.

Das neue Verfahren wird erprobt mit lokalen Daten, mit Abschnittsdaten aus infrastrukturseitiger Fahrzeugwiedererkennung sowie mit Linkgeschwindigkeiten aus Positionsmeldungen einer Taxiflotte. Als Historie werden Daten verwendet, die im Laufe von drei Monaten in einem Netzausschnitt im Münchner Norden gesammelt wurden. Die Ergebnisse der Regressionsanalyse werden differenziert nach Datenart von Ein- und Ausgangsgröße, nach Verkehrszustand sowie nach zeitlichem Versatz zwischen Ein- und Ausgangsgröße für verschiedene Varianten der Regressionsanalyse untersucht. Diese Varianten beleuchten neben der einfachen die gewichtete lineare Regression sowie die nach Tagestyp und -zeit differenzierte Auswertung. Im Allgemeinen bestätigen sich Erwartungen, die auf der Kenntnis des Fundamentaldiagramms und auf statistischen Vorüberlegungen basieren. Lokale Belegungsgrade und Routenreisezeiten erweisen sich als besonders aussagekräftige Eingangsgrößen.

Zur Bewertung zusammengeführter Schätzergebnisse werden ein geeignetes Fehlermaß und die Detektionsrate von Perzentilereignissen verwendet, angewandt auf einen Zeitraum von einer Woche. Varianten des entwickelten Ansatzes werden mit einem Ganglinienverfahren, mit der Trivialprognose und mit einer pragmatischen Kombination aus beiden verglichen. Für das neue Verfahren werden bei der Schätzung infrastrukturseitig erhobener Daten fast durchgängig die besten Kennwerte ermittelt. Messwerte dienen hierbei als Referenz. Für geschätzte Linkreisezeiten fehlt eine solche Referenz. Verschiedene Untersuchungsansätze zeigen aber, dass qualitativ und quantitativ plausible Schätzergebnisse erzielt werden.

Segmented Regression Analysis for Estimation of Traffic Characteristics – Abstract

The growing variety of data sources and data types to describe the traffic situation is complicating the integration of all available information in existing methods for traffic state estimation. Hence these methods can suffer significant disadvantages in the quality of their estimates. This dissertation presents a new method allowing the consideration of all relevant data and the continuous estimation of all input characteristics independent of their temporal availability. It applies a regression approach and therefore does not use a traffic flow model.

Approximation functions are calculated by a univariate segmented regression approach using historic data, in case the distance between the detection locations of the correlated characteristics is not too far. These functions are used together with currently available data to estimate current traffic characteristics. The approach allows the description of interdependencies which are not completely linear and the qualitative evaluation of interdependencies differentiated by segments. Bringing together numerous estimates for a single characteristic only the results of the strongest interdependencies are used. The consolidated estimates complete the current data set and can be used to check the measurements' plausibility. In the case of integrating historic vehicle probe data, the current network-wide traffic state can be estimated.

The new method is tested with local data, section data and link-related speeds derived from position reports of a taxi fleet. For regression analysis data are used which were collected during a period of three months in a road network section in Munich. Regression results are investigated for multiple variations of the approach differentiated by data types of input and output characteristics, by traffic states and by temporal offsets between input and output data. The variations include the simple and the weighted linear regression as well as the differentiation by day types and time of day. In general, expectations based on the knowledge of the fundamental diagram and on statistical pre-considerations are confirmed. Local occupancy rates and section data prove to be especially meaningful input data.

To evaluate the consolidated estimates, a suitable error index and detection rates of percentile events are used, applied to the period of one week. Variations of the new method are compared to a method using time variation curves, trivial predictions and a pragmatic combination of both. The new method nearly always shows better results for the estimation of local and section data. Measurements are used as reference. Due to a lack of comparable references for the estimates of link-related speeds only the general plausibility of these results can be checked. Different investigations show that the new method also generates qualitatively and quantitatively reasonable results for these characteristics.

Inhaltsverzeichnis

1.	Einleitung.....	1
2.	Verkehrsdatenerfassung und -aufbereitung.....	5
2.1	Lokale Verkehrsdaten	6
2.2	Abschnittsdaten – Reisezeitmessung durch Fahrzeugwiedererkennung.....	8
2.2.1	Videodetektion: Kennzeichenerfassung	9
2.2.2	Induktivschleifen: Signaturanalyse	9
2.3	Linkbezogene Verkehrskenngrößen aus Positionsdaten.....	13
2.3.1	Map Matching mit räumlicher Indizierung	15
2.3.2	Routensuche.....	21
2.3.3	Ableitung und Aggregation von Linkreisezeiten	25
3.	Verkehrslageschätzung – Stand der Technik.....	27
3.1	Ansätze mit Umlegungs- oder Verkehrsflussmodell.....	28
3.1.1	Makroskopische Ansätze.....	28
3.1.2	Mikroskopische Ansätze.....	30
3.2	Ansätze ohne Umlegungs- oder Verkehrsflussmodell.....	32
3.2.1	Ganglinienverfahren und Zeitreihenanalysen.....	32
3.2.2	Regressions- und musterbasierte Ansätze.....	32
3.2.3	Fuzzy-Logik und Neuronale Netze.....	37
3.3	Anforderungen an die zu entwickelnde Methode	39
4.	Neue Methode zur Schätzung von Verkehrskenngrößen.....	43
4.1	Konzeptioneller Ansatz	43
4.2	Angepasster Regressionsansatz	46
4.2.1	Einfache lineare Regression	46
4.2.2	Gewichtete lineare Regression	48
4.2.3	Orthogonale lineare Regression	49
4.2.4	Abschnittsweise lineare Regression	50
4.3	Zusammenführung mehrerer Schätzungen	53
4.4	Kalibrierung zusammengeführter Schätzungen.....	54
5.	Datenqualität und Suchraumbeschränkung.....	57
5.1	Verfügbarkeit von Positionsdaten und Linkreisezeiten.....	57
5.2	Verfügbarkeit infrastrukturseitig erhobener Daten.....	60
5.3	Datengüte	61
5.4	Beschränkung des Raums zur Suche nach Zusammenhängen.....	63

6.	Bewertung der ermittelten Näherungsfunktionen.....	65
6.1	Konzept zur Bewertung der Näherungsfunktionen.....	65
6.2	Näherungsfunktionen aus einfacher linearer Regression	74
6.2.1	Lokale Daten als Eingangsgröße	74
6.2.2	Abschnittsdaten als Eingangsgröße	83
6.2.3	Aufbereitete Positionsdaten als Eingangsgröße	85
6.2.4	Näherungsfunktionen mit zeitversetzten Ein- und Ausgangsdaten	88
6.3	Näherungsfunktionen aus gewichteter linearer Regression	93
6.4	Zeitliche Differenzierung bei der linearen Einfachregression	94
6.5	Rückschlüsse für die Zusammenführung von Einzelschätzwerten.....	98
7.	Zusammenführung der Einzelschätzungen und Bewertung der Schätzergebnisse.....	101
7.1	Konzept zur Bewertung der Schätzergebnisse	101
7.2	Zusammenführung der Einzelschätzungen und Kalibrierung	104
7.3	Bewertung geschätzter lokaler Kenngrößen.....	108
7.4	Bewertung geschätzter Abschnittsdaten.....	117
7.5	Bewertung geschätzter Linkgeschwindigkeiten aus Positionsmeldungen.....	118
7.6	Weitergehende Sensitivitätsanalysen	124
7.6.1	Schätzung lokaler Kenngrößen ohne Eingangsgrößen vom selben Link...	127
7.6.2	Schätzung von Linkgeschwindigkeiten ausschließlich aus lokalen Daten oder ausschließlich aus Routenreisezeiten	128
8.	Resümee.....	131
	Literaturverzeichnis.....	139
	Abkürzungsverzeichnis	145
	Abbildungsverzeichnis	147
	Tabellenverzeichnis	151
	Anhang 1: Vorabveröffentlichungen	153
	Anhang 2: Überblick über Verfahren zur Verkehrslageschätzung	155

1. Einleitung

Umstände wie die steigende Verkehrsnachfrage – sowohl in der Gesamtbetrachtung als auch insbesondere lokal –, begrenzte Verfügbarkeit von Verkehrsflächen, sich entwickelnde Ansprüche an das Verkehrsangebot und nicht zuletzt die Belange der Umwelt führen zu einer zunehmenden Bedeutung von Einrichtungen und Maßnahmen des Verkehrsmanagements. Diese stellen ein umfangreiches Instrumentarium zur Verfügung, um damit die Verkehrsnachfrage sicherer, umweltverträglicher und wirtschaftlicher abzuwickeln und dabei die Leistungsfähigkeit der vorhandenen Infrastruktur bestmöglich auszunutzen. Dieses Instrumentarium umfasst statische und dynamische Elemente. Es kann zur kollektiven Steuerung und Regelung des Verkehrsablaufs sowie zur kollektiven oder individuellen Verbreitung von Empfehlungen und Informationen genutzt werden. Diese Empfehlungen oder Informationen können dem Verkehrsteilnehmer vor Fahrtbeginn oder während der Fahrt mit geeigneten Medien zugespielt werden.

Grundlage für Entscheidungen über die Nutzung der dynamischen Werkzeuge des Verkehrsmanagements auf taktischer und operativer Ebene ist die Kenntnis der aktuellen Verkehrsverhältnisse im Netz. Dabei eröffnet ein höherwertiges, räumlich-zeitlich feiner aufgelöstes Verkehrslagebild verbesserte Möglichkeiten für gezielte und damit erfolgreiche Eingriffe des Verkehrsmanagements ins Verkehrsgeschehen. Dem Erreichen heutiger und zukünftiger verkehrlicher und damit verbundener weiterer Ziele wird also durch die verbesserte Kenntnis der aktuellen Verkehrssituation Vorschub geleistet.

Die Ermittlung der aktuellen Verkehrslage in einem Straßennetz stellt nach wie vor eine verkehrstechnische Herausforderung dar. Dabei steht dem räumlich-zeitlich fein aufgelösten Verkehrslagebild die begrenzte Verfügbarkeit aktueller Verkehrsdaten im Wege. Verfahren zur Verkehrslageschätzung versuchen daher, aus den verfügbaren Messwerten möglichst viele Informationen zu gewinnen. So verbessern sich die Möglichkeiten zur Verkehrslageschätzung mit dem Umfang aktueller Messwerte und mit der Qualität des eingesetzten Schätzverfahrens.

Durch ständige technische Innovationen hinsichtlich der Verkehrsdetektion vergrößert sich der Umfang aktueller verkehrlich relevanter Informationen. Diese Entwicklung bietet bei der Verkehrslageschätzung einerseits zusätzliche Chancen, die Güte und die räumlich-zeitliche Auflösung der Ergebnisse verwendeter Verfahren zu verbessern. Andererseits führt sie zu Herausforderungen bei der Integration der neu erhobenen Datenarten mit ihren ebenso neuen Eigenschaften und Besonderheiten in bestehende Algorithmen und Abläufe.

Traditionelle aktuell erhobene Eingangsgrößen in Verfahren zur Verkehrslageschätzung stammen aus lokalen Detektoren, deren Messwerte nicht unmittelbar in die Verkehrssteuerung einfließen. Sie können auf einem oder benachbarten Fahrstreifen an einem Querschnitt

Fahrzeuge zählen, deren Geschwindigkeit messen und die anteilige Belegungszeit des Detektors ermitteln. Weitere lokale Detektoren finden sich im Zu- und Abflussbereich von Lichtsignalanlagen. Die hier oft ausschließlich durchgeführte Fahrzeugzählung kann direkt in die Anpassung der Lichtsignalisierung einfließen.

In den vergangenen Jahren wurden Sensoren entwickelt, die die zuverlässige Wiedererkennung von Fahrzeugen an unterschiedlichen Querschnitten und somit die infrastrukturseitige Messung routenbezogener Reisezeiten ermöglichen. Diese hochwertigen Verkehrsdaten wurden bisher als alleinstehender Indikator für den Verkehrszustand auf den untersuchten Routen verwendet. Sie wurden noch nicht in netzweite Verfahren eingebracht, die verschiedenartige Verkehrsdaten zusammenführen und somit die Kenntnisse zur aktuellen Verkehrssituation aus heterogenen Datenquellen kombinieren.

Auch die Integration von Informationen aus Positionsdaten in Verfahren zur netzweiten Schätzung von Verkehrskenngrößen ist nicht endgültig gelöst. Ihr herausragendes Merkmal gegenüber infrastrukturseitig erhobenen Daten ist die mögliche komplette Abdeckung eines untersuchten Straßennetzes. Je nach Größe, Fahrleistung und üblichen Routen der eingesetzten Flotte sowie Größe des Netzes erfolgt diese Abdeckung jedoch besser oder schlechter, in kürzeren oder längeren zeitlichen Intervallen. So können die Informationen aus aktuellen Positionsdaten unterschiedlich große Beiträge zur Verkehrslageschätzung leisten.

In Kapitel 2, das sich mit der Erfassung und Aufbereitung von Verkehrsdaten beschäftigt, wird schwerpunktmäßig ein Verfahren zur Gewinnung verkehrlich relevanter Informationen aus Positionsdaten einer Fahrzeugflotte entwickelt und beschrieben. Es zeigt die Herausforderungen beim Gewinnen von linkbezogenen Geschwindigkeiten aus Positionsdaten einer Taxiflotte auf und enthält diesbezügliche Lösungswege. Neben einem Verfahren zum Abgleich der Positionsdaten mit einer digitalen Straßenkarte wird ein Verfahren zur Routensuche vorgestellt, das bei Übertragungsintervallen von mehr als einer halben bis hin zu mehreren Minuten, bei üblichen Abschlägen hinsichtlich der Genauigkeit von übertragenen Positionen und bei einer hohen Netzdichte Fahrtverläufe zuverlässig und korrekt im Netz rekonstruiert.

Die Zusammenstellung gesichteter Verfahren zur Verkehrslageschätzung ist Gegenstand von Kapitel 3. Die grundlegenden Ideen dieser Verfahren werden erläutert und zu einer Kategorisierung der verschiedenen Ansätze genutzt.

Nach den Innovationen der vergangenen Jahre hinsichtlich der Verkehrsdatenerhebung kann auch für die Zukunft davon ausgegangen werden, dass neue Erhebungstechnologien zusätzliche Daten mit bisher unbekanntem Eigenschaften liefern, die für die Schätzung der Verkehrslage im Netz hilfreiche Informationen bereitstellen. So ist ein wesentliches Merkmal gegenwärtiger und zukünftiger Verfahren zur Schätzung von Verkehrskenngrößen in Netzen, ob und mit welchem Aufwand sie die zusätzliche Verarbeitung neuartiger verkehrlich rele-

vanter Daten erlauben. Ziel dieser Arbeit ist daher unter Berücksichtigung des Stands der Technik gemäß Kapitel 3 und daraus abgeleiteter Anforderungen die Entwicklung einer neuen Methode, die einen offenen Ansatz verfolgt und so die Schwelle für die Integration weiterer Datenquellen niedrig hält. Sie soll Daten unterschiedlicher Qualität, also variierender Verfügbarkeit und Güte, problemlos verarbeiten können. Dabei sollen die spezifischen Vorteile unterschiedlicher Datenarten wie beispielsweise die hohe zeitliche Verfügbarkeit lokaler Daten oder die gute Netzabdeckung von Informationen aus Positionsmeldungen genutzt werden. Es soll keine Rolle spielen, ob Kennwerte in städtischem Umfeld oder außerorts geschätzt werden.

Die neue Methode zur Schätzung verkehrlicher Kenngrößen, die in Kapitel 4 beschrieben wird, verfolgt einen statistischen Ansatz (grundsätzlicher Ablauf in Bild 1.1) und verzichtet damit auf den Einsatz eines Verkehrsflussmodells. Die eingesetzte lineare Regressionsanalyse wird auf die Eigenschaften von Zusammenhängen zwischen verkehrlichen Kenngrößen adaptiert. Zahlreiche Formen dieser Adaption werden diskutiert, einige werden zur weiteren Untersuchung ausgewählt. Näherungsfunktionen, die aus der Regressionsanalyse historischer Daten resultieren, sollen zusammen mit aktuell verfügbaren Daten zur Schätzung aktueller Kennwerte verwendet werden. Hauptziele sind die Vervollständigung und die Plausibilisierung aktueller Daten, die Prognose spielt eine untergeordnete Rolle.

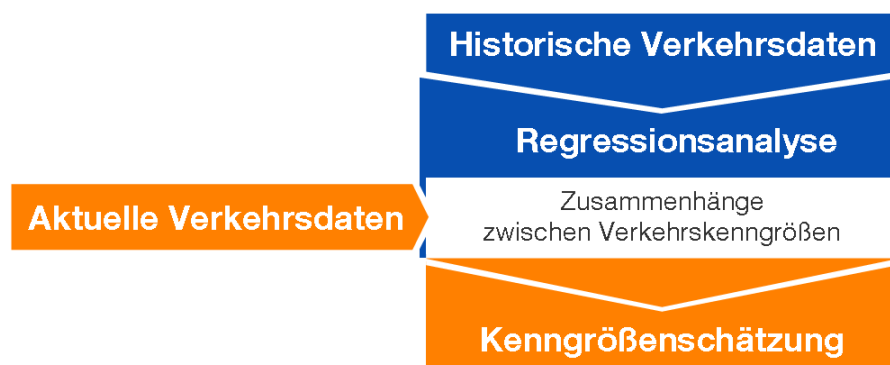


Bild 1.1 Grundsätzlicher Ablauf der Schätzung von Verkehrskenngrößen bei Verwendung eines regressionsbasierten Ansatzes

Die Erprobung der neuen Methode erfolgt unter Verwendung von lokalen und abschnittsbezogenen Daten sowie von Informationen aus Positionsmeldungen. Dabei kommen Daten aus einem Netzausschnitt im Norden Münchens zum Einsatz. Gegenstand von Kapitel 5 ist die Beschreibung des Netzausschnitts, der darin eingesetzten Verkehrssensorik und der Qualität der damit erhobenen Verkehrsdaten. Das Kapitel beschreibt zudem einen pragmatischen Ansatz zur Beschränkung des Suchraums nach Zusammenhängen, ohne dabei aussagekräftige Zusammenhänge auszuschließen.

Die entwickelte Methode sieht die Regressionsanalyse für eine Anzahl von Zusammenhängen vor, die die Nutzung sämtlicher daraus hervorgehender Ergebnisse für die Schätzung

überflüssig macht, im Gegenzug gar eine gezielte Auswahl nahe legt. Auch die Erwartung, dass mitunter wenig hilfreiche Zusammenhänge zwischen verkehrlichen Kenngrößen ermittelt werden, spricht für das Verwerfen der zugehörigen Näherungsfunktionen vor der Zusammenführung der Schätzwerte. Hierfür ist eine geeignete Entscheidungsgrundlage auszumachen. Kapitel 6 geht dieser Fragestellung nach. Es erlaubt für ausgewählte Varianten der Regressionsanalyse die Entwicklung eines Verständnisses für Umstände und Abläufe innerhalb des Verfahrens. Dieses Verständnis ist vor dem Hintergrund zahlreicher Ein- und Ausgangsgrößen und Teilberechnungen, die wegen ihrer Vielzahl nicht umfassend eingesehen werden können, bei der Ergebnisinterpretation erforderlich. Gezielt aggregierte statistische Auswertungen werden hierbei ebenso eingesetzt wie zahlreiche Beispiele für Zusammenhänge zwischen verkehrlichen Kenngrößen samt den zugehörigen statistischen Berechnungen, die die Wirkungen der Varianten der Regressionsanalyse veranschaulichen. Die Ergebnisse für zeitlich versetzte Ein- und Ausgangsgrößen ermöglichen hierbei, die zeitliche Gültigkeit erhobener Verkehrsdaten statistisch zu untersuchen.

Auf Grundlage der Ergebnisse aus Kapitel 6 wird in Kapitel 7 das Vorgehen beschrieben, mit dem für jede Kenngröße im Netz verschiedene Einzelschätzungen zu einem Schätzwert zusammengeführt werden. Ein geeignetes Konzept zur Bewertung von geschätzten Verkehrskenngrößen wird vorgestellt und anschließend auf die zusammengeführten Schätzergebnisse verschiedener Varianten der neuen Methode angewendet. Zum Vergleich werden Schätzungen eines Trivial- und eines Ganglinienverfahrens herangezogen. Abschließende Sensitivitätsuntersuchungen zielen auf Aussagen zur Bedeutung verschiedener Eingangsdatenarten für die Schätzergebnisse und analysieren exemplarisch die Konsequenzen für den Fall, dass bestimmte Eingangsdatenarten nicht zur Verfügung stehen.

2. Verkehrsdatenerfassung und -aufbereitung

Grundlage für jede Abbildung der gegenwärtigen Verkehrslage sind aktuelle verkehrlich relevante Daten. Neben Verkehrsdaten sind dies im Wesentlichen Umfelddaten, die im Rahmen dieser Arbeit nicht weiter berücksichtigt werden sollen.

Mikroskopische Verkehrsdaten beschreiben die verkehrlichen Merkmale von Einzelfahrzeugen wie beispielsweise Geschwindigkeiten und räumliche oder zeitliche Fahrzeugabstände. Makroskopische Verkehrsdaten hingegen werden durch die Betrachtung einer Menge von Fahrzeugen, also eines Verkehrsstroms, ermittelt und umfassen somit die Kenngrößen Verkehrsstärke, Verkehrsdichte und mittlere Geschwindigkeit.

Ein Überblick über Datenquellen und -arten zur Beschreibung des motorisierten Individualverkehrs:

- Infrastrukturseitig erhobene Daten
 - Lokale Verkehrsdaten – automatische Erhebung z.B. mit Induktivschleifen, Infrarot-, Radar-, Ultraschall-, Magnetfeldsensoren, Kombinations-, Laserdetektoren oder Videokameras
 - Abschnittsdaten – automatische Reisezeitermittlung von einem Start- zu einem Zielquerschnitt durch Fahrzeugwiedererkennung basierend auf Kennzeichenerfassung oder auf dem Abgleich von Fahrzeugmerkmalen, die an lokalen Detektoren erhoben werden (z.B. an Induktivschleifen oder an Laserdetektoren)
- Fahrzeuggenerierte Daten
 - Positionsdaten – während einer Fahrt mit einem globalen Ortungssystem erhobene Koordinaten, häufig bezeichnet als Floating Car Data (FCD)
 - Daten aus fahrzeugseitiger Sensorik wie z.B. Rad-, Regen- oder Radarsensoren für Abstandsregeltempomaten, gekoppelt mit Positionsdaten – häufig bezeichnet als Extended FCD (XFCD)
- Andere mobil erhobene Daten
 - Luftbildaufnahmen
 - Geräteortungsdaten von Mobiltelefonen (vgl. [FRIEDRICH ET AL., 2008]), häufig bezeichnet als Mobile Phone Data oder Floating Phone Data (FPD)
- Modellerte Daten – Daten aus Verkehrsmodellen jeglicher Art
- Verkehrsmeldungen – z.B. Baustellen, Störungen, Sperrungen oder Veranstaltungen betreffend

Zur Bearbeitung der vorliegenden Fragestellung lagen vor:

- Lokale Verkehrsdaten
- Abschnittsdaten
- Positionsdaten einer Taxiflotte

Daher wird im weiteren Verlauf des Kapitels auf die Erfassung und Aufbereitung ausschließlich dieser Daten eingegangen.

2.1 Lokale Verkehrsdaten

Lokale Detektoren messen den Verkehr meist auf einem, mitunter auch auf mehreren Fahrstreifen. Um Querschnittsdaten zu ermitteln, müssen im Fall mehrstreifiger Querschnitte meist also mehrere Detektoren eingesetzt werden.

Detektoren erfassen Fahrzeugmerkmale, die zur Fahrzeugerkennung und zur Bestimmung der mikroskopischen Verkehrskenngrößen Fahrzeuggeschwindigkeit und Fahrzeugtyp eingesetzt werden (Klassifizierung gemäß den Technischen Lieferbedingungen für Streckenstationen (TLS) [BMVBW, 2002]). Verwendete Merkmale sind je nach Erfassungstechnologie beispielsweise Fahrzeuglänge, -breite, -höhe, Längs-, Querprofil, elektromagnetische Signatur, Achsenanzahl, Masse, Reifendruck, Achslast, akustische oder optische Merkmale. Detektoren erfassen zu einem Fahrzeug weitere mikroskopische Verkehrskenngrößen wie Geschwindigkeit, Detektorbelegungszeit, Zeitlücke zum Vorderfahrzeug o.ä. Die mikroskopischen Verkehrskenngrößen werden in der Regel in Streckenstationen zeitlich zu den makroskopischen Verkehrskenngrößen Verkehrsstärke, lokale mittlere Geschwindigkeit und Belegungsgrad über Zeiträume meist zwischen einer und mehreren Minuten aggregiert, bevor sie beispielsweise zu Verkehrsmanagementzentralen übertragen werden.

Detektionstechnologien unterscheiden sich hinsichtlich ihrer Fehleranfälligkeit, ihrer Robustheit gegen Umwelteinflüsse, ihrer Kosten bei Anschaffung, Montage und Unterhaltung, ihrer mobilen Einsetzbarkeit und hinsichtlich der Anzahl der erfassbaren Fahrstreifen.

Üblicherweise werden zur lokalen Verkehrsdetektion Induktivschleifen, Infrarot-(IR-), Ultraschall-, Radar- und Videosensoren eingesetzt. Druckschläuche, Piezo-Koaxialkabel und faseroptische Detektoren können als historisch eingestuft werden [MESCHENDÖRFER, 2004]. Die Verbreitung von Magnetfelddetektoren nimmt in der jüngeren Vergangenheit zu. Akustische Sensoren werden alleinstehend kaum eingesetzt.

Induktivschleifen sind Spulen im Fahrbahnbelag, die mit einem Kondensator zu einem Schwingkreis geschaltet werden. Ein Fahrzeug über der Schleife verstimmt mit seiner Metallmasse den Schwingkreis. Diese Verstimmung ist messbar und kann zur Fahrzeugerkennung und -klassifizierung eingesetzt werden. Die Messung der Geschwindigkeit ist insbesondere bei einer Ausführung als Doppelschleife möglich. Positiven Eigenschaften wie den

niedrigen Kosten für den Detektor, der hohen Qualität bei Fahrzeugzählung und -klassifizierung und der Robustheit gegen Umwelteinflüsse stehen die aufwendige Installation und Abstriche hinsichtlich der Haltbarkeit der Schleife gegenüber.

Ultraschallsensoren erfassen per Laufzeitmessung von Schallwellen exakte Höhenprofile und Fahrzeuglängen zur Fahrzeugerkennung und -klassifizierung sowie zur Geschwindigkeitsmessung. Sie sind preiswert und verschleißfrei, werden durch Umweltbedingungen jedoch stark beeinflusst. Echolotgeräte in der Nautik basieren ebenfalls auf der Ultraschalltechnologie.

Passive IR-Detektoren werden über der Fahrbahn montiert und registrieren den Temperaturunterschied zwischen der Fahrbahn und einem Fahrzeug, das den Sensor passiert. Aktive IR- oder Laserdetektoren erfassen mit der Laufzeitmessung eines Lichtstrahls Fahrzeughöhen-, je nach Ausführung auch -querprofile. Ihre Abtastfrequenz liegt weit höher als die von Ultraschallsensoren. Dies führt zu detaillierter dargestellten Profilen. Dabei sind Laserdetektoren jedoch deutlich teurer.

IR-Detektoren sind verschleißfrei und lichtunempfindlich. Ihre Qualität leidet jedoch bei Regen, Schnee oder Nebel. Zudem erfolgt die Klassifizierung bei passiven Sensoren recht ungenau über die Schätzung der Fahrzeuglänge.

Radar (Radio Detection And Ranging)-Detektoren nutzen den Dopplereffekt, der die Frequenzänderung einer Welle bei der Reflexion an einem bewegten Objekt beschreibt. Radargeräte erlauben eine sehr genaue Geschwindigkeitsmessung und die Überwachung mehrerer Fahrstreifen, erfordern jedoch einen exakten Aufbau und erfassen stehende Fahrzeuge nicht.

Bei der Verkehrsdatenerhebung durch Videodetektion kommen Verfahren wie das Differenzbildverfahren, die Hintergrundsubtraktion, das Verfahren des optischen Flusses, die Stereoskopie oder die Grauwertanalyse zum Einsatz. Dem hohen Potential steht aufwendige Algorithmen gegenüber.

Kombinationsdetektoren verwenden mehrere Sensoren unterschiedlicher Technologien und nutzen jeweils deren Vorteile. So werden beispielsweise ein Passiv-IR-Sensor zur Fahrzeugerkennung, ein Ultraschallsensor zur Klassifizierung und ein Radarsensor zur Geschwindigkeitsmessung kombiniert. Die Kombination von Akustik- und Radarsensoren führen gekoppelt zur sicheren Erkennung von Fahrzeugen mit Heckmotor und damit zur verbesserten Klassifizierung von Bussen.

2.2 Abschnittsdaten – Reisezeitmessung durch Fahrzeugwiedererkennung

Die Reisegeschwindigkeit gemäß FGSV [2000] und die zu ihrer Ermittlung benötigte Reisezeit auf einer Strecke oder auf einer Route sind wesentliche Informationen für Infrastrukturbetreiber und Verkehrsteilnehmer. Sie ist nicht nur ein Indikator für freien oder gestörten Verkehrsfluss, sondern liefert im Falle einer Störung einen Anhaltspunkt für deren Charakteristik.

Reisezeitmessung durch Fahrzeugwiedererkennung erweitert die Grundlage für eine Verkehrslageschätzung somit erheblich. Dafür ist eine hohe Wiedererkennungsrate nicht zwingend erforderlich. Es genügen auf einer Route bereits einige Fahrzeuge pro Stunde, um eine wesentliche Verbesserung der Informationsgrundlage zu erreichen. Zudem können Verfahren zur Fahrzeugwiedererkennung Quelle-Ziel-Informationen liefern.

Infrastrukturseitige Reisezeitmessung wird ermöglicht durch die Erhebung von charakteristischen Fahrzeug- oder Fahrzeugpulkinformationen an mehreren Querschnitten. In einer zentralen Auswerteeinheit werden diese Merkmale abgeglichen. Bei ausreichender Übereinstimmung werden sie demselben Fahrzeug oder Fahrzeugpulk zugeordnet. Durch Subtraktion der Querschnittsdurchgangszeiten wird die Reisezeit ermittelt. Zur Erhebung der Fahrzeug- oder Fahrzeugpulkinformationen sind unter anderem folgende Technologien denkbar:

- Videodetektion: Erfassung der Fahrzeuggeometrie, der Farbe oder des Kennzeichens
- Induktivschleifendetektoren: Erfassung der Fahrzeugsignatur
- Laserscanner: Erstellung fein aufgelöster dreidimensionaler Fahrzeugprofile
- Verarbeitung von Gerätesignalen aus den Fahrzeugen: z.B. Feststellung der Seriennummer eines Bluetooth-Geräts im Fahrzeug, von Radio Frequency Identification (RFID)-Signalen, von Gerätecodes in drahtlosen Netzwerken (vgl. [LEONHARDT, 2008]) oder von Transpondern zur Dedicated Short Range Communication (DSRC) entsprechend der Mauterhebungstechnologie z.B. in Österreich für Lkw basierend auf Mikrowellen

Im Folgenden wird auf Reisezeitmessung basierend auf zwei Erhebungstechnologien eingegangen:

- Kennzeichenerfassung durch Videodetektion – so erhobene Daten liegen zur Bearbeitung der Fragestellung vor
- Analyse von Signaturen aus Induktivschleifendetektoren – die hierfür erforderliche Sensorik ist in großen Stückzahlen im Straßennetz installiert

2.2.1 Videodetektion: Kennzeichenerfassung

Das amtliche Kennzeichen eines Fahrzeugs ist einmalig. Daher stellt es ein sehr gut geeignetes Merkmal für die Fahrzeugwiedererkennung und somit für die Reisezeitmessung dar.

Geeignete Algorithmen sind in der Lage, Bildaufnahmen nach amtlichen Fahrzeugkennzeichen zu durchsuchen und deren Zeichenfolge auszulesen. Durch die Verwendung von Infrarotkameras ist die Qualität der Aufnahmen unabhängig von den Lichtverhältnissen, nur bei starker Kennzeichenverschmutzung oder bei Kennzeichenverdeckung z.B. durch Schnee leiden die Ergebnisse. Eine Kamera kann entweder mobil auf Stativen oder beispielsweise an Lichtsignalmasten oder Brücken fest montiert eingesetzt werden. Sie überträgt ihre Aufnahmen an einen Rechner, der die Kennzeichen ausliest. Zur Einhaltung des gesetzlichen Datenschutzes codiert der Rechner die ausgelesenen Kennzeichen und sendet sie für Online-Anwendungen an einen Server. Dort werden die codierten Kennzeichen von verschiedenen Querschnitten gesammelt, abgeglichen und in der Folge Reisezeiten abgeleitet. Die zeitlich diskretisierte Ausgabe von einzelnen Reisezeiten kann durch eine Aggregation erzielt werden.

Gemäß GRÜBER und RÖHR [2007] kann die Kennzeichenerkennungsrate bei nahezu 100% liegen. Die Kameras sind in der Lage, einen oder mehrere Fahrstreifen zu überwachen. Hierbei kann der Verkehr wahlweise zu- oder abfließend beobachtet werden. Aufnahmedistanz, Aufnahmewinkel und übliche Geschwindigkeiten wirken sich auf die Datenqualität aus. Eine Falschzuordnung zweier unterschiedlicher Kennzeichen bleibt jedoch extrem unwahrscheinlich.

SPANGLER [2009] sieht die Reisezeit auf Routen als wichtige Kenngröße und stellt fest, dass die Fahrzeugwiedererkennung durch automatische Kennzeichenerfassung neben der Erkennung von Transpondern aufgrund der hohen Genauigkeit und der vergleichsweise umfangreichen gelieferten Stichprobe die besten Rahmenbedingungen zur Reisezeitmessung bietet. Er bewertet die räumliche Aussagekraft der gewonnenen Reisezeiten positiv, erwähnt aber auch, dass die Messwerte schon bei ihrer Ermittlung vergangene Verkehrszustände dokumentieren und dabei keine räumlich fein aufgelösten Informationen liefern. Durch den Einsatz geeigneter Filterkriterien stellt er sicher, dass einzelne gemessene Reisezeiten von weiteren Betrachtungen ausgeschlossen werden, die beispielsweise aufgrund eines Zwischenhalts den Verkehrszustand auf der untersuchten Route nicht repräsentativ darstellen.

2.2.2 Induktivschleifen: Signaturanalyse

Die am weitesten verbreitete Detektorart ist nach wie vor die Induktivschleife. Eine Induktivschleife ist physikalisch betrachtet ein Schwingkreis. Sie besteht aus einem Kondensator und einer Spule. Metall, das sich im Umfeld eines Schwingkreises befindet, beeinflusst die Induktivität der Spule. Diese Induktivitätsänderung ist mit entsprechenden Detektoren

messbar. Durch die unterschiedliche Metallmassenverteilung an verschiedenen Fahrzeugen erzeugen Fahrzeuge bei der Überfahrt einer Induktivschleife charakteristische Verläufe der Induktivitätsänderung, eine so genannte Signatur (vgl. Bild 2.1 ff.).

Diese Signatur variiert von Fahrzeug zu Fahrzeug mitunter erheblich, so dass sie zusätzlich zur Klassifizierung der Fahrzeuge am Querschnitt auch zur Fahrzeugwiedererkennung an einem anderen Querschnitt herangezogen werden kann.

BÖHNKE [1980] und PFANNERSTILL [1983] beschreiben ein Verfahren zur Wiedererkennung von Fahrzeugen mit Hilfe ihrer Signaturen über Induktivschleifen, das die Signaturen einer Hauptachsentransformation zuführt. Mit den resultierenden fünf wesentlichsten Merkmalen einer Signatur wird das zugehörige Fahrzeug einer Klasse zugeteilt. Fahrzeugpuls können dann durch die Betrachtung der Klassifizierungsergebnisse in einer Fahrzeugfolge wiedererkannt werden.

SUN ET AL. [1999] und OH ET AL. [2001, 2002] ziehen zusätzlich zur Signatur die Fahrzeuggeschwindigkeit als Merkmal zur Fahrzeugwiedererkennung heran, wohingegen COIFMAN [1998a, 1998b, 2001] lediglich Fahrzeuglänge und Position im Fahrzeugpulk verwendet. Sein Verfahren ist für Autobahnabschnitte mit gestörtem Verkehrsfluss vorgesehen. Für freien Verkehrsfluss liefert es keine Ergebnisse, da die dann vorherrschenden Durchmischungsvorgänge eines der Merkmale unbrauchbar machen. Coifman argumentiert, dass Daten jedoch ohnehin nur benötigt werden, wenn der Verkehrsfluss ins Stocken gerät.

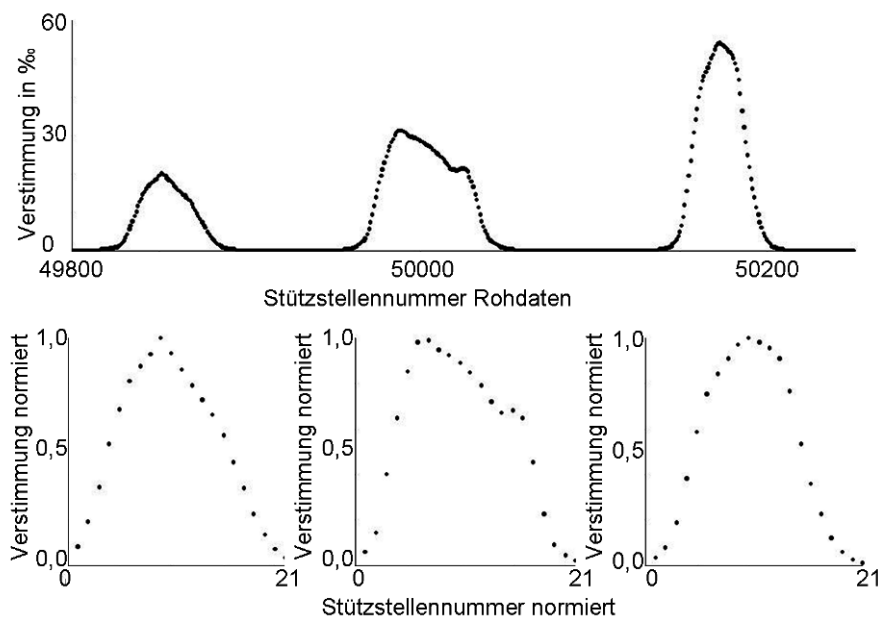


Bild 2.1 Fahrzeugsignaturen als Rohdaten (oben) und nach der Normierung

Ein eigenes Verfahren (vgl. [MAIER UND ROTH, 2008]) verarbeitet mit 125Hz aufgezeichnete Fahrzeugsignaturen. Nach dem Herauslösen der Fahrzeugsignaturen aus den aufgereihten

Verstimmungswerten werden sie zunächst einer Normierung unterzogen, bevor sie jeweils mit Signaturen von einem Nachbarquerschnitt abgeglichen werden und im Übereinstimmungsfall einander zugeordnet werden. Aus den Zeitstempeln wird abschließend die Reisezeit ermittelt.

Die Normierung der Signaturen hat die Eliminierung der Einflüsse systematischer Fehler bei der Datenerhebung zum Ziel. Sie kommen durch unterschiedlich empfindliche Detektoren und unterschiedliche Geschwindigkeiten derselben Fahrzeuge an unterschiedlichen Querschnitten zustande. Der variierenden Sensorempfindlichkeit kann durch eine Kalibrierung begegnet werden. Veränderliche Geschwindigkeiten führen zu Signaturen mit veränderlichen Stützstellenanzahlen. Ihre Vereinheitlichung kann durch eine lineare, kubische oder höhergradige Interpolation erreicht werden (Bild 2.1 unten). Neben dem Zeitstempel wird die fahrzeugspezifische maximale Verstimmung des Schwingkreises vermerkt, da sie unter Berücksichtigung der Sensorkalibrierung ein wichtiges Fahrzeugmerkmal darstellt. Zudem werden die Signaturen als mehr (v.a. Lkw-Signaturen) oder weniger markant (v.a. Pkw-Signaturen) klassifiziert. Dies ermöglicht bei der späteren Zuordnung eine differenzierte Anforderung an die Übereinstimmung der Signaturen.

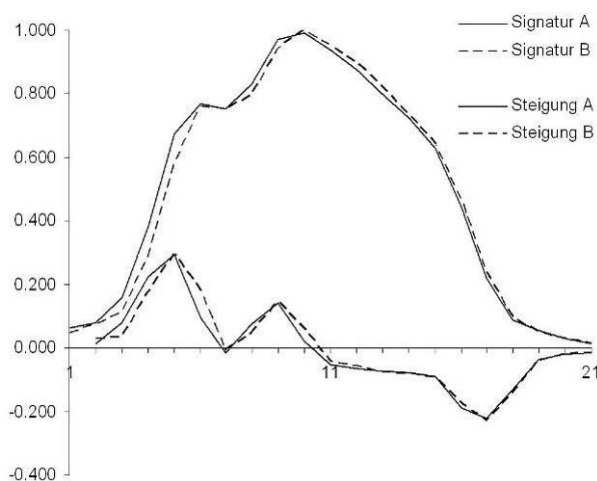


Bild 2.2 Zwei einander zugeordnete Pkw-Signaturen und ihr Steigungsverlauf

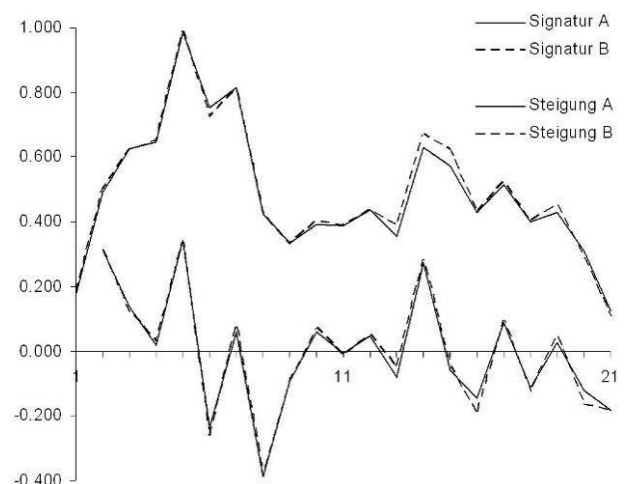


Bild 2.3 Zwei einander zugeordnete Lkw-Signaturen und ihr Steigungsverlauf

Verschiedene Ähnlichkeitsmaße wie der vertikale Abstand oder die Korrelation wurden auf ihre Eignung als Metrik zum Abgleich zweier Signaturen getestet. Als Merkmal konnten entweder die Signaturen selbst, ihr Steigungsverlauf (vgl. Bild 2.2 f.) oder die Fourier-Transformation der Signaturen herangezogen werden. Es zeigte sich, dass der Steigungsverlauf in Verbindung mit dem folgenden Ähnlichkeitsmaß zu sehr aussagekräftigen Werten bezüglich der Ähnlichkeit zweier Signaturen führt:

$$U = \frac{1}{n} \sum_n \frac{\min(y_{A, n}; y_{B, n})}{\max(y_{A, n}; y_{B, n})}$$

Gleichung 2-1

U...	Ähnlichkeit (Uniformeness)
y...	normierte Verstimmung
A,B...	Querschnittsindex
n...	Index über die Stützstellen

Eine Multiplikation des ermittelten Werts mit der entsprechend berechneten Ähnlichkeit der kalibrierten maximalen Verstimmungen zweier Signaturen von zwei Querschnitten verbessert die Aussagekraft weiter.

Werden Signaturpaarungen schon vor dem Abgleich als ungeeignet herausgefiltert, werden die späteren Zuordnungsergebnisse verbessert und zusätzlich die Berechnung der Ähnlichkeiten beschleunigt. Nach dem Abgleich bewirkt ein Filter nur noch die Verbesserung der Zuordnungsergebnisse. Filterkriterien vor dem Abgleich sind beispielsweise unplausible Zeitdifferenzen zwischen den Fahrzeugpassagen an beiden untersuchten Querschnitten, große Abweichungen der kalibrierten maximalen Verstimmungen oder unterschiedliche Klassifizierung bezüglich der Signaturmarkantheit. Nach dem Abgleich können ein zu geringer Ähnlichkeitswert oder ein lokal zu großer Abstand zwischen den Signaturen oder ihrem Steigungsverlauf zum Verwerfen einer Paarung herangezogen werden.

Zwei Signaturen werden letztlich einander zugeordnet, wenn ihr Ähnlichkeitswert höher ist als die Ähnlichkeitswerte aller Paarungen, zu denen eine der beiden Signaturen gehören könnte. Durch das zulässige Zeitfenster liegt diese Information erst mit einem gewissen zeitlichen Versatz vor.

Bei einem Feldtest mit über 85% Zu- und Abflüssen zwischen Start- und Zielquerschnitt und entsprechend harten Filterkriterien wurde ungefähr ein Achtel der durchfahrenden Fahrzeuge wiedererkannt. 38 richtigen Zuordnungen standen sieben falsche entgegen, die v.a. bei Signaturen unterschiedlicher Exemplare desselben Fahrzeugtyps über beiden Detektoren zu beobachten waren. Die verwendeten Fahrzeugsignaturen wurden mit LSA-Detektoren erhoben.

Bei einem Einsatz im Rahmen von AMONES [2010] an mehreren Querschnitten in Bremerhaven wies das Verfahren seine Eignung zur Erhebung von Reisezeiten über mehrere Tage nach.

Weitere Verbesserungen sind für sinkende Zu- und Abflüsse und einen steigenden Lkw-Anteil und damit einhergehend einem höheren Anteil sehr differenzierter Signaturen zu erwarten. Zudem sollten strategische Detektoren mit einem gewissen Abstand zum nächsten Knotenpunkt stabilere Signaturen liefern als LSA-Detektoren, da in Halteliniennähe zahlreiche Beschleunigungs- und Verzögerungsvorgänge die Signaturen verzerren.

2.3 Linkbezogene Verkehrskenngrößen aus Positionsdaten

Globale Ortungssysteme (Global Positioning Systems (GPS)) ermöglichen dem Nutzer eine unkomplizierte Positionsdatenermittlung. Die US-Marine begann mit der Entwicklung des ersten satellitenbasierten Systems TRANSIT bereits Ende der 1950er Jahre. Das System war von 1960 bis 1998 in Betrieb und konnte ab 1973 auch zivil genutzt werden, was insbesondere zu Verbesserungen bei der Schiffsnavigation führte. Weiterentwicklungen des Systems wurden unter anderem aufgrund begrenzter Genauigkeit und lückenhafter räumlich-zeitlicher Verfügbarkeit erforderlich. Dies führte letztlich zur Entwicklung des NAVSTAR-GPS (Navigation System with Timing And Ranging - Global Positioning System) im Auftrag des US-Verteidigungsministeriums ab 1970. Das System war 1995 komplett betriebsbereit und liefert seit dem Jahr 2000 auch zivile Signale mit einem Fehler von maximal 10m in 90% der Messungen (vgl. [GRÜLLER, 1997], [BÜHLER, 2006]). Die Genauigkeit kann durch Verwendung von Verfahren des Differential GPS gesteigert werden (vgl. [HANKE, 1999]). Diese Verfahren nutzen exakt vermessene Referenzpunkte.

Das kommerzielle Ortungssystem Galileo der Europäischen Union wird seit 1994 ausschließlich für den zivilen Einsatz entwickelt und wird so in diesem Bereich erstmals Daten mit einem Fehler von weniger als 1m liefern. Es soll im Laufe der nächsten Jahre in Betrieb gehen (vgl. [RÜFFER ET AL., 2005], [BÜHLER, 2006]).

Fahrzeuggenerierte Positionsdaten mit der gegenwärtigen Genauigkeit aus dem NAVSTAR-GPS oder künftig aus Galileo erlauben eine Aufbereitung zur Ableitung von streckenbezogenen verkehrlichen Kenngrößen. Durch die Betrachtung mehrerer zeitlich aufeinander folgender Positionsdatensätze aus einer Fahrt, die dabei beobachtete örtliche Veränderung und deren Inbezugsetzung zur zwischenzeitlich verstrichenen Zeit können Reisezeiten und Geschwindigkeiten gewonnen werden. Seitens der Erhebungsgeräte sind die Zeitintervalle zwischen den Positionsbestimmungen nahezu beliebig klein, üblich sind Intervalle von einer Sekunde und weniger. Zur Offline-Analyse können so zeitfein erhobene Positionsdaten ohne weiteres verwendet werden. Sie ermöglichen die problemlose Verfolgung einer Fahrt durchs Straßennetz unter genauer Kenntnis der räumlich-zeitlich fein aufgelösten Geschwindigkeiten. Mögliche Darstellungsformen sind Zeit-Weg-, Zeit-Geschwindigkeits- oder Weg-Geschwindigkeitsdiagramme. Auch lässt sich nach einem Abgleich der Positionsdaten mit dem Straßennetz die befahrene Route mit geschwindigkeitsabhängig eingefärbten Streckenabschnitten auf einer Straßenkarte darstellen.

Für Online-Anwendungen ist zu beachten, dass die erhobenen Positionsdaten an eine zentrale Stelle übermittelt werden müssen. Neben einer Übermittlung per Mobilfunk aus beliebigen Fahrzeugen kommen Positionsdaten aus Flottenfahrzeugen in Frage, die ohnehin ständig mit einer Zentrale in Kontakt stehen (z.B. Taxis, Busse des Öffentlichen Personennahverkehrs (ÖPNV), Paketdienste). Diese Art der Datenübertragung hat den Vorteil, dass sie durch die Nutzung ohnehin vorhandener Kommunikationswege wie Betriebsfunkkanäle

deutlich kostengünstiger ist. Zudem sind derartige Flottenfahrzeuge meist überproportional häufig im Straßennetz unterwegs, so dass relativ wenige ausgestattete Fahrzeuge relativ viele Daten liefern können.

Positionsdaten werden auf den genannten Kommunikationswegen nicht in Sekundenintervallen übertragen, üblicher sind Intervalle zwischen 30 und 300s. Folge ist, dass zwischen der Übertragung zweier Positionsdatensätze ein Fahrzeug mehrere Knotenpunkte passieren kann und dadurch nicht mehr offensichtlich ist, welcher Link oder welche Links befahren wurden. Um Rückschlüsse auf die Verkehrsqualität auf den befahrenen Streckenabschnitten ziehen zu können, ist also neben der Herstellung des Bezugs zum Straßennetz durch ein Map-Matching-Verfahren die Ermittlung der am wahrscheinlichsten zwischenzeitlich befahrenen Streckenabschnitte erforderlich, die durch einen Routenwahlalgorithmus bewerkstelligt werden kann.

Bei Bussen des ÖPNV kann die Ortung auch lokal statt zeitlich fixiert sein, wenn sie durch eine An- oder Abmeldung an einer Bake statt durch Übermittlung eines GPS-Datensatzes erfolgt. Da die Linienführung aber festgelegt und die Position der Bake im Kartenmaterial verankert ist, erübrigen sich für diesen Anwendungsfall Map Matching und Routensuche. Daher gleicht diese Art der FCD-Erhebung eher einer abschnittswisen Verkehrsdatenerhebung ausschließlich für Fahrzeuge des ÖPNV. Nach FGSV [2003] werden Meldepunkte oder Türschließsignale auch zur Synchronisation verwendet, um unter Kenntnis der Route und unter Verwendung von Radsensoren den genauen Fahrtverlauf eines Fahrzeugs zu verfolgen.

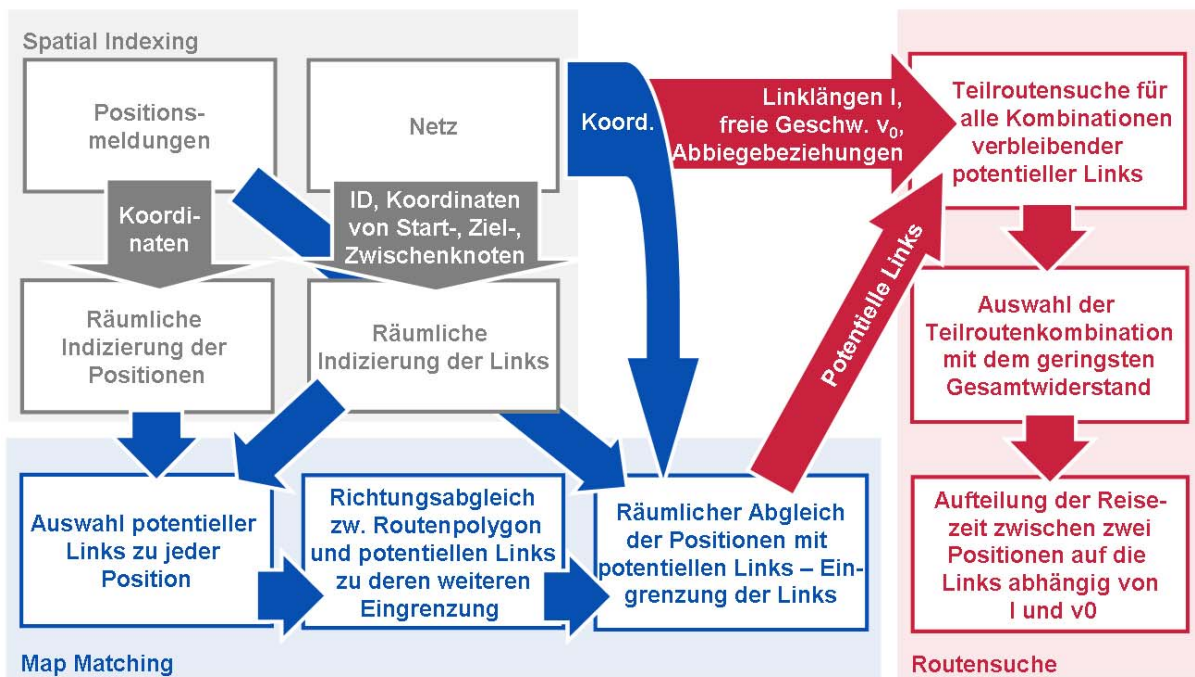


Bild 2.4 Ableitung von linkbezogenen Kenngrößen aus Flottenpositionsdaten – Überblick

Nachfolgend wird ein Offline-Verfahren zur Ableitung von streckenbezogenen Kenngrößen aus Flottenpositionsdaten mit Übertragungsintervallen im Minutenbereich entwickelt (Überblick in Bild 2.4). Die daraus gewonnenen Informationen können im Gegensatz zu Daten aus der infrastrukturseitigen Datenerhebung für sämtliche Streckenabschnitte im Netz gelten. Damit liefern sie einen essentiellen Beitrag zur netzweiten Verkehrslageschätzung. Dieser Umstand und die vergleichsweise aufwendige Datenaufbereitung rechtfertigen eine eingehende Beschreibung des Verfahrens. Der im Grundsatz aus der Literatur bekannte, jedoch selten umfassend dargestellte Ablauf wird durch neue Aspekte erweitert insbesondere im Bereich „Teilroutensuche“ und „Auswahl der Teilroutenkombination“ zur Sicherstellung der korrekten Routenreproduktion und beim „Richtungsabgleich zwischen Routenpolygon und Links“ zum verbesserten Map Matching und zur Eingrenzung des Aufwands für die Teilroutensuche.

2.3.1 Map Matching mit räumlicher Indizierung

Ziel beim Map Matching ist die Zuordnung eines Positionsdatensatzes zunächst zu einem Link und nachfolgend zu einem Punkt auf dem Link. Hierbei ist zu beachten, dass digitale Straßenkarten und erhobene Positionsdaten wie alle Messdaten einem Messfehler unterliegen, der beim Map Matching nicht zum Versagen führen darf. Entsprechend vorherigen Ausführungen beträgt die Abweichung der gemessenen von der tatsächlichen Position meist weniger als 10m.

Bei digitalen Karten wird unterschieden nach geometrischen Karten, die die einzelnen Links darstellen, und topologischen Karten, die lediglich die Knotenpunkte enthalten. Links entstehen in digitalen Karten durch das Aufspannen einer Strecke zwischen zwei Knotenpunkten. So sind lediglich die Koordinaten der Knotenpunkte hinterlegt, der Link wird durch eine Geradengleichung mit einer Einschränkung des Gültigkeitsbereichs beschrieben. Enthält eine digitale Karte auch Strecken, die durch Polygonzüge dargestellt werden, so ist die Berücksichtigung der Stützstellenkoordinaten zum Aufspannen der Polygonzüge erforderlich. Weitere Merkmale von Links sind Länge, durchschnittliche Geschwindigkeit bei freiem Verkehrsfluss (freie Geschwindigkeit), Längsneigung, Kapazität, Fahrtrichtung und zulässige Verkehrsmittel. Die erweiterte Beschreibung von Knotenpunkten kann die vorhandenen Abbiegebeziehungen oder Angaben zur Kapazität umfassen. Widerstände, die die Routenwahl (Abschnitt 2.3.2) beeinflussen, können sowohl für Links als auch für Knotenpunkte festgesetzt werden.

Beim Map Matching existieren neben dem Ansatz „Probieren“ globale Ansätze sowie inkrementelle Suchmethoden. Beim globalen Map Matching wird der schnelle gezielte Zugriff auf Kartenobjekte durch Baumstrukturen ermöglicht. Baumstrukturen beginnen mit einer Wurzel, verzweigen sich vielmals und enthalten in den Blättern, der letzten Verzweigungsebene, die Karteninformationen. Die Anzahl der Verzweigungsebenen wird durch die Höhe des

Baumes erfasst. Verschiedene Baumstrukturen unterscheiden sich hinsichtlich des Speicherraumbedarfs und der Eigenschaften beim Einfügen, Löschen und Auffinden von Datensätzen.

Eine weitere globale Suchmethode beruht auf einer Rasterung der digitalen Karte (vgl. Bild 2.5). Elemente der Karte wie Knotenpunkte oder Links und Positionsdaten werden jeweils Rastersegmenten zugeordnet, also räumlich indiziert. Daher wird diese Methode Spatial Indexing genannt [SCHIMANDL ET AL., 2009]. Liegt eine Position im selben Rastersegment wie ein Netzelement, so kann eine Zuordnung erwogen werden. Der Vorteil des Spatial Indexing gegenüber den Baumstrukturen liegt in einer deutlichen Beschleunigung, da es statt der Suche durch mehrere Verzweigungsebenen lediglich einen Zugriff erforderlich macht.

Inkrementelle Methoden sehen eine schrittweise Suche ausgehend von einer zuvor bestimmten Position vor. Schnell kann ein inkrementelles Vorgehen unter der Verwendung von Ausschlusskriterien und von Merkmalen des Wegenetzes zum Ziel führen: Ist die Position vom vorangegangenen Erhebungsintervall bekannt, so ist der Radius eingegrenzt, in dem die aktuelle Position liegen kann. Folglich führt eine Suche ausgehend von der alten Position schnell zu einem Sucherfolg.

Nach BRAKATSOULAS ET AL. [2005] führen globale Suchverfahren zu geringfügig besseren Ergebnissen, wohingegen inkrementelle Ansätze bei immer noch sehr guten Ergebnissen deutlich kürzere Suchzeiten aufweisen.

Fein aufgelöste Straßenkarten mit einer hohen Linkdichte und mit zahlreichen Links, die durch Polygonzüge dargestellt sind, führen hingegen zu einer Verlangsamung inkrementeller Suchverfahren, weil dann mehr Streckenabschnitte in einem plausiblen Radius um die vorherige Position mit der aktuellen Position abgeglichen werden müssen.

Nachfolgend wird ein selbst entwickeltes, zuverlässiges Verfahren zum Map Matching basierend auf dem Spatial Indexing beschrieben, das weitgehend unabhängig von Netzeigenschaften eine absolut und relativ hohe Leistungsfähigkeit aufweist. Entsprechend der physikalischen Definition wird unter Leistung die Arbeit je Zeiteinheit verstanden, wobei die Arbeit hier nicht in Joule, sondern in der Anzahl der im Netz zugeordneten Positionsdaten gemessen wird. Die Leistungsfähigkeit ist insbesondere relevant für Online-Anwendungen und zunehmend für größere Netze und eine wachsende FCD-Flotte.

Zunächst ist festzulegen, ob Knotenpunkte oder Links räumlich indiziert werden. Knotenpunkte sind leichter zu indizieren, da sie eindeutig einem Rastersegment zugeordnet werden können. Entsprechend verhält es sich für Positionsmeldungen. Gemeldete Positionen liegen jedoch häufig auf einem Link weit vom nächsten Knotenpunkt entfernt, während der Knotenpunkt eines ungeeigneten Links in der Nähe liegt, so dass es in der Folge zu Falschzuordnungen kommen kann. Durch Einfügen von Zwischenknoten auf den Links in einem fest-

zulegenden Höchstabstand voneinander kann dieser Fehler in vielen Fällen umgangen werden. Soll die gemeldete Position später exakt auf einem Link verortet werden, so bleibt zu beachten, dass an einen Knotenpunkt mehrere Links angeschlossen sind, die mit der Positionsmeldung abgeglichen werden müssen. Dies bedeutet einen erhöhten Rechenaufwand.

Die räumliche Indizierung von Links ist schwieriger, da ein Link häufig durch mehrere Rastersegmente führt und folglich mehrfach indiziert werden muss (Beispiel siehe Bild 2.5). Nach Indizierung einer Positionsmeldung können Links, zu denen die Positionsmeldung gehören könnte, jedoch schneller und gezielter ausgewählt werden als über den Umweg, zunächst nach Knoten oder Zwischenknoten zu suchen. Daher sieht das vorgestellte Verfahren die Indizierung von Links vor. Da die gemeldeten Positionen ebenso wie Links unglücklich an der Grenze zwischen zwei Rastersegmenten liegen können, müssen neben dem Rastersegment einer Positionsmeldung auch die umliegenden Segmente nach Links durchsucht werden.

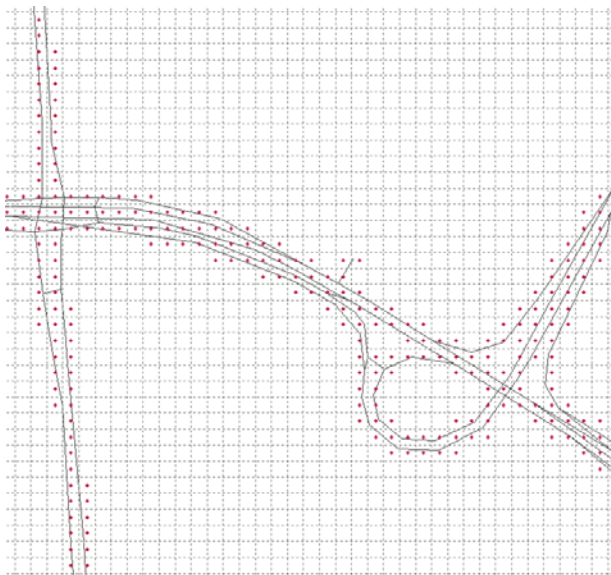


Bild 2.5 Gerasterte digitale Karte mit räumlich indizierten Links

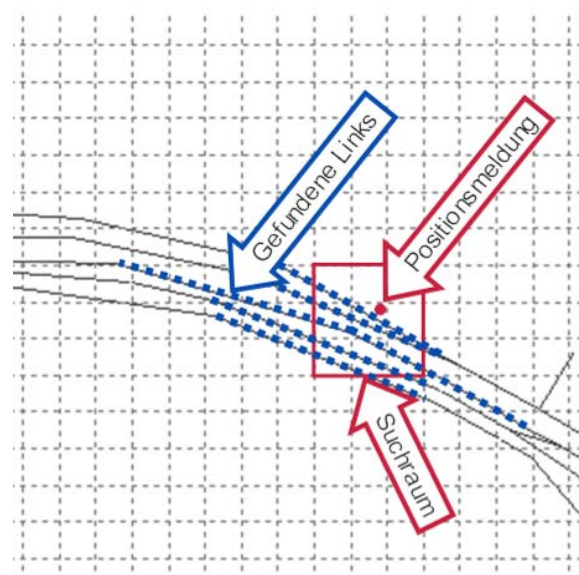


Bild 2.6 Abgleich einer Position mit dem Netz basierend auf Spatial Indexing

Wie eingangs dieses Abschnitts erwähnt, zielt das Map Matching auf die rasche Zuordnung einer gemeldeten Position zunächst zu einem Link und nachfolgend zu einem Punkt auf dem Link. Das Spatial Indexing kann aber mehrere Links zu einer Position liefern (vgl. Bild 2.6).

Die Anzahl der ausgewählten Links sollte vor der Routensuche so weit wie möglich reduziert werden. Es muss davon ausgegangen werden, dass die definitive Festlegung auf einen Link erst bei der Routensuche erfolgen kann (siehe Kapitel 2.3.2). Ein Abgleich der Richtungen der gefundenen Links mit den Richtungen der beiden anliegenden Abschnitte des Routen-

polygons kann jedoch schon jetzt zum Verwerfen ungeeigneter Links herangezogen werden. Überschreitet die Differenz für beide Abschnitte einen Schwellenwert, so gilt die Fahrtrichtung auf dem Link für die Route als ungeeignet. Der Link kann verworfen werden. Der Schwellenwert sollte so gewählt werden, dass bei Abbiegevorgängen nicht zu viele Links verworfen werden. Daher sollten Abweichungen von mindestens 90° zulässig sein.

Polygone Links bedürfen einer besonderen Aufbereitung, da ihnen nicht eindeutig eine Richtung zuzuweisen ist: Weicht die Richtung eines einzelnen Abschnitts um mehr als einen Schwellenwert (z.B. 90°) von der Richtung der Geraden durch die erste und die letzte Stützstelle des Polygonzugs ab, so wird dem Link keine Richtung zugeordnet und er kann beim Abgleich mit Routenpolygonen nicht verworfen werden. Dies ist häufig bei Rampen an kleeblattförmigen planfreien Knotenpunkten der Fall. Wird dieser Schwellenwert nie überschritten, so wird dem Polygonlink die Richtung der Geraden durch die erste und die letzte Stützstelle zugeordnet; so kann der Link beim Abgleich mit dem Routenpolygon verworfen werden. Dies ist beispielsweise bei Autobahnen der Fall, deren Links sich über mehrere 100m erstrecken und aufgrund ihrer Trassierung als Polygone mit nur geringen Richtungsänderungen dargestellt sind.

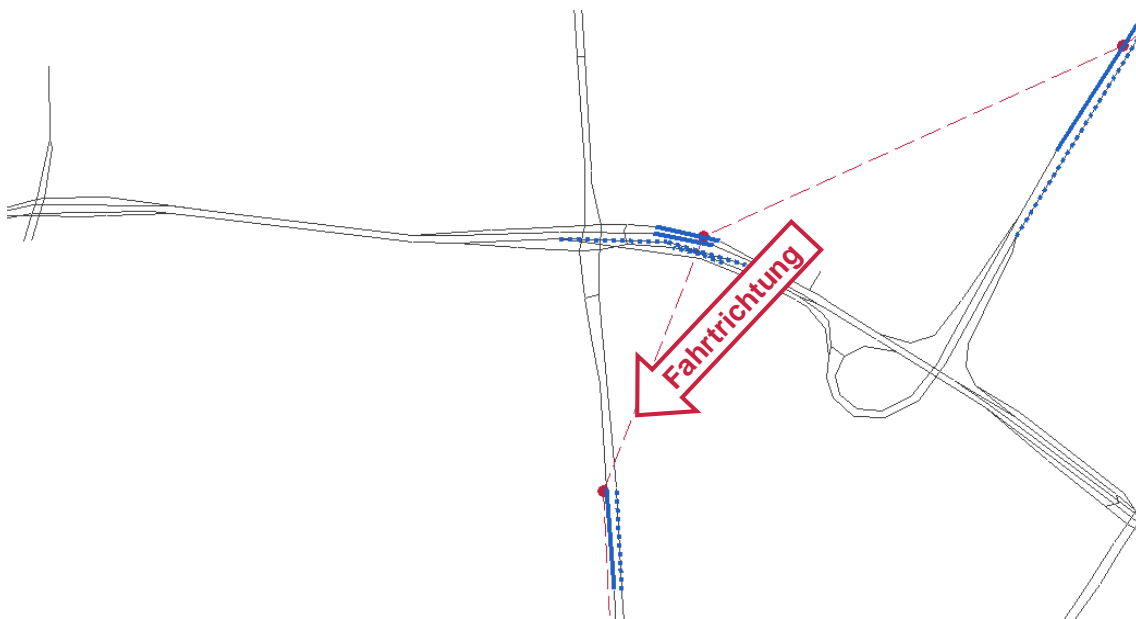


Bild 2.7 Positionen mit Polygonzug einer Route – über Spatial Indexing ausgewählte eventuell zugehörige Links, nicht verworfene Links durchgezogen eingezeichnet

Bild 2.7 zeigt, dass durch den Richtungsabgleich zwischen Links und Routenpolygon die Anzahl der Links, die zu den Positionen der Route gehören können, erfolgreich reduziert werden kann.

Nun werden die gemeldeten Positionen mit den jeweils geeigneten Links abgeglichen, um die genaue Lage der Positionen auf den Links festzustellen.

Unterschreitet der Abstand einer gemeldeten Position zu einer der Stützstellen eines Links einen Schwellenwert (z.B. 10m), so wird darauf geschlossen, dass die Position auf dieser Stützstelle liegen kann. Der resultierende Fehler von weniger als 10m kann aufgrund der Größe der Übertragungsintervalle und der im Straßenverkehr üblichen Geschwindigkeiten akzeptiert werden.

Ist dies nicht der Fall, so erfolgt die Bestimmung des Abstands zwischen Position und Link sowie die genaue Verortung auf einem Link wie folgt:

Ein Link, der von Knoten K1 über Zwischenknoten Z1 zu Knoten K2 führt, wurde zum Abgleich mit der Position P1 ausgewählt. Erste Bedingung für eine mögliche Zuordnung von Position P1 zu einem der beiden Linkabschnitte ist, dass ihr Abstand zu den jeweiligen Startpunkten nicht größer ist als die Länge des jeweiligen Abschnitts. Daher wird um K1 ein Kreis mit der Länge des angrenzenden Abschnitts gezogen (vgl. Bild 2.8).

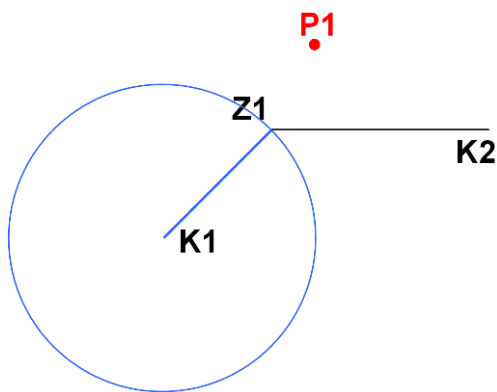


Bild 2.8 Abstandsbestimmung Position-Link und genaue Verortung (1)

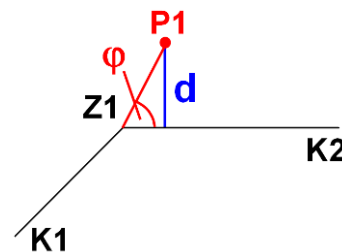


Bild 2.9 Abstandsbestimmung Position-Link und genaue Verortung (2)

Nur wenn P1 innerhalb des Kreises liegt, ist eine Zuordnung zum Abschnitt möglich. Ist dies nicht der Fall, so wird die Position mit dem nächsten Teilabschnitt des Polygonlinks abgeglichen.

Die Länge des Teillinks und die Entfernung zwischen einer Position und einem Knotenpunkt können mit den Beträgen der beiden Vektoren ermittelt werden. Wird die erste Bedingung erfüllt, so gilt

$$\left| \overrightarrow{K_{FN} P_n} \right| \leq \left| \overrightarrow{K_{FN} K_{TN}} \right|$$

Gleichung 2-2

K_{FN} ... Startknoten, FromNode

K_{TN} ... Zielknoten, ToNode

P_n ... gemeldete Position

K_{TN} entspricht in Bild 2.8 Z1.

Zweite Bedingung ist, dass für den Winkel φ zwischen den beiden zuvor verwendeten Vektoren gilt:

$$|\varphi| \leq 90^\circ \text{ mit } \cos \varphi = \frac{\overrightarrow{K_{FN} P_n} \cdot \overrightarrow{K_{KN} K_{TN}}}{\left| \overrightarrow{K_{FN} P_n} \right| \cdot \left| \overrightarrow{K_{FN} K_{TN}} \right|} \quad \text{Gleichung 2-3}$$

Wird diese Bedingung eingehalten, so liegt P_n im Halbkreis um K_{FN} mit Radius $\left| \overrightarrow{K_{FN} K_{TN}} \right|$ in Blickrichtung K_{TN} .

Dritte Bedingung für eine Zuordnung ist, dass der Abstand d zwischen gemeldeter Position und Link (vgl. Bild 2.9) einen Schwellenwert nicht überschreitet (z.B. 10m). Da gilt

$$\sin \varphi = \frac{d}{\left| \overrightarrow{K_{FN} P_n} \right|} \text{ und } \cos \varphi = \frac{\overrightarrow{K_{FN} P_n} \cdot \overrightarrow{K_{KN} K_{TN}}}{\left| \overrightarrow{K_{FN} P_n} \right| \cdot \left| \overrightarrow{K_{FN} K_{TN}} \right|},$$

berechnet sich der Abstand d zu

$$d = \sin \left(\arccos \left(\frac{\overrightarrow{K_{FN} P_n} \cdot \overrightarrow{K_{FN} K_{TN}}}{\left| \overrightarrow{K_{FN} P_n} \right| \cdot \left| \overrightarrow{K_{FN} K_{TN}} \right|} \right) \right) \cdot \left| \overrightarrow{K_{FN} P_n} \right| \quad \text{Gleichung 2-4}$$

Die zugeordnete Position auf dem Link ergibt sich für den Fall, dass alle drei Bedingungen erfüllt werden, zu

$$P_{n,zug.} = K_{FN} + \frac{\cos \varphi \cdot \left| \overrightarrow{K_{FN} P_n} \right|}{\left| \overrightarrow{K_{FN} K_{TN}} \right|} \cdot \overrightarrow{K_{FN} K_{TN}} \quad \text{Gleichung 2-5}$$

oder durch die Näherung $\cos \varphi \approx 1$ für kleine φ näherungsweise zu

$$P_{n,zug.} = K_{FN} + \frac{\left| \overrightarrow{K_{FN} P_n} \right|}{\left| \overrightarrow{K_{FN} K_{TN}} \right|} \cdot \overrightarrow{K_{FN} K_{TN}} \quad \text{Gleichung 2-6}$$

Für das Beispiel ergibt sich also

$$P_{1,zug.} = Z_1 + \frac{\cos \varphi \cdot \left| \overrightarrow{Z_1 P_1} \right|}{\left| \overrightarrow{Z_1 K_2} \right|} \cdot \overrightarrow{Z_1 K_2}$$

Links, auf denen keine Position der gemeldeten Position zugeordnet werden kann, werden verworfen.

Ergebnis des entwickelten Verfahrens zum Map Matching ist eine Liste sämtlicher Links zu sämtlichen gemeldeten Positionen einer Route samt exakter möglicher Verortung der Positionsmeldungen, die als plausibel erachtet wurden und die einen ausreichend geringen Abstand zu den gemeldeten Positionen aufweisen. Diese Liste ist Grundlage für das anschließende Routensuchverfahren.

2.3.2 Routensuche

Routensuchalgorithmen suchen nach der besten Route zwischen zwei Punkten in einem Wegenetz. Sie benötigen als Grundlage eine digitale geometrische Karte sowie auf die Karte referenzierte Angaben zu Quelle und Ziel einer Route. Diese Referenzierung kann sich sowohl auf Knotenpunkte als auch auf Links beziehen. Die beste Route vom Start- zum Zielpunkt ist die mit dem geringsten Widerstand. Als Widerstand wird beispielsweise die zurückgelegte Strecke, die erforderliche Reisezeit oder eine Kombination aus beiden verwendet. Routensuchalgorithmen werden zur Ermittlung der Routenempfehlungen in Navigationssystemen, in der Verkehrsplanung im Rahmen des Vier-Stufen-Algorithmus beim Verfahrensschritt Umlegung sowie bei der Gewinnung von Informationen über den Verkehrszustand aus Positionsdaten genutzt. In dieser Arbeit ist der letztgenannte Punkt relevant.

Die bekannten Algorithmen werden nach uninformierten und informierten Suchverfahren unterschieden. Die beiden Hauptvertreter sind Dijkstra- (uninformiert) und A*-Algorithmus (informiert). Uninformierte Verfahren gehen unter Kenntnis der Netzinformationen von der Quelle einer Fahrt aus, informierte Verfahren berücksichtigen zusätzlich die Ortsveränderung während der Routensuche in Relation zum Fahrtziel.

Am Startpunkt der Route beträgt der kumulierte Widerstand 0. Der Dijkstra-Algorithmus ermittelt vom Startpunkt aus die Links zu den umliegenden Knotenpunkten. Liegt der Zielpunkt an den Links, so ist die Routenwahl bereits abgeschlossen. Ist dies nicht der Fall, addiert der Algorithmus zum bisherigen Widerstand den Widerstand des jeweils befahrenen Links. Die Routensuche wird gemäß den Regeln fortgesetzt:

- Die Routensuche wird immer an dem Suchast fortgesetzt, der den geringsten kumulierten Widerstand aufweist. Der Suchast teilt sich auf die anknüpfenden Links auf. So entstehen neue Suchäste, deren kumulierte Widerstände aus dem kumulierten Widerstand des bisherigen Suchasts und den einzelnen Widerständen der anknüpfenden Links aufsummiert werden.
- Enthält einer der anknüpfenden Links oder der Zielknoten des Links das Ziel der Route, so ist eine mögliche Route gefunden. Diese wird mit dem ermittelten kumu-

lierten Widerstand vermerkt. Die Suche ist aber noch nicht beendet, da noch andere Routen mit geringerem Widerstand gefunden werden können.

- Ein Suchast wird nicht fortgesetzt, wenn er auf einen Knoten trifft, zu dem bereits ein anderer Suchast mit geringerem Widerstand führt.
- Ein Suchast wird nicht fortgesetzt, wenn die Route bereits einen höheren Widerstand aufweist als eine gefundene abgeschlossene Route zum Zielpunkt.
- Ein Suchast wird nicht fortgesetzt, wenn ein maximal möglicher Widerstand zum Ausgangspunkt überschritten wird.
- Die Suche ist beendet, wenn an jedem Suchast eines der Kriterien zum Abbruch der Suche greift.

Dijkstra- und A*-Algorithmus sind sich grundsätzlich sehr ähnlich. Der A*-Algorithmus berücksichtigt neben dem Widerstand zwischen Start und Ende eines Suchasts zusätzlich die Schätzung des Widerstands zwischen dem Ende des Suchasts und dem Ziel der Route. Wird die Luftliniendistanz zwischen Ende des Suchasts und Zielposition herangezogen, so kann diese Distanz mit einem Faktor >1 multipliziert werden, um zu berücksichtigen, dass die Zielposition in der Regel über netzgeometrieverursachte Umwege erreicht wird. Eine weitere Variante ermittelt den Restwiderstand über eine Zelleinteilung [WELCH UND BISHOP, 2004]. Die Auswirkungen aus diesen Modifikationen sind jedoch gering und können sich sowohl positiv als auch negativ auf die Leistungsfähigkeit des Verfahrens und auf die Qualität der Ergebnisse auswirken.

Als Konsequenz durchsucht der A*-Algorithmus den Raum in einer Ellipse um Quelle und Ziel, während der Dijkstra-Algorithmus den Raum in einem Kreis um die Quelle analysiert. So greifen beim A*-Algorithmus Kriterien zum Beenden der Routensuche an falschen Suchästen früher. Dadurch läuft der A*-Algorithmus deutlich schneller und effektiver ab.

Gemäß SANDERS UND SCHULTES [2005] kann die Leistungsfähigkeit der Algorithmen durch eine bidirektionale Suche sowohl von Start- als auch von Zielpunkt ausgehend gesteigert werden. So werden im Fall des Dijkstra-Algorithmus näherungsweise zwei Kreise mit der halben Luftlinienentfernung als Radius statt einer mit der ganzen Luftlinienentfernung als Halbmesser durchsucht. Da der Radius quadratisch in die Fläche eingeht, halbiert sich die durchsuchte Fläche. Für den A*-Algorithmus werden mit der bidirektionalen Umsetzung entsprechend verkleinerte Ellipsen durchsucht. Ein weiterer Ansatz zur Beschleunigung der Routensuche über lange Distanzen wie beispielsweise bei der Fahrzeugnavigation sieht die Berücksichtigung des fein aufgelösten Netzes nur in unmittelbarer Nähe von Start- und Zielpunkt vor. Dazwischen werden ausschließlich Strecken mit regionaler oder überregionaler Verbindungsfunktion als Teil der Route in Betracht gezogen.

Ein Vergleich beider Algorithmen gemäß Bild 2.10 bei der Auswertung einer Beispielroute, von der sechs Positionen bekannt sind, zeigt, dass beide Algorithmen dieselbe Route fin-

den. Jedoch durchsucht der Dijkstra-Ansatz 20 Knotenpunkte (durch blaue Kreise markiert) und 72 Links, während der A*-Algorithmus lediglich elf Knotenpunkte und 31 Links analysiert. Dieser Unterschied vergrößert sich weiter für dichtere Straßennetze und für wachsende Abstände zwischen gemeldeten Positionen.

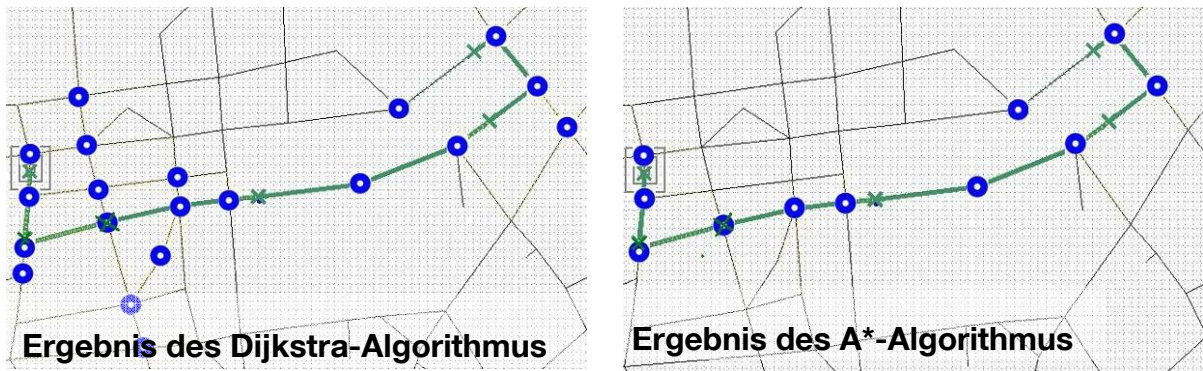


Bild 2.10 Vergleich Dijkstra-Algorithmus – A*-Algorithmus

Beim A*-Algorithmus bleibt aber die Schwierigkeit der Schätzung des Widerstands zwischen Ende eines Suchraums und Ziel. Werden Streckenlängen als Widerstand herangezogen, so gelingt die Schätzung durch eine Luftlinienverbindung gut und einfach. Werden jedoch freie Reisezeiten herangezogen und liegen Links mit unterschiedlichen freien Geschwindigkeiten im Suchraum vor, so muss die geringste der freien Linkgeschwindigkeiten im Suchraum für die Umrechnung der geometrischen Distanz in eine Reisezeit herangezogen werden, um mit Sicherheit keinen Fehler einzugehen. Variieren die freien Geschwindigkeiten stark, wird der Suchraum dadurch erheblich vergrößert. Dadurch verliert der A*-Algorithmus wieder einen Teil seines Geschwindigkeitsvorteils.

Beide Algorithmen liefern zuverlässig plausible Ergebnisse der Routensuche. Ist bei einer Anwendung die Geschwindigkeit das entscheidende Kriterium, so spricht vieles für den A*-Algorithmus. Bei einer Offline-Anwendung spricht nichts gegen die Verwendung des Dijkstra-Ansatzes, der bei der Variation der Messeinheit des Widerstands weniger anfällig ist. So wird als Grundlage für die weiteren Überlegungen im Rahmen des vorliegenden Vorhabens der Dijkstra-Algorithmus gewählt.

Das Verfahren zum Map Matching gemäß Kapitel 2.3.1 kann zu einer Position jeweils mehrere mögliche Links oder Knotenpunkte liefern. Entscheidet sich der Routensuchalgorithmus bei der Routensuche von Position i zu Position $i+1$ bei Position $i+1$ für den falschen Link und wird dieser beim Schritt zu Position $i+2$ als Startlink herangezogen, so kann dies zu erheblichen Fehlern bei der Routensuche führen (rot gestrichelte Route in Bild 2.11 zwischen den Positionen $i+1$ und $i+2$; der Knotenpunkt ungefähr in Bildmitte ist teilplanfrei ausgeführt; die Position $i+1$ liegt auf der Ausfahrtrampe, wird aber der Hauptfahrbahn zugeordnet). Folge wäre die Ableitung von falschen Geschwindigkeiten für nicht befahrene Links. Der Fehler könnte im abgebildeten Beispiel durch die Berücksichtigung der Position $i+2$ korrigiert wer-

den. Es könnten auch Kriterien zum Verwerfen einer Route entwickelt werden wie zu unterschreitende Schwellenwerte für einen Umwegfaktor oder für eine maximale plausible Geschwindigkeit. Dieser Ansatz steht jedoch der Zielsetzung entgegen, aus den gegebenen Daten die maximal mögliche Menge an Informationen zu gewinnen. Die Rückwärtssuche liefert in vielen Fällen eine Verbesserung, da Wartezeiten eher vor als nach einem Knotenpunkt anfallen und folglich strittige Positionen häufiger vor einem Knotenpunkt gemeldet werden. Jedoch kann der beobachtete Fehler auch bei der Rückwärtssuche auftreten. Eine weitere Verbesserung bringt der Abgleich von Vorwärts- und Rückwärtsroutensuche. Auch das garantiert jedoch nicht die fehlerfreie Bestimmung einer Route aus Positionsmeldungen.

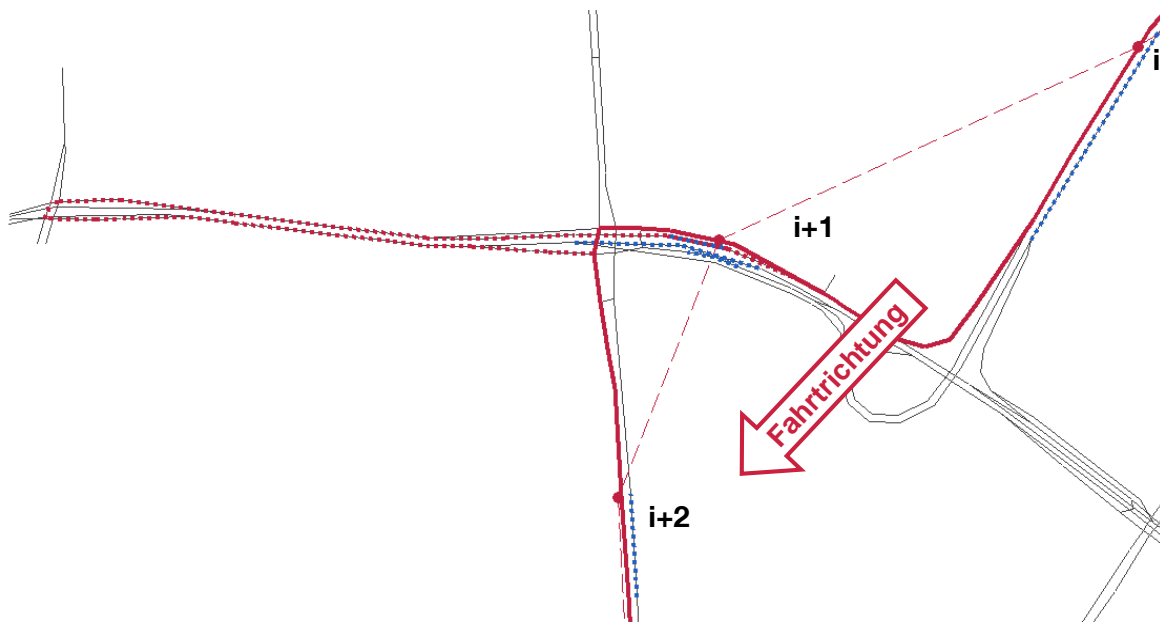


Bild 2.11 Ergebnis der Routensuche ohne und mit optimierter Teilroutenkombination

Grundsätzlich ist zu bemerken, dass der beschriebene Fehlertyp für engmaschigere Netze und kürzere Intervalle zwischen Positionsmeldungen häufiger auftritt, da dann mehr gemeldete Positionen in unmittelbarer Knotenpunktnähe liegen.

Für den Fall eines Rückstaus auf der Ausfahrt am Knotenpunkt kann sich auch die Situation einstellen, dass j weitere Positionen auf der Rampe liegen, die genauso gut der Hauptfahrbahn zugewiesen werden könnten. So würde der Fehler bei der Routensuche erst beim Schritt von Position $i+j+1$ nach Position $i+j+2$ offensichtlich werden. Folge wäre, dass der erkannte Rückstau zwischen den Positionen $i+1$ und $i+j+1$ den falschen Links zugeordnet würde. Dieser Fehler ist also unbedingt zu vermeiden.

Die Anzahl der j weiteren Positionen auf der Rampe ist jedoch nicht bekannt. Da das vorgestellte Verfahren für die Offline-Analyse von Positionsmeldungen optimiert sein soll, bietet sich als Lösung an, Teilrouten zwischen den m möglichen Links zu einer Position und den n möglichen Links zur folgenden Position zu suchen. So werden $m \cdot n$ mögliche Teilrouten

gefunden. Werden diese Teilrouten für sämtliche Positionsmeldungsintervalle einer Route verknüpft, so kann nach der Bestimmung des Gesamtwiderstands für sämtliche Teilroutenkombinationen die plausibelste Route am geringsten Gesamtwiderstand erkannt werden (durchgezogene gezeichnete Route in Bild 2.11).

Dieses Vorgehen bedeutet zwar einen erhöhten Rechenaufwand, da der Routensuchalgorithmus zahlreiche Teilrouten ermitteln muss, gewährleistet jedoch eine annähernd fehlerfreie Routensuche bei einer Anzahl von verworfenen Routen nahe 0. Somit kann die größtmögliche Menge an verkehrlichen Informationen aus den Positionsdaten gewonnen werden.

2.3.3 Ableitung und Aggregation von Linkreisezeiten

Aus Map Matching und Routensuche sind neben der zurückgelegten Strecke die Links bekannt, die zwischen den Übertragungszeitpunkten zweier Positionen befahren wurden. Eine grobe Schätzung der linkfeinen Geschwindigkeiten kann erfolgen mit

$$v_L = \frac{\Delta s}{\Delta t} \quad \text{Gleichung 2-7}$$

v_L ... Reisegeschwindigkeit auf sämtlichen Links zwischen zwei gemeldeten Positionen

Δs ... zurückgelegte Strecke zwischen zwei Positionen

Δt ... verstrichene Zeit zwischen zwei Positionsmeldungen

Damit berechnet sich die linkfeine Reisezeit zu

$$t_L = \frac{l_L}{v} = \frac{\Delta t}{\Delta s} \cdot l_L \quad \text{Gleichung 2-8}$$

t_L ... Reisezeit auf Link L

l_L ... Länge des Links L

Häufig enthält das Kartenmaterial zu jedem Link die Information über die freie Geschwindigkeit, die sich bei ungestörtem Verkehr als mittlere Geschwindigkeit einstellt. So kann die freie Reisezeit zwischen zwei Positionen verbessert geschätzt werden zu

$$t_{0,R} = \sum_{L=1}^n \frac{a_L \cdot l_L}{v_{o,L}} \quad \text{Gleichung 2-9}$$

- $t_{0,R}$... freie Reisezeit zwischen zwei Positionen
 a_L ... Anteil von Link L, der zwischen zwei Positionen befahren wurde
 $v_{0,L}$... freie Geschwindigkeit auf Link L
 n ... Anzahl der Links, die zwischen zwei Positionen befahren wurden

Diese freie Reisezeit auf der Route zwischen zwei gemeldeten Positionen ermöglicht die verbesserte Berechnung der aktuellen linkbezogenen Geschwindigkeit zu

$$v_L = v_{0,L} \cdot \frac{t_{0,R}}{\Delta t} \quad \text{Gleichung 2-10}$$

und der linkbezogenen Reisezeit zu

$$t_L = \frac{s_L}{v_L} = \frac{s_L \cdot \Delta t}{v_{0,L} \cdot t_{0,R}} \quad \text{Gleichung 2-11}$$

Liegen in einem Zeitschritt mehrere Linkreisezeiten aus Positionsmeldungen für einen Link vor, weil in diesem Zeitschritt mehrere Positionen meldende Fahrzeuge diesen Link befahren, so müssen die unterschiedlichen Linkreisezeiten zu einem Wert vermittelt werden. Im Fall von Mikrosimulationsdaten kann der Mittelwert verwendet werden, da hier alle Fahrer-Fahrzeug-Einheiten regulär und kontinuierlich am Verkehr teilnehmen und nicht nebenbei telefonieren, trödeln, einen Parkplatz suchen, zwischendurch kleine Einkäufe tätigen oder Mitfahrer ein- bzw. aussteigen lassen usw., wie es im realen, insbesondere innerörtlichen Verkehr der Fall ist. Geringe Linkgeschwindigkeiten aus Positionsmeldungen auf Autobahnen sind daher zuverlässiger als solche aus innerörtlichen Netzausschnitten.

Im Fall von Realdaten sind Linkreisezeiten aus einer Fahrt nicht zu berücksichtigen, wenn die Fahrt für die Verkehrsverhältnisse im befahrenen Netzausschnitt nicht repräsentativ ist. Dies ist zunächst für unangemessen langsame Fahrzeuge der Fall. Gegen die Berücksichtigung ausschließlich des schnellsten Fahrzeugs spricht, dass auf einem Link auf einem Abbiegefahrstreifen ein Rückstau vorliegen kann, während in Richtung Geradeaus hohe Geschwindigkeiten vorherrschen können, so dass der Rückstau nicht abgebildet würde, obwohl eventuell sogar mehrere Fahrzeuge im Rückstau diesen dokumentieren würden. Abhilfe schaffen würde eine Betrachtung des Abbiegefahrstreifens als separaten Link, was jedoch vor allem bei wenigen Fahrzeugen im Netz, die Positionen melden, die Netzabdeckung weiter verringert.

Im Fall dieser Arbeit konnten nicht repräsentative Fahrausschnitte durch geeignete Filterkriterien verworfen werden. Aufgrund der geringen Verfügbarkeit der vorliegenden Daten (Kapitel 5.1) genügt die Mittelung räumlich-zeitlich redundanter Linkinformationen.

3. Verkehrslageschätzung – Stand der Technik

Bestehende Arbeiten, Ansätze und Verfahren, die mit der netzweiten Verkehrslageschätzung eine ähnliche Zielsetzung verfolgen wie die vorliegende Arbeit, firmieren häufig unter dem Begriff der Datenfusion. Hierzu heißt es in FGSV [2003]: „Die Datenfusion umschreibt einen Vorgang, bei dem Signale bzw. Daten aus mehreren Sensoren automatisch detektiert, assoziiert, kombiniert und geschätzt werden. Es wird hierbei das Ziel verfolgt, die reale Welt mit höherer Genauigkeit und Zuverlässigkeit als mit einem Sensor allein identifizieren zu können.“

In der Literatur wird in manchen Fällen bereits die Zusammenführung der Daten aus benachbarten gleichartigen Verkehrsdetektoren als Verkehrsdatenfusion beschrieben. Meist beschreibt der Begriff jedoch die Zusammenführung von Daten aus verschiedenartigen Informationsquellen in einem netzweiten Kontext. Für diesen Ansatz der Datenfusion steht der Begriff Multi-Sensor-Datenfusion.

Gemäß SAAKE ET AL. [2000] wird der „[...] Begriff der Datenfusion [...] gegenwärtig in verschiedenen Anwendungsgebieten in teilweise unterschiedlicher Bedeutung verwendet. Grundsätzlich wird darunter jedoch die Kombination von Daten aus verschiedenen (heterogenen) Quellen verstanden. Konkrete Problemstellungen sind u.a.: Multi-Sensor Fusion, die sich mit der Verbindung von Messdaten verschiedener Geräte befasst, Multiple Source Integration mit dem Ziel der Zusammenführung von Informationen aus Daten- oder Wissensbanken und Image Fusion als Kombination von Bildern einer Szene mit dem Ziel einer verbesserten Interpretation.“ Das später beschriebene und bewertete Verfahren ist demzufolge als Multi-Sensor-Datenfusion einzuordnen. Der Ansatz der Multi-Sensor-Datenfusion wirft zahlreiche Fragen auf wie beispielsweise nach der Art und Weise der Zusammenführung von Daten, die unterschiedliche Strukturen aufweisen, die unterschiedliche Messgrößen enthalten, die in unterschiedlichen, nicht zwangsläufig regelmäßigen Intervallen vorliegen und die sich hinsichtlich Güte, Aktualität sowie räumlicher und zeitlicher Verfügbarkeit unterscheiden. So ist in Betracht zu ziehen, dass eine mögliche Gewichtung unterschiedlicher Eingangsdaten bei der Datenfusion eine wichtige Rolle spielen kann.

Aufgrund des vielfältigen Begriffsverständnisses wird der Begriff der Datenfusion im weiteren Verlauf der Arbeit vermieden.

VORTISCH UND ORTGIESE [2005] weisen darauf hin, dass für einen Link gleichzeitig lokale Messdaten und fahrzeuggenierte Daten vorliegen können, die dann in den seltensten Fällen völlig übereinstimmen, sich gelegentlich sogar widersprechen. Daraus wird die Notwendigkeit eines Zuverlässigkeitswerts abgeleitet, der jedem erhobenen Datum zugeordnet wird. Als Beispiel wird angeführt, dass eine fahrzeugseitig detektierte Information zuverlässiger wird, wenn sie von mehreren Fahrzeugen sehr ähnlich erhoben wird, eine lokale Messung

jedoch verliert an Zuverlässigkeit mit zunehmender Entfernung vom Detektor. Bei der Zusammenführung verschiedener Daten ist es meist nicht sinnvoll, einen Mittelwert zu bilden, sondern den Wert zu verwenden, der mit einer höheren Zuverlässigkeit vorliegt.

LISTL UND DAMMANN [2009] stellen für Frankfurt fest, dass eine Flotte von 550 Taxis, die Positionsmeldungen zur Verfügung stellen, bei weitem nicht für eine netzabdeckende Verkehrslageermittlung ausreicht. Hierzu müsste die Flotte überschlägig auf 3.500 Fahrzeuge erweitert werden. So wird die infrastrukturseitige Datenerhebung vorerst unerlässlich bleiben, wobei Informationen aus Positionsmeldungen eine wichtige Rolle bei der Verbesserung der Verkehrsdatenlage zugeschrieben wird.

WAGNER [2007] stellt für Taxiflotten ebenfalls fest, dass sie einen wichtigen Beitrag zur Verkehrslageerfassung leisten können, aber als alleinige Informationsquelle für Anwendungen des Verkehrsmanagements nicht ausreichen.

Die Verfahren zur Schätzung von Verkehrskenngrößen, die nachfolgend genannt und in ihren Grundzügen beschrieben werden, können danach unterschieden werden, ob sie auf ein Umlegungs- bzw. Verkehrsflussmodell zurückgreifen oder nicht. Nicht jedes Verfahren greift ausschließlich auf eine dieser beiden Varianten zurück. Im Fall kombinierter Ansätze wird die Zuordnung letztlich nach dem Gewicht eingesetzter Teilverfahren vorgenommen.

3.1 Ansätze mit Umlegungs- oder Verkehrsflussmodell

3.1.1 Makroskopische Ansätze

Makroskopische Verkehrsflussmodelle beschreiben die Eigenschaften von Verkehrsströmen auf Strecken. Meist basieren sie auf physikalischen Modellen. Nach SCHNABEL UND LOHSE [1997] ist die Adaption der Kontinuumstheorie auf den Straßenverkehr nach LIGHTHILL UND WHITHAM [1955] die bekannteste Anwendung. Das hierbei ebenfalls diskutierte Fundamentaldiagramm wird in CREMER [1979] zur Modellierung des Verkehrsflusses auf mehrstreifigen Fahrbahnen mit Anschlussstellen und damit auch zur Beschreibung der Entstehung einer Störung des Verkehrsflusses herangezogen. Es wird festgestellt, dass „... dies ein entscheidendes Charakteristikum [ist], das den Verkehrsfluss von anderen Verteilprozessen mit beschränkter Transportkapazität unterscheidet: wenn einmal ein gewisser Grenzwert der Verkehrsdichte überschritten wird, kommt es zum Zusammenbruch des Prozesses, wobei die dann sich einstellende Verkehrsstärke weit unter der maximal möglichen zu liegen kommt. Eine Wiederherstellung der vollen Leistungsfähigkeit der Straße erfordert dann eine längere Erholungsphase.“

Das Cell Transmission Model nach DAGANZO [1993, 1994, 1995] setzt einen räumlich-zeitlich diskretisierten Ansatz basierend auf der Kontinuumstheorie um. Es zielt gemäß FRIEDRICH UND ALMASRI [2005] auf die Analyse und Prognose von „... dynamische[n] Vorgänge[n] wie

den Aufbau, das Wandern und den Abbau von Dichtewellen bzw. Stauungen“ und „... setzt eine vereinfachte Version des Fundamentaldiagramms in Form eines Trapez ein“. Dazu wird in regelmäßigen Zeitschritten abhängig von den Verkehrsverhältnissen in je zwei benachbarten Streckenabschnitten, den Zellen, eine bestimmte Anzahl von Fahrzeugen von der einen in die nächste Zelle versetzt. Zur Darstellung der Verkehrsabläufe an Knotenpunkten müssen Abbiegeanteile beispielsweise aus einer Umlegung angegeben werden. KIM [2002] erweitert das Cell Transmission Model um eine Klassifizierung der Zellen zur Vermeidung negativ oder unlogisch geschätzter Verkehrsdichten und um eine Dynamisierung des vereinfachten Fundamentaldiagramms.

Wie im Verfahren ESE (Erweiterte Situationserkennung) nach BUSCH ET AL. [1994] werden im Verfahren AZTEK (Automatische zeitlich-räumliche Störungserkennung) nach MEIER UND WEHLAN [2000] lokale Messwerte von je zwei benachbarten Querschnitten mit den Schätzwerten eines makroskopischen Verkehrsflussmodells abgeglichen. Treten Unterschiede auf, wird im Modell zwischen den Erhebungsquerschnitten ein hypothetischer Verkehrsfluss ein- und wieder ausgeleitet, der die Anpassung der Modellwerte an die Messergebnisse bewirkt. Zur Dimensionierung des hypothetischen Flusses wird ein Kalman-Filter (Kapitel 3.2.2) eingesetzt. Der Ansatz ist insbesondere für den Einsatz auf Autobahnen ausgerichtet. SCHÖBER [2009] erweitert AZTEK um topologische Aspekte wie Längsneigung und Kurvigkeit und um eine dynamische Kompensation von Zu- und Abflüssen an nicht-detektierten Anschlussstellen. Zudem zielt er mit der Einbindung von Beschränkungen auf eine Steigerung der Robustheit des Verfahrens.

MATSCHKE ET AL. [2004] beschreiben ein Verfahren, das aus Detektoren in Lichtsignalanlagen und Signalzeitenplänen Informationen zu Abbiegeanteilen und Rückstaulängen, Reisezeitverlusten sowie Verkehrsstärken ermittelt. Die Fahrzeugströme werden auf nachfolgende Links propagiert. Diese Zwischenergebnisse gehen mit einer darauf basierenden ersten Schätzung einer Quelle-Ziel-Matrix in eine Verkehrsumlegung ein, die eine verbesserte Quelle-Ziel-Matrix liefert. Reisezeiten aus Positionsmeldungen können zur Kalibrierung ermittelter Reisezeiten oder der Quelle-Ziel-Matrix eingesetzt werden.

HENNINGER [2001] stellt ein Verfahren zur Schätzung von Verkehrsdichten, Abbiegeanteilen und Störungen aus lokalen Daten und LSA-Steuerungsdaten vor, die mit einem makroskopischen Modell gekoppelt werden. Die Messdaten werden ca. 100 bis 150m vor Lichtsignalanlagen erhoben und in 3s-Intervallen aggregiert. Die Ergebnisse können zur Optimierung der LSA-Steuerung sowie zu Verkehrsprognose und -information verwendet werden.

FRIEDRICH UND VORTISCH [2005] unterscheiden Ansätze der dynamischen Verkehrsumlegung nach Verfahren mit wiederkehrendem Zustand – z.B. werktäglich oder wöchentlich – , mit veränderlichem Systemzustand – Verkehrsnachfrage und -angebot beeinflussen sich gegenseitig im Laufe der Zeit – und nach Spontanumlegungen. Letztgenannte ermöglichen die Berücksichtigung spontaner Ereignisse wie Unfälle oder Sperrungen durch eine Anpassung

des Verkehrsangebots, die sich im Fall der Verkehrsteilnehmerinformation in einer veränderten Nachfrage niederschlagen kann. Unter den drei genannten prinzipiellen Ansätzen ist dies der einzige, der sich für den Einsatz in Verkehrsleitzentralen eignet.

Der Path Flow Estimator (PFE) nach BELL ET AL. [1997] ist ein Verfahren zur Schätzung von linkbezogenen Verkehrsstärken und Reisezeiten, die zur verkehrsabhängigen Steuerung von Lichtsignalanlagen und zur dynamischen Routenführung eingesetzt werden können. Es kombiniert Daten aus einer Verkehrsumlegung mit lokalen Online-Daten. Zur Umlegung wird ein Logit-Modell eingesetzt. Aus den Unterschieden zwischen Umlegungsergebnis und lokalen Daten werden Faktoren ermittelt, mit denen die Umlegung mit der Detektion in Übereinstimmung gebracht wird. Für Links ohne lokale Datenerhebung wird die Reisezeit aus der Bildung eines Gleichgewichts zwischen Angebot und Nachfrage berechnet.

Der Dynamische Netzwerkmonitor (DINO) nach LOGI UND ULLRICH [2001] (Bewertung in [SCHENDZIELORZ, 2003]) verwendet zur Anpassung der Umlegungsergebnisse an aktuelle Messwerte das endogene Verfahren zur Schätzung von Verkehrsbeziehungen nach VAN ZUYLEN [1980]. Durch die iterative Bestimmung von Ausgleichsfaktoren wird die Nachfragematrix angepasst, die als Grundlage für die Umlegung herangezogen wird. Die dabei ermittelten linkbezogenen Verkehrsstärken können zur Ableitung von Verkehrsdichten, mittleren Geschwindigkeiten, Reisezeiten, Warteschlangenlängen oder zur Bestimmung von Qualitätsstufen des Verkehrsablaufs herangezogen werden. DINO gilt als Weiterentwicklung des Netzwerkmonitors (NEMO) nach HOOPS ET AL. [1999] um eine iterative Umlegungsberechnung und um die Berücksichtigung einer zeitlichen Veränderung von streckenbezogenen Widerständen.

VORTISCH [2006] kombiniert mehrere Teilverfahren: In einem Vorverarbeitungsschritt werden für Tagesklassen aus historischen Messwerten und Matrizen stundenfeine Nachfragematrizen erstellt und abgelegt. Die Online-Anwendung wählt unter Berücksichtigung aktueller lokaler Messwerte in definierten Zeitschritten eine geeignete Nachfragematrix aus, die nachfolgend auf das Netz umgelegt wird. Die hierbei ermittelten Ergebnisse für die Routenwahl werden zur Propagierung aktueller lokaler Messwerte durchs Netz herangezogen. Unter zusätzlicher Verwendung von tagesstypischen Ganglinien lokaler Messwerte ermöglicht das Propagierungsverfahren eine kurzfristige Verkehrsprognose für das gesamte Netz.

3.1.2 Mikroskopische Ansätze

Neben makroskopischen Verkehrsmodellen, die Betrachtungen für Fahrzeugmengen und damit für Verkehrsströme anstellen, können auch solche zur Verkehrslageschätzung eingesetzt werden, die das Verhalten einzelner Fahrzeuge nachbilden. Diese mikroskopischen Modellansätze greifen auf die Fahrzeugfolgetheorie zurück (siehe z.B. [WIEDEMANN, 1974], kompakte Beschreibung auch in [MAIER, 2004]) oder verwenden zellulare Automaten ([NAGEL

UND SCHRECKENBERG, 1992], [GÜRTLER, 2005]), die aufgrund höherer Berechnungsgeschwindigkeiten Vorteile für eine Online-Anwendung aufweisen.

Im Fall des zellularen Automaten werden die Strecken des Straßennetzes in Zellen aufgeteilt, so dass eine solche Zelle im Fall eines Staus genau ein Fahrzeug aufnehmen kann. Eine Zelle kann also die Zustände „leer“ und „belegt“ annehmen. Die Geschwindigkeit der Fahrzeuge wird in ganzzahligen Anzahlen von Zellen pro Zeitschritt modelliert. Beschleunigungs- und Verzögerungsvorgänge werden entsprechend diskretisiert dargestellt. Von Zeitschritt zu Zeitschritt bewegt sich also ein Fahrzeug im Netz um eine entsprechende Anzahl von Zellen nach vorn. Zur zeitlichen Diskretisierung üblich sind 1s-Intervalle. Sie ermöglichen eine zweckmäßige Darstellung bis hin zu Geschwindigkeiten, wie sie auf Autobahnen üblich sind. Für sinkende Geschwindigkeiten lässt die Darstellungsgenauigkeit nach.

Ein Fahrzeug im Netz weist die Eigenschaften Beschleunigungs- und Bremsvermögen sowie Wunschgeschwindigkeit auf. Ein Fahrzeug beschleunigt bis zum Erreichen seiner Wunschgeschwindigkeit. Sind vor dem Fahrzeug jedoch nicht ausreichend viele Zellen unbelegt, um die aktuelle Geschwindigkeit aufrecht zu erhalten, so wird sie reduziert. Durch eine Randomisierung werden zufällige Geschwindigkeitsschwankungen nachgebildet, die beispielsweise durch spontane Änderungen im Fahrverhalten hervorgerufen werden.

Das psycho-physische Wahrnehmungsmodell nach WIEDEMANN [1974] teilt einer Fahrer-Fahrzeug-Einheit eine Wunschgeschwindigkeit zu, die im Falle unbeeinflussten Fahrens auch erreicht wird. Treten vorübergehend Beeinflussungen durch vorausfahrende Verkehrsteilnehmer ein, wird die Geschwindigkeit ebenso vorübergehend reduziert. Zu deren Quantifizierung wird eine Fahrzeugfolgegleichung nach GAZIS ET AL. [1961] verwendet, in die unter anderem der Abstand und der Geschwindigkeitsunterschied zum Vorderfahrzeug sowie die Reaktionszeit und ein Empfindlichkeitsfaktor als Parameter eingehen. Der Einsatz von Wahrnehmungsschwellen führt dazu, dass das Pendeln des Folgefahrzeugs um einen mittleren Abstand zum Vorderfahrzeug nachgebildet wird. Das psycho-physische Wahrnehmungsmodell sieht ebenso wie das Modell nach GIPPS [1981] neben der Wunschgeschwindigkeit Beschleunigungs- und Verzögerungseigenschaften als Parameter für eine Fahrer-Fahrzeug-Einheit vor.

Makroskopische Kenngrößen können aus mikroskopischen Verkehrsmodellen durch Aggregation entsprechend Kapitel 2 abgeleitet werden. Im Fall einer Online-Anwendung können real erhobene Detektorwerte zur Kalibrierung des Modells herangezogen werden. Mikroskopische Modelle liefern keine Routeninformationen. Zur Darstellung von Verkehrsabläufen in komplexeren Netzen müssen sie folglich mit Informationen aus makroskopischen Ansätzen angereichert werden.

3.2 Ansätze ohne Umlegungs- oder Verkehrsflussmodell

Ansätze, die ohne Umlegungs- oder Verkehrsflussmodell auskommen, greifen auf Erfahrungswerte zurück, die aus historischen makroskopischen Daten abgeleitet werden können.

3.2.1 Ganglinienverfahren und Zeitreihenanalysen

Ganglinienverfahren gehören zu den einfachsten Methoden der Verkehrsschätzung basierend auf historischen Daten. Sie beschreiben den typischen Verlauf verkehrlicher Kenngrößen in der Regel über einen Tag für verschiedene Tagestypen. Solche Tagestypen können beispielsweise anhand von Kalendermerkmalen oder durch eine Clusteranalyse festgelegt werden. Gemäß PINKOFSKY [2006] werden Ganglinienverfahren im Rahmen der Verkehrssteuerung aufgrund des zu geringen Detaillierungsgrads nicht eingesetzt. Für größere Prognosehorizonte gewinnen sie jedoch an Bedeutung und werden somit auch bei der Planung verkehrstechnischer Anwendungen verwendet. Zudem kommen sie als Bestandteil musterbasierter Ansätze zum Einsatz.

JANKO [1994] prognostiziert Verkehrskenngrößen aus jungen historischen Daten mit Hilfe einer mehrfachen exponentiellen Glättung und vergleicht die Ergebnisse mit anderen Verfahren der Zeitreihenanalyse. Hierzu bietet auch SPANGLER [2009] einen Überblick. Er entwickelt einen zeitreihenbasierten Filter zum Verwerfen einzelner unplausibler gemessener Routenreisezeiten. Ebenso zeitreihenbasiert ermittelt er eine Kurzfristprognose für Routenreisezeiten, die zur Störungserkennung herangezogen wird.

Ganglinienverfahren und Zeitreihenanalyse zielen insbesondere auf die Prognose von Daten, die bis zum gegenwärtigen Zeitpunkt durchgängig erhoben wurden. Somit verfolgen sie eine Fragestellung, die von der vorliegenden deutlich abweicht.

3.2.2 Regressions- und musterbasierte Ansätze

Regressions- und musterbasierte Ansätze verwenden in der Vergangenheit beobachtete Zusammenhänge zur Schätzung von Verkehrskenngrößen aus aktuell verfügbaren Daten.

In FGSV [2003] wird es zur Zusammenführung unterschiedlich strukturierter Daten als sinnvoll erachtet, „... alle Messdaten in eine merkmalsbasierte Beschreibung zu überführen, indem die durch Messdaten repräsentierten Merkmale bzw. Merkmalsvektoren als Punkte in einem Merkmalsraum abgebildet werden. Über den euklidischen Abstand der einzelnen Punkte im Merkmalsraum kann ein Ähnlichkeitsmaß bestimmt werden, welches eine Klasseneinteilung durch räumliches Aufteilen des Merkmalsraumes in Gebiete ermöglicht. Anders ausgedrückt, es wird versucht, ‚Punktwolken‘ zu identifizieren, anhand derer eine Aufteilung des Merkmalsraumes in verschiedene Gebiete sinnvoll erscheint. Als Klassen sollen

hier die zu unterscheidenden Verkehrszustände ‚gestauter‘, ‚zähfließender‘, ‚dichter‘ und ‚freier Verkehr‘ gelten.“

BERMANN UND BUSCH [2006] beschreiben ein Verfahren zur Online-Verkehrslageschätzung, das aktuelle lokale Verkehrsdaten und Informationen aus Positionsmeldungen einer Taxiflotte berücksichtigt und mit historischen Daten aus denselben Quellen koppelt. Im ersten Schritt wird auf Basis lokaler Daten für sämtliche Erhebungsquerschnitte ein lokaler Level of Service (LOS) ermittelt, der Grundlage für die Ermittlung eines linkbezogenen LOS ist. Wo Positionsmeldungen vorliegen, werden die Informationen zur Bestimmung eines linkbezogenen LOS herangezogen. Die netzweiten LOS aus lokalen und aus fahrzeuggenerierten Daten werden zusammengeführt und zum Abgleich mit historischen Situationen genutzt. Nachfolgend werden die ähnlichsten historischen Verkehrsszenarien zur Vervollständigung der aktuellen Verkehrslageschätzung herangezogen. Als Endergebnis liefert das Verfahren eine netzweite linkfeine LOS-Bestimmung.

In GÖRZ ET AL. [2000] werden instanzbasierte Lernverfahren als wiederaufgenommene und weiterentwickelte „Nächste Nachbarn“-Verfahren bezeichnet. Das Lernen besteht zunächst nur aus dem Ablegen historischer Informationen. Die Verfahren beginnen erst zu arbeiten, wenn es etwas zu schätzen gibt, weshalb sie auch als „lazy learners“ bezeichnet werden. Durch eine Klassifizierung der abgelegten Informationen können instanzbasierte Lernverfahren an „Faulheit“ verlieren. Für Verkehrsdaten kann eine solche Klassifizierung beispielsweise nach Kalenderinformationen erfolgen.

LEONHARDT [2008, 2010] beschreibt ein instanzbasiertes Lernverfahren zur Schätzung von lokalen Daten und Reisezeiten aus lokalen Daten und Kalenderinformationen. Aktuell verfügbare Daten werden mit historischen räumlich-zeitlichen Verkehrsmustern abgeglichen. Eine Maßzahl für die Abstände zwischen historischen und dem aktuellen Muster ermöglicht die Auswahl historischer Muster, die zur Datenvervollständigung und zur Prognose der aktuellen Situation herangezogen werden. In diese Maßzahl gehen die Unähnlichkeit verkehrlicher Kenngrößen, ein Distanzmaß aufgrund unterschiedlicher Kalendermerkmale und ein Strafterm für gegenwärtig nicht vorliegende Messwerte ein. Die Schätzung selbst wird anhand der Verkehrsstärkeschätzung aus lokalen mittleren Geschwindigkeiten erläutert: Zu einem aktuellen Geschwindigkeitsmesswert wird eine bestimmte Anzahl historischer Messwerte ausgewählt, die die größte Ähnlichkeit zum aktuellen Wert aufweisen. Die zugehörigen historischen Verkehrsstärkemesswerte liefern einfach oder gewichtet gemittelt den Schätzwert für die aktuelle Verkehrsstärke. Das Verfahren führt zu besseren Ergebnissen als Ganglinienverfahren oder die Trivialprognose „Alles bleibt so, wie es gerade noch war“.

MIN ET AL. [2007] passen zur Schätzung und Prognose von Verkehrskenngrößen ein multivariates räumlich-zeitliches Autoregressionsmodell an. Das Verfahren wird ausschließlich mit lokalen Daten getestet.

THOMAS [1998] wählt zur Störungserkennung auf Hauptstraßen einen multiplen Regressionsansatz. Zur Auswahl historischer Störungsmuster bewertet er deren Ähnlichkeit mit Informationen zur aktuellen Situation in Anlehnung an das Bayestheorem. Er wendet das Verfahren auf lokale Daten und Informationen aus Positionsmeldungen an.

KATES ET AL. [2005] stellen fest, dass sich Verfahren zur Verkehrsprognose hinsichtlich ihres Aktualisierungsintervalls, ihres Prognosehorizonts, ihrer Modellgüte, ihres Verhaltens bei fehlenden oder unplausiblen Daten sowie ihrer Aussage zur eigenen Unschärfe unterscheiden können. Eine geringe Unschärfe stellt hierbei ein Qualitätsmerkmal dar. Liegen für eine Kenngröße verschiedene Schätzungen aus verschiedenen Verfahren vor, ist die Aufgabe „... der Datenfusion [...], eine Aussage zu erzeugen, die gemäß einer statistisch fundierten Modellvorstellung mit einer geringen Ungenauigkeit behaftet ist“. Grundgedanke des entwickelten Lösungsansatzes „... ist die Gewichtung verschiedener unabhängiger Informationen mit einem Indikator der Verlässlichkeit dieser Information. [Diese erfolgt] durch die Bildung des mit der reziproken Varianz gewichteten Mittelwertes aus unabhängigen Stichproben einer Zufallsgröße“. Unterschiedlichen Aktualisierungsintervallen kann mit einem Zwischenzeitschritt für das langsamer getaktete Verfahren begegnet werden.

Mehrere Arbeiten (z.B. [EID, 2001], [BÖKER, 2000]) befassen sich mit der Verwendung eines Kalman-Filters zur Zusammenführung gleich- und verschiedenartiger Verkehrsdaten. Der von Rudolf E. Kalman 1960 vorgestellte Ansatz wird in zahlreichen Varianten in verschiedenen Disziplinen eingesetzt. Kalman-Filter arbeiten zeitdiskret nach der Methode der quadratischen Fehlerminimierung und werden zur Datenvervollständigung, zur Prognose und zur Messwertglättung eingesetzt. Lineare Zusammenhänge zwischen sämtlichen denkbaren Kenngrößen werden bei der Erstellung einer Kovarianzmatrix hergestellt. Abweichungen von der linearen Näherung sowie Messfehler und Messrauschen werden normalverteilt angenommen.

Beim Erweiterten Kalman-Filter (EKF) wird die Kovarianzmatrix durch eine mit partiellen Ableitungen befüllte Matrix (Jacobi-Matrix) ersetzt. Dadurch wird auch die Abbildung nichtlinearer Zusammenhänge ermöglicht. Durch die nichtlineare Abbildung ist aber die Annahme der Normalverteilungen nicht mehr zu halten [WELCH ET AL., 2004].

FRÖHLICH [2006], SEYDEL [2008] sowie EHMKE UND MEISEL [2008] setzen Clusterverfahren zur Verkehrsdatenanalyse ein. Verfahren der Clusteranalyse ermöglichen die automatische Klassifizierung umfangreicher Messdaten, die komplexe Strukturen aufweisen. Festzulegende Parameter sind relevante Ähnlichkeits- oder Distanzmaße, der zu verwendende Clusteralgorithmus und die Anzahl der zu bildenden Cluster (Überblick in [SEYDEL, 2008]). Typische Ganglinien abhängig von der Tageskategorie können aus einer Datenhistorie für verschiedenste Verkehrskenngrößen ermittelt werden. Somit können Links v.a. zu Planungszwecken nach typischen linkbezogenen Geschwindigkeitsganglinien aus fahrzeuggenerierten Daten

geclustert werden (vgl. [EHMKE UND MEISEL, 2008]). SEYDEL [2008] schlägt auch die Planung von Signalzeitenplänen basierend auf Erkenntnissen aus fahrzeuggenerierten Daten vor.

SEYDEL [2008] führt die Datenhistorie zusätzlich einer Hauptkomponentenanalyse zu. Durch eine Hauptachsentransformation angewandt auf den Merkmalsraum werden „... grundlegende linear unabhängige Komponenten aus dem stark korrelierten Datenmaterial [extrahiert], so dass wesentliche Charakteristika des zu Grunde liegenden Prozesses durch weniger Dimensionen als die gemessenen Größen abgeschätzt werden können“. Basierend auf diesem Ansatz wird ein Konzept zur Prognose auf Basis von Hauptkomponenten vorgestellt.

BERNARD ET AL. [2006] untersuchen Korrelationen zwischen Linkreisezeiten aus GPS-Daten, die in Intervallen von einer Sekunde im Kanton Zürich gesammelt und offline ausgewertet wurden. Dabei stellen sie fest, dass sämtliche Anwendungen zur Bestimmung oder Schätzung linkbezogener Reisezeiten oder Geschwindigkeiten auf historisch so beobachtete Zusammenhänge zurückgreifen sollten. Ansonsten verzichten die Verfahren auf die Berücksichtigung wichtiger, aussagekräftiger Information und laufen Gefahr, schlechte oder gar falsche Ergebnisse zu liefern.

Ausführliches Beispiel für ein musterbasiertes Verfahren zur Prognose der Störungsausbreitung

Neben der Kenntnis der aktuellen netzweiten Verkehrslage ist die räumlich-zeitliche Prognose Voraussetzung für die zielgerichtete und rechtzeitige Einleitung geeigneter Maßnahmen des Verkehrsmanagements.

Städtische Netze zeichnen sich durch kurze Knotenpunktabstände und häufig lichtsignalgesteuerte Knotenpunkte aus und stellen deshalb eine besondere Herausforderung dar. Ursachen hierfür sind die damit assoziierte Relativierung der Verkehrsflusstheorie und der maßgebliche Einfluss lokaler Netzeigenschaften auf die Entwicklung einer Störung.

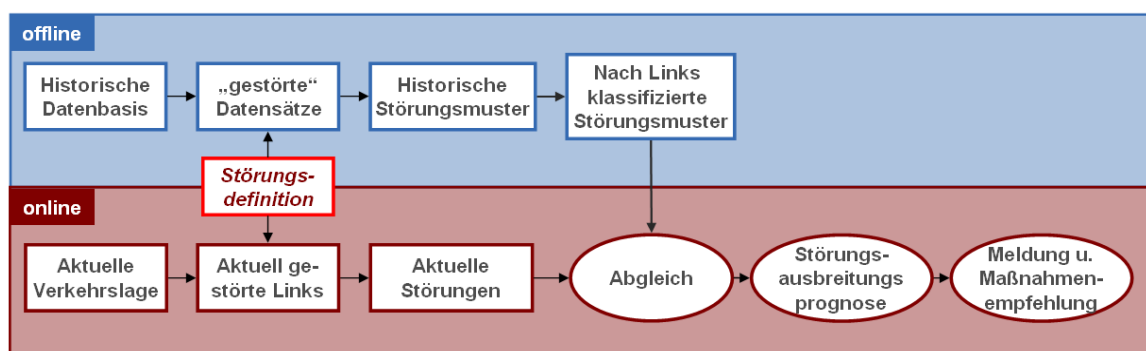


Bild 3.1 Überblick über das Verfahren zur Störungsausbreitungsprognose

Ein statistischer Ansatz zur Störungsausbreitungsprognose nach MAIER ET AL. [2008] benötigt als Basis eine Historie zeitdiskretisierter verkehrlicher Kenngrößen für sämtliche Links

eines Netzes. Er wurde mit Verkehrslagedaten der Verkehrsmanagementzentrale Berlin entwickelt und erprobt, die mit dem Messwertpropagierungsverfahren nach VORTISCH [2006] modelliert wurden. Der Ansatz ist sowohl für inner- als auch für außerörtliche Netze anwendbar (Überblick: Bild 3.1).

Unter Verwendung eines Störungsindikators und eines zugehörigen Schwellenwerts werden aus den historischen Verkehrslagedaten diejenigen Datensätze ausgelesen, die zeitintervallbezogen verkehrlich gestörte Links beschreiben. Ein geeigneter Störungsindikator ist der Quotient aus link- und zeitintervallbezogener mittlerer Reisegeschwindigkeit und linkbezogener freier Geschwindigkeit mit dem zu unterschreitenden Schwellenwert von 0,4, der Bereich zwischen 0,4 und 0,5 dient als Warnstufe. Weitere diesbezügliche Untersuchungen finden sich in BRILON UND SCHNABEL [2003] und TRB [2000]. Die ausgelesenen Datensätze werden dann zu Störungsmustern zusammengefügt, wenn die zugehörigen Links räumlich und die Zeitintervalle zeitlich benachbart sind (Bild 3.2, Rückgang des verwendeten Störungsindikators von Gelb über Rot bis Blau).

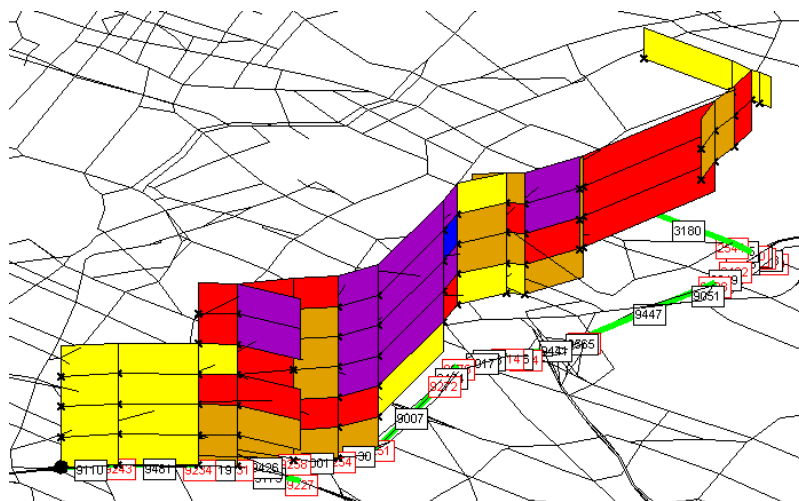


Bild 3.2 Störungsmuster, Zeit von unten nach oben fortschreitend

Um bei einem aktuellen Störungsereignis schnell zur Prognose geeignete historische Störungsmuster auswählen zu können, werden jene einer Klassifizierung unterzogen. Die Klassifizierung beispielsweise nach Merkmalen wie Tageszeit, Gesamtverlustzeit oder Gesamtausdehnung wird nicht als zielführend erachtet, da der Einfluss lokaler Netzstrukturen hierbei vernachlässigt wird und beispielsweise zwei Vormittagsstörungen in unterschiedlichen Netzausschnitten keinem gemeinsamen Muster folgen müssen.

Anders für eine Klassifizierung nach Links: Zu jedem Link wird eine Klasse angelegt. Dieser Klasse werden sämtliche Muster zugeordnet, die auf dem Link in der Vergangenheit beobachtet wurden. Folglich wird eine historische Störung mehreren Klassen zugeordnet, wenn sie sich über mehrere Links erstreckt hat. Wenn im gegenwärtigen Intervall der Verkehrslageschätzung nun eine definitionsgemäße Störung vorliegt, so können durch diese Klassifi-

zierung sämtliche Störungsmuster aus der verwendeten Historie aufgerufen werden, die mindestens einen gemeinsamen Link mit der aktuellen Störung aufweisen. Jedoch eignen sich manche Muster besser als andere für eine Prognose, da sie der aktuellen Störung stärker ähneln. Daher ist es notwendig, die Ähnlichkeit zwischen der aktuellen und den historischen Störungen zu quantifizieren, um nachfolgend für die Prognose besonders geeignete historische Störungen auswählen zu können. Der diesbezügliche Musterabgleich berücksichtigt räumliche und zeitliche sowie verkehrliche Merkmale. Er teilt jeder historischen Störung einen Ähnlichkeitswert mit der aktuellen Störung zu, der für die Entscheidung über die Berücksichtigung des historischen Musters und nachfolgend über seine Gewichtung bei der Ausbreitungsprognose für die aktuelle Störung herangezogen werden kann.

Die Ausbreitungsprognose für die gegenwärtige Störung wird durch eine gewichtete Mittelung der ausgewählten historischen Störungen ermittelt. Sie gibt nach Links und nach künftigen Zeitschritten differenziert eine Störungswahrscheinlichkeit und einen verkehrsqualitativen Kennwert an. Es kann der Fall eintreten, dass einige Links von Ausbreitungsprognosen zu mehreren Störungen betroffen sind, so dass in diesen Fällen mehrere Ausbreitungsprognosen in geeigneter Weise zusammengeführt werden müssen.

Die Störungsausbreitungsprognose kann zur Generierung von Warnmeldungen eingesetzt werden, die die frühzeitige Einleitung von Maßnahmen des Verkehrsmanagements zur Reduzierung negativer verkehrlicher Wirkungen ermöglichen. Prognosebasierte Meldungen sollen Netzausschnitte ausweisen, in denen sich die größten Veränderungen hinsichtlich des Verkehrsflusses gegenüber der aktuellen Verkehrslage ergeben werden. So wird für jede Störung die prognostizierte verkehrliche Änderung unter Verwendung der Linklängen, der Störungswahrscheinlichkeiten und der Veränderung der verkehrsqualitativen Kennwerte quantifiziert. Somit liegt ein Kriterium für die Auswahl von Störungen vor, die einem Operator eventuell zusammen mit Maßnahmenempfehlungen angezeigt werden sollen.

3.2.3 Fuzzy-Logik und Neuronale Netze

Fuzzy-Logik und Neuronale Netze zählen zu den Formen künstlicher Intelligenz. Die zugehörigen Verfahren versuchen, menschliche Denkweisen und Lernmethoden zu imitieren.

Bei Fuzzy-Systemen erlauben sprachlich interpretierte Messgrößen durch den Einsatz von Wenn-Dann-Regeln Rückschlüsse auf andere Kenngrößen [OFFERMANN, 2001]. Hierbei können mehrere Bedingungen durch „und“, „oder“ oder „nicht“ verknüpft werden. Dieses Vorgehen ähnelt der menschlichen Denkweise und führt zu vergleichsweise transparenten Verfahrensabläufen. Fuzzy-Systeme ermöglichen gemäß ZIMMERMANN [1999] die Modellierung von Unsicherheiten, die Reduktion der Verfahrenskomplexität und eine Relaxierung im Sinne des Ergänzens und Ersetzens ausschließlich scharfer mathematischer Regeln durch Elemente der unscharfen Entscheidungsfindung. Hierbei kommt die unscharfe Klassifizierung von Elementen nach ZADEH [1965] zum Einsatz, die den Elementen Zugehörigkeitsgrade

zwischen 0 und 1 zu verschiedenen Klassen zuweist. OFFERMANN [2001] bietet einen Überblick über Anwendungen von Fuzzy-Systemen in der Verkehrstechnik und hält fest, dass durch die Notwendigkeit der händischen Regelerstellung der Aufwand für die Erstellung eines Fuzzy-Systems erheblich ist.

„Fuzzy-Mengen werden zur Repräsentation vager Daten oder unscharfer Konzepte meist auf einer rein intuitiven Basis benutzt. [...] Eine Fuzzy-Menge [...] wird durch die Angabe der Zugehörigkeitsgrade [...] für jedes [Element der Menge] festgelegt.“ [BORGELT ET AL., 2003]

EID [2001] entwickelt ein Verfahren zur Bestimmung des Verkehrszustands in einem Hauptstraßennetz mit recht geringer Detektionsinfrastruktur basierend auf einem Fuzzy-Logik-Ansatz. Als Eingangsgrößen setzt er lokale mittlere Geschwindigkeiten und Belegungsgrade ein.

BUSCH UND GHIO [1994] führen lokale Messgrößen von je zwei benachbarten Messquerschnitten einem Fuzzy-Regler zur schnellen Störfallerkennung auf Autobahnen zu. In einer ersten Stufe wird die Wahrscheinlichkeit für eine Pulkbildung ermittelt, nach zwei weiteren Stufen wird die Wahrscheinlichkeit für eine Störung quantifiziert.

Fuzzy-Systeme bilden die menschliche Entscheidungsfindung nach, jede Regel muss ihnen jedoch einzeln beigebracht werden. Zur Simulation menschlicher Lernprozesse werden Neuronale Netze eingesetzt, die die Weitergabe und -verarbeitung von Informationen durch eine große Anzahl von Neuronen imitieren, die parallel geschaltet und in mehreren Schichten angelegt sind [BORGELT ET AL., 2003].

„Die Neuronalen Netze bilden ein Teilgebiet der Neuroinformatik, deren Gegenstand die Erforschung biologischer Informationsverarbeitung mit den Methoden der Informatik und Informationstechnologie ist. [...] Der Zweck biologischer Informationsverarbeitung ist es, Verhalten (und interne Regulationsprozesse) in Abhängigkeit von der jeweiligen Umweltsituation so zu organisieren, dass das Lebewesen sich behauptet.“ [GÖRZ ET AL., 2000]

Die Schichten aus Neuronen sind von einer Eingangsdatenschicht und einer Ausgangsdatenschicht umgeben. Wie die Neuronen Informationen weiterleiten, erlernt das Neuronale Netz vor der Anwendung selbst unter Verwendung von Trainingsdaten.

„Wir können Neuronale Netze immer dann verwenden, wenn uns Trainingsdaten zur Verfügung stehen. [...] Andererseits lässt sich die Lösung durch das Neuronale Netz nicht interpretieren. [...] Damit lässt sich auch nicht überprüfen, ob die Lösung überhaupt plausibel ist.“ [BORGELT ET AL., 2003]

MATHIAS [1999] beschreibt ein Verkehrsumlegungsmodell auf der Basis von Rekurrenten Neuronalen Netzen. Er bildet das Straßennetz in einem Neuronalen Netz ab, Zu- und Abflüsse fungieren als Aktivierungsfunktionen in der Eingabe- und in der Ausgabeschicht. Ein

weiteres Neuronales Netz dient der Einstellung von Steuerungsparametern wie Abbiegeanteilen.

Während die Lernfähigkeit und die Fehlertoleranz zu den Vorzügen Neuronaler Netze zählen, liegen Nachteile des Ansatzes in der geringen Transparenz, im hohen Rechenaufwand und in den Schwierigkeiten beim Einbringen vorhandenen Wissens.

Die Komplementarität der Vor- und Nachteile von Fuzzy-Systemen und Neuronalen Netzen legt eine Kombination der beiden nahe, wie sie für Neuro-Fuzzy-Systeme vollzogen wird.

OFFERMANN [2001] setzt diese Idee durch die Kalibrierung eines Fuzzy-Reglers mit Hilfe eines Neuronalen Netzes um. Die Kalibrierung erfolgt offline, während bei der Online-Schätzung von Geschwindigkeiten aus lokalen Daten ausschließlich der Fuzzy-Regler zum Einsatz kommt.

LINAUER ET AL. [2006] zielen auf die Schätzung von Reisezeiten auf Routen aus lokalen Daten. Historische Reisezeiten werden hierbei aus Positionsmeldungen einer Taxiflotte gewonnen. Unter Verwendung eines Neuro-Fuzzy-Modells werden Zusammenhänge beobachtet, die zur Schätzung aktuell nicht vorliegender Reisezeiten aus vorliegenden lokalen Daten verwendet werden können. Gemessene und geschätzte Kenngrößen werden je zur Ableitung einer Qualitätsstufe des Verkehrs herangezogen, die wiederum zu einem Wert zusammengeführt werden können.

3.3 Anforderungen an die zu entwickelnde Methode

Unter Berücksichtigung der Kenntnisse zu gesichteten Arbeiten und der verfügbaren Daten soll eine neue Methode zur Schätzung von Verkehrskenngrößen entwickelt werden. Bild 3.3 bietet einen Überblick über die gesichteten Verfahren.

Verfahren, die ein Umlegungs- oder Verkehrsflussmodell verwenden, weisen umfangreiche und damit zeitaufwendige Berechnungen auf. Zudem liegen Quelle-Ziel-Informationen häufig nicht vor, wie sie für eine Umlegung erforderlich sind. Die Übertragung von Verkehrsflussmodellen basierend auf Kenntnissen aus dem Fundamentaldiagramm auf städtische Netze ist zumindest nicht unkritisch. Darüber hinaus bleibt es eine Herausforderung, neue Datenarten zur Kalibrierung in Verfahren mit Umlegung oder Verkehrsflussmodell zu integrieren.

Ganglinienverfahren und Zeitreihenanalyse scheiden aufgrund der Eindimensionalität als alleinstehende Lösung für die vorliegende Aufgabenstellung aus.

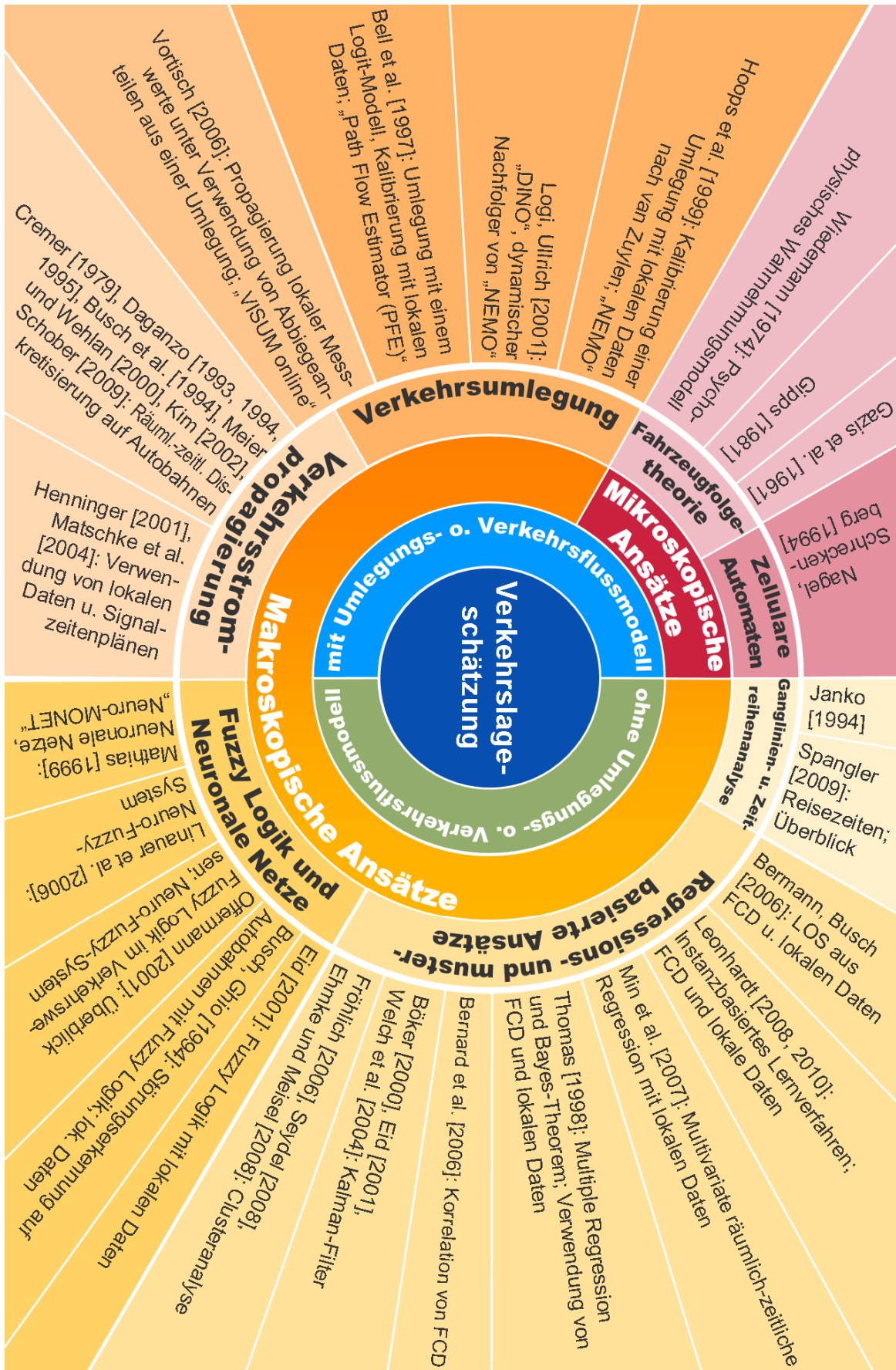


Bild 3.3 Zusammenstellung der gesichteten Verfahren zur Verkehrslageschätzung

Ansätze, die historische räumlich-zeitliche Verkehrsmuster zur Schätzung der aktuellen Verkehrslage heranziehen, haben beim Musterabgleich einen nicht zu vernachlässigenden Rechenaufwand zur Konsequenz. Darüber hinaus bleibt die Beantwortung der Frage nach der räumlich-zeitlichen Abgrenzung der Muster kritisch. Für regelmäßig wiederkehrende Ereignisse verbessern sich die Ergebnisse mit größeren Mustern, während die Schätzung weniger üblicher Ereignisse mutmaßlich mit kleineren Mustern besser gelingen sollte.

Ansätze der multivariaten Regression ermöglichen die Schätzung einer oder mehrerer Kenngrößen unter Berücksichtigung mehrerer Eingangsgrößen. Hierzu ist jedoch die Verfügbarkeit sämtlicher einbezogener Eingangsgrößen erforderlich, die nicht immer gegeben ist. Zudem kann durch die Betrachtung zu vieler Eingangswerte die Diagnose statistischer Zusammenhänge erschwert werden. Auch wird die Untersuchung, welche Eingangsgrößen eine besondere Eignung zur Schätzung anderer Kenngrößen aufweisen, beeinträchtigt.

Verfahren der künstlichen Intelligenz gehen Kompromisse zwischen Genauigkeit, geringem Aufwand für die manuelle Kalibrierung und Transparenz ein.

Abschnittsdaten, wie sie für diese Arbeit vorliegen, wurden in keinem der gesichteten Ansätze als Eingangsgröße hinzugezogen. Zudem wurden aus Positionsmeldungen – so weit das aus der Literatur korrekt hervorgeht – meist Geschwindigkeiten für längere Routen und nicht für Links abgeleitet, wie sie im Netz zwischen Knotenpunkten aufgespannt werden.

Somit ergeben sich für ein neues Verfahren zur Verkehrslageschätzung folgende Anforderungen:

- Die Methode soll einen Ansatz verfolgen, der die Verarbeitung sämtlicher verkehrlich relevanter Daten ermöglicht. Dadurch soll sie auch in der Lage sein, hier nicht betrachtete Datenarten zu verarbeiten, und soll zusätzlich nicht auf Datenarten angewiesen sein, wie sie für diese Arbeit vorliegen.
- Das Verfahren soll die Berücksichtigung von Qualitätsindikatoren ermöglichen, die zu Eingangsdaten vorliegen können.
- Die neue Methode soll auch mit lückenhaften Eingangsdaten problemlos umgehen.
- Das Verfahren soll Schätzungen zur Plausibilisierung verfügbarer Daten und insbesondere zur Vervollständigung nicht verfügbarer Daten liefern. Die Prognose von Verkehrskenngrößen spielt bei der Zielsetzung eine nachgeordnete Rolle.
- Wissen, das aus einer Datenhistorie gewonnen werden kann, soll für das zu entwickelnde Schätzverfahren genutzt werden.
- Das Verfahren soll in der Lage sein, auch für große Netze eine Verkehrslageschätzung anzubieten bei gleichzeitig geringem Aufwand für die Kalibrierung. Die große Anzahl von Schätzgrößen soll nicht dazu führen, dass die Verfahrensabläufe undurchschaubar werden.

- Das Verfahren soll eine schnelle Verkehrslageschätzung ermöglichen, um bei einer Online-Anwendung die Zeit zwischen Dateneingang und Verkehrslageschätzung kurz zu halten. Online-Rechenoperationen sind folglich in einem möglichst kompakten Rahmen zu halten. Offline-Auswertungen können jedoch großzügig mit Rechenleistung umgehen.

4. Neue Methode zur Schätzung von Verkehrskenngrößen

Unter Berücksichtigung der Vorgaben aus Kapitel 3.3 wurde eine Methode zur Schätzung von Verkehrskenngrößen entwickelt, die die netzweite Verkehrslageschätzung ermöglicht. Primäres Ziel des Verfahrens ist die Vervollständigung und Plausibilisierung aktueller Daten und weniger die Prognose zukünftiger Verkehrszustände. Die Methode wird im Rahmen dieses Kapitels beschrieben. Ihre Erprobung wird unter Verwendung von Realdaten vorgenommen, deren Qualität in Kapitel 5 behandelt wird. Die Bewertung der Schätzmethode erfolgt in den Kapiteln 6 und 7.

4.1 Konzeptioneller Ansatz

Der entwickelten Methode liegt die Annahme zugrunde, dass sich verkehrliche Muster in Netzen oder Netzausschnitten in ähnlicher Form wiederholen. Demzufolge entwickeln sich zahlreiche Verkehrskenngrößen im Netz abhängig von anderen Kenngrößen. Die Muster resultieren aus typischen Verläufen der Verkehrsnachfrage, die sich im Netz immer wieder in ähnlicher Form darstellt, aber auch aus Störungsursachen oder aus der Umfeldsituation, die u.a. durch Lufttemperatur und -feuchtigkeit, Fahrbahnzustand oder Sichtverhältnisse beschrieben wird. Verkehrskenngrößen, in denen sich diese Muster niederschlagen können, sind in Kapitel 2 beschrieben. Die neue Methode verfolgt einen statistischen Ansatz und verwendet kein Verkehrsflussmodell.



Bild 4.1 Konzeptioneller Ansatz zur Schätzung verkehrlicher Kenngrößen

Zusammenhänge zwischen Verkehrskenngrößen können unter Verwendung der Datenhistorie festgestellt und beschrieben werden. Hierfür wird ein regressionsanalytischer Ansatz gewählt. Die Daten sämtlicher Quellen sind folglich einer einheitlichen zeitlichen Diskretisierung zu unterziehen, so dass die Messwerte verschiedener Kenngrößen zeitlich einander eindeutig zugeordnet werden können. Verfügbare Informationen zur Datengüte (vgl. Kapitel 5.3) sollten mitgeführt werden, da sie bei der Ermittlung der Zusammenhänge relevant sein können. Aufgrund der Tatsache, dass sämtliche Kenngrößen auf Querschnitte, Abschnitte oder Links referenziert sind, verfügen sie grundsätzlich über einen Netzbezug. Die Zusammenhänge können durch geeignete Näherungsfunktionen beschrieben werden. Die Ermittlung dieser Näherungsfunktionen wird in Kapitel 4.2 behandelt.

Gegenwärtig verfügbare Daten und aus der Historie bekannte Zusammenhänge liefern Schätzwerte für aktuelle Daten. Aufgrund der Vielzahl von Zusammenhängen können für denselben Wert zahlreiche Schätzwerte vorliegen. Diese Werte sind in geeigneter Form zusammenzuführen. Dies kann durch Selektion von Schätzwerten oder durch ihre gewichtete oder ungewichtete Mittelung erfolgen. Auch eine Kombination dieser Werkzeuge ist denkbar. Hierbei sollte die Qualität der beobachteten Zusammenhänge berücksichtigt werden.

Die zusammengeführten Schätzwerte liegen nun zu gegenwärtig sowohl erhobenen als auch nicht erhobenen – also zu sowohl verfügbaren als auch nicht verfügbaren – Kennwerten vor. Ist der Kennwert gegenwärtig erhoben, so kann der Schätzwert zum Plausibilisieren oder zum Ersetzen des Messwerts herangezogen werden. Auch die Vermittlung beider Werte ist möglich. Kann der Kennwert aktuell nicht erhoben werden, so vervollständigt der Schätzwert die Sammlung der Messwerte. Durch die Schätzung aktuell nicht erhobener Linkreisezeiten kann so eine netzweite Verkehrslageschätzung zustande kommen.

Bild 4.1 stellt den Ansatz grafisch dar. Entsprechend den Anforderungen ist er für sämtliche weitere Datenarten (in der Abbildung „weitere Daten“) offen und ermöglicht so deren problemlose Integration. Ebenso ist er anforderungsgemäß nicht auf die Verfügbarkeit der Datenarten lokale, Abschnitts- und Positionsdaten angewiesen. Jedoch kann die netzweite Verkehrslageschätzung nur unter Verwendung der flächendeckend erhobenen Positionen und der abgeleiteten Linkreisezeiten gelingen.

Bild 4.2 zeigt eine Messwert-Matrix für das Beispielnetz gemäß Bild 4.3 und für fünf Zeitschritte. Für jeden Zeitschritt wird ein Datenvektor – eine Spalte in der Matrix – erstellt, für jede mögliche Kenngröße ist eine Zeile in der Matrix vorgesehen. Im vorliegenden Netz wird die Verkehrsstärke an Link 1 sowie infrastrukturseitig die Reisezeit auf den Links 3-4 gemessen. Linkreisezeiten aus Positionsmeldungen können für sämtliche Links vorliegen – wie in den Zeitschritten T_1 und T_2 gemäß Bild 4.2 –, fehlen aber häufig wie in den Zeitschritten T_3 , T_4 und T_5 und führen dann zu einer leeren Stelle im Vektor. Liegen in mehreren Zeitschritten zu einem Kenngrößenpaar beide Werte vor (gestrichelte Doppelpfeile in Bild 4.2), so können diese Zeitschritte zur Suche nach Zusammenhängen zwischen den Kenngrößen

herangezogen werden. Fehlt in einem Zeitschritt einer der beiden Werte, so kann der verfügbare mit dem beobachteten Zusammenhang zur Schätzung des fehlenden genutzt werden (durchgezogene Pfeile in Bild 4.2, Plausibilisierung im Beispiel nicht dargestellt).

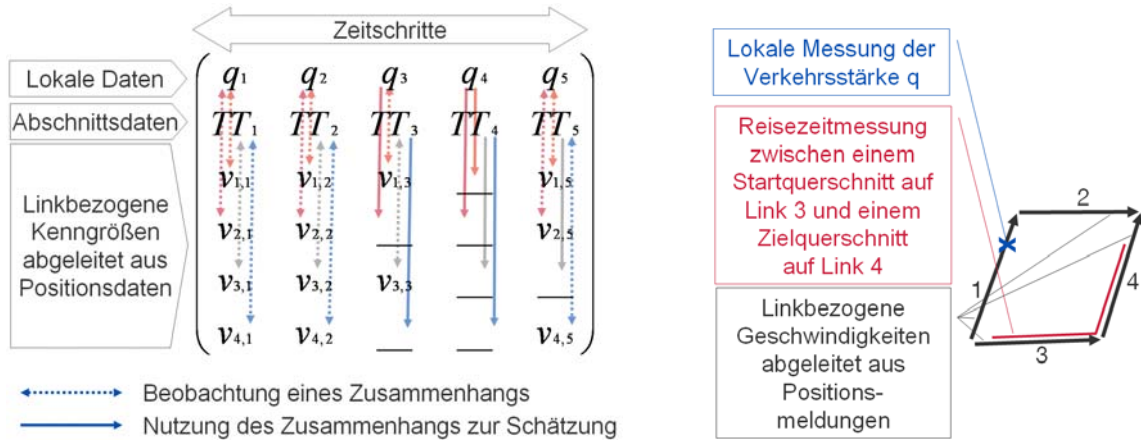


Bild 4.2 Messwertmatrix für das Netz gemäß Bild 4.3 für fünf Zeitschritte mit Schätzung fehlender Daten unter Verwendung beobachteter Zusammenhänge

Bild 4.3 Beispielnetz

Die Suche nach Zusammenhängen kann zusätzlich über einen oder mehrere Zeitschritte hinweg sowohl für dieselbe (Autoregression, horizontale Pfeile in Bild 4.4) als auch für unterschiedliche Kenngrößen (diagonale Zusammenhänge, diagonale Pfeile) erfolgen. Da bei der Ermittlung der Zusammenhänge gemäß Kapitel 4.2 auch Kenngrößen für die Qualität der Zusammenhänge ermittelt werden, erlaubt dieses Vorgehen eine Beurteilung der zeitlichen Verwendbarkeit von Messwerten für das entwickelte Verfahren. Darüber hinaus verbreitert es die Datenbasis insbesondere für die Schätzung aus zeitlich nicht immer verfügbaren Kenngrößenarten wie Routenreisezeiten oder Linkreisezeiten aus Positionsmeldungen. Die so beobachteten Zusammenhänge können auch die Kurzfristprognose der Kennwerte ermöglichen (Vektor rechts in Bild 4.4).

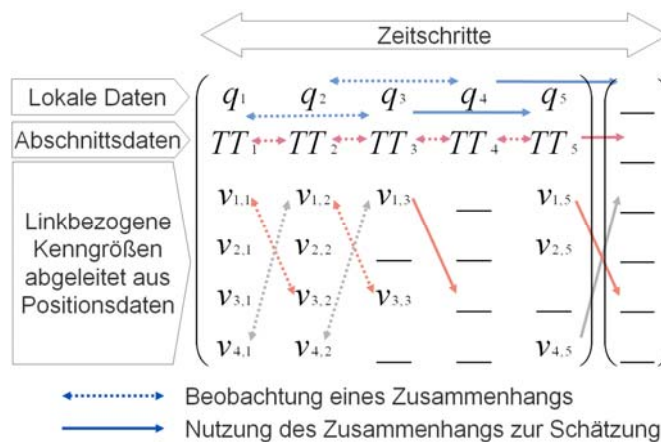


Bild 4.4 Diagonale Beobachtung von Zusammenhängen zwischen Verkehrskenngrößen und Autoregression

Sollte in der Matrix auch die räumlich-zeitliche Ausbreitung von Verkehrszuständen enthalten sein, so kann sie mithilfe der Zusammenhänge gemäß Bild 4.4 dargestellt werden.

Zusammengefasst kann der beschriebene Ansatz folgende Ergebnisse liefern:

- Gegenwärtig nicht verfügbare Daten können geschätzt werden. Werden Linkreisezeiten aus Positionsmeldungen im Verfahren eingesetzt, so kann eine netzweite Schätzung der Verkehrslage gelingen.
- Gegenwärtig verfügbare Daten können durch Schätzungen aus anderen verfügbaren Daten plausibilisiert werden.
- Temporär gemessene Daten können fortlaufend geschätzt werden.
- Zusammenhänge gemäß Bild 4.4 liefern Aussagen zur zeitlichen Verwendbarkeit von Daten und können zur Kurzfristprognose der Verkehrslage herangezogen werden.

4.2 Angepasster Regressionsansatz

Bei der Regressionsanalyse werden zwei Messgrößen daraufhin untersucht, ob zwischen ihnen ein Zusammenhang besteht. Dieser Zusammenhang wird mit einer Näherungsfunktion beschrieben. Gemäß den Anforderungen aus Abschnitt 3.3 sollen mit der zu entwickelnden Methode Zusammenhänge differenziert und robust dargestellt werden. Dabei soll auch die Möglichkeit bestehen, für einzelne Datenausschnitte die Qualität der Näherung festzustellen.

Die Überlegungen zur Entwicklung des eingesetzten angepassten Regressionsansatzes werden in den Abschnitten 4.2.1 und 4.2.2 mit frei gewählten Beispieldatensätzen erläutert. Im Abschnitt 4.2.3 werden Datensätze aus einer beliebigen kubischen Funktion sowie eine Punktwolke aus einem Messquerschnittsprotokoll verwendet, der 2806 Minutenwerte für die fahrestreifenbezogene Verkehrsdichte und die querschnittsbezogene Verkehrsstärke an einem dreistreifigen Autobahnquerschnitt zugrunde liegen.

4.2.1 Einfache lineare Regression

Bei der linearen Regression wird ein linearer Zusammenhang zwischen zwei Größen x und y unterstellt. Dabei wird ein normalverteilter Zufallsfehler der beiden Größen angenommen. Ziel der linearen Regression ist die Bestimmung einer Näherungs- oder Ausgleichsgeraden, zu der die Gesamtheit aller Abweichungen der Wertepaare minimal ist.

Eine Gerade, die die Größe y in Abhängigkeit von der Größe x darstellt, wird allgemein beschrieben durch

$$y = m \cdot x + c$$

Gleichung 4-1

m ... Steigung

c ... Achsenabschnitt

Für sämtliche Wertepaare $(x_i|y_i)$ aus Messungen kann der Fehler, der durch die lineare Näherung eingegangen wird, bestimmt werden zu

$$\varepsilon_i = m \cdot x_i + c - y_i \tag{Gleichung 4-2}$$

$\varepsilon_i \dots$ Abweichung des Messwerts y_i von der Ausgleichsgeraden

Nach dem Prinzip der kleinsten Quadrate ergibt sich die beste Ausgleichsgerade gemäß GRÄNICHER [1996] für ein Minimum von

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_i (m \cdot x_i + c - y_i)^2 \tag{Gleichung 4-3}$$

$n \dots$ Anzahl der Wertepaare, Stichprobenumfang

Diese wird bestimmt durch das lineare Gleichungssystem

$$\frac{\delta \sum \varepsilon_i^2}{\delta m} = 0 \tag{Gleichung 4-4}$$

$$\frac{\delta \sum \varepsilon_i^2}{\delta c} = 0 \tag{Gleichung 4-5}$$

Damit ergeben sich die beiden Parameter für die Näherungsgerade zu

$$m = \frac{n \cdot \sum_{i=1}^n (x_i \cdot y_i) - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \tag{Gleichung 4-6}$$

$$c = \frac{1}{n} \left(\sum_{i=1}^n y_i - m \cdot \sum_{i=1}^n x_i \right) = \frac{1}{n} \left(\sum_{i=1}^n y_i - \sum_{i=1}^n x_i \cdot \frac{n \cdot \sum_{i=1}^n (x_i \cdot y_i) - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \right) \tag{Gleichung 4-7}$$

Nun gelingt die Näherung durch die Gerade für Wertepaare besser, die tatsächlich nahezu auf einer Gerade liegen, als für solche, die völlig zufällig im untersuchten zweidimensionalen Raum verteilt sind. Gemäß HERZ ET AL. [1992] ist „... ein erstes Maß für die Güte der Anpassung der Beobachtungswerte durch Regressionsgleichungen [...] das Bestimmtheitsmaß“.

$$r^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad \text{Gleichung 4-8}$$

$$\text{Mit } \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

r ... Korrelationskoeffizient, quadriert ergibt sich das Bestimmtheitsmaß

\hat{y}_i ... Schätzwert für y_i durch die Näherungsgerade an der Stelle x_i

„Mit statistischen Testverfahren werden die Ergebnisse von Stichproben daraufhin geprüft, ob die Abweichungen noch zufallsbedingt sein könnten oder nicht. Wenn Abweichungen in der vorliegenden Größe nicht mehr rein zufällig entstanden sein könnten, bezeichnet man sie als signifikant“ [HERZ ET AL., 1992]. Ein solcher Signifikanztest ist gemäß BÁRDOSSY [1996] der Hypothesentest des Korrelationskoeffizienten. Dabei wird der Wert der Teststatistik \hat{t} nach Student verwendet.

$$\hat{t} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad \text{Gleichung 4-9}$$

Hier geht neben dem Bestimmtheitsmaß die Anzahl der Messpunkte in die Berechnung ein. Je höher der Wert liegt, desto besser nähert sich die Gerade den Punkten an. Liegt der Wert außerhalb des Verwerfungsbereichs, ist der lineare Zusammenhang zwischen den beiden Messgrößen signifikant. Der Verwerfungsbereich liegt unterhalb eines Wertes t . Der Wert t ist ein Quantil der t-Verteilung, der für das geforderte Signifikanzniveau aus entsprechenden Tabellenwerken entnommen werden kann.

Das Bestimmtheitsmaß r^2 und der Wert der Teststatistik \hat{t} sind also Größen zur Bewertung einer Näherungsgeraden.

4.2.2 Gewichtete lineare Regression

Messungen unterschiedlicher Qualität können den Bedarf an einer Gewichtung verschiedener Messpunkte und damit an einer gewichteten linearen Regression wecken. Sie wird erreicht durch das Berücksichtigen einer Gewichts p_i für jedes Wertepaar i in Gleichung 4-3.

$$\sum_{i=1}^n p_i \varepsilon_i^2 = \sum_i p_i (m \cdot x_i + c - y_i)^2 \quad \text{Gleichung 4-10}$$

p_i ... Gewicht des Messwertpaars i bei der gewichteten linearen Regression

Dies führt zu den angepassten Gleichungen für die Parameter der Näherungsgeraden

$$m = \frac{\sum_{i=1}^n p_i \cdot \sum_{i=1}^n (p_i x_i y_i) - \sum_{i=1}^n p_i x_i \cdot \sum_{i=1}^n p_i y_i}{\sum_{i=1}^n p_i \cdot \sum_{i=1}^n p_i x_i^2 - \left(\sum_{i=1}^n p_i x_i\right)^2} \tag{Gleichung 4-11}$$

$$c = \frac{1}{\sum_{i=1}^n p_i} \left(\sum_{i=1}^n y_i - m \cdot \sum_{i=1}^n x_i \right) = \frac{\sum_{i=1}^n p_i y_i}{\sum_{i=1}^n p_i} - \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i} \cdot \frac{\sum_{i=1}^n p_i \cdot \sum_{i=1}^n (p_i x_i y_i) - \sum_{i=1}^n p_i x_i \cdot \sum_{i=1}^n p_i y_i}{\sum_{i=1}^n p_i \cdot \sum_{i=1}^n p_i x_i^2 - \left(\sum_{i=1}^n p_i x_i\right)^2}$$

Gleichung 4-12

Bild 4.6 zeigt das Ergebnis der gewichteten linearen Regression für einen Beispieldatensatz. Die Gewichtung der Messpunkte ist durch die Größe der Kreuze dargestellt. Abgesehen von der Gewichtung liegt der Berechnung derselbe Datensatz zugrunde wie der Darstellung in Bild 4.5 für das Ergebnis aus der einfachen linearen Regression.

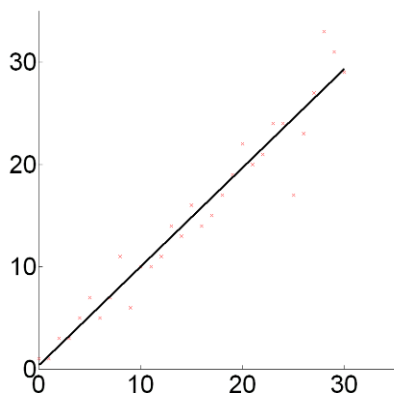


Bild 4.5 Näherungsgerade aus einfacher linearer Regression

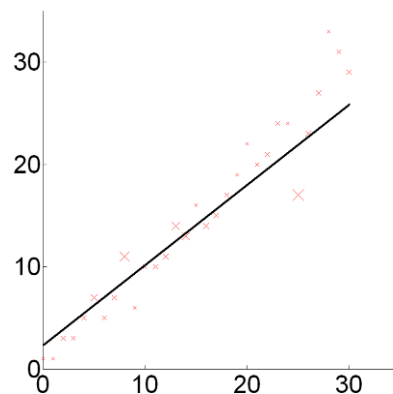


Bild 4.6 Näherungsgerade aus gewichteter linearer Regression

4.2.3 Orthogonale lineare Regression

Die bisher beschriebenen Regressionsansätze berücksichtigen lediglich den vertikalen Fehler zwischen Punkt und Näherungsgerade. Da hier aber jeweils zwei Verkehrskenngrößen in Zusammenhang gebracht werden, die beide – wenn auch unterschiedlich – fehlerbehaftet sein können, so ist auch die Berücksichtigung des horizontalen Fehlers und somit der Einsatz der orthogonalen Regression in Erwägung zu ziehen. Dabei wird statt der vertikalen Abweichung des Punkts von der Geraden die orthogonal zur Geraden anfallende Abweichung berücksichtigt. Dazu schreibt Stahel (1995):

„Wenn aber die Maßeinheit von X und Y geändert wird, ändert sich die Gerade in einer Weise, die schwierig interpretierbar ist. [...] Man soll diese Art der Regression daher nur auf geeignet standardisierte Daten anwenden. Wenn X und Y auf Standardabweichung 1 transformiert werden, ergibt sich aber immer eine Steigung +1 oder -1 für die optimale Gerade, unabhängig von der Stärke des Zusammenhangs. [...] Eine Rechtfertigung der orthogonalen Regression ergibt sich, wenn die X- und Y-Variable mit Messfehlern behaftet und so standardisiert sind, dass die Ungenauigkeit (Standardabweichung der Messfehler) für beide Varianten gleich groß ist. Das Modell läuft unter dem Namen ‚errors-in-variables regression‘ oder ‚functional relationship‘. Es ist nicht für die Prognose von Y aus X geeignet.“

Für die verwendeten Daten wird die Bedingung der gleich großen Ungenauigkeiten der zu korrelierenden Kenngrößen häufig nicht erfüllt. Zudem zielt die beschriebene Methode auf die Schätzung bzw. die Prognose von Y aus X. Der Ansatz der orthogonalen Regression ist für das hier verfolgte Vorhaben also ungeeignet.

4.2.4 Abschnittsweise lineare Regression

Zusammenhänge zwischen Messgrößen müssen weder linear noch über sämtliche Wertebereiche mit derselben Qualität feststellbar sein.

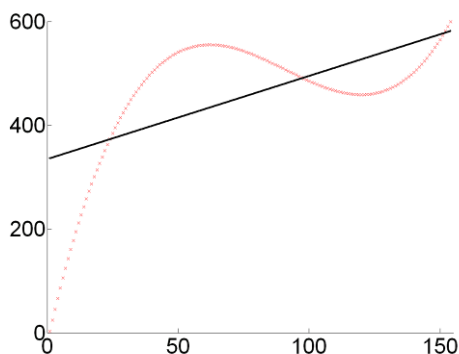


Bild 4.7 Einfache lineare Regression für Stützstellen aus einer kubischen Funktion

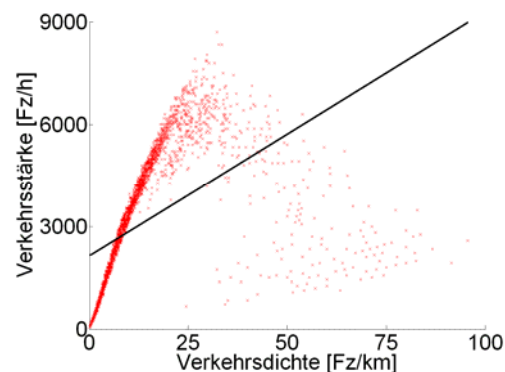


Bild 4.8 Einfache lineare Regression für Stützstellen aus einer q-k-Punktvolke

Bild 4.7 veranschaulicht, dass die lineare Regression in diesem Fall für kleine x deutlich schlechter annähert als für mittlere und große. In Bild 4.8 ist zu erkennen, dass die Punkte für kleine Verkehrsdichten k bei weitem zahlreicher sind und somit die ungewichtete lineare Regression dominieren. Der aus dem Fundamentaldiagramm [FGSV, 2005] bekannte Zusammenhang mit einer unter Leistungsfähigkeitsaspekten optimalen Verkehrsdichte und der zugehörigen maximalen Verkehrsstärke, dem Ansteigen des Graphen für geringere und dem Abfallen für größere Verkehrsdichten ist durch eine Gerade ohnehin nicht zu beschreiben.

Eine Idee zur Verbesserung dieser Umstände ist die abschnittsweise lineare Regression. Durch geschickte Abschnittsbildung können sich verändernde Zusammenhänge differenziert analysiert und mit abschnittsweise gültigen Näherungsgeraden dargestellt werden. Zudem veranschaulicht die q-k-Punktwolke, dass eine abschnittsweise lineare Regression für manche Wertebereiche sehr gute, für andere deutlich schlechtere Annäherungen an die Messwerte durch eine Gerade erlauben. Bereiche, die unbrauchbare Ergebnisse liefern, können als solche gekennzeichnet werden.

Um durchgehend Näherungsgeraden ähnlicher Qualität zu ermitteln, ist eine Optimierung nach dem Bestimmtheitsmaß r^2 oder nach dem Wert der t-Statistik \hat{t} denkbar. Die Abschnitte werden im Rahmen eines Optimierungsverfahrens also so gewählt, dass einer der Kennwerte für sämtliche Abschnitte ungefähr konstant ist.

In den folgenden Abbildungen links werden jeweils fünf Abschnitte gebildet, rechts sind es jeweils sieben.

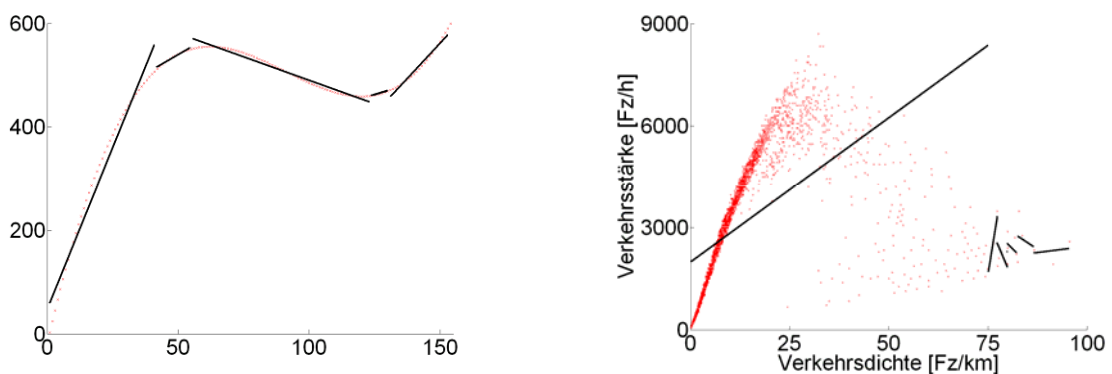


Bild 4.9 Abschnittsweise lineare Regression – optimiert nach dem Bestimmtheitsmaß

Die Ergebnisse für die Optimierungen nach Bestimmtheitsmaß und Wert der t-Statistik sind fast identisch. Bild 4.9 links zeigt gute Schätzungen der Messpunkte durch die abschnittswiseen Näherungsgeraden, rechts zeigt sich aber, dass aufgrund der zahlreichen Messpunkte für geringe Verkehrsdichten k und dem dort guten linearen Zusammenhang mit der Verkehrsstärke der erste Abschnitt weit nach rechts ausgedehnt wird, um die Kennwerte r^2 und \hat{t} in den unterschiedlichen Abschnitten einander anzugleichen. Auch nach der Optimierung sind die Werte r^2 und \hat{t} für den ersten Abschnitt noch deutlich höher als für die anderen Abschnitte. Diese können aber aufgrund einer Mindestanzahl der Messpunkte in einem Abschnitt nicht weiter verkleinert werden. So führt die Optimierung nach r^2 oder \hat{t} nicht zur Ausweisung von Abschnitten mit schlechtem Zusammenhang zwischen den Kenngrößen. Zudem gelingt die Darstellung des Zusammenhangs zwischen Verkehrsstärke und -dichte gemäß Fundamentaldiagramm in Bild 4.9 rechts nur unzureichend. Daher wird eine Abschnittsbildung entweder mit einer konstanten Anzahl von Messpunkten je Abschnitt oder

mit äquidistanten Abschnitten (für jeden Abschnitt gilt: $x_{\max} - x_{\min} = \text{const}$) in Erwägung gezogen.

Der Ansatz einer Abschnittsbildung mit einer konstanten Anzahl von Messpunkten je Abschnitt wird wieder verworfen, da bei Zusammenhängen zwischen verkehrlichen Größen ähnlich wie in Bild 4.9 rechts damit zu rechnen ist, dass meist freier Verkehrsfluss herrscht und somit für freien Verkehrsfluss zahlreiche Messwerte vorliegen. So würden zahlreiche Abschnitte für freien Verkehrsfluss gebildet, für gestörten Verkehrsfluss stünden vergleichsweise wenige Abschnitte zur Verfügung (Bild 4.10 rechts).

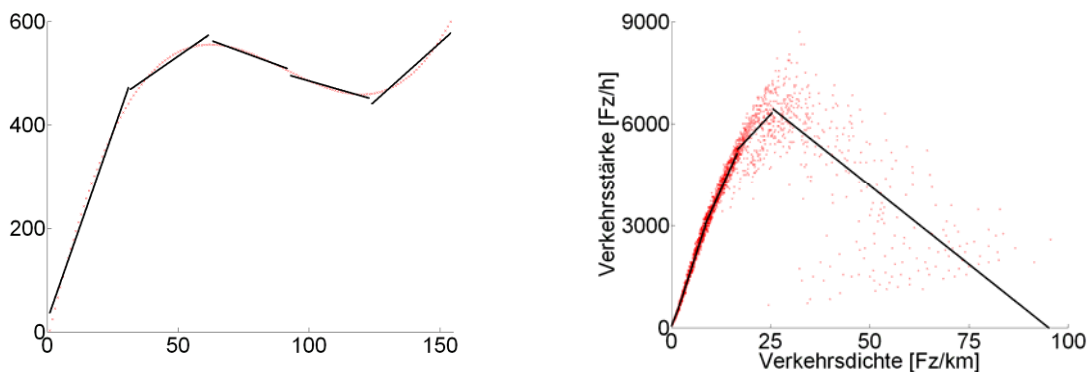


Bild 4.10 Abschnittsweise lineare Regression mit einer konstanten Anzahl von Messpunkten

Die abschnittsweise lineare Regression mit äquidistanten Abschnitten liefert für den Zusammenhang zwischen Verkehrsdichte und -stärke ein Ergebnis (Bild 4.11), das den auf dem Fundamentaldiagramm basierenden Erwartungen deutlich näher kommt. Jedoch kommt es an den Abschnittsgrenzen teilweise zu Unstetigkeiten.

Bild 4.10 f. jeweils links sind identisch, da hier die Werte x_i gleichverteilt sind.

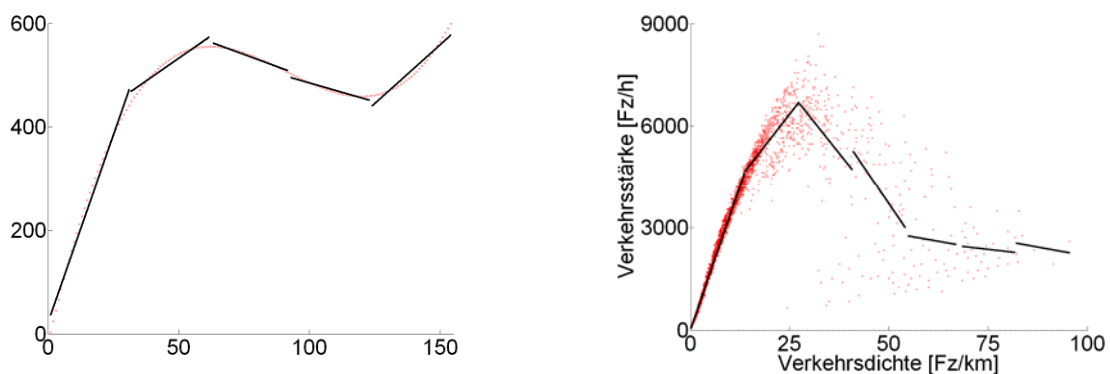


Bild 4.11 Abschnittsweise lineare Regression mit äquidistanten Abschnitten

Um die Unstetigkeiten zu reduzieren, können die Abschnitte überlappend gewählt werden. So werden bei einer Überlappung von beispielsweise 0,2 die Abschnitte an beiden Seiten um das 0,2-fache ihrer ursprünglichen Größe ausgedehnt. Bild 4.12 zeigt, dass die ge-

wünschten Effekte eintreten: Mit wachsender Überlappung ist eine Reduzierung der Unstetigkeiten zu beobachten. Jedoch kann auf den linken Darstellungen auch beobachtet werden, dass für wachsende Überlappungen vergleichsweise kleinteilige Veränderungen der Messreihe wie in der Nähe des lokalen Maximums weniger differenziert dargestellt werden.

So wird die abschnittsweise lineare Regression mit äquidistanten Abschnitten und Überlappungen an den Abschnittsgrenzen für die neue Methode zur Schätzung von Verkehrskenngrößen verwendet. Die geschilderten Ansätze zur gewichteten linearen Regression und zur abschnittswise linearen Regression können zur gewichteten abschnittswise linearen Regression kombiniert werden.

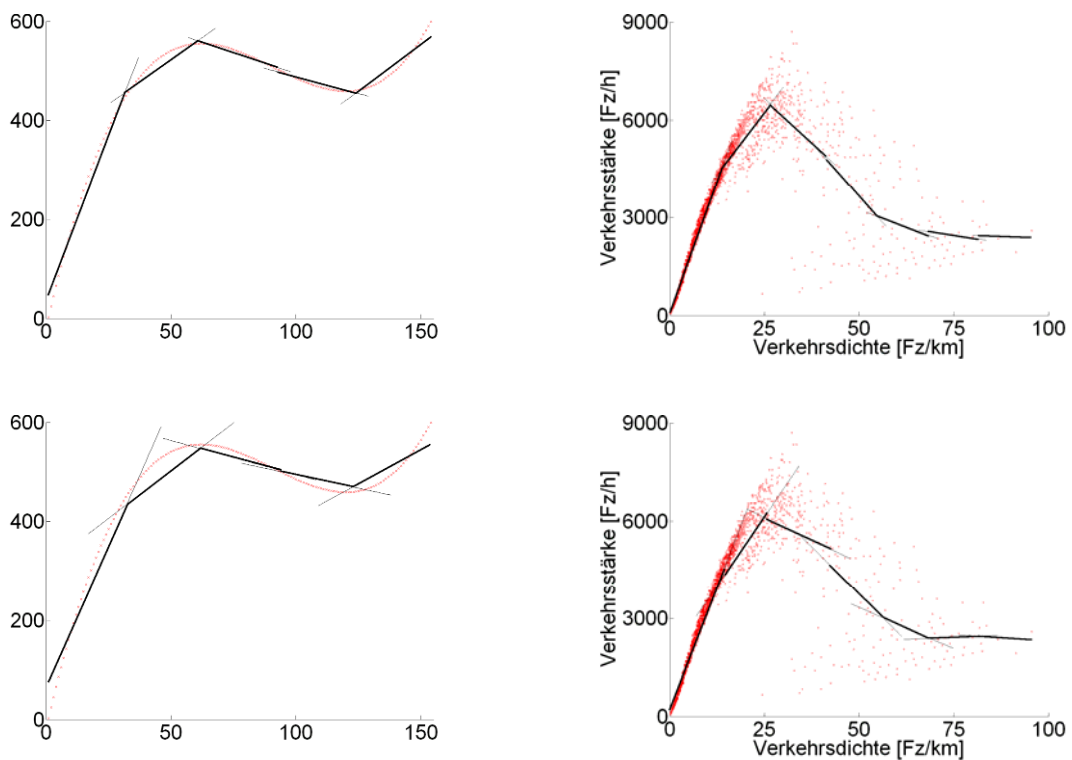


Bild 4.12 Lineare Regression für äquidistante Abschnitte und Überlappungen von 0,2 (oben) und 0,5 (unten)

4.3 Zusammenführung mehrerer Schätzungen

Die Zusammenhänge können für alle Paarungen von verkehrlichen Kenngrößen im Netz untersucht werden. Grafisch kann das Ergebnis wie in Bild 4.13 dargestellt werden. Die Darstellung enthält keine Zusammenhänge zwischen Kenngrößen mit zeitlichem Versatz. In diesem Fall käme für jeden untersuchten Versatz eine weitere Matrix aus Näherungsfunktionen wie in der Abbildung hinzu, bei der auch die Diagonale besetzt wäre.

Liegen in einem Zeitschritt einige Messwerte vor, können im Beispiel daraus sämtliche Kenngrößen im Netz mehrfach geschätzt werden. Diese Schätzungen unterliegen einer Ver-

teilung, da nicht jede Schätzung eines Kennwerts dasselbe Ergebnis liefert. Aus diesen Verteilungen oder durch eine Aggregation können endgültige Schätzwerte bestimmt werden.

Um bei der Zusammenführung einzelner Schätzwerte ausschließlich gute Einzelschätzungen zu verwenden und damit die Verteilungen möglichst schmal zu halten, müssen die Einzelschätzungen hinsichtlich ihrer Qualität bewertet werden. Dabei kann die Qualität des vorhandenen Datums (Kapitel 5) und die Qualität der verwendeten Näherungsfunktion berücksichtigt werden. Die Qualität der Näherungsfunktion kann beispielsweise mit dem Bestimmtheitsmaß, mit dem Korrelationskoeffizienten oder mit dem Wert der Teststatistik nach Student beurteilt werden. Höhere Werte stehen für einen zuverlässigeren Zusammenhang.

Die endgültigen Schätzwerte können die in Kapitel 4.1 beschriebenen Ergebnisse liefern. Da sämtliche Linkreisezeiten aus Positionsmeldungen gemessen oder geschätzt oder aus beidem kombiniert vorliegen können, ist die Schätzung der netzweiten Verkehrslage möglich.

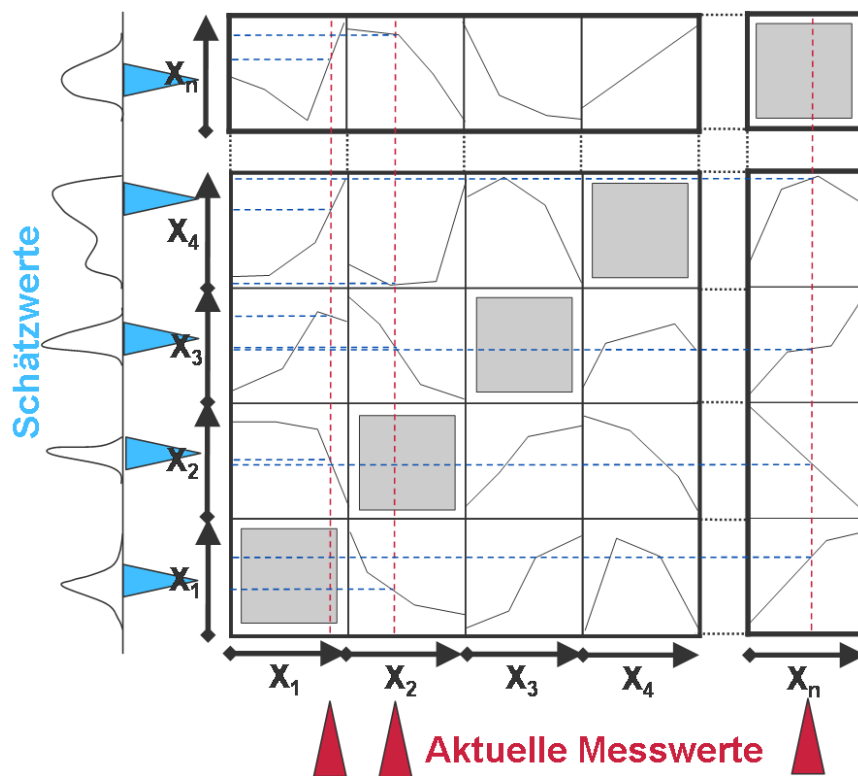


Bild 4.13 Zusammenhänge zwischen zahlreichen Kenngrößen und Schätzung von Kenngrößen basierend auf aktuellen Messwerten

4.4 Kalibrierung zusammengeführter Schätzungen

Unter den ermittelten Näherungsgeraden können solche sein, die einen starken Zusammenhang zwischen zwei Kenngrößen mit einer großen Steigung dokumentieren, jedoch auch solche, die – möglicherweise auch aufgrund einer vergleichsweise diffusen Punktwolke – nur einen schwachen Zusammenhang mit einer geringen Steigung beschreiben. Weil Störungen

in aller Regel weit seltener auftreten als ungestörte Verkehrsverhältnisse, führen solche diffusen Punktwolken zu vorsichtigen Schätzfunktionen, die also über den gesamten Wertebereich der Eingangsgröße für die Ausgangsgröße ungestörte Verkehrsverhältnisse ausgeben.

Es wird erwartet, dass Einzelschätzwerte, die bei der Zusammenführung zu einem aggregierten Schätzwert verwendet werden, häufig aus denselben Näherungsfunktionen stammen. Grund dafür ist die Auswahl der Näherungsfunktionen anhand von Eigenschaften, die sich wie die Näherungsfunktionen selbst nicht kurzfristig verändern. So treffen häufig dieselben Kombinationen aus direkten und eher defensiven, also vorsichtigen Einzelwertschätzern zusammen. Über die Historie kann beobachtet werden, ob diese Kombinationen die jeweiligen Kenngrößen verlässlich berechnen oder systematisch unter- bzw. überschätzen. Dabei ist in Betracht zu ziehen, dass sich eine solche Systematik für unterschiedliche Skalenbereiche differenziert darstellt, die Schätzfehler also beispielsweise für große und kleine Ausgangswerte unterschiedlich ausfallen. Dies wird im weiteren Verlauf der Arbeit mit einer zweiten abschnittswisen Regressionsanalyse für Messwerte als Eingangsgröße und aggregierte Schätzwerte als Ausgangsgröße untersucht. Ergeben sich hier stabile Näherungsfunktionen, die wesentlich von der Mittelgeraden abweichen, so liefern diese einen Term zur Verbesserung des aggregierten Schätzwerts. In diesem Fall würde der aggregierte Schätzwert zunächst lediglich einen Indikator darstellen, der erst zusammen mit einem Korrekturwert aus der Verbesserungsfunktion den endgültigen Schätzwert für eine Kenngröße liefert.

5. Datenqualität und Suchraumbeschränkung

Zur Entwicklung und Erprobung der neuen Methode liegen ein digitales Netz sowie historische Verkehrsdaten für den Norden Münchens vor. Das fein aufgelöste Netz enthält Straßen für den motorisierten Individualverkehr, Schienen-, Rad-, Fuß- und Wasserwege und besteht aus 47.772 Links und 18.110 Knoten (Bild 5.1). Das Kapitel befasst sich mit der Qualität der vorliegenden Verkehrsdaten und mit der Ausdünnung des Netzes, die von der räumlichen Datenverfügbarkeit mit beeinflusst wird. Die dennoch erforderliche Beschränkung des Suchraums nach Zusammenhängen wird am Ende des Kapitels erläutert.

Datenqualität wird hier als Oberbegriff von Datenverfügbarkeit und Datengüte verstanden.

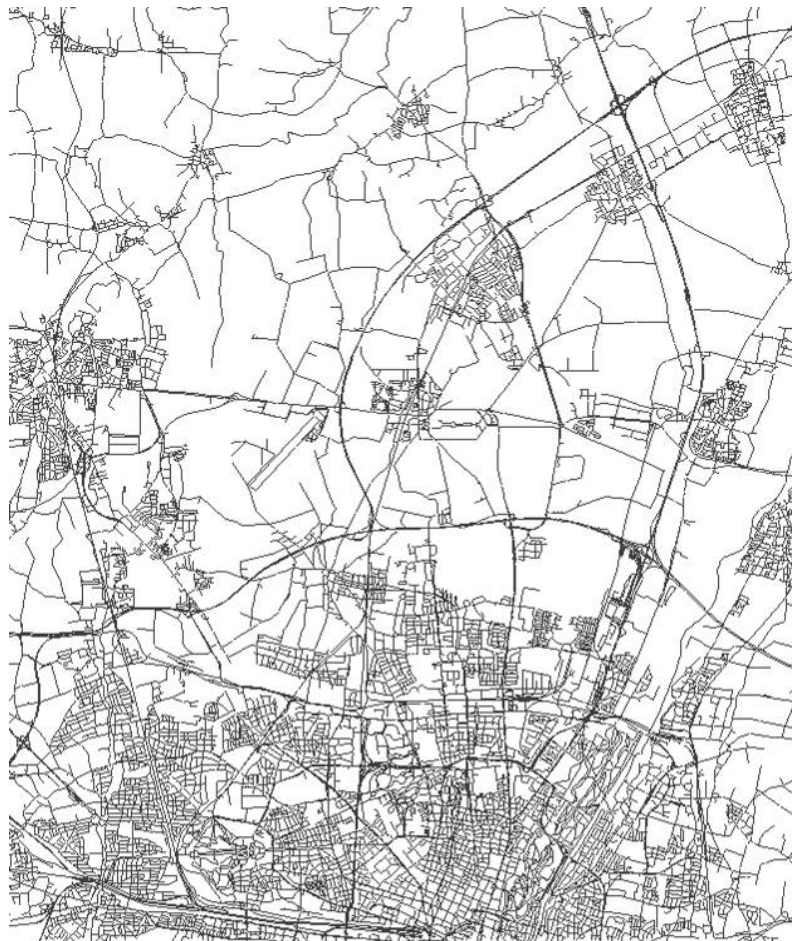


Bild 5.1 München Nord – ursprüngliches Netz

5.1 Verfügbarkeit von Positionsdaten und Linkreisezeiten

Die Münchner Taxizentrale IsarFunk stellt dem Lehrstuhl für Verkehrstechnik die Positionsmeldungen ihrer Fahrzeuge zur Verfügung. Im Zeitraum 10.12.2008 bis 22.3.2009 wurden am 12.12.2008 die meisten Positionsmeldungen übermittelt (67.026 Meldungen aus 10.615

Fahrten). An manchen Tagen läuft nur ungefähr ein Achtel dieser Datenmenge auf. Die Änderungen der Datenmengen erfolgen abrupt, meist für einen längeren Zeitraum. Verkehrliche Ursachen dafür können ausgeschlossen werden.

Um das Netz für die Analyse der Positionsdaten aufzubereiten, wurde es mit dem räumlich differenzierten Positionsdatenaufkommen abgeglichen. Netzausschnitte, in denen demzufolge keine oder sehr wenige Befahrungen durch die Taxi-Flotte stattfinden, wurden daraufhin entfernt. Davon waren auch Schienen-, Fuß-, Rad- und Wasserwege betroffen. Straßen mit Verbindungsfunktion wurden aber im Netz belassen, auch wenn das dortige Positionsdatenaufkommen nicht dafür sprach. Das resultierende Netz (Grundlage für Bild 5.2) wurde zur Aufbereitung der Positionsdaten herangezogen. Es besteht aus 15.170 Links und 5.827 Knoten. Nach Abzug der Links in Einbahnstraßen in Gegenrichtung bleiben 11.299 befahrene Links.

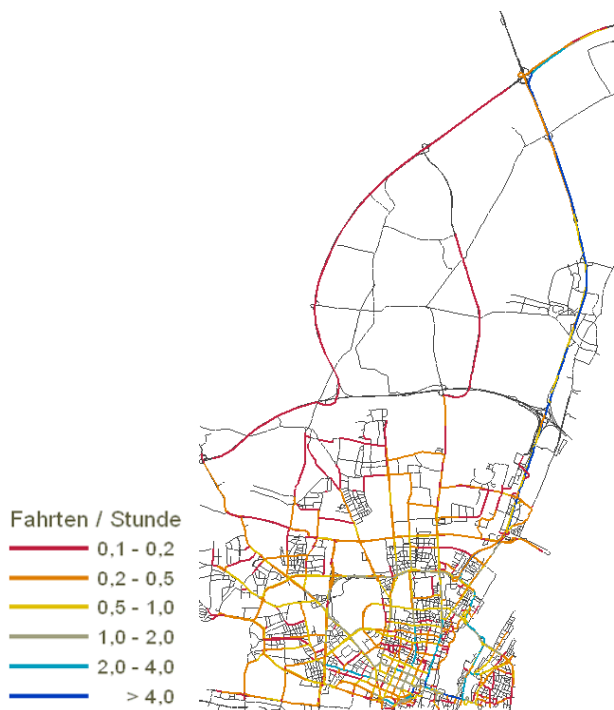


Bild 5.2 Verfügbarkeit von Linkgeschwindigkeiten aus Positionsmeldungen

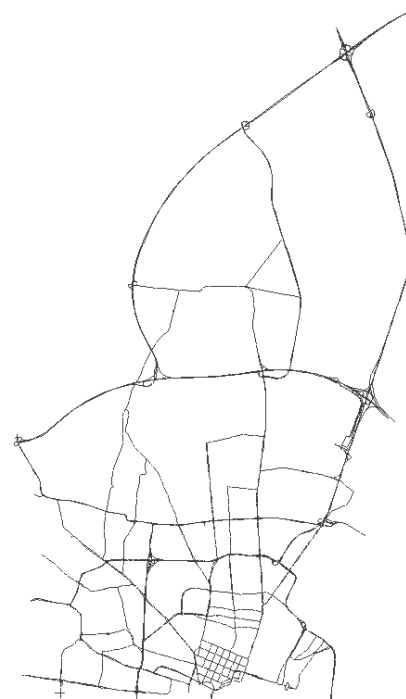


Bild 5.3 Netz zur Verkehrslageschätzung

Die zeitlichen Abstände zwischen den Positionsmeldungen eines Fahrzeugs aus der untersuchten Flotte sind nicht konstant. Am 12.12.2008 betragen 30% dieser Abstände maximal 60s, 60% maximal 120s und 90% maximal 150s. So wurden nur Positionsdaten mit einem maximalen Übertragungsintervall von 150s ausgewertet, wodurch ca. 10% der Positionsmeldungen verworfen wurden. Damit eine Fahrt ausgewertet wurde, mussten zu ihr mindestens fünf Positionsmeldungen vorliegen. Positionen wurden nur berücksichtigt, wenn ihr räumlicher Abstand zum nächsten Link 20m nicht überschritt. Zudem wurden ermittelte

Linkgeschwindigkeiten verworfen, wenn sie mehr als das Doppelte der linkbezogenen freien Geschwindigkeit betragen, die im Netz hinterlegt ist.

Für den 12.12.2008 ergeben sich so 56.340 ermittelte Linkreisezeiten, nach der Aggregation auf Fünf-Minuten-Intervalle bleiben 52.159 Werte.

Als Grundlage für die weitere Abstraktion des Netzes für die Verkehrslageschätzung wurde die räumliche Verteilung der Linkreisezeiten im Zeitraum 15.-19.12.2008 herangezogen. Jedoch wurden auch hier Straßen mit Verbindungsfunktion nicht gelöscht. Das resultierende Netz weist 6.664 Links, davon 4.150 befahren, und 2.794 Knoten auf (Bild 5.3).

Die Aufbereitung der Positionsdaten erlaubt eine Untersuchung der Verfügbarkeit von Linkreisezeiten aus Positionsmeldungen wie in Bild 5.2 auch für eine feinere zeitliche Unterteilung wie z.B. in Stundenintervalle. Es zeigen sich interessante Muster hinsichtlich einer verstärkten Taxi-Nachfrage insbesondere zum Flughafen für den frühen Morgen vor 5 Uhr und damit zu Zeiten eines vergleichsweise spärlichen ÖPNV-Angebots sowie nachts (ca. 1-2 Uhr) nach Betriebsschluss der U- und Straßenbahnen innerhalb des Mittleren Rings.

Eine Darstellung sämtlicher aus Positionsdaten netzweit ermittelter Quotienten aus Linkgeschwindigkeit und linkbezogener freier Geschwindigkeit einer Woche über die Zeit wie in Bild 5.4 zeigt die verstärkte Taxi-Nachfrage werktags zwischen ungefähr 7 und 20 Uhr, das vergleichsweise häufige Absinken des Geschwindigkeitsindikators während der Hauptverkehrszeiten, die über die Woche zunehmende Nachtaktivität sowie die werktäglichen Höchstgeschwindigkeiten insbesondere nachts. Die ausgewerteten Positionsdaten folgen also einem Muster. Ähnliche Abbildungen können folglich auch für andere Netze, Streckenabschnitte oder Zeitbereiche als Grundlage für Planungsverfahren herangezogen werden.

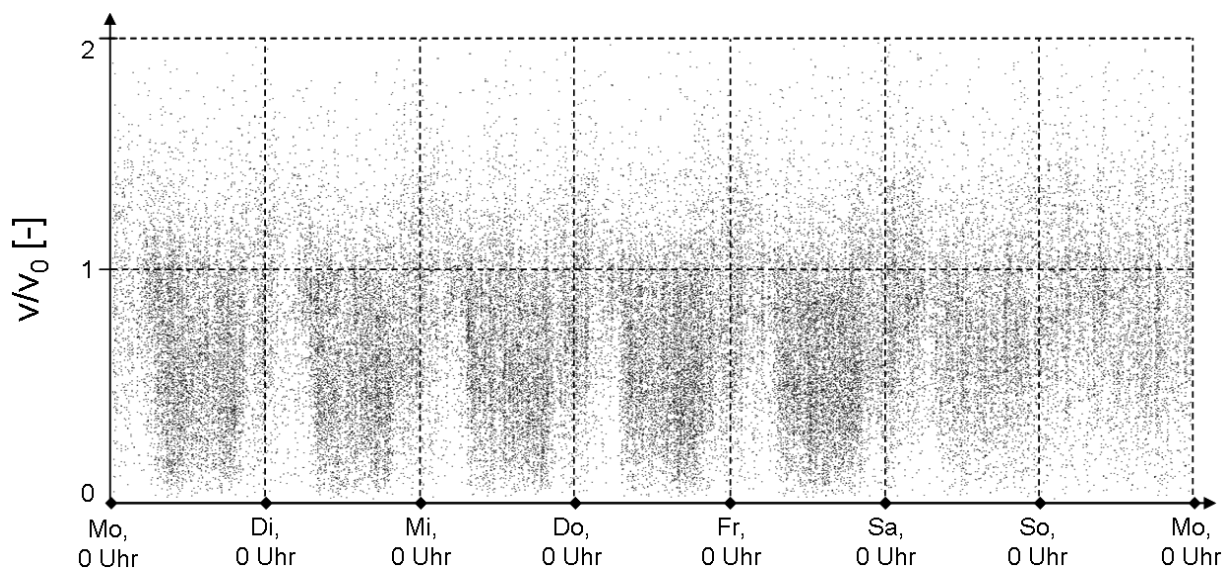


Bild 5.4 Quotienten v/v_0 aus linkbezogenen Geschwindigkeiten (netzweit aus Positionsdaten ermittelt) und linkbezogenen freien Geschwindigkeiten über die Zeit

5.2 Verfügbarkeit infrastrukturseitig erhobener Daten

Die Stadt München stellt dem Lehrstuhl für Verkehrstechnik Daten aus ihren lokalen Messungen zur Verfügung. Die Daten werden mit strategischen Doppel- und Einfachschleifen (im Netzausschnitt 99 und 68 Detektoren) sowie LSA-Einfachschleifen (485 Detektoren) an insgesamt 229 Zählstellen detektiert. Neben der Verkehrsstärke liefern nur die Doppelschleifen Durchschnittsgeschwindigkeit und Belegungsgrad. Die Daten liegen meist zu Zwei-Minuten-Intervallen aggregiert vor.

Bild 5.5 zeigt, dass insbesondere der Mittlere Ring und mit Abstrichen auch der Frankfurter Ring mit Doppelschleifen ausgestattet sind. Darüber hinaus verfügen der Streckenzug Frankfurter Ring/Dachauer Straße und die Ingolstädter Straße über zahlreiche Detektoren an Lichtsignalanlagen. Die vergleichsweise wenigen Einfachschleifen sind einigermaßen über den Netzausschnitt verteilt. Es ist darauf hinzuweisen, dass im Bereich Leopoldstraße und westlich davon, wo die FCD-Verfügbarkeit gemäß Bild 5.2 vergleichsweise hoch ist, fast keine Detektoren vorliegen. Im Auswertungszeitraum lieferten 556 der insgesamt 652 lokalen Detektoren Daten. Unter den Detektoren, zu denen keine Daten vorliegen, sind sämtliche strategische Einfachschleifen.

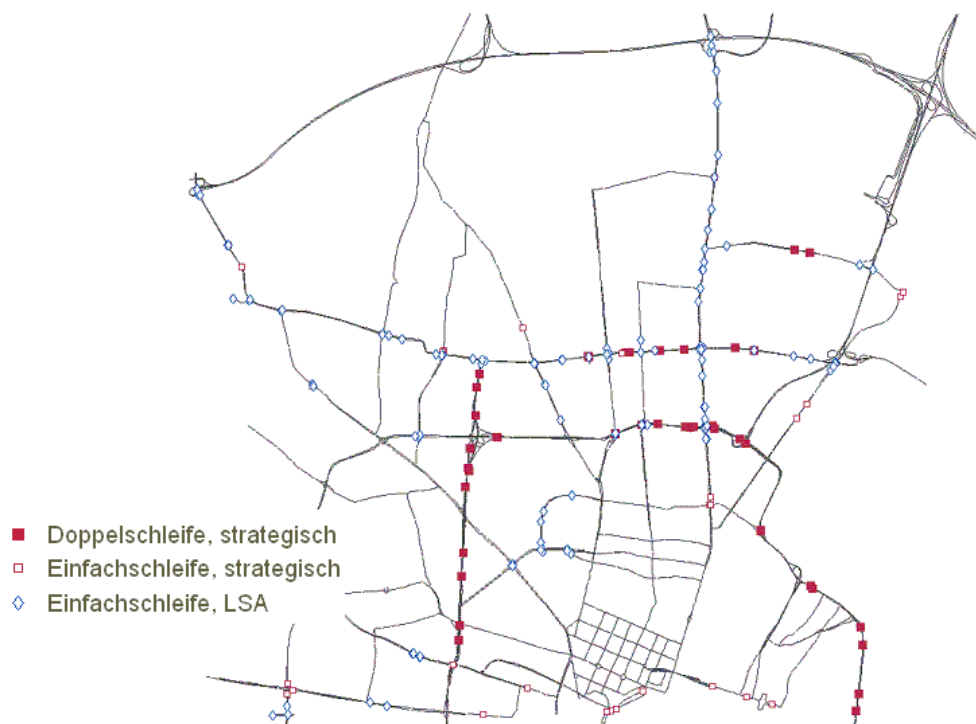


Bild 5.5 Lage der lokalen Detektoren im Netzausschnitt

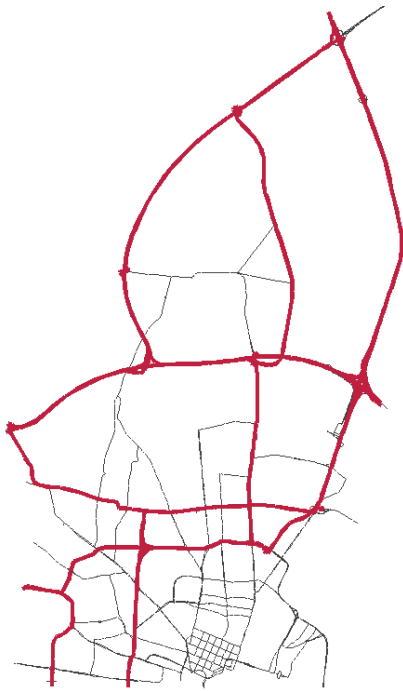


Bild 5.6 Netzabdeckung der Routen mit infrastrukturseitiger Reisezeitmessung für den vorliegenden Netzausschnitt

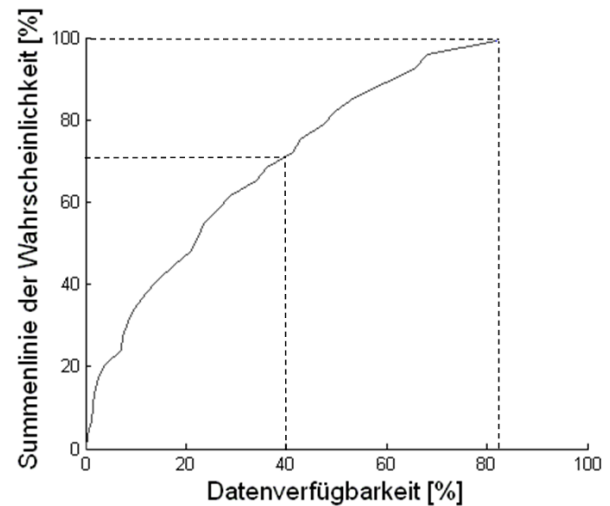


Bild 5.7 Summenlinie der Datenverfügbarkeit für infrastrukturseitig gemessene Reisezeiten in 5-min-Intervallen über 73 Routen

Im Rahmen eines Forschungsprojekts wurden in München zahlreiche Querschnitte mit Kameras zur Kennzeichenanalyse installiert, um auf ausgewählten Routen Reisezeiten durch Fahrzeugwiedererkennung messen zu können (vgl. [SPANGLER, 2009]). Für den vorliegenden Netzausschnitt betrifft dies insgesamt 73 Routen auf dem Mittleren Ring und auf Verbindungsstraßen außerhalb des mittleren Rings. Auch die Autobahnabschnitte liegen im Verlauf der erhobenen Routen. Die Netzabdeckung der reisezeitdetektierten Abschnitte ist Bild 5.6 zu entnehmen. Bild 5.7 gibt einen Anhaltspunkt für die zeitliche Verfügbarkeit von Reisezeitdaten auf den 73 Routen. Sie zeigt, dass für diese Datenart nicht mit einer durchgängigen Datenverfügbarkeit gerechnet werden darf. So weisen 70% der Routen eine Datenverfügbarkeit von maximal 40% auf, auf keiner wird eine Datenverfügbarkeit von 85% erreicht.

5.3 Datengüte

Unter Datengüte wird hier verstanden, inwiefern ein gemessener Wert die Verkehrssituation auf dem Link, der Route oder am Querschnitt richtig beschreibt, an dem oder der er erhoben wurde.

Der Ansatz der gewichteten linearen Regression gemäß Kapitel 4.2.2 ermöglicht, die Güte von verkehrlichen Kenngrößen bei der Suche nach Näherungsfunktionen zur Darstellung der Zusammenhänge mit anderen Kenngrößen zu berücksichtigen. Zudem können bei der Zu-

sammenführung von Schätzwerten solche Schätzwerte stärker gewichtet werden, die auf Messwerten mit einer höheren Güte basieren.

Die Datengüte kann beispielsweise beeinflusst werden durch

- die Datenart (vgl. Kapitel 2)
- die Erhebungstechnologie
- zufällige Eigenschaften des einzelnen Sensors wie Fertigungstoleranzen oder Justierung
- Umfeldbedingungen
- das Alter einer Information
- das Zeitintervall, über das Daten aggregiert werden

Zusätzlich wird die Güte von Linkgeschwindigkeiten aus Positionsdaten beeinflusst durch

- Ort und Inhalt einer Information: Die Güte der Information über eine geringe Geschwindigkeit auf der Autobahn ist höher als die entsprechende Information von einer Einkaufsstraße
- Positionsmeldungsintervalle: Je kürzer der zeitliche Abstand zwischen zwei Positionsmeldungen, umso höher die Güte der abgeleiteten Linkgeschwindigkeiten

Darüber hinaus steigt die Güte aggregierter Werte mit der Anzahl der einzelnen Werte, die zur Aggregation vorliegen. Aggregierte Werte sind je nach Datenverfügbarkeit lokale mittlere Geschwindigkeit, Routenreisezeiten und Linkgeschwindigkeiten aus Positionsdaten. Auf die Güte aggregierter Werte wirkt sich positiv aus, wenn sich die einzelnen Werte sehr ähnlich sind. Entsprechend beeinflussen widersprüchliche Einzelwerte die Güte des aggregierten Werts negativ.

BUSCH ET AL. [2006] beschreiben zahlreiche Verfahren zur Kontrolle der Messwerte an Querschnitten, sowohl Verkehrs- als auch Umfelddaten betreffend. Diese Verfahren dienen dem Verwerfen unplausibler Messwerte. Vorliegende Messwerte werden räumlich und zeitlich mit anderen Messwerten, Ersatzwerten oder Schwellenwerten abgeglichen. Zudem werden Muster, die aus zahlreichen Messwerten bestehen, durch Abgleich mit historischen Mustern auf Plausibilität überprüft.

SPANGLER [2006] ermittelt fünfminütlich aggregierte Routenreisezeiten aus mindestens drei Einzelwerten und weist eine Güte von 100 zu, wenn dafür nicht weiter als fünf Minuten zurückgeblenkt werden muss. Bedarf es dazu einer größeren Zeitspanne, so wird ein geringerer Gütewert zugeordnet. Finden sich in den letzten 15 Minuten weniger als drei Werte, werden historische Werte eingesetzt und die Güte wird mit 50 angegeben.

Zur Bewertung der Schätzergebnisse aus der gewichteten im Vergleich zur einfachen linearen Regression in Kapitel 6.3 werden zu den Linkreisezeiten aus Positionsdaten Güteindika-

toren ausschließlich abhängig vom Zeitintervall zwischen zwei Positionsmeldungen ermittelt. Beträgt das Intervall genau oder weniger als 60s, so ist die Güte der Linkreisezeiten 100, beträgt es 150s, so ist die Güte 50. Für dazwischen liegende Werte wird der Güteindikator linear interpoliert.

5.4 Beschränkung des Raums zur Suche nach Zusammenhängen

Ein Straßennetz weist für das beschriebene Verfahren zahlreiche Kenngrößen auf. Neben Kennwerten aus infrastrukturseitiger Datenerhebung wird aufgrund der Auswertungen von Positionsmeldungen für jeden Link ein Geschwindigkeitskennwert angelegt. So verfügt der Netzausschnitt über etwas mehr als 5.000 Kenngrößen, so dass mehr als 25.000.000 Zusammenhänge untersucht werden könnten. Diese wiederum sind aufgrund der Abschnittsbildung bei der Regression und der Suche nach Zusammenhängen auch für zeitversetzte Intervalle mit weiteren Faktoren zu multiplizieren, um die Gesamtanzahl der Schätzfunktionen zu ermitteln. Es müssen also Wege gefunden werden, um die Anzahl der untersuchten Zusammenhänge zu reduzieren.

Bestehen starke Zusammenhänge zwischen zwei Kenngrößen, so nimmt mit zunehmender Entfernung zwischen den Erhebungsorten der Einfluss von regelmäßig wiederkehrenden Ereignissen zu und der Einfluss weniger typischer Ereignisse wie Unfälle oder Sperrungen ab. Beispiele für wiederkehrende Ereignisse sind nachfrageverursachte Störungen des Verkehrsablaufs in der Hauptverkehrszeit. Diese erlauben, dass von einem Kennwert im Süden einer Stadt auf einen Kennwert im Norden geschlossen werden kann. Tritt aber nun im Süden eine Störung in Folge eines Unfalls ein, so kann aufgrund der Veränderung des südlichen Kennwerts nicht mehr auf den nördlichen geschlossen werden. Da die Ursache für die Störung dem Kennwert aber nicht hinzugefügt werden kann, entstünde folglich eine fehlerhafte Schätzung. Daher sind Zusammenhänge von der Untersuchung auszuschließen, die in zu großer Entfernung voneinander erhoben werden. Als Kriterien können die Luftlinienentfernung, die Entfernung unter Berücksichtigung der Netzgeometrie, die Anzahl der Knotenpunkte oder die freie Reisezeit unter Verwendung der freien Geschwindigkeit gemäß der Netzdatei zwischen den Erhebungsorten herangezogen werden, hier ergänzt um einen Abgleich der Fahrrichtungen. Von der Verwendung der Knotenanzahl als Kriterium wird abgesehen, da die Linklängen im Netz sehr unterschiedlich sind und daher der Einfluss des Zufalls groß wäre. Ebenso scheinen Luftlinienentfernung und Entfernung unter Berücksichtigung der Netzgeometrie ungeeignet, da dann vergleichsweise viele Zusammenhänge in Wohngebieten, wo die Links geringe freie Geschwindigkeiten aufweisen, betrachtet werden, gleichzeitig aber Beziehungen zwischen Kenngrößen auf Autobahnen ausgeschlossen werden, die in geringem zeitlichen Abstand durch dieselbe Fahrzeuggruppe beeinflusst werden können und so durchaus in unmittelbarem Zusammenhang stehen können, aber eine gewisse räumliche Distanz aufweisen. So bleibt die freie Reisezeit zwischen den Erhebungsort-

ten zweier Kenngrößen als Kriterium für die Entscheidung, ob der Zusammenhang betrachtet wird oder nicht.

Wird diese freie Reisezeit auf 150s begrenzt und werden Vorgänger- und Nachfolgerlinks zu einem Link ausgeschlossen, deren Fahrtrichtung um mehr als 120° von der Richtung dieses Links abweichen, reduziert sich die Anzahl der Zusammenhänge im Netz von über 2,5 Mio. auf ca. 728.000. Hierbei wird der Nachbarschaftsgrad in Sekunden zwischen den Links ermittelt. Die Nachbarschaftsgrade zu einem Detektor entsprechen denen zu dem Link, an dem der Detektor installiert ist. Die Nachbarschaftsgrade zu einer Route entsprechen dem geringsten Nachbarschaftsgrad zu allen Links, die auf der Route liegen.

Betrachtet man ausschließlich die Nachbarschaften zwischen Links, so sind dies ca. 472.000. Ein feineres Netz lässt diese Zahl rapide ansteigen. So steigt sie für das Netz zur Auswertung von Positionsdaten auf über 2 Mio., obwohl sich die Anzahl der Links nur etwas mehr als verdoppelt.

Ergebnis dieser Berechnungen ist für jeden Link eine Zu- und Abflussspinne, in der nach Zusammenhängen zwischen dort erhobenen Kenngrößen und Kenngrößen, die für den untersuchten Link erhoben werden, gesucht wird (vgl. Bild 5.8).



Bild 5.8 Suchraum nach Zusammenhängen zu einem untersuchten Link

6. Bewertung der ermittelten Näherungsfunktionen

Gegenstand dieses Kapitels ist die Bewertung der aus der Datenhistorie ermittelten Näherungsfunktionen. Die bei der Regressionsanalyse verwendeten verkehrlichen Kenngrößen liegen aggregiert zu Fünf-Minuten-Intervallen vor.

Die zu bewertende Schätzmethode basiert auf der Annahme, dass bestimmte verkehrliche Muster aufgrund wiederkehrender Nachfrageszenarien ebenfalls wiederkehren. Informationen über diese Nachfrageszenarien liegen jedoch nicht in Datenform vor, ausschließlich ihre verkehrlichen Konsequenzen sind in Form von verkehrlichen Kenngrößen bekannt. Neben der Bewertung unter Verwendung real erhobener Daten kann auch eine Bewertung mit Daten aus einer mikroskopischen Verkehrsflusssimulation in Betracht gezogen werden. Jedoch sind hierfür genau die Informationen über die Nachfrageszenarien erforderlich, die nicht vorliegen. So basieren sämtliche nachfolgende Bewertungen auf real erhobenen Daten.

6.1 Konzept zur Bewertung der Näherungsfunktionen

Im Folgenden werden die Parameter zur Ermittlung der Näherungsfunktionen aufgezählt und ihre möglichen Auswirkungen erläutert. Zudem wird dargestellt, auf welcher Datengrundlage diese Auswirkungen im weiteren Verlauf untersucht werden sollen.

- (1) Anzahl der Abschnitte für die abschnittsweise lineare Regression und Überlappung der Abschnitte:

Eine steigende Abschnittszahl für die abschnittsweise lineare Regression erlaubt eine differenziertere Darstellung der Zusammenhänge und deren Güte zwischen den untersuchten Kenngrößen. Sie erhöht jedoch den Rechenaufwand und – was schwerer wiegt – das Risiko, dass in manchen Abschnitten zu wenige Punkte vorliegen, um einen signifikanten Zusammenhang ermitteln zu können. Unglücklicherweise ist dieses Risiko für Abschnitte besonders hoch, die Informationen zu gestörtem Verkehrsfluss enthalten und somit zur Schätzung besonderer Verkehrsverhältnisse benötigt werden. Ursache dafür ist primär, dass solche Zustände in der Historie im Vergleich zu ungestörten Verkehrszuständen selten vorliegen. Darüber hinaus muss davon ausgegangen werden, dass die Zusammenhänge in diesen Bereichen weniger eindeutig ausgeprägt sein können und daher mehr Punkte erforderlich sind, um die Signifikanz dieser Zusammenhänge nachzuweisen. Insbesondere sind hier Probleme für Zusammenhänge zu erwarten, in die Daten mit einer geringen Verfügbarkeit eingehen. Dies betrifft vor allem Linksgeschwindigkeiten aus Positionsmeldungen und in geringerem Maße infrastrukturseitig gemessene Reisezeiten. Vor dem Hintergrund der Charakteristik erwarteter Zusammenhänge sind mindestens drei Abschnitte zu bilden. Um zu überprüfen, ob mehr Abschnitte die Schätzergebnisse positiv beeinflussen, werden Untersuchungen mit drei und fünf Abschnitten angestellt.

Die Überlappung der Abschnitte soll bezwecken, dass die Übergänge zwischen den Abschnitten verstetigt werden, ohne dabei die differenzierte Darstellung der Zusammenhänge zu unterdrücken. Sie ist daher auf die Breite der Abschnitte abzustimmen. Für die nachfolgenden Untersuchungen wird daher basierend auf Erfahrungen aus Kapitel 4.2.4 und nach Betrachtung zahlreicher Grafiken wie Bild 4.12 bei fünf Abschnitten ein Überlappungsgrad von 0,2 gewählt. Da der Überlappungsgrad relativ zur Abschnittsbreite festgelegt wird, wird er bei drei Abschnitten auf 0,13 reduziert.

(2) Differenzierung nach Tageskategorie und Tageszeit:

Es kann ohne Prüfung nicht festgestellt werden, ob es Vorteile bringt, wenn die Zusammenhänge zwischen verkehrlichen Kenngrößen differenziert nach Tageskategorie und Tageszeit ermittelt werden anstatt in dieser Hinsicht undifferenziert über alle Zeiträume. Für eine Differenzierung spricht, dass die Zusammenhänge in diesem Fall sicher nicht schwächer sind. Dagegen spricht, dass dadurch weniger Daten vorliegen, was zu Nachteilen beim Signifikanznachweis führen könnte. Wie beim zuvor erläuterten Sachverhalt würde dies insbesondere Zusammenhänge betreffen, die Störungssituationen dokumentieren. Zur Untersuchung dieser Fragestellung werden zum einen Zusammenhänge über alle vorliegenden Daten ermittelt, zum anderen über Daten, die montags bis freitags zwischen sechs und zehn Uhr gesammelt wurden.

(3) Diagonale und horizontale Zusammenhänge:

Unter diagonalen und horizontalen Zusammenhängen werden gemäß Bild 4.4 Zusammenhänge mit einem bestimmten zeitlichen Versatz zwischen der Erhebung der korrelierten Kenngrößen verstanden. Die Bewertung der Schätzfunktionen wird für einen zeitlichen Versatz von bis zu 30 Minuten vorgenommen. Auch die Kurzfristprognose nach VORTISCH [2006] betrachtet diesen zeitlichen Horizont. Zudem liegt dieser Einschränkung die pragmatische Annahme zugrunde, dass Messwerte nach einer halben Stunde keine oder nur noch eine sehr geringe Aussagekraft haben.

(4) Einfache oder gewichtete lineare Regression:

Die Effekte der gewichteten Regression werden unter Verwendung aller Daten, also zeitlich nicht differenziert, und für drei Abschnitte bei der Abschnittsbildung mit einer Überlappung von 0,13 untersucht. So ist die Vergleichbarkeit mit den Ergebnissen aus der ungewichteten linearen Regressionsanalyse gegeben.

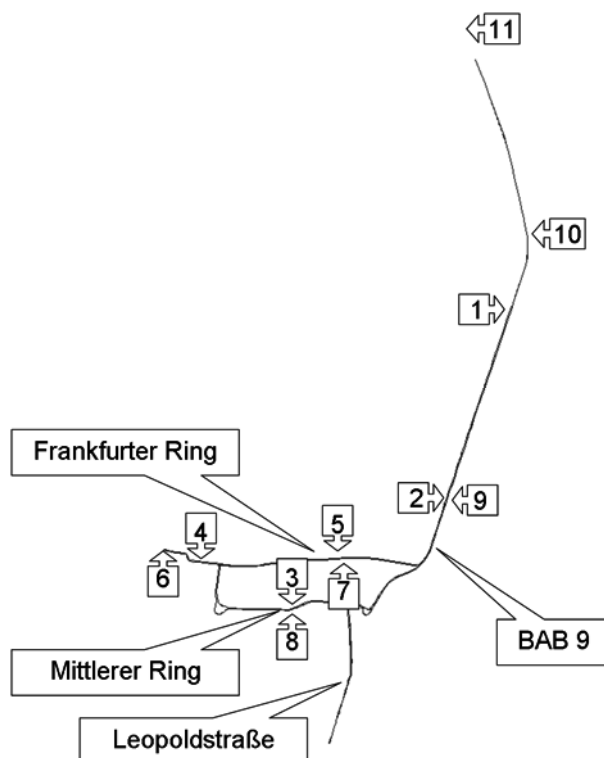
(5) Statistische Signifikanz:

Damit eine ermittelte Näherungsgerade nicht verworfen wird, muss sie ihre statistische Signifikanz mit einem Signifikanzniveau von 0,9 im t-Test nachweisen.

Zur Bewertung der Schätzfunktionen werden ihre Verfügbarkeit sowie das Bestimmtheitsmaß und der Wert der t-Statistik \hat{t} verwendet.

Um die Übersichtlichkeit zu wahren, wird im weiteren Verlauf ein weiter reduziertes Netz verwendet. Folgende Vorgaben waren bei der Netzreduktion maßgeblich:

- (1) Das Bewertungsnetz soll Routen mit infrastrukturseitiger Reisezeiterhebung aus verschiedenen verkehrlichen Umfeldern enthalten.
- (2) Das Netz soll Streckenabschnitte mit strategischen Detektoren enthalten.
- (3) Das Netz soll Streckenabschnitte mit LSA-Detektoren enthalten.
- (4) Das Netz soll innerörtliche Streckenzüge mit planfreien Knotenpunkten enthalten.
- (5) Das Netz soll innerörtliche Streckenzüge mit plangleichen, lichtsignalgeregelten Knotenpunkten enthalten.
- (6) Das Netz soll einen Streckenzug enthalten, der über keine infrastrukturseitige Verkehrsdatenerfassung verfügt, auf dem aber möglichst viele Positionsmeldungen gesammelt werden.



Infrastrukturseitige Reisezeitmessung im Bewertungsnetz:

1-2 2-3 3-4 5-4

6-7 6-8 8-9 9-10 10-11

Route 10-11 wird wegen der Charakteristik dort erhobener Daten mit betrachtet, obwohl Querschnitt 11 etwas außerhalb des Testfelds liegt.

Streckenzug „Leopoldstraße“ enthält auch die Ludwigstraße.

„Frankfurter Ring“ enthält auch Moosacher Straße und Triebstraße.

„Mittlerer Ring“ enthält auch Teile der Landshuter Allee.

Die Lage lokaler Detektoren ist Bild 5.5 zu entnehmen.

Bild 6.1 Bewertungsnetz mit Querschnitten zur Fahrzeugwiedererkennung und Benennung von Streckenzügen

Bild 6.1 zeigt das für die Bewertung ausgewählte Netz. Der Streckenzug „Frankfurter Ring“ mit den Reisezeitrouten 5-4 und 6-7 erfüllt die Anforderungen (3) und (5). Die Anforderungen (2) und (4) werden vom „Mittleren Ring“ mit den Reisezeitrouten 3-4 und 6-8 erfüllt. Der Streckenzug „Leopoldstraße“ weist wie in (6) gefordert keine infrastrukturseitige Datenerhebung auf. Reisezeiten werden infrastrukturseitig sowohl auf Autobahnen (1-2, 9-10, 10-11) als auch innerorts erhoben, dort sowohl auf lichtsignalgeregelten Abschnitten (5-4, 6-7) als auch auf Abschnitten mit überwiegend planfreien Knotenpunkten (3-4, 6-8). Die Abschnitte 2-3 und 8-9 liegen im Übergang zwischen Innerortsstraßen und Bundesautobahnen. So ist auch Anforderung (1) erfüllt.

Zur Bewertung der Näherungsfunktionen werden neben grafischen Darstellungen Tabellen eingesetzt. Tab. 6.1 gibt differenziert nach Ein- und Ausgangsgrößenart sowie nach ihrem zeitlichen Versatz an

- Die Anzahl der Beziehungen, die nach der Suchraumbeschränkung gemäß Kapitel 5.4 untersucht wurden
- Absolute Häufigkeiten der Näherungsfunktionen für diese Konstellation von Datenarten bei Bildung von fünf Abschnitten
- Absolute Häufigkeiten der Näherungsfunktionen für diese Konstellation von Datenarten bei Bildung von drei Abschnitten
- Relative Häufigkeiten der Näherungsfunktionen für diese Konstellation von Datenarten bezogen auf die Summe aller Häufigkeiten für denselben zeitlichen Versatz bei Bildung von drei Abschnitten

Tab. 6.2 gibt die absolute Häufigkeit der ermittelten statistisch signifikanten Näherungsgeraden für die Bildung von drei Abschnitten an differenziert nach

- Datenart von Ein- und Ausgangsgröße
- Zeitversatz zwischen Ein- und Ausgangsgröße
- kleinen, mittleren und großen Eingangswerten

Tab. 6.3 enthält entsprechend differenziert die durchschnittlichen Bestimmtheitsmaße der verzeichneten statistisch signifikanten Schätzfunktionen.

In Tab. 6.1 kann die Häufigkeit der Näherungsfunktionen größer sein als die Anzahl der untersuchten Beziehungen, da für eine Beziehung mehrere Abschnitte gebildet werden können. Hier wie in Tab. 6.2 ist zu erwarten, dass die Anzahl der Zusammenhänge mit wachsendem zeitlichen Versatz abnimmt, da durch den zeitlichen Abstand zweier zu korrelierender Kenngrößen mögliche Zusammenhänge schwächer ausfallen sollten. Vergleichbares ist für einen wachsenden räumlichen Abstand zwischen den Erhebungsorten der beiden Kenngrößen zu erwarten. Auch in Tab. 6.3 gilt die Annahme, dass die Qualität der Zusammen-

hänge für wachsende zeitliche Versätze abnimmt. Dies würde ein Absinken der Werte implizieren.

In Tab. 6.2 werden absolute und nicht auf die Anzahl der untersuchten Beziehungen bezogene Werte angeführt, um weiter direkt eine Aussage zum mutmaßlichen Einfluss der aus diesen Näherungsfunktionen ermittelten Schätzungen bei der Verkehrslageschätzung zu ermöglichen.

Bild 6.2 veranschaulicht die Verknüpfungen zwischen Tab. 6.1 ff., die im Lauf des Kapitels ausführlich analysiert werden. Hierbei sind die nachfolgend aufgeführten Abkürzungen zu beachten.

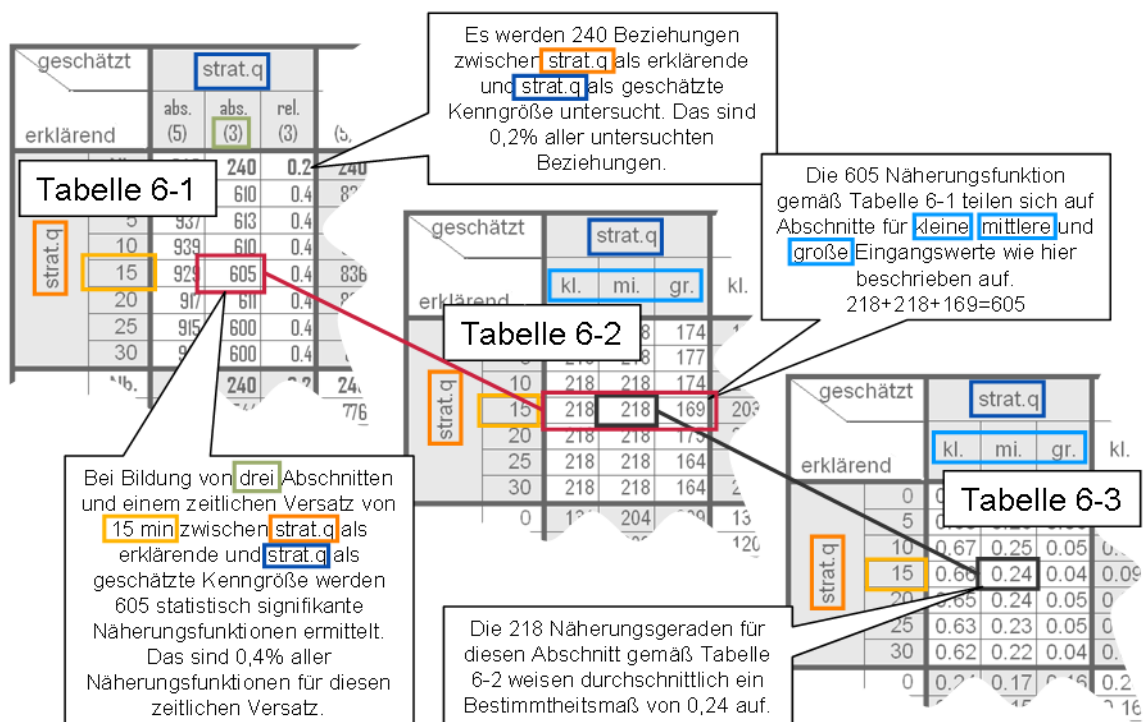


Bild 6.2 Erläuterung zum Verständnis und zur Verknüpfung von Tab. 6.1, Tab. 6.2 und Tab. 6.3

Verwendete Abkürzungen in den Tabellen und Abbildungen des Kapitels:

strat.q	Verkehrsstärke, mit einem strategischen Detektor erhoben [Fz/h]
strat.v	mittlere lokale Geschwindigkeit, mit einem strategischen Detektor erhoben [km/h]
strat.occ	Belegungsgrad, mit einem strategischen Detektor erhoben [%]
LSA q	Verkehrsstärke, mit einem LSA-Detektor erhoben [Fz/h]
Routen-TT	Routenreisezeit durch Kennzeichenanalyse [min]
Link-v/v0	Quotient aus linkbezogener Geschwindigkeit aus Positionsmeldungen und aus linkbezogener freier Geschwindigkeit [-]
Nb.	Anzahl der Beziehungen, die nach der Suchraumbeschränkung gemäß Kapitel 5.4 untersucht wurden („Nb.“ für Nachbarschaft)
0...30	Zeitlicher Versatz zwischen Ein- und Ausgangsgröße [min]
abs. (5)	Absolute Häufigkeit der Näherungsfunktionen bei Bildung von fünf Abschnitten
abs. (3)	Absolute Häufigkeit der Näherungsfunktionen bei Bildung von drei Abschnitten
rel. (3)	Relative Häufigkeit bezogen auf alle untersuchten Beziehungen bzw. sämtliche Näherungsfunktionen bei Bildung von drei Abschnitten für diesen zeitlichen Versatz
kl.	Abschnitt für kleine Eingangswerte, also linker Abschnitt bei grafischer Darstellung
mi.	Abschnitt für mittlere Eingangswerte, also mittlerer Abschnitt
gr.	Abschnitt für große Eingangswerte, also rechter Abschnitt
r	Reisewiderstand zwischen zwei Erhebungsorten
Egw.	Eingangswert
Δt	Zeitversatz
r ²	Bestimmtheitsmaß
n	Stichprobenumfang
\hat{t}	Wert der t-Statistik
0-24u	Verwendung aller Daten (7 Tage/Woche, 24 Stunden/Tag) ungewichtete lineare Regression
0-24g	wie 0-24u, jedoch gewichtete lineare Regression
6-10u	Verwendung der Daten zwischen Montag und Freitag, jeweils 6-10 Uhr, ungewichtete lineare Regression

geschätzt		strat.q			strat.v			strat.occ			LSA q			Routen-TT			Link-v/v0		
		abs. (5)	abs. (3)	rel. (3)	abs. (5)	abs. (3)	rel. (3)	abs. (5)	abs. (3)	rel. (3)	abs. (5)	abs. (3)	rel. (3)	abs. (5)	abs. (3)	rel. (3)	abs. (5)	abs. (3)	rel. (3)
erklärend																			
strat.q	Nb.	240	240	0.2	240	240	0.2	240	240	0.2	1585	1585	1.1	305	305	0.2	5471	5471	3.9
	0	962	610	0.4	838	564	0.3	925	600	0.4	6113	3970	2.4	1037	699	0.4	9817	8294	4.9
	5	937	613	0.4	846	549	0.4	916	592	0.4	6068	3940	2.7	1056	708	0.5	10015	8297	5.6
	10	939	610	0.4	846	554	0.4	909	583	0.4	6075	3939	2.9	1056	699	0.5	10018	8354	6.1
	15	929	605	0.4	836	542	0.4	899	588	0.4	6014	3950	2.9	1032	709	0.5	9797	8195	6.0
	20	917	611	0.4	830	548	0.4	899	588	0.4	5994	3919	2.9	1048	711	0.5	9668	8173	6.0
	25	915	600	0.4	831	540	0.4	901	586	0.4	6036	3927	2.9	1069	706	0.5	9823	8218	6.0
	30	912	600	0.4	819	545	0.4	893	583	0.4	5981	3917	2.9	1050	713	0.5	9565	8257	6.1
strat.v	Nb.	240	240	0.2	240	240	0.2	240	240	0.2	1585	1585	1.1	305	305	0.2	5471	5471	3.9
	0	826	544	0.3	776	528	0.3	814	541	0.3	5235	3371	2.0	779	568	0.3	7337	5963	3.5
	5	834	559	0.4	756	524	0.4	791	541	0.4	5222	3396	2.3	762	571	0.4	7306	5955	4.0
	10	835	552	0.4	746	517	0.4	799	532	0.4	5207	3403	2.5	772	564	0.4	7231	5814	4.2
	15	819	555	0.4	735	517	0.4	782	523	0.4	5174	3447	2.5	767	564	0.4	7265	5935	4.3
	20	836	549	0.4	732	519	0.4	790	526	0.4	5187	3414	2.5	776	563	0.4	7407	5932	4.4
	25	826	542	0.4	703	520	0.4	769	516	0.4	5248	3416	2.5	746	552	0.4	7513	5982	4.4
	30	816	555	0.4	706	504	0.4	776	525	0.4	5182	3409	2.5	772	557	0.4	7317	5728	4.3
strat.occ	Nb.	240	240	0.2	240	240	0.2	240	240	0.2	1585	1585	1.1	305	305	0.2	5471	5471	3.9
	0	612	413	0.2	595	437	0.3	621	435	0.3	3410	2511	1.5	651	484	0.3	7305	5999	3.6
	5	591	404	0.3	579	412	0.3	593	421	0.3	3319	2447	1.7	650	482	0.3	7215	6069	4.1
	10	589	388	0.3	588	422	0.3	596	416	0.3	3366	2436	1.8	634	489	0.4	7120	5941	4.3
	15	557	387	0.3	590	414	0.3	572	402	0.3	3349	2459	1.8	660	480	0.4	7014	5992	4.4
	20	551	384	0.3	554	402	0.3	549	411	0.3	3275	2422	1.8	652	493	0.4	7278	5963	4.4
	25	546	374	0.3	517	394	0.3	552	397	0.3	3295	2406	1.8	627	475	0.3	7189	6061	4.5
	30	553	379	0.3	521	383	0.3	532	378	0.3	3228	2441	1.8	617	477	0.4	7060	5973	4.4
LSA q	Nb.	833	833	0.6	833	833	0.6	833	833	0.6	4668	4668	3.3	979	979	0.7	16998	16998	12.1
	0	1977	1349	0.8	1775	1272	0.8	1927	1352	0.8	12161	8297	4.9	2947	2013	1.2	27888	22885	13.6
	5	1976	1347	0.9	1775	1270	0.9	1912	1339	0.9	11949	8167	5.5	2987	2062	1.4	28151	22847	15.5
	10	1969	1339	1.0	1742	1272	0.9	1906	1348	1.0	11968	8159	5.9	3074	2094	1.5	27536	22692	16.5
	15	1982	1337	1.0	1755	1258	0.9	1903	1334	1.0	11917	8215	6.0	3062	2084	1.5	27711	22722	16.6
	20	2000	1349	1.0	1752	1264	0.9	1903	1355	1.0	11818	8152	6.0	3086	2082	1.5	27628	22642	16.6
	25	1963	1331	1.0	1729	1269	0.9	1860	1343	1.0	11814	8143	6.0	3068	2118	1.6	27430	22646	16.7
	30	1945	1337	1.0	1738	1263	0.9	1830	1356	1.0	11760	8067	6.0	3089	2103	1.6	27323	22540	16.8
Routen-TT	Nb.	152	152	0.1	152	152	0.1	152	152	0.1	820	820	0.6	168	168	0.1	3749	3749	2.7
	0	340	225	0.1	289	206	0.1	317	222	0.1	1988	1328	0.8	407	285	0.2	3875	3283	1.9
	5	337	218	0.1	297	214	0.1	323	214	0.1	2027	1331	0.9	409	300	0.2	3897	3273	2.2
	10	345	227	0.2	301	225	0.2	327	225	0.2	2029	1313	1.0	406	282	0.2	3951	3229	2.3
	15	334	233	0.2	292	211	0.2	314	224	0.2	2028	1321	1.0	404	290	0.2	3853	3239	2.4
	20	329	229	0.2	279	213	0.2	310	219	0.2	1976	1311	1.0	401	294	0.2	4115	3229	2.4
	25	329	211	0.2	286	202	0.1	304	214	0.2	1995	1319	1.0	400	289	0.2	4014	3322	2.4
	30	319	206	0.2	289	203	0.2	301	206	0.2	1947	1259	0.9	373	296	0.2	3915	3257	2.4
Link-v/v0	Nb.	2167	2167	1.5	2167	2167	1.5	2167	2167	1.5	10975	10975	7.8	3542	3542	2.5	65113	65113	46.3
	0	4733	3677	2.2	3211	2550	1.5	4075	3202	1.9	24418	18906	11.2	4406	3662	2.2	71711	57324	34.0
	5	4648	3704	2.5	3226	2576	1.7	3993	3125	2.1	24294	18730	12.7	4316	3580	2.4	43091	36782	24.9
	10	4672	3641	2.6	3167	2558	1.9	3974	3171	2.3	23996	18550	13.5	4344	3508	2.5	30887	27736	20.1
	15	4707	3589	2.6	3151	2509	1.8	3969	3112	2.3	24051	18525	13.6	4302	3484	2.5	29428	26697	19.5
	20	4620	3575	2.6	3019	2431	1.8	3929	3091	2.3	23755	18299	13.4	4264	3508	2.6	29973	26798	19.7
	25	4524	3543	2.6	3042	2485	1.8	3859	3133	2.3	23933	18466	13.6	4197	3560	2.6	29597	26200	19.3
	30	4570	3555	2.6	3136	2485	1.9	3798	3039	2.3	23780	18403	13.7	4163	3418	2.5	27906	25356	18.9

Tab. 6.1 Absolute Häufigkeit von Näherungsfunktionen bei drei und fünf gebildeten Abschnitten sowie relative Häufigkeit bei drei Abschnitten für verschiedene Arten von erklärenden und geschätzten Kenngrößen und für unterschiedliche zeitliche Versätze

geschätzt erklärend		strat.q			strat.v			strat.occ			LSA q			Routen-TT			Link-v/v0		
		kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.
strat.q	0	218	218	174	197	213	154	218	218	164	1483	1480	1007	262	268	169	3536	3369	1389
	5	218	218	177	196	211	142	218	217	157	1478	1478	984	262	272	174	3578	3319	1400
	10	218	218	174	203	212	139	217	218	148	1482	1479	978	264	268	167	3537	3368	1449
	15	218	218	169	203	211	128	218	218	152	1480	1480	990	262	268	179	3541	3371	1283
	20	218	218	175	203	211	134	217	217	154	1482	1473	964	264	267	180	3456	3322	1395
	25	218	218	164	203	210	127	218	218	150	1482	1474	971	265	265	176	3513	3324	1381
	30	218	218	164	206	210	129	218	218	147	1480	1473	964	266	267	180	3528	3373	1356
strat.v	0	131	204	209	130	194	204	134	196	211	676	1283	1412	113	201	254	641	2145	3177
	5	141	209	209	120	196	208	132	197	212	701	1278	1417	122	197	252	682	2215	3058
	10	134	208	210	120	196	201	125	195	212	697	1286	1420	110	197	257	720	2085	3009
	15	136	210	209	117	197	203	114	197	212	729	1302	1416	113	196	255	678	2186	3071
	20	134	206	209	119	197	203	116	199	211	696	1303	1415	107	200	256	738	2268	2926
	25	128	205	209	120	198	202	111	196	209	706	1298	1412	100	197	255	802	2176	3004
	30	142	205	208	106	196	202	117	198	210	719	1272	1418	99	200	258	672	2105	2951
strat.occ	0	216	103	94	213	123	101	217	114	104	1492	566	453	277	122	85	4392	1078	529
	5	216	91	97	212	110	90	217	110	94	1491	528	428	276	116	90	4443	1125	501
	10	216	84	88	211	116	95	217	103	96	1493	516	427	277	126	86	4350	1082	509
	15	216	90	81	211	113	90	218	98	86	1491	538	430	278	121	81	4421	1063	508
	20	216	91	77	212	111	79	217	108	86	1490	516	416	277	130	86	4332	1027	604
	25	216	83	75	212	106	76	218	95	84	1491	528	387	278	119	78	4348	1113	600
	30	216	86	77	212	108	63	217	95	66	1492	534	415	278	117	82	4358	1090	525
LSA q	0	750	387	212	723	363	186	748	407	197	4584	2482	1231	948	711	354	12028	7643	3214
	5	750	397	200	722	364	184	747	398	194	4594	2418	1155	949	744	369	11862	7493	3492
	10	750	377	212	717	380	175	746	394	208	4591	2404	1164	937	770	387	11869	7529	3294
	15	750	376	211	715	361	182	747	381	206	4598	2429	1188	943	750	391	11986	7469	3267
	20	750	369	230	719	365	180	746	383	226	4587	2368	1197	940	758	384	12026	7250	3366
	25	750	380	201	720	363	186	748	396	199	4602	2366	1175	946	768	404	11913	7457	3276
	30	750	383	204	718	366	179	750	398	208	4591	2355	1121	955	773	375	11987	7360	3193
Routen-TT	0	124	71	30	111	65	30	116	72	34	793	399	136	159	88	38	2249	743	291
	5	124	72	22	112	69	33	116	70	28	792	377	162	156	92	52	2265	714	294
	10	124	73	30	112	68	45	116	71	38	790	352	171	157	79	46	2275	718	236
	15	124	76	33	111	69	31	116	73	35	792	365	164	159	82	49	2237	728	274
	20	124	72	33	111	68	34	116	67	36	799	351	161	156	86	52	2255	707	267
	25	124	60	27	111	67	24	116	70	28	797	374	148	158	86	45	2292	724	306
	30	124	57	25	111	70	22	116	66	24	795	332	132	159	88	49	2212	767	278
Link-v/v0	0	1290	1638	749	908	1210	432	1137	1470	595	6297	8737	3872	1245	1509	908	16216	23305	17803
	5	1313	1641	750	898	1247	431	1126	1452	547	6284	8651	3795	1192	1530	858	11719	15090	9973
	10	1267	1649	725	915	1222	421	1115	1488	568	6166	8655	3729	1148	1504	856	9326	11656	6754
	15	1261	1640	688	876	1217	416	1069	1462	581	6283	8625	3617	1143	1507	834	8336	11162	7199
	20	1272	1631	672	856	1186	389	1050	1452	589	6123	8575	3601	1132	1529	847	8381	11129	7288
	25	1260	1617	666	872	1195	418	1090	1464	579	6219	8574	3673	1148	1560	852	8584	11010	6606
	30	1266	1606	683	862	1161	462	1099	1396	544	6250	8526	3627	1137	1479	802	8022	10772	6562

Tab. 6.2 Aufteilung der absoluten Häufigkeiten gemäß Tab. 6.1 auf die Abschnitte bei Bildung von drei Abschnitten

geschätzt erklärend		strat.q			strat.v			strat.occ			LSA q			Routen-TT			Link-v/v0		
		kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.
strat.q	0	0.69	0.29	0.07	0.09	0.09	0.05	0.39	0.16	0.06	0.48	0.13	0.04	0.12	0.07	0.06	0.05	0.05	0.08
	5	0.68	0.26	0.06	0.09	0.08	0.04	0.37	0.15	0.05	0.48	0.13	0.04	0.12	0.07	0.06	0.05	0.05	0.09
	10	0.67	0.25	0.05	0.09	0.08	0.03	0.37	0.14	0.04	0.48	0.12	0.04	0.12	0.07	0.07	0.05	0.05	0.08
	15	0.66	0.24	0.04	0.09	0.08	0.03	0.36	0.14	0.04	0.48	0.12	0.04	0.12	0.07	0.06	0.05	0.05	0.08
	20	0.65	0.24	0.05	0.09	0.08	0.03	0.36	0.14	0.03	0.47	0.12	0.04	0.12	0.07	0.06	0.05	0.05	0.08
	25	0.63	0.23	0.05	0.09	0.07	0.03	0.35	0.13	0.04	0.47	0.12	0.03	0.12	0.07	0.06	0.05	0.05	0.09
	30	0.62	0.22	0.04	0.09	0.07	0.02	0.34	0.13	0.03	0.46	0.12	0.03	0.12	0.07	0.05	0.05	0.06	0.08
strat.v	0	0.24	0.17	0.16	0.20	0.19	0.11	0.20	0.18	0.15	0.11	0.10	0.11	0.21	0.08	0.06	0.24	0.07	0.05
	5	0.21	0.15	0.16	0.16	0.14	0.09	0.16	0.15	0.14	0.12	0.10	0.10	0.24	0.09	0.06	0.24	0.07	0.05
	10	0.20	0.15	0.15	0.15	0.13	0.08	0.14	0.14	0.13	0.11	0.09	0.10	0.25	0.08	0.06	0.22	0.07	0.05
	15	0.19	0.14	0.15	0.12	0.11	0.08	0.14	0.13	0.12	0.11	0.09	0.10	0.24	0.07	0.06	0.22	0.07	0.05
	20	0.19	0.14	0.14	0.11	0.10	0.08	0.11	0.11	0.12	0.11	0.09	0.09	0.22	0.07	0.06	0.25	0.05	0.05
	25	0.19	0.14	0.14	0.10	0.09	0.07	0.11	0.11	0.11	0.11	0.09	0.09	0.19	0.07	0.06	0.23	0.06	0.05
	30	0.19	0.13	0.13	0.10	0.08	0.07	0.11	0.10	0.11	0.10	0.09	0.09	0.18	0.06	0.06	0.24	0.06	0.05
strat.occ	0	0.61	0.13	0.21	0.23	0.20	0.19	0.44	0.21	0.20	0.41	0.09	0.13	0.15	0.12	0.21	0.10	0.12	0.26
	5	0.59	0.10	0.17	0.21	0.15	0.16	0.41	0.13	0.15	0.41	0.10	0.12	0.16	0.11	0.19	0.10	0.11	0.28
	10	0.58	0.08	0.17	0.21	0.11	0.18	0.39	0.10	0.16	0.41	0.10	0.12	0.16	0.12	0.19	0.10	0.11	0.28
	15	0.57	0.08	0.15	0.21	0.11	0.16	0.38	0.11	0.15	0.40	0.09	0.12	0.16	0.12	0.18	0.10	0.11	0.30
	20	0.56	0.08	0.17	0.20	0.11	0.13	0.37	0.12	0.12	0.40	0.09	0.12	0.16	0.11	0.21	0.10	0.12	0.26
	25	0.55	0.09	0.16	0.20	0.11	0.10	0.36	0.11	0.11	0.40	0.09	0.12	0.16	0.08	0.19	0.10	0.11	0.28
	30	0.54	0.07	0.14	0.19	0.10	0.12	0.35	0.11	0.11	0.39	0.09	0.11	0.16	0.11	0.21	0.10	0.12	0.27
LSA q	0	0.61	0.16	0.12	0.12	0.08	0.15	0.34	0.11	0.13	0.45	0.15	0.13	0.11	0.06	0.06	0.07	0.06	0.14
	5	0.60	0.16	0.11	0.12	0.08	0.14	0.34	0.11	0.11	0.44	0.13	0.11	0.11	0.06	0.05	0.07	0.06	0.14
	10	0.59	0.15	0.12	0.12	0.08	0.13	0.33	0.10	0.13	0.44	0.13	0.12	0.11	0.06	0.07	0.07	0.07	0.14
	15	0.58	0.15	0.10	0.12	0.08	0.11	0.32	0.10	0.12	0.43	0.13	0.12	0.11	0.06	0.06	0.07	0.06	0.14
	20	0.58	0.15	0.16	0.12	0.08	0.11	0.32	0.10	0.16	0.43	0.12	0.12	0.11	0.06	0.06	0.07	0.07	0.14
	25	0.57	0.14	0.11	0.12	0.07	0.12	0.31	0.10	0.10	0.42	0.12	0.11	0.11	0.06	0.05	0.07	0.06	0.14
	30	0.56	0.15	0.10	0.12	0.07	0.12	0.31	0.10	0.11	0.42	0.12	0.10	0.11	0.06	0.05	0.07	0.06	0.14
Routen-TT	0	0.23	0.08	0.30	0.12	0.13	0.23	0.20	0.10	0.26	0.17	0.07	0.29	0.10	0.12	0.23	0.08	0.16	0.33
	5	0.23	0.08	0.28	0.11	0.10	0.24	0.19	0.09	0.25	0.17	0.06	0.26	0.10	0.12	0.27	0.08	0.18	0.32
	10	0.22	0.08	0.33	0.11	0.09	0.25	0.19	0.08	0.29	0.16	0.06	0.27	0.10	0.12	0.25	0.08	0.17	0.31
	15	0.21	0.08	0.35	0.11	0.08	0.24	0.18	0.08	0.29	0.16	0.07	0.27	0.10	0.10	0.32	0.07	0.17	0.32
	20	0.20	0.07	0.35	0.10	0.07	0.26	0.17	0.06	0.33	0.15	0.07	0.26	0.09	0.11	0.32	0.08	0.17	0.35
	25	0.19	0.07	0.38	0.10	0.07	0.27	0.16	0.06	0.37	0.15	0.07	0.28	0.09	0.09	0.25	0.07	0.16	0.32
	30	0.18	0.08	0.38	0.09	0.07	0.24	0.16	0.06	0.35	0.14	0.07	0.28	0.08	0.09	0.26	0.07	0.17	0.31
Link-v/v0	0	0.07	0.08	0.09	0.08	0.04	0.09	0.08	0.06	0.10	0.07	0.07	0.10	0.17	0.08	0.15	0.22	0.23	0.33
	5	0.07	0.08	0.09	0.08	0.04	0.09	0.09	0.06	0.10	0.07	0.07	0.10	0.18	0.08	0.13	0.21	0.18	0.28
	10	0.07	0.08	0.09	0.08	0.04	0.08	0.08	0.06	0.11	0.07	0.07	0.10	0.18	0.07	0.13	0.22	0.19	0.29
	15	0.07	0.07	0.10	0.07	0.04	0.09	0.08	0.06	0.11	0.07	0.07	0.10	0.17	0.07	0.14	0.23	0.19	0.28
	20	0.07	0.07	0.10	0.07	0.04	0.09	0.07	0.06	0.11	0.07	0.07	0.10	0.18	0.08	0.12	0.23	0.18	0.29
	25	0.07	0.07	0.10	0.06	0.04	0.10	0.07	0.06	0.10	0.07	0.07	0.11	0.16	0.08	0.13	0.23	0.18	0.27
	30	0.07	0.07	0.09	0.06	0.04	0.09	0.07	0.05	0.09	0.07	0.06	0.10	0.15	0.07	0.13	0.24	0.19	0.27

Tab. 6.3 Mittleres Bestimmtheitsmaß der Näherungsgeraden gemäß Tab. 6.2

6.2 Näherungsfunktionen aus einfacher linearer Regression

6.2.1 Lokale Daten als Eingangsgröße

LSA-Detektoren erheben im vorliegenden Netz ausschließlich lokale, fahrstreifenbezogene Verkehrsstärken, strategische Detektoren erheben auch mittlere lokale Geschwindigkeiten und Belegungsgrade als verkehrliche Kenngrößen. Geringe Verkehrsstärken haben – wie im Fundamentaldiagramm ersichtlich wird – die Eigenschaft, dass sie sowohl bei absolut ungestörtem Verkehrsfluss auftreten können als auch im Überlastungsfall und damit während erheblicher verkehrlicher Störungen. Auch wenn das Fundamentaldiagramm in seiner gewohnten Form auf Erhebungen insbesondere auf Autobahnen basiert, ist der Sachverhalt auch auf innerörtlichen Straßen in vergleichbarer Form zu erwarten. Weil Zustände freien Verkehrsflusses weit häufiger vorliegen als Überlastungen, ist damit zu rechnen, dass die ungestörten Verkehrszustände eine Näherungsfunktion für die Schätzung einer anderen verkehrlichen Kenngröße dominieren.

Im Gegensatz zu Verkehrsstärken sind lokale mittlere Geschwindigkeiten und Belegungsgrade eindeutige Indikatoren für einen Verkehrszustand.

Die Bewertung der Näherungsfunktionen mit lokalen Daten als Eingangsgröße wird wie folgt gegliedert:

- Strategisch erhobene Verkehrsstärke (strat.q) als Eingangsgröße
- Strategisch erhobene lokale mittlere Geschwindigkeit (strat.v) als Eingangsgröße
- Strategisch erhobener lokaler Belegungsgrad (strat.occ) als Eingangsgröße
- Mit LSA-Detektoren erhobene Verkehrsstärke (LSA q) als Eingangsgröße
- Schätzung aus Daten vom Nachbarfahrstreifen

Strategisch erhobene Verkehrsstärke (strat.q) als Eingangsgröße

Für einen sehr großen Anteil der untersuchten Beziehungen zwischen Verkehrsstärken, die mit strategischen Detektoren erhoben wurden, und anderen Verkehrskenngrößen konnten statistisch signifikante Zusammenhänge ermittelt werden. Weniger Näherungsfunktionen für wachsende Zeitversätze ergeben sich nur für die kleineren beobachteten Zeitversätze bei Schätzung lokaler Geschwindigkeiten und von Verkehrsstärken an LSA-Detektoren. Für die Schätzung sämtlicher lokaler Daten kann aber ein Absinken des Bestimmtheitsmaßes für wachsende zeitliche Versätze festgestellt werden. So hohe durchschnittliche Bestimmtheitsmaße wie für die Funktionen zur Schätzung von strategisch erhobenen Verkehrsstärken aus anderen strategisch erhobenen Verkehrsstärken werden für keine andere Datenartkonstellation erreicht.

Lokale Daten haben eine sehr hohe Verfügbarkeit. Offensichtlich führt dies zu einem häufigen Bestehen des Signifikanztests. Die Bestimmtheitsmaße sinken stets für ansteigende Verkehrsstärken (von kleine über mittlere bis große Eingangswerte), die Schätzung von Linkgeschwindigkeiten ausgenommen. Ursache hierfür ist deren geringe Datenverfügbarkeit, besonders für selten auftretende Verkehrszustände, also für Störungen. Somit ist der Stichprobenumfang für die Ableitung zugehöriger Näherungsfunktionen gering, so dass das Bestimmtheitsmaß hoch sein muss, damit der Signifikanztest bestanden wird. Dieser Test ist aber offensichtlich eine Schwelle, wie an den vergleichsweise wenigen Funktionen zur Schätzung von Linkgeschwindigkeiten aus großen Verkehrsstärken für alle zeitlichen Versätze abgelesen werden kann. Auch für die Schätzung der anderen Datenarten ist die Verringerung der Anzahl der Schätzgeraden aus hohen Verkehrsstärken feststellbar, jedoch im Verhältnis zur Anzahl der Schätzfunktionen für ungestörte Verkehrszustände in geringerem Maße.

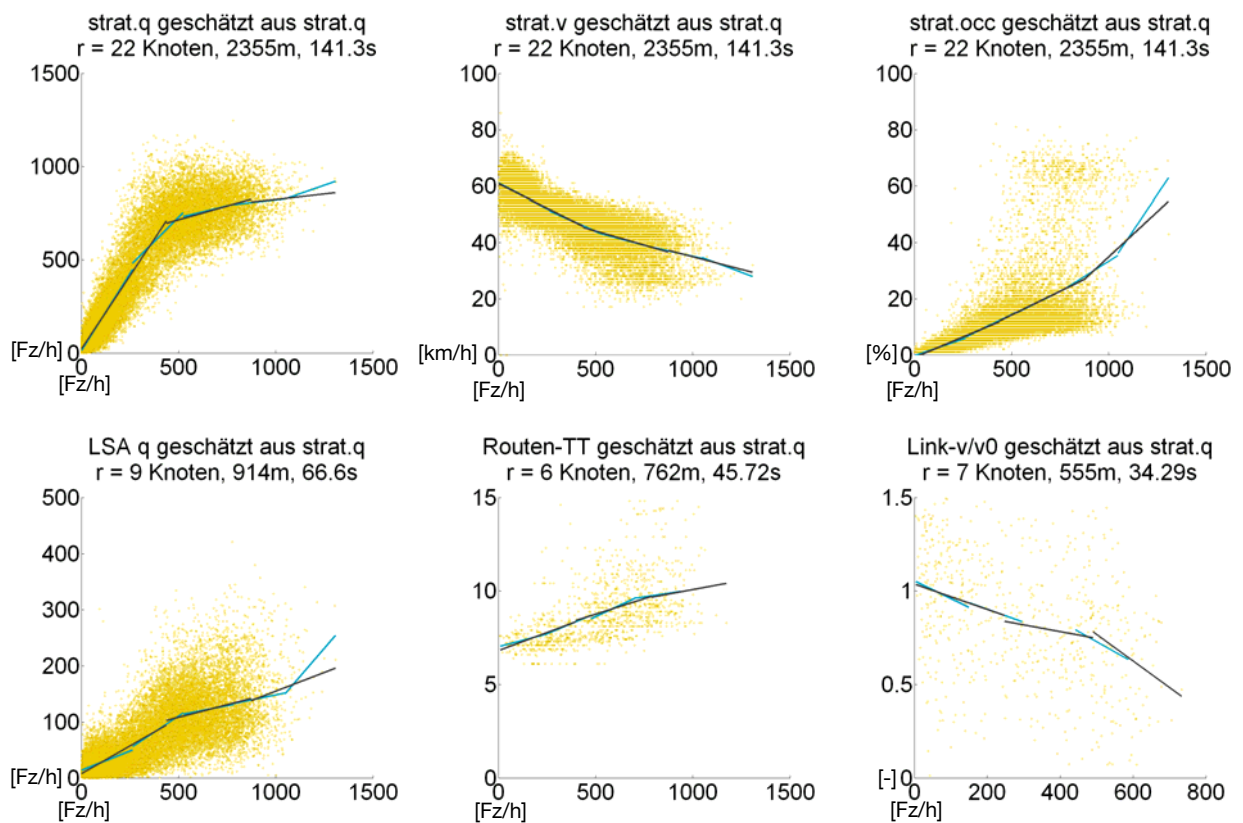


Bild 6.3 Beispiele für Näherungsfunktionen mit Verkehrsstärken aus strategischen Detektoren als erklärende Kenngröße, jeweils für drei (schwarz) und für fünf (blau) Abschnitte, ohne Zeitversatz; Distanz r zwischen den Erhebungsorten in unterschiedlichen Maßeinheiten: Anzahl der dazwischen liegenden Knotenpunkte, Fahrentfernung [m] und freie Reisezeit [s]

Bild 6.3 zeigt ermittelte Näherungsfunktionen mit den Messwert-Punktwolken, die ihnen zu Grunde liegen, jeweils für drei und für fünf gebildete Abschnitte. Die oberen Diagramme zeigen die Zusammenhänge mit den Messwerten aus einem anderen strategischen Detektor in recht großer Entfernung. Da am zu schätzenden Detektor keine sehr geringen Geschwindig-

keiten gemessen wurden, wirkt sich hier die Uneindeutigkeit der Verkehrsstärke nicht in Form einer zweigeteilten Punktwolke aus. Für den Belegungsgrad zeigt sich jedoch, dass sowohl mittlere als auch hohe Werte zum Großteil im selben Abschnitt liegen, die Schätzfunktion aber von den weit zahlreicheren mittleren Werten dominiert wird. Eine solche Näherungsfunktion ist bei der Schätzung einer verkehrlichen Störung nur begrenzt dienlich.

Der dargestellte Zusammenhang zwischen Verkehrsstärken erklärend aus einem strategischen und geschätzt aus einem LSA-Detektor ist unbestreitbar, jedoch kann eine Funktion die recht breit streuende Messwertpunktwolke nur mäßig annähern. Dies schlägt sich in einem vergleichsweise geringen Bestimmtheitsmaß für die einzelnen Näherungsgeraden nieder. Ähnliches gilt für geschätzte Routenreisezeiten und Linkgeschwindigkeiten. Jedoch muss hier ins Gedächtnis gerufen werden, dass die eingehenden Linkgeschwindigkeiten aus Positionsmeldungen für den aktuellen Verkehrszustand am Erhebungsort weniger repräsentativ sein können als andere Kenngrößen. Näherungsgeraden wie dargestellt können aber einen zusätzlichen Teil an Repräsentativität liefern, da sie auf einer Vielzahl von Messungen basieren. So kann der Fall eintreten, dass der Schätzwert sogar besser ist als der Messwert, der als Referenz herangezogen wird.

Eine größere Anzahl von Abschnitten bei der Berechnung der Näherungsgeraden ermöglicht die differenziertere Darstellung von nichtlinearen Zusammenhängen und die nach Datenbereichen differenzierte Untersuchung der Qualität eines Zusammenhangs. Die sich ergebenden Unterschiede hinsichtlich des Verlaufs der Näherungsfunktionen für die dargestellten Punktwolken sind aber gering. Schwerer wiegt der Nachteil für eine höhere Abschnittszahl, dass die Abschnitte schmaler werden und in der Folge weniger Punkte enthalten. Dadurch wird das Bestehen des Signifikanztests schwieriger. Als Konsequenz können weniger Abschnitte zu einer besseren Abdeckung des gesamten Datenbereichs mit Näherungsfunktionen führen. Dies kann in Bild 6.3 für die Schätzung von Routenreisezeiten und von Linkgeschwindigkeiten beobachtet werden, ohne dass die Näherungsfunktionen erkennbar an Qualität verlieren.

Strategisch erhobene mittlere lokale Geschwindigkeit (strat.v) als Eingangsgröße

Die Anzahl der abschnittsbezogenen Näherungsgeraden mit strategisch erhobenen mittleren lokalen Geschwindigkeiten als Eingangsgröße liegt gemäß Tab. 6.1 durchgehend niedriger als für strategisch erhobene Verkehrsstärken als Eingangsgröße. Tab. 6.2 deckt weitere Unterschiede auf: Anders als im Fall zuvor geht die Anzahl der Näherungsgeraden für alle geschätzten Datenarten bereits für mittlere Geschwindigkeiten zurück. Noch deutlicher ist der Rückgang für kleine Geschwindigkeiten, also für Störungssituationen. Bereits für hohe Eingangswerte werden weniger Näherungsfunktionen gebildet als zuvor für geringe Verkehrsstärken, die dem Signifikanztest standhalten.

Hohe Verkehrsstärken treten noch vor geringen lokalen Geschwindigkeiten auf und liegen somit häufiger vor. Zur Schätzung von Linkgeschwindigkeiten aus großen strategisch erhobenen Verkehrsstärken liegen bei den signifikanten Zusammenhängen durchschnittlich 104 Messwertpaare vor, für Schätzungen aus kleinen Geschwindigkeiten sind es nur 39. Für die Schätzung von Routenreisezeiten beträgt das Verhältnis 1.002:149. So kommt als Ursache für die kleinere Anzahl der Näherungsfunktionen mit geringen lokalen Geschwindigkeiten als Eingangsgröße insbesondere ein geringerer Stichprobenumfang in Frage. Dies wird durch Tab. 6.3 bestätigt: Sie dokumentiert durchgehend steigende durchschnittliche Bestimmtheitsmaße für sinkende Geschwindigkeiten.

Als weitere Ursache dafür, dass weniger Schätzfunktionen mit lokalen Geschwindigkeiten als mit Verkehrsstärken als Eingangsgröße ermittelt werden, können weniger deutliche Zusammenhänge erwogen werden.

Für steigende zeitliche Versätze gilt fast durchgängig, dass das Bestimmtheitsmaß abnimmt. Ausnahmen gibt es nur für die Schätzung von Routenreisezeiten und Linkgeschwindigkeiten aus geringen lokalen Geschwindigkeiten. Hier verharrt es auf konstantem Niveau.

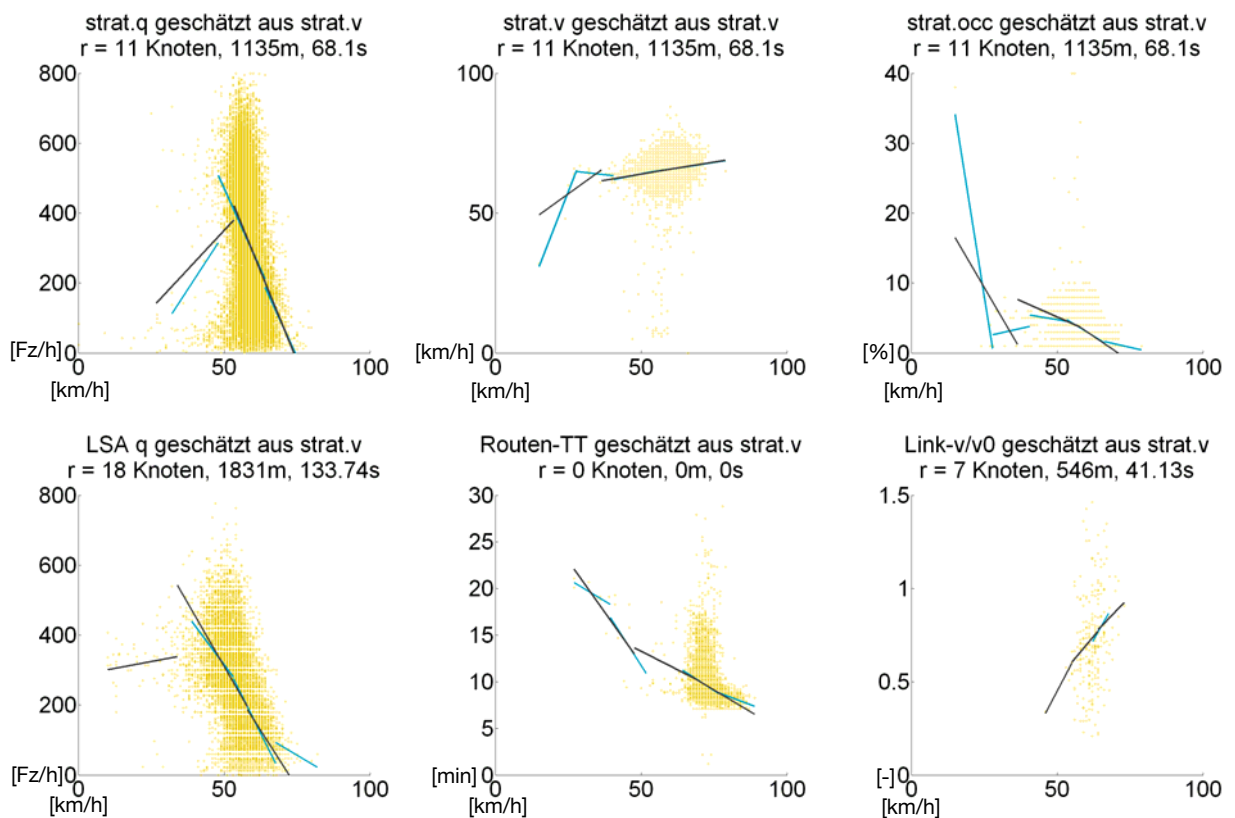


Bild 6.4 Beispiele für Näherungsfunktionen mit lokalen mittleren Geschwindigkeiten aus strategischen Detektoren als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz

Den bisherigen Beobachtungen entsprechend zeigen sich in den Beispielen gemäß Bild 6.4 nur sehr wenige Messungen geringer lokaler Geschwindigkeiten. Dies führt zu weniger Näherungsfunktionen mit höherem Bestimmtheitsmaß für geringe Geschwindigkeiten als für hohe Verkehrsstärken. Kritischer ist jedoch zu bewerten, dass die zu schätzenden Kenngrößen auch für hohe und mittlere Geschwindigkeiten recht breit streuen. Dies ist auch an Tab. 6.3 ablesbar: Die mittleren Bestimmtheitsmaße sind für ungestörten Verkehr grundsätzlich niedriger als für Schätzungen aus anderen lokalen Verkehrsdatenarten. (Hinweis: hier muss der Abschnitt „Kleine Eingangswerte“ für Geschwindigkeiten mit dem Abschnitt „Große Eingangswerte“ für Verkehrsstärke und Belegungsgrad verglichen werden.)

Die Funktionen zur Schätzung von Verkehrsstärken zeigen wie erwünscht, jedoch in unterschiedlicher Ausprägung einen Rückgang der Schätzwerte für geringe lokale Geschwindigkeiten. Außerdem beschreibt der Verlauf der Funktionen zur Schätzung der Linkgeschwindigkeiten einen eindeutigen Zusammenhang und stetige Übergänge an den Abschnittsgrenzen. Ein Grund für die recht diffuse Punktwolke ist die begrenzte Repräsentativität der ermittelten Linkgeschwindigkeiten.

Die Bildung von fünf Abschnitten liefert in den Beispielen keinen erkennbaren Mehrwert gegenüber der Darstellung mit drei Abschnitten.

Strategisch erhobener lokaler Belegungsgrad (strat.occ) als Eingangsgröße

Die Anzahl der Funktionen zur Schätzung lokaler Daten aus strategisch erhobenen lokalen Belegungsgraden geht für steigende Zeitversätze im Grundsatz zurück, auch wenn es dabei zu einigen Schwankungen kommt (siehe Tab. 6.1). Tab. 6.2 zeigt, dass dieser Rückgang überwiegend die Schätzungen aus hohen und mit Abstrichen auch aus mittleren Belegungsgraden betrifft, die Störungssituationen dokumentieren und vergleichsweise selten auftreten. Die Anzahl der Näherungsfunktionen sinkt über die Abschnitte deutlicher ab als für die zuvor diskutierten Eingangsgrößen. Schon für mittlere Belegungsgrade werden deutlich weniger Näherungsfunktionen ermittelt als für kleine, anders als sich das für Verkehrsstärken und Geschwindigkeiten (beim Übergang von großen zu mittleren Geschwindigkeiten) darstellt. Dies ist darauf zurückzuführen, dass der Belegungsgrad die meiste Zeit auf recht geringem Niveau verharrt und nur für sehr seltene Verkehrssituationen deutlich ansteigt. Folge ist ein im Vergleich noch geringerer Stichprobenumfang bei der Suche nach einer Näherungsfunktion für mittlere und besonders für hohe Belegungsgrade, der zu Schwierigkeiten beim Bestehen des Signifikanztests führt. Diese Begründung wird durch die Entwicklung der mittleren Bestimmtheitsmaße in Tab. 6.3 untermauert.

Bild 6.5 zeigt Beispiele für Näherungsfunktionen mit strategisch erhobenen Belegungsgraden als Eingangsgröße. Insbesondere wenn die Belegungsgrade einen großen Skalenbereich abdecken, sind mitunter deutliche Unstetigkeiten an den Abschnittsübergängen zu verzeichnen. Die Schätzfunktionen für Verkehrsstärken geben beide ein Absinken der Ver-

kehrstärke für hohe Belegungsgrade wider, wie es qualitativ den Erwartungen entspricht. Dies ist insbesondere hinsichtlich der jeweils erheblichen Entfernungen zwischen den Erhebungsorten beachtlich. Die wechselnden Vorzeichen für die Steigungen der Abschnittsgeraden zur Schätzung lokaler Geschwindigkeiten und Belegungsgrade entsprechen nicht den Erwartungen. Mögliche Ursachen hierfür sind ein schlechter Zusammenhang zwischen den untersuchten Kenngrößen oder die wenigen Messwerte für mittlere Belegungsgrade. Der Umstand tritt bei der Bildung von fünf Abschnitten ebenso auf wie im Fall von drei Abschnitten. Nur bei der Schätzung von Verkehrsstärken zeigt sich eine Verbesserung durch die zahlreicheren Abschnitte. Dies ist darauf zurückzuführen, dass geringe Verkehrsstärken sowohl bei freiem als auch bei gestörtem Verkehrsfluss auftreten können und die Zusammenhänge daher im Vergleich stärker einer Parabel ähneln als sonstige Zusammenhänge. Zur Annäherung einer Parabel sind drei Abschnitte aber recht knapp bemessen. Jedoch führen auch hier die zusätzlichen Abschnitte mit den zurückgehenden Stichprobenumfängen zu Schwierigkeiten beim Signifikanztest und damit dazu, dass die Näherungsfunktionen einen geringeren Teil der Skala für Eingangswerte abdecken.

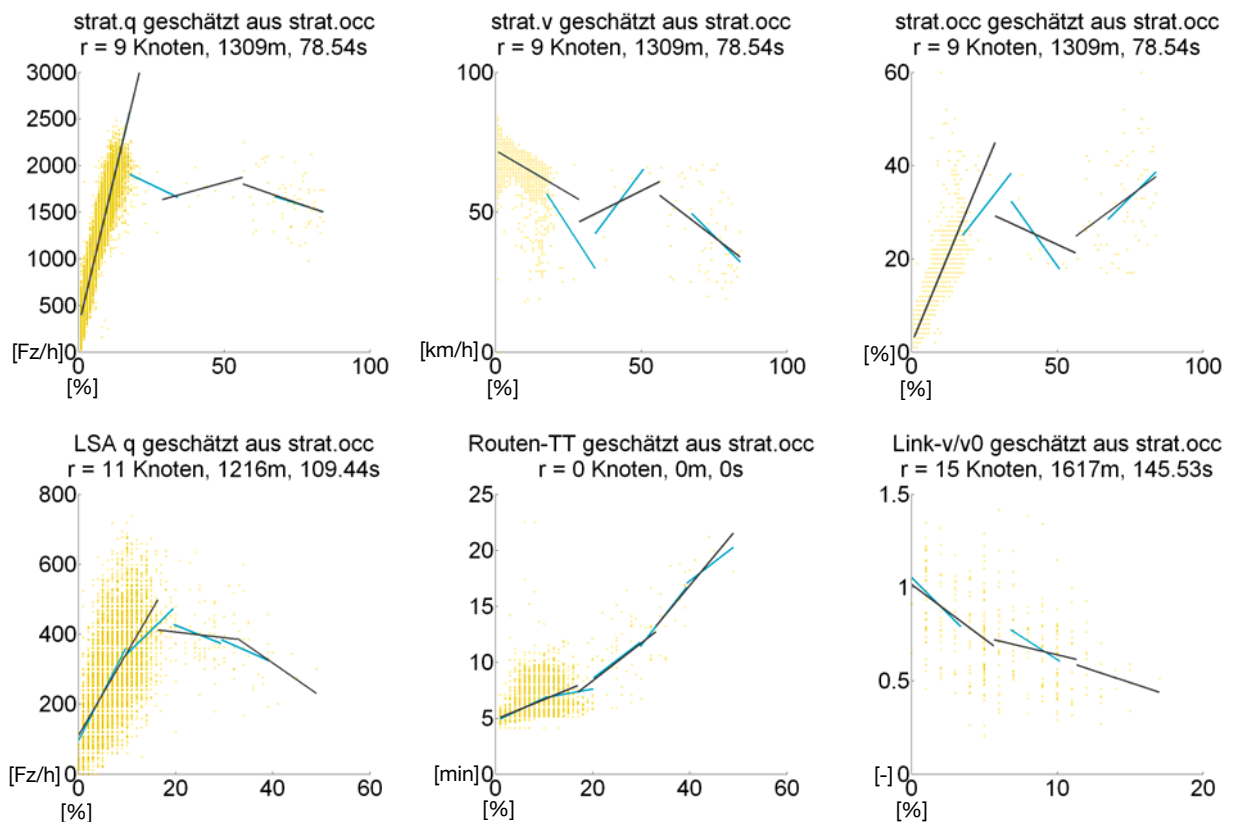


Bild 6.5 Beispiele für Näherungsfunktionen mit Belegungsgraden aus strategischen Detektoren als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz

Die Regressionsanalyse liefert für die Schätzung von Routenreisezeiten und Linkgeschwindigkeiten im Beispiel gemäß Bild 6.5 erfreuliche Ergebnisse. Der strategische Detektor liegt

auf der reisezeitdetektierten Route und erlaubt offensichtlich eine gute Schätzung der Routenreisezeit aus dem lokalen Belegungsgrad, bemerkenswerterweise auch für hohe Routenreisezeiten. Der recht stetige und durchgängig fast lineare Zusammenhang zwischen Belegungsgrad und Linkgeschwindigkeit ist vor dem Hintergrund der großen Entfernung zwischen den Erhebungsorten positiv zu bewerten.

Mit LSA-Detektoren erhobene Verkehrsstärke (LSA q) als Eingangsgröße

Die erwarteten Muster für Schätzungen aus Verkehrsstärken, die mit LSA-Detektoren erhoben wurden, liegen in Tab. 6.1 und Tab. 6.2 nur sehr begrenzt vor. Die Anzahl der Näherungsgeraden sinkt mit Ausnahme der Schätzung von Routenreisezeiten für die ersten ein bis drei ausgewerteten Zeitversätze leicht ab. Dieses Absinken verteilt sich für unterschiedliche geschätzte Kenngrößenarten unterschiedlich auf Abschnitte für kleine, mittlere oder große Verkehrsstärken als Eingangsgröße. Im Vergleich zu Schätzungen aus strategisch erhobenen Verkehrsstärken ist bemerkenswert, dass die Anzahl statistisch signifikanter Zusammenhänge schon für mittlere Verkehrsstärken deutlich gegenüber der für geringe Verkehrsstärken zurückgeht. Dies gilt für sämtliche geschätzte Datenarten. Der Rückgang für große Verkehrsstärken fällt nochmals drastisch aus.

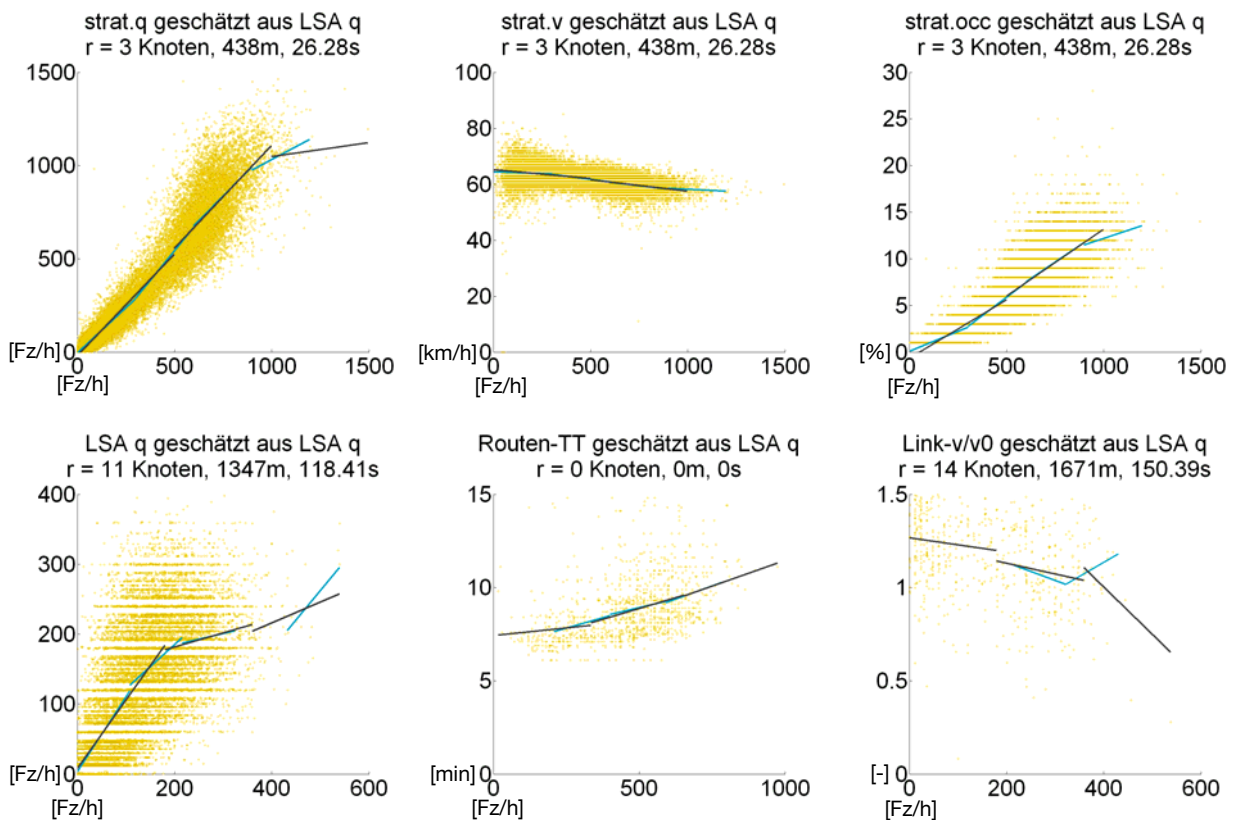


Bild 6.6 Beispiele für Näherungsfunktionen mit Verkehrsstärken aus LSA-Detektoren als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz

Die Zusammenhänge, die in Bild 6.6 als Beispiele dargestellt sind, zeigen Ergebnisse, die nicht schlechter ausfallen als für Schätzungen aus Verkehrsstärken, die mit strategischen Detektoren erhoben werden. Hervorzuheben sind die Funktionen zur Schätzung von strategisch erhobenen lokalen Daten. Der Näherungsfunktion zur Schätzung der Verkehrsstärke aus einem LSA-Detektor liegt zwar eine recht diffuse Punktwolke zu Grunde, dennoch sind die Abschnittsübergänge fast stetig. Ähnliches gilt für die Schätzung von Routenreisezeiten und von Linkgeschwindigkeiten. Bei letztgenannten kann aufgrund der zu hinterfragenden Repräsentativität des Schätzwerts die Schätzung wie erläutert einen besseren Wert liefern als die Messung.

Die Vorteile der Bildung von fünf statt drei Abschnitten sind auch hier begrenzt: Zum größten Teil liegen die Näherungsgeraden für beide Fälle nahezu übereinander. Bei der Abdeckung der Skalen für Eingangswerte finden sich hier neben Nach- auch Vorteile, da für die Schätzung von strategisch erhobenen lokalen mittleren Geschwindigkeiten und Belegungsgraden für hohe Eingangswerte bei fünf Abschnitten eine bessere Abdeckung erreicht wird.

Schätzung aus Daten vom Nachbarfahrstreifen

Strategische Detektoren auf benachbarten Fahrstreifen sollten zur gleichen Zeit recht ähnliche Messwerte liefern. Daher sollte die Schätzung der Werte des einen Detektors aus denen des benachbarten gut gelingen. Bild 6.7 zeigt die ermittelten Näherungsfunktionen für einen solchen Fall. Die Näherungsgeraden oben links, mittig und unten rechts werden als Mittelgeraden – also als Ursprungsgerade mit der Steigung 1 – erwartet, da hier geschätzte und gemessene Kenngrößenart dieselben sind. Direkte Zusammenhänge zwischen den beiden Detektoren unterstellt und den Belegungsgrad als linear abhängig von der Verkehrsdichte angenommen, könnten sich für die Konstellationen

- Geschwindigkeit geschätzt aus Verkehrsstärke
- Geschwindigkeit geschätzt aus Belegungsgrad
- Verkehrsstärke geschätzt aus Belegungsgrad

Punktwolken ergeben, wie sie die Darstellungen im Fundamentaldiagramm liefern, auch wenn dieses im Gegensatz zu den ausgewerteten Daten insbesondere für Autobahnen ermittelt wurde.

Die drei noch ausstehenden Diagramme in Bild 6.7 sollten sich aus den Darstellungen im Fundamentaldiagramm ergeben, jedoch jeweils mit vertauschten Ein- und Ausgangsgrößen.

Beim Betrachten der Bereiche, die die erhobenen Daten auf ihren Skalen jeweils abdecken, fällt auf, dass an Detektor 1 im Wesentlichen keine Geschwindigkeiten kleiner als 20 km/h erhoben wurden. Zudem liegen die Maximalwerte für Verkehrsstärke und Belegungsgrad an Detektor 1 höher als an Detektor 2. Ob dies auf die Verkehrsverhältnisse oder auf Detektor-

eigenschaften zurückzuführen ist, kann hier nicht geklärt werden. Der Umstand wirkt sich aber auf die ermittelten Näherungsfunktionen aus.

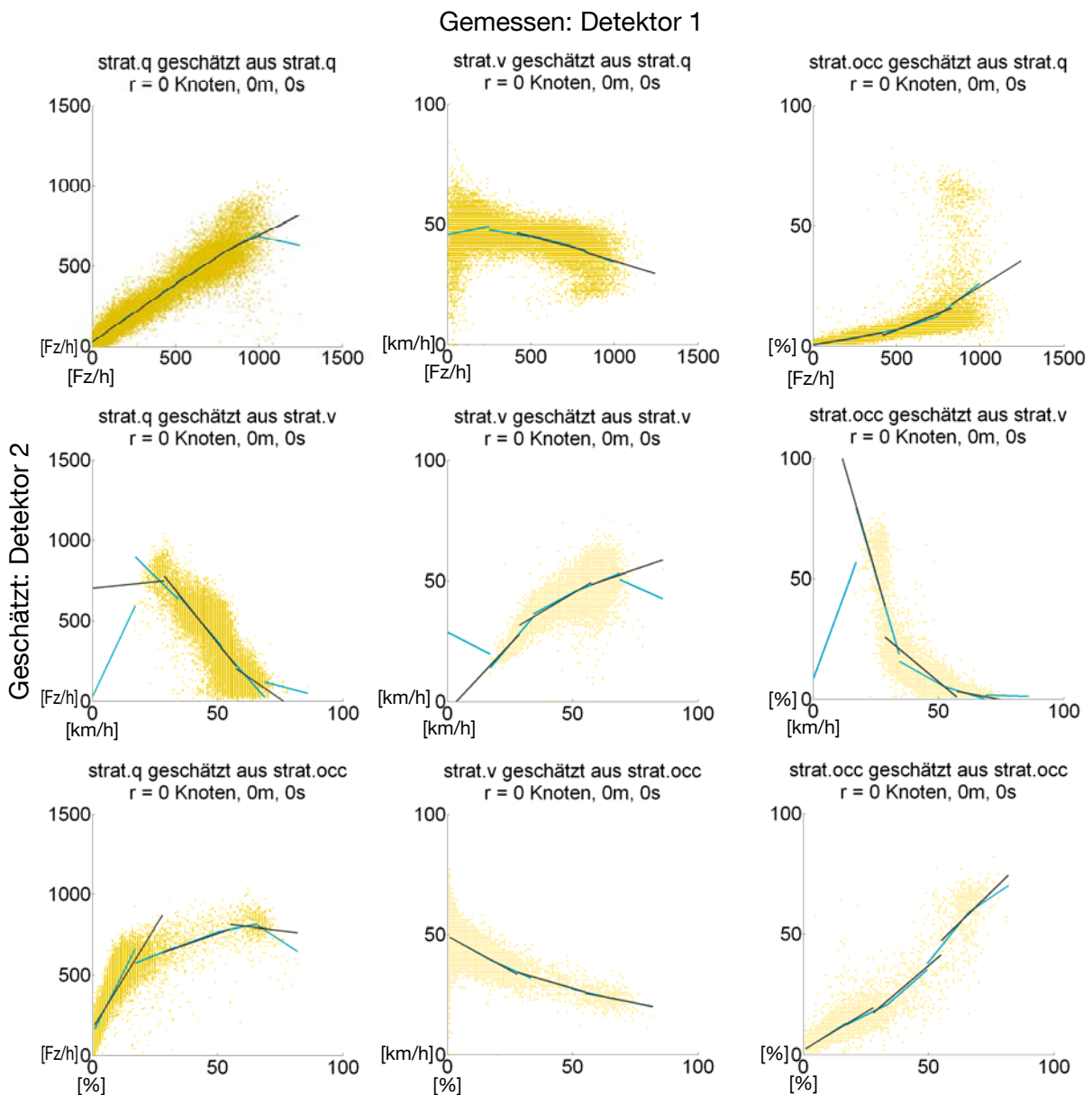


Bild 6.7 Näherungsfunktionen für strategisch erhobene lokale Kenngrößen mit den Messwerten eines strategischen Detektors auf dem benachbarten Fahrstreifen als Eingangsgröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte

Die Näherungsfunktionen in der Diagonalen zeigen im Wesentlichen die erwartete Charakteristik. Da wie erwähnt an Detektor 1 höhere Verkehrsstärken und Belegungsgrade als an Detektor 2 erhoben wurden, liegen die Verläufe für die Näherung derselben Datenart an Detektor 2 aus Detektor 1 etwas unterhalb der Mittelgeraden.

Die Punktwolke für die Schätzung der Geschwindigkeit aus der Verkehrsstärke in Bild 6.7 zeigt, dass auch innerorts die Verkehrsstärke kein eindeutiger Indikator für die Qualität des

Verkehrsablaufs ist, da sie sowohl für freien Verkehr bei hohen Geschwindigkeiten als auch für gestörten Verkehrsfluss bei sehr geringen Geschwindigkeiten einen geringen Wert annehmen kann. Die Häufigkeit des gestörten Verkehrsflusses, wie er im Beispiel dokumentiert ist, reicht offensichtlich aus, um für den betroffenen Skalenabschnitt bei Bildung von drei Abschnitten das Scheitern am Signifikanztest zu erwirken. Treten Störungen des Verkehrsflusses sehr selten auf, so dominieren die Punkte für ungestörten Verkehrsfluss den Abschnitt in einer Art und Weise, dass die geringen Geschwindigkeitswerte bei der Regressionsanalyse näherungsweise übergangen werden.

Auch die darüber hinaus ermittelten Näherungsfunktionen bestätigen grundsätzlich die Erwartungen. Jedoch fällt der Rückgang der geschätzten Verkehrsstärken für hohe Belegungsgrade recht gering aus. Ursache hierfür könnte sein, dass Störungen in städtischen Netzen in den meisten Fällen durch Engpässe entstehen und so auch im Fall einer Störung eine gewisse Verkehrsstärke weiter erreicht wird. Ist beispielsweise eine Lichtsignalanlage nicht überstaut und verursacht ihrerseits einen Rückstau, so sind in ihrem Zufluss sowohl für die Verkehrsstärke als auch für den Belegungsgrad hohe Werte zu erwarten.

Der Rückgang der Geschwindigkeiten für steigende Belegungsgrade stellt sich im Beispiel erwartungsgemäß ein. Er nimmt für steigende Belegungsgrade ab.

Auch hier bestätigt sich, dass die Bildung von fünf statt drei Abschnitten sich ausschließlich bei der Schätzung von Verkehrsstärken aus Geschwindigkeiten oder Belegungsgraden positiv auf die Ergebnisse auswirkt.

6.2.2 Abschnittsdaten als Eingangsgröße

Für Näherungsfunktionen zur Schätzung aus Routenreisezeiten (Routen-TT), die infrastrukturseitig durch Fahrzeugwiedererkennung erhoben wurden, werden in Tab. 6.1 die erwarteten Muster nicht deutlich erkennbar. Tab. 6.2 zeigt jedoch insbesondere für große Routenreisezeiten und größere zeitliche Versätze sinkende Anzahlen der ermittelten signifikanten Näherungsgeraden. Darüber hinaus ist in Tab. 6.3 für kleinere Routenreisezeiten fast durchgängig ein Absinken des mittleren Bestimmtheitsmaßes erkennbar. Für mittlere Reisezeiten treten beide Effekte gelegentlich auf. Die Zahlen untermauern die bereits zuvor diskutierten Eigenschaften von Kombinationen aus Stichprobenumfang und Bestimmtheitsmaß, die für die Signifikanz eines Zusammenhangs maßgeblich sind.

Die Beispiele in Bild 6.8 oben zeigen für kleine und mittlere Reisezeiten brauchbare Näherungsgeraden. Die Funktionen im dritten Abschnitt verlaufen nicht erwartungsgemäß. Es ist jedoch anzumerken, dass bereits Reisezeiten im mittleren Sektor sehr selten auftreten und damit besondere verkehrliche Situationen dokumentieren. Die angegebene Distanz zwischen beiden Kenngrößen von 0 gibt wider, dass der Querschnitt, dessen Daten auf der vertikalen Achse angetragen sind, auf der reisezeitdetektierten Route liegt.

Es wird ersichtlich, dass für den Fall der Bildung von fünf Abschnitten die Näherungsfunktionen für lokale Daten den Erwartungen besser entsprechen. Jedoch wird der Signifikanztest in relevanten Abschnitten nicht bestanden, zudem treten die Verbesserungen vor allem in Bereichen auf, in denen sehr wenige Daten vorliegen. Somit könnten die Verbesserungen der Schätzfunktionen nur in sehr seltenen Fällen ihre Wirkung entfalten. Die Form der Punktwolken rechtfertigt auch hier die Bildung von fünf Abschnitten nur für die Schätzung von Verkehrsstärken.

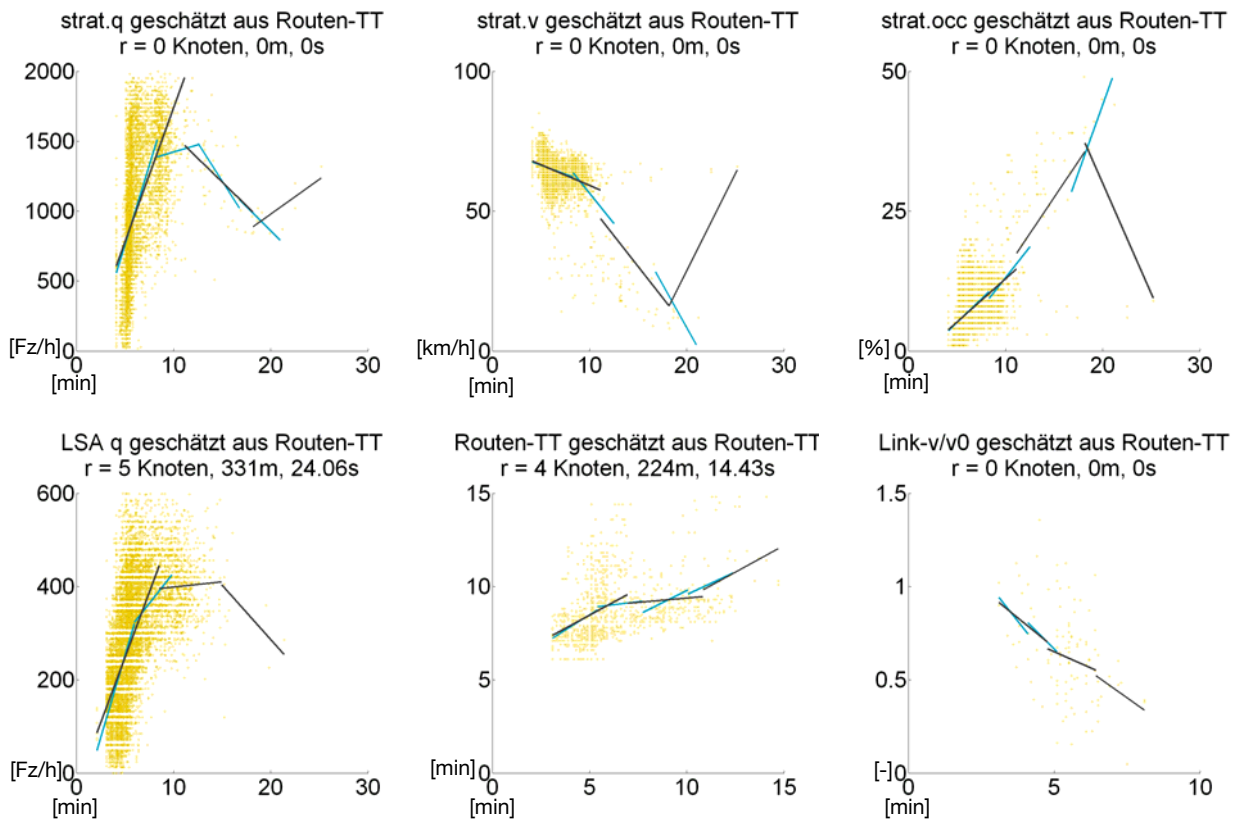


Bild 6.8 Beispiele für Näherungsfunktionen mit Routenreisezeiten als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz

Die Punktwolken in den Beispielen gemäß Bild 6.8 unten sind recht diffus, liefern jedoch Näherungsfunktionen, wie sie erwartet und gewünscht sind. Links ist zu erkennen, dass für sehr hohe Reisezeiten die Verkehrsstärken wieder zurückgehen. Das rechte Beispiel dokumentiert auf einem Link der Route einen Rückgang der Geschwindigkeiten für den Fall, dass die Reisezeit auf der Route ansteigt. Das mittige Diagramm zeigt, dass auch für Routen, die sich nicht überlappen, ein signifikanter Zusammenhang zwischen den Reisezeiten bestehen kann.

6.2.3 Aufbereitete Positionsdaten als Eingangsgröße

Wie bereits mehrfach erwähnt, ist die Repräsentativität von Linkgeschwindigkeiten, die aus Positionsmeldungen abgeleitet sind (bezogen auf linkbezogene freie Geschwindigkeiten: $\text{Link-v}/v_0$), stärker zu hinterfragen als die infrastrukturseitig erhobener aggregierter Werte. Jedoch ist unter den verfügbaren ausschließlich diese Datenart netzweit auf jedem Link erhebbar, so dass ohne Verkehrsflussmodell nur unter Verwendung dieser fahrzeuggenerierten Informationen eine netzweite Verkehrslage bestimmt werden kann. Aufgrund der eingeschränkten Verfügbarkeit werden Linkgeschwindigkeiten häufiger Ausgangs- statt Eingangsgröße einer Schätzung sein, dennoch soll auch ihre Eignung als erklärende Kenngröße untersucht werden.

Tab. 6.1 zeigt für alle zu schätzende Datenarten durchgängig für steigende zeitliche Versätze den erwarteten Rückgang der Anzahl der Näherungsfunktionen, auch wenn es dabei mitunter zu Schwankungen kommt. Die Verteilung auf Abschnitte für kleine, mittlere und große Linkgeschwindigkeiten gemäß Tab. 6.2 folgt keinem strengen Muster. Rückgänge des mittleren Bestimmtheitsmaßes für wachsende zeitliche Versätze finden sich in Tab. 6.3 nur vereinzelt.

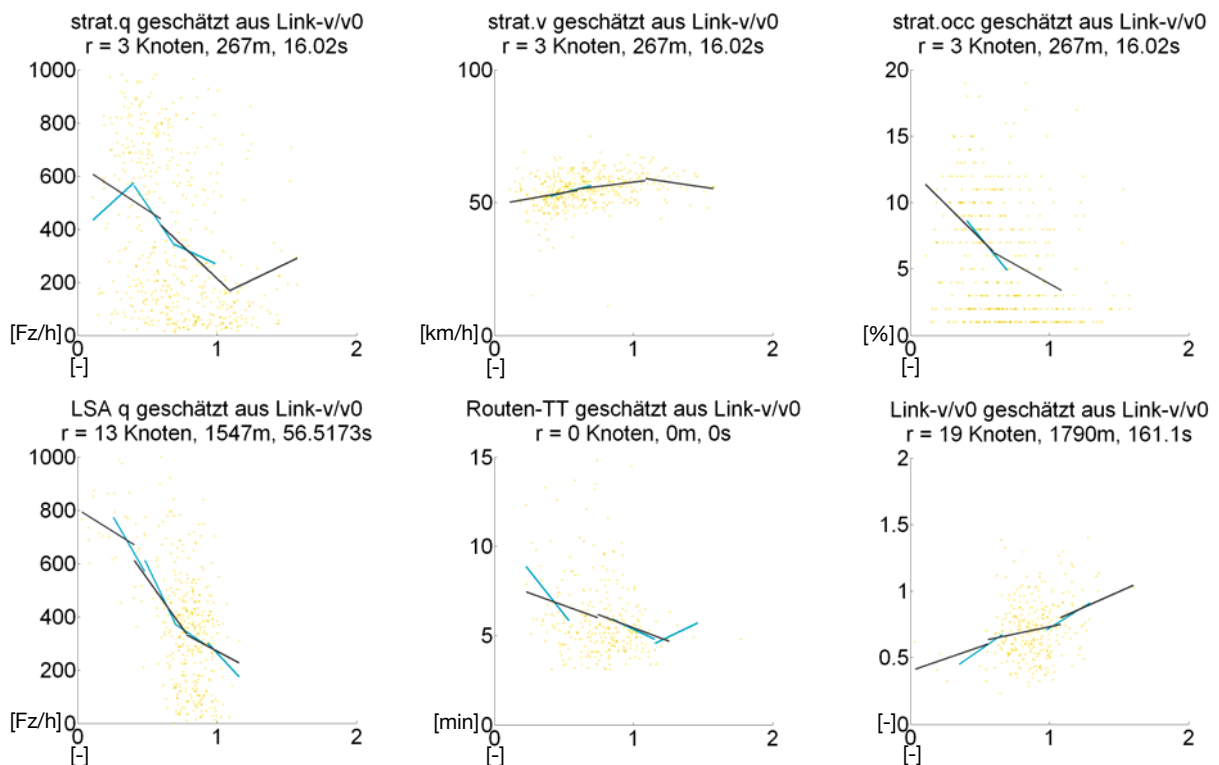


Bild 6.9 Beispiele für Näherungsfunktionen mit Linkgeschwindigkeiten als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz

Der aufgrund der geringen Verfügbarkeit von Linkgeschwindigkeiten geringere Stichprobenumfang führt in den Beispielen (Bild 6.9) wiederholt dazu, dass nur für zwei der drei gebilde-

ten Abschnitte eine signifikante Näherungsgerade bestimmt werden kann. Dieser Umstand verstärkt sich bei der Bildung von fünf Abschnitten wie erwartet. Die Punktwolken streuen weiter als in den zuvor angeführten Beispielen, liefern jedoch dennoch den Erwartungen entsprechende Ergebnisse der Regressionsanalyse mit nahezu stetigen Übergängen an den Abschnittsgrenzen. Einzig die steigende strategisch erhobene Verkehrsstärke für steigende Linkgeschwindigkeiten im rechten Abschnitt widerspricht der Theorie. Der Verlauf der Näherungsgeraden zur Schätzung von Linkgeschwindigkeiten ist vor dem Hintergrund der recht großen Entfernung zwischen den Erhebungsorten beachtlich. In allen Abschnitten weisen sie eine deutliche Steigung auf. Die Übergänge an den Abschnittsgrenzen sind nahezu stetig. Die diffuse Punktwolke führt zu Abzügen bei der Ermittlung des Bestimmtheitsmaßes.

Jedoch weist die Charakteristik der vorliegenden Linkgeschwindigkeiten für die Regressionsanalyse und die Ersatzwertbildung gewisse Tücken auf. So ist für ein Fahrzeug, das seine Position meldet, bei der Datenübermittlung nicht die Begrenzung von Links, sondern das Fortschreiten der Zeit relevant. Allen Links, die während eines Übertragungsintervalls befahren wurden, wird nachfolgend derselbe Quotient aus linkbezogener Geschwindigkeit und linkbezogener freier Geschwindigkeit zugewiesen. Da die allermeisten zu Fünf-Minuten-Intervallen aggregierten Werte auf nur einem Fahrzeug basieren, hat dies zur Konsequenz, dass auch aggregierte Werte für nahe beieinander liegende Links im selben Zeitschritt denselben Wert annehmen. Dadurch treten folgende Umstände ein, die es bei der weiteren Bewertung zu berücksichtigen gilt:

- Es ergeben sich zahlreiche Punktwolken und Näherungsfunktionen wie in Bild 6.10, bei denen sehr viele Punkte auf der Mittelgeraden liegen, da für beide Kenngrößen derselbe Wert abgelegt wurde. Dies führt zu vergleichsweise hohen Werten für Bestimmtheitsmaß und Wert der Teststatistik \hat{t} .
- Es können mehrere Schätzungen aus demselben Eingangswert erfolgen, der für verschiedene Links vorliegt, der aber nicht unabhängig voneinander für die verschiedenen Links erhoben wurde. Dies ist insbesondere für die Zusammenführung unterschiedlicher Schätzwerte von Belang.
- Bei der Plausibilisierung kann sich ein linkbezogener Schätzwert aus dem Kennwert für den Nachbarlink selbst bestätigen.

Bei der Regressionsanalyse mit zeitversetzt erhobenen Daten stellt sich diese Charakteristik nicht mehr ein, lediglich für die erste Stufe mit fünf Minuten Versatz kann er die Ergebnisse in deutlich weniger Fällen beeinflussen, wenn zwischen zwei Positionsmeldungen nicht nur mehrere Links, sondern auch mehrere Aggregierungsintervalle abgedeckt werden. Da der maximale zeitliche Abstand zwischen zwei Positionsmeldungen, die zu einer Fahrt gehören, zu 150s festgesetzt wurde, können maximal zwei Zeitintervalle betroffen sein.

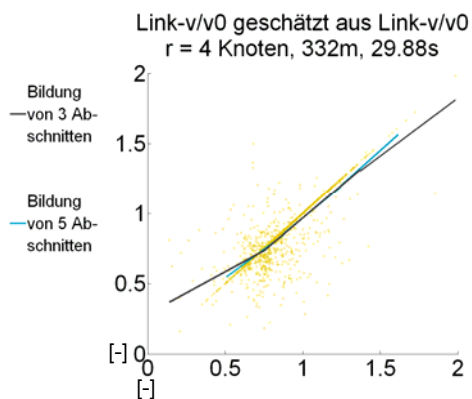


Bild 6.10 Näherungsfunktion für Linkgeschwindigkeiten aus Linkgeschwindigkeiten bei nahegelegenen Links

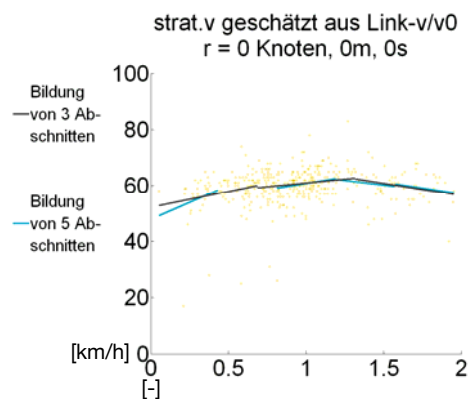


Bild 6.11 Mäßiger Zusammenhang zwischen lokaler und linkbezogener Geschwindigkeit

Die Wirkungen auf die Verhältnisse in Tab. 6.1 sind erheblich: Die Anzahl der gefundenen statistisch signifikanten Zusammenhänge zur Schätzung von Linkgeschwindigkeiten aus Linkgeschwindigkeiten geht von 0 zu 5 min Zeitversatz drastisch, von 5 zu 10 min nochmals deutlicher als für alle anderen Konstellationen von Kenngrößenarten zurück. Da sich bei größeren Zeitversätzen recht konstante Verhältnisse für die relativen Häufigkeiten der Näherungsgeraden unterschiedlicher Konstellationen einstellen, kann der Schluss gezogen werden, dass bei der Regressionsanalyse ohne zeitlichen Versatz 25.000 bis 30.000 der Geraden vom zuvor geschilderten Umstand beeinflusst werden.

Die relative Häufigkeit der Näherungsgeraden liegt nur für die Schätzung von Linkgeschwindigkeiten aus Linkgeschwindigkeiten deutlich unter dem Anteil der diesbezüglich untersuchten Zusammenhänge. Dies ist zu einem gewissen Anteil sicher Folge der geringeren Repräsentativität vorliegender Werte. Viel deutlicher wirkt sich hier jedoch die geringe Datenverfügbarkeit aus: Werden zwei selten erhobene Kenngrößen auf einen Zusammenhang hin untersucht, so liegen sie noch viel seltener im selben Zeitschritt vor.

Bild 6.11 zeigt einen wenig zufriedenstellenden Zusammenhang zwischen Linkgeschwindigkeiten und am selben Link erhobenen lokalen Geschwindigkeiten. Folgende Ursachen können hierfür in Betracht gezogen werden:

- Die Verkehrsverhältnisse am Link sind sehr konstant, die Schwankungen der weniger repräsentativen Linkgeschwindigkeiten sind rein zufällig.
- Die lokalen Daten werden an einem Querschnitt erhoben, der sehr selten überstaut wird. Jedoch kommt es auf anderen Teilen des Links immer wieder zu Reisezeitverlusten, beispielsweise vor dem Zielknotenpunkt des Links.
- Die lokalen Daten werden auf einem Fahrstreifen erhoben, für dessen Abbiegebeziehung nie Reisezeitverluste anfallen. Die geringen Linkgeschwindigkeiten können sich aber auf alle Abbiegebeziehungen beziehen, die für den Link vorliegen. Somit können sie – zumindest zu einem großen Teil – unabhängig von den lokalen Daten sein.

6.2.4 Näherungsfunktionen mit zeitversetzten Ein- und Ausgangsdaten

Nach Kapitel 4.1 bezwecken Näherungsfunktionen mit zeitversetzten Ein- und Ausgangsgrößen

- Das Verfolgen der Ausbreitung von Verkehrszuständen im Netz
- Die Kurzzeit-Prognose der Kennwerte
- Die Verbreiterung der Datenbasis insbesondere zur Schätzung aus zeitlich nicht immer verfügbaren Kenngrößen
- Autoregression, also die Schätzung einer Kenngröße aus ihrem vorherigen Wert
- Die Beurteilung der zeitlichen Verwendbarkeit von Messwerten

Da der erstgenannte Punkt insbesondere mit linkbezogenen Daten gelingen könnte, deren Verfügbarkeit aber nachweislich und ihre Repräsentativität für den Verkehrszustand maßlich im vorliegenden Fall Grenzen unterliegen, wird diese mögliche Wirkung der Näherungsfunktionen mit Zeitversatz nicht weiter untersucht.

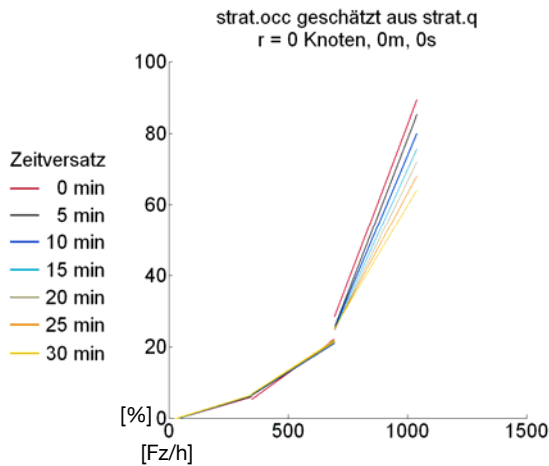
Da Verkehrszustände in Netzausschnitten sich in der Regel nicht in kurzen Abständen ständig ändern, sondern häufig über zahlreiche Fünf-Minuten-Intervalle in ähnlicher Weise anhalten, ist zu erwarten, dass sich die Zusammenhänge zwischen zwei Kenngrößen ohne und mit einem geringen Zeitversatz stark ähneln. Insbesondere wenn sich die Kenngrößen über einen langen Zeitraum nur wenig ändern – z.B. Linkgeschwindigkeiten auf Links, auf denen sich keine Störungen abspielen – ändert sich auch für große zeitliche Versätze wenig. Wo die Kenngrößen verursacht durch veränderte Verkehrsverhältnisse jedoch häufig variieren, ist damit zu rechnen, dass mit wachsendem zeitlichem Versatz die Ähnlichkeiten zur zeitgleichen Regressionsanalyse abnehmen, die Zusammenhänge schwächer ausfallen, das Bestimmtheitsmaß abnimmt und sich die Näherungsfunktionen einebnen. Ihre Steigungen nehmen also ab, die Schätzungen tendieren folglich zum Bereich ungestörten Verkehrsflusses. Das sinkende Bestimmtheitsmaß führt dazu, dass der Signifikanztest für manche Zusammenhänge nicht mehr bestanden wird und somit für wachsende zeitliche Versätze weniger signifikante Näherungsgeraden ermittelt werden.

Für Näherungsfunktionen aus der Autoregression ist also zu erwarten, dass sie in aller Regel ein Verharren der Kenngrößen auf dem zuletzt gemessenen Niveau prognostizieren, jedoch mit abnehmender Zuverlässigkeit für steigende zeitliche Versätze. Dies würde sich in Näherungsfunktionen nahe der Mittelgeraden niederschlagen. Zudem sollten sich auch hier Schätzfunktionen mit Eingangsdaten, die seltene Verkehrsverhältnisse dokumentieren, für wachsende Zeitversätze zunehmend von dieser Mittelgeraden entfernen und mit zurückgehendem Bestimmtheitsmaß eher die Schätzung üblicher, ungestörter Verkehrsverhältnisse herbeiführen.

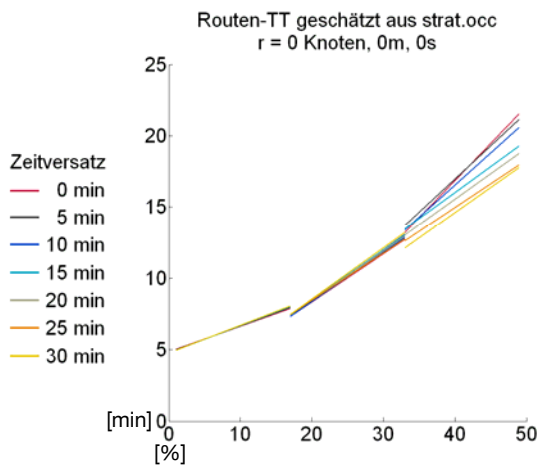
Untersuchungen zu Näherungsfunktionen mit zeitversetzten Ein- und Ausgangsgrößen werden angestellt für

- Lokale Daten als Eingangsgröße
- Abschnittsdaten als Eingangsgröße
- Beispiele der Autoregression (lokale, abschnittsbezogene und fahrzeuggenerierte Eingangsdaten)

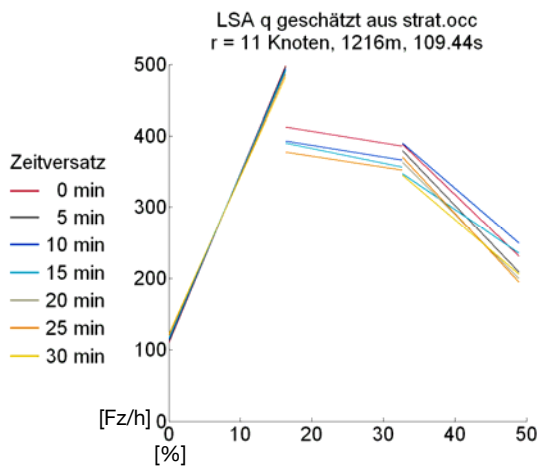
Lokale Daten als Eingangsgröße



Δt	Kleine Egw.			Mittlere Egw.			Große Egw.		
	r^2	n	\hat{t}	r^2	n	\hat{t}	r^2	n	\hat{t}
0	0.36	9762	73.46	0.34	10385	72.79	0.44	1666	36.33
5	0.33	9667	69.19	0.22	10321	54.36	0.37	1643	31.00
10	0.35	9566	71.49	0.20	10346	51.35	0.31	1651	27.05
15	0.32	9507	67.29	0.18	10337	48.51	0.26	1649	24.32
20	0.32	9495	66.67	0.18	10303	47.01	0.23	1646	22.10
25	0.29	9539	62.27	0.18	10322	47.72	0.19	1644	19.64
30	0.26	9580	58.66	0.17	10320	46.42	0.16	1637	17.72



Δt	Kleine Egw.			Mittlere Egw.			Große Egw.		
	r^2	n	\hat{t}	r^2	n	\hat{t}	r^2	n	\hat{t}
0	0.23	7980	48.26	0.52	182	13.98	0.50	33	5.58
5	0.24	8004	50.18	0.53	180	14.20	0.50	31	5.34
10	0.25	8018	52.07	0.53	179	14.18	0.35	33	4.05
15	0.26	8013	52.93	0.50	179	13.43	0.22	32	2.90
20	0.27	8014	53.85	0.43	179	11.68	0.19	32	2.62
25	0.27	8002	53.99	0.40	177	10.92	0.13	33	2.16
30	0.27	7977	54.21	0.42	177	11.25	0.15	32	2.33



Δt	Kleine Egw.			Mittlere Egw.			Große Egw.		
	r^2	n	\hat{t}	r^2	n	\hat{t}	r^2	n	\hat{t}
0	0.45	14513	109.42	0.01	253	1.49	0.15	38	2.52
5	0.45	14505	107.92				0.20	38	2.96
10	0.44	14428	105.91	0.01	253	1.56	0.13	38	2.29
15	0.43	14386	103.44	0.01	253	1.92	0.15	38	2.48
20	0.41	14352	100.99				0.18	38	2.85
25	0.41	14349	99.62	0.01	253	1.46	0.20	38	2.99
30	0.39	14348	96.32				0.22	38	3.17

Bild 6.12 Beispiele für Näherungsfunktionen mit verschiedenen Zeitversätzen aus lokalen Daten samt statistischen Kennwerten

Bild 6.12 oben zeigt die Funktionen zur Schätzung eines lokalen Belegungsgrads mit der Verkehrstärke als Eingangsgröße, die auf dem benachbarten Fahrstreifen am selben Querschnitt erhoben wird. Dieses Diagramm bestätigt alle Erwartungen: Sehr geringe Unterschiede der Näherungsgeraden für geringe Zeitversätze, deutlich größere Unterschiede für große Eingangsgrößen. Die Unterschiede schlagen sich in einer Schätzung geringerer Belegungsgrade und damit einer weniger schwerwiegenden Störung für größere zeitliche Versätze nieder. Auf allen Abschnitten nimmt das Bestimmtheitsmaß für wachsende Zeitversätze konsequent ab. Bei großen Versätzen liegen die Werte der t-Statistik durchgängig und die Bestimmtheitsmaße insbesondere für große Eingangswerte niedriger als für kleine.

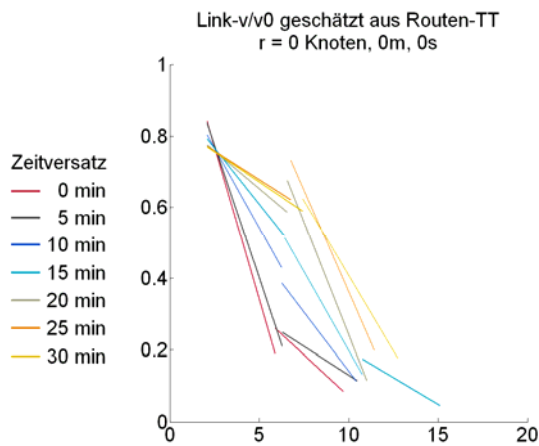
Ähnlich stellen sich die Verhältnisse in Bild 6.12 Mitte für die Schätzung der Routenreisezeit aus dem Belegungsgrad dar. Der Detektor liegt hier im Verlauf der Route. Auch das untere Beispiel gemäß Bild 6.12 führt zu defensiveren Schätzungen für größere Zeitversätze. Die statistischen Kennwerte nehmen für kleine Eingangswerte ab, für mittlere und große stellt sich dieser Zustand nicht ein. Jedoch ergibt sich für mittlere Eingangswerte und tendenziell für wachsende Versätze ein zunehmendes Problem beim Bestehen des t-Tests.

Abschnittsdaten als Eingangsgröße

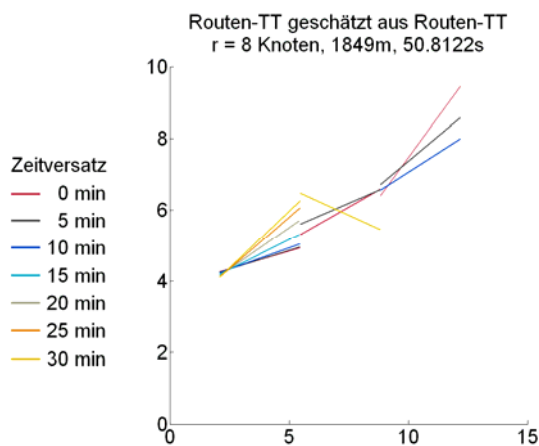
Das obere Beispiel in Bild 6.13 bestätigt die Erwartung: Üblicherweise ist der Verkehrsfluss ungestört, die Geschwindigkeit recht hoch. Entsprechend führen die Näherungsgeraden mit steigenden zeitlichen Versätzen mit derselben Reisezeit auf einer Route als Eingangswert zu höheren Schätzwerten. Zudem sinken die Bestimmtheitsmaße und tendenziell auch der Wert der t-Statistik für kleine eingehende Reisezeiten mit wachsendem Versatz ab.

Für mittlere Eingangswerte ist dieses Verhalten der statistischen Kennwerte jedoch nicht feststellbar. Für große Eingangswerte wird der t-Test nur für einen Zeitversatz von 15 min bestanden. Hierbei wirken sich erwartungsgemäß schlechtere Zusammenhänge zwischen Kenngrößen, geringere Stichproben und weiter abnehmende Zusammenhänge für wachsende zeitliche Versätze negativ aus. Dies sollte zum Nichtbestehen des t-Tests bevorzugt für große Versätze führen. Dieses Verhalten ist für das Beispiel in Bild 6.13 unten festzustellen. Für mittlere und große Eingangsgrößen zeigt sich auch hier der erwartete Rückgang der statistischen Kenngrößen. Eine recht diffuse Punktwolke für kleine Eingangsgrößen könnte Ursache dafür sein, dass sich das erwartete Muster hier nicht einstellt.

Die unterschiedlichen Abschnittsgrenzen in Bild 6.13 oben kommen durch die dynamische Abschnittswahl zustande. Sie ist abhängig von der Skalenabdeckung der verfügbaren Datenpaare. Da insbesondere die geschätzten Linkgeschwindigkeiten eher gering verfügbar sind, kommt es hierbei zu sichtbaren Unterschieden.



Δt	Kleine Egw.			Mittlere Egw.			Große Egw.		
	r^2	n	\hat{t}	r^2	n	\hat{t}	r^2	n	\hat{t}
0	0.26	258	9.47	0.26	13	1.96			
5	0.28	263	10.17	0.15	14	1.44			
10	0.16	279	7.24	0.37	17	2.95			
15	0.08	267	4.85	0.40	17	3.19	0.15	13	1.37
20	0.04	262	3.45	0.67	19	5.83			
25	0.02	248	2.43	0.35	22	3.30			
30	0.03	247	2.67	0.17	23	2.07			



Δt	Kleine Egw.			Mittlere Egw.			Große Egw.		
	r^2	n	\hat{t}	r^2	n	\hat{t}	r^2	n	\hat{t}
0	0.01	829	3.22	0.08	67	2.34	0.14	43	2.55
5	0.02	808	3.45	0.03	71	1.56	0.06	45	1.63
10	0.02	789	3.90				0.05	45	1.42
15	0.04	773	5.35						
20	0.07	767	7.75						
25	0.11	764	9.76						
30	0.12	755	10.07	0.03	88	1.53			

Bild 6.13 Beispiele für Näherungsfunktionen mit verschiedenen Zeitversätzen aus Abschnittsdaten samt statistischen Kennwerten

Beispiele der Autoregression

Die Ergebnisse der Autoregression bestätigen die Vorüberlegungen ebenfalls: Im Beispiel für Linkgeschwindigkeiten (Bild 6.14 oben) ergeben sich Abweichungen von der eingezeichneten Mittelgeraden für sehr seltene große Geschwindigkeiten und zunehmend bei steigenden zeitlichen Versätzen für kleine Geschwindigkeiten. Die Schwankungen beim Stichprobenumfang für kleine Eingangswerte kommen durch die dynamische Abschnittswahl zustande. Das Absinken der statistischen Kennwerte zeigt sich tendenziell für kleine Eingangswerte.

Durchgängig ergibt es sich beim Beispiel für die Autoregression mit Routenreisezeiten (Bild 6.14 Mitte). Eine Prognose basierend auf diesen Ergebnissen würde abnehmende Reisezeiten mit geringerer Bestimmtheit für mittlere Eingangswerte und einen ansteigenden Prognosehorizont besagen. Für große Eingangswerte reicht die Datenbasis mit einer Ausnahme offensichtlich nicht aus, um statistisch signifikante Aussagen zu ermöglichen.

Dieser Umstand ergibt sich auch für kleine Geschwindigkeiten als Eingangsgröße und den größten untersuchten Zeitversatz gemäß Bild 6.14 unten. Die statistischen Muster zeigen

sich für kleine und mittlere Eingangswerte. Auch hier führen die Ergebnisse aus der Regressionsanalyse zu zunehmend defensiven Prognosen für steigende Versätze im Fall einer aktuell vorliegenden Störung, die durch geringe Geschwindigkeiten dokumentiert ist.

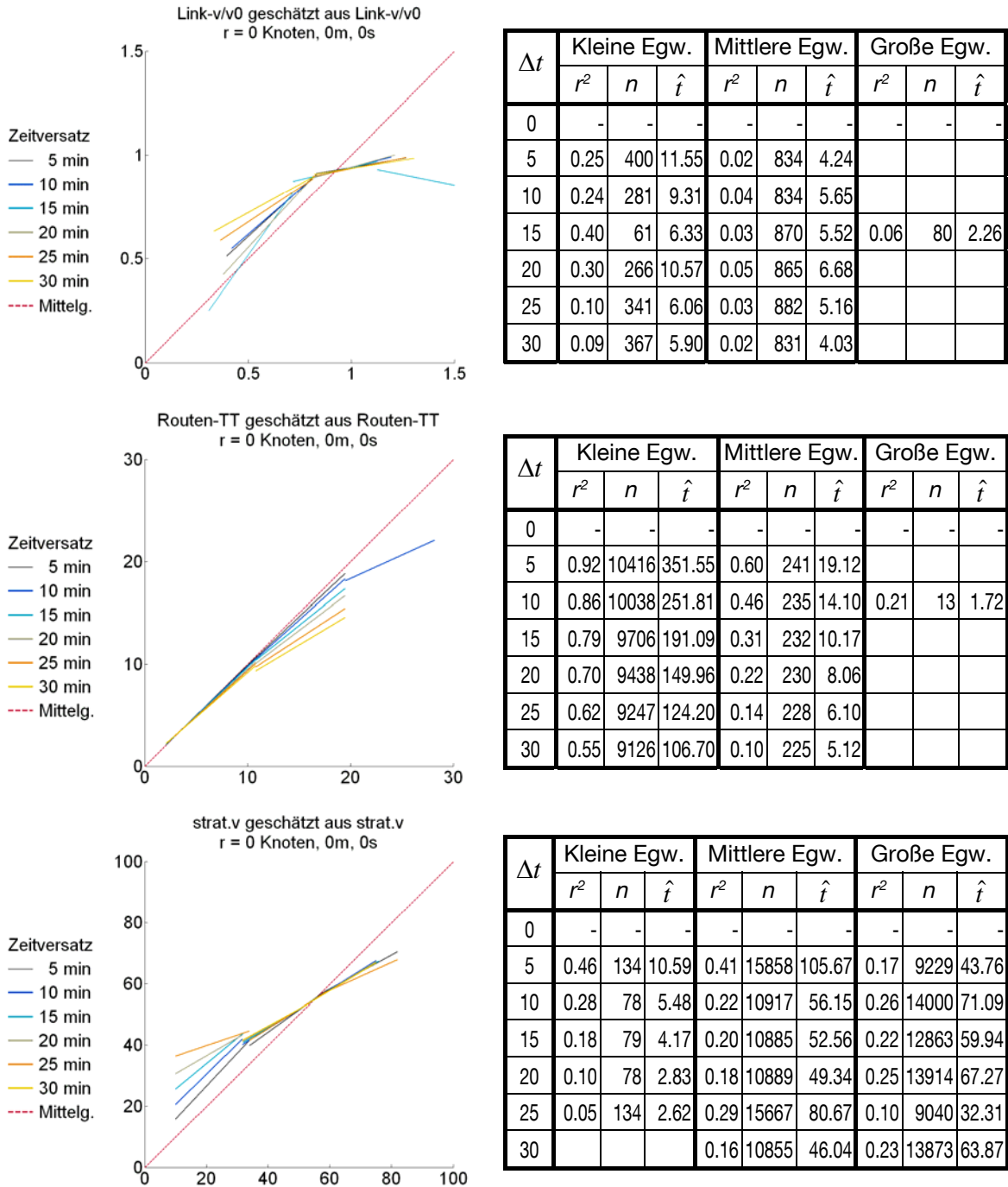


Bild 6.14 Beispiele für Ergebnisse der Autoregression mit verschiedenen Zeitversätzen samt statistischen Kennwerten

6.3 Näherungsfunktionen aus gewichteter linearer Regression

Um die Wirkungen der Gewichtung bei der linearen Regression zu ermitteln, wurden den Linkgeschwindigkeiten aus Positionsmeldungen und den infrastrukturseitig erhobenen Reisezeitmessungen Güteindikatoren wie in Kapitel 5.3 beschrieben zugeteilt. Lokalen Daten wurde grundsätzlich eine Güte von 100 zugewiesen. Um das Gewicht eines Messwertpaars zu ermitteln, wurden beide Güteindikatoren miteinander multipliziert.

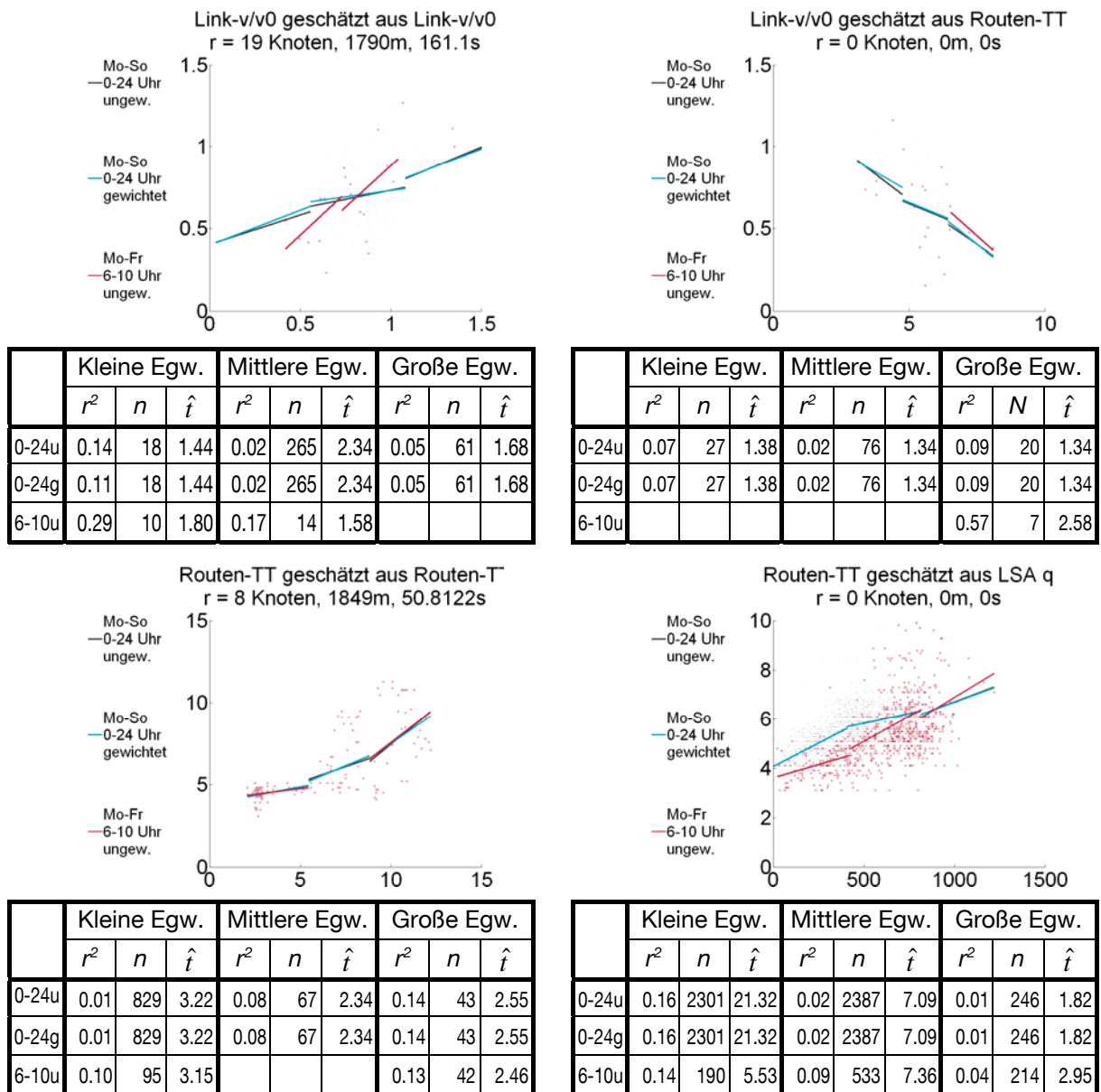


Bild 6.15 Vergleich der Näherungsfunktionen aus der Regressionsanalyse zeitlich undifferenziert (ungewichtet und gewichtet) sowie für Mo-Fr, 6-10 Uhr (ungewichtet) – Beispiele für die Schätzung von Linkgeschwindigkeiten und Routenreisezeiten

Bild 6.15 enthält Beispiele für die Ergebnisse der gewichteten Regressionsanalyse (blau) im Vergleich zur einfachen (schwarz) samt statistischen Kennwerten. Zahlreiche weitere Bei-

spiele wurden gesichtet. Sie alle zeigen sehr geringe bis nicht sichtbare Unterschiede zwischen den Ergebnissen, häufig ist die schwarze Näherungsgerade mit der blauen exakt deckungsgleich. Wo geringe Unterschiede sichtbar sind, sind keine Merkmale feststellbar, aufgrund derer offensichtlich die eine oder die andere Näherungsgerade das bessere Ergebnis darstellt.

So wird der Ansatz der gewichteten linearen Regression im Rahmen dieser Arbeit nicht weiter verfolgt, da kein Mehrwert im Vergleich zur einfachen linearen Regression festzustellen ist, zumal die Ermittlung der statistischen Kennwerte für diesen Sonderfall nicht eindeutig geklärt ist; sie wurden hier wie für die einfache lineare Regression ermittelt. Die Wirkung anders ermittelter Güteindikatoren wird hier nicht untersucht. Die vorliegenden Ergebnisse legen aber die Vermutung nahe, dass sich Messwerte unterschiedlicher Güte bei der linearen Regression zu einem erheblichen Anteil ausgleichen, und dass Messwerte geringer Güte sich in einem zurückgehenden Bestimmtheitsmaß niederschlagen, wenn sie einen großen Anteil aller Messwerte ausmachen.

6.4 Zeitliche Differenzierung bei der linearen Einfachregression

Die Vermutung liegt nahe, dass sich übliche verkehrliche Muster in der morgendlichen Hauptverkehrszeit anders einstellen als in der abendlichen, an einem Sonn- oder Feiertag anders als an einem Werktag. Für den entwickelten Ansatz bedeutet dies, dass die Regressionsanalyse differenziert nach Tageskategorie und -zeit durchzuführen ist, da sich Zusammenhänge im Netz folglich für die verschiedenen Zeitbereiche unterschiedlich darstellen.

Da die Datenhistorie jedoch dieselbe bleibt, wird die Stichprobe bei der beschriebenen Differenzierung für jeden Zeitbereich kleiner. Dies lässt in großem Umfang Näherungsgeraden erwarten, die den Signifikanztest nicht bestehen, insbesondere für seltene Verkehrszustände mit den damit einhergehenden geringen Stichprobenumfängen. Um die Wirkung der zeitlichen Differenzierung bei der Regressionsanalyse exemplarisch abschätzen zu können, wird mit Montag bis Freitag, 6 bis 10 Uhr ein Zeitfenster ausgewählt, in dem sich erwartungsgemäß aufgrund der morgendlichen Hauptverkehrszeit ein vergleichsweise breites Spektrum an verkehrlichen Situationen einstellt und die beschriebene Schwierigkeit für seltene Verkehrszustände, sprich Störungen, sich gedämpft darstellt.

Dennoch geht die Anzahl der Näherungsgeraden für hohe Verkehrsstärken, Belegungsgrade und Reisezeiten sowie für geringe Geschwindigkeiten gemäß Tab. 6.4 wesentlich deutlicher zurück als für das jeweils andere Ende der Skala. Während die Anzahl der Näherungsgeraden für die Schätzung von infrastrukturseitig erhobenen Daten aus infrastrukturseitig erhobenen Daten selten um mehr als 30% und nie um mehr als 40% abnimmt, beträgt der Rückgang für Schätzungen aus oder von Linkgeschwindigkeiten immer wieder mehr als 60%, für die Schätzung von Linkgeschwindigkeiten aus Linkgeschwindigkeiten sogar

durchgehend über 70%. So werden die Schätzergebnisse insbesondere in Bereichen erheblich dezimiert, in denen keine infrastrukturseitige Detektion erfolgt.

geschätzt erklärend	strat.q			strat.v			strat.occ			LSA q			Routen-TT			Link-v/v0		
	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.	kl.	mi.	gr.
strat.q	100.0	98.7	85.6	91.1	96.8	85.4	97.0	98.5	91.8	98.5	96.2	78.6	75.9	89.5	79.1	34.7	44.8	77.4
strat.v	66.1	73.6	101.8	80.8	77.8	100.2	76.1	77.8	100.8	60.0	71.0	98.0	73.8	75.2	84.7	67.0	39.2	50.5
strat.occ	100.7	90.4	76.9	98.9	72.8	65.7	99.9	80.1	62.6	98.6	69.9	65.5	88.4	103.9	74.8	47.1	75.6	92.9
LSA q	96.6	159.5	139.7	94.0	154.8	140.8	96.5	153.6	146.3	95.2	149.1	136.9	86.0	98.1	87.3	37.3	53.3	82.7
Routen-TT	99.5	62.6	96.4	99.4	66.4	68.9	100.0	63.0	78.4	96.2	71.3	68.6	82.4	106.0	75.9	39.5	75.3	90.0
Link-v/v0	31.6	42.0	70.9	49.1	50.5	95.7	41.0	43.3	78.4	33.2	45.2	72.8	45.8	51.4	69.4	27.2	29.9	27.9

Tab. 6.4 Anzahl der Näherungsgeraden aus dem Zeitfenster werktags 6-10 Uhr relativ zur Anzahl aus der zeitlich nicht differenzierten Regressionsanalyse gemittelt über alle untersuchten zeitlichen Versätze [%]

Bemerkenswert ist, dass die Anzahl der Näherungsgeraden für Schätzungen von infrastrukturseitig erhobenen Daten aus Verkehrsstärken an LSA-Detektoren häufig sogar zunimmt. Für diese Konstellationen liegen aufgrund der hohen Datenverfügbarkeiten immer noch große Stichprobenumfänge vor, so dass der Nachweis der Signifikanz zumindest daran nicht scheitert. Offensichtlich werden die ausgewerteten Punktwolken durch die zeitliche Differenzierung geschärft, so dass die Zusammenhänge klarer erkennbar sind. So kann die zeitliche Differenzierung für solche Fälle das Bestehen des Signifikanztests auch begünstigen.

Die Unterschiede zwischen den Ergebnissen der zeitlich differenzierten und nicht differenzierten Regressionsanalyse können in den Beispielen gemäß Bild 6.15 und Bild 6.16 häufig leicht erkannt werden. Nur selten sind rote und schwarze Gerade nahezu deckungsgleich. Überwiegend liegt das Bestimmtheitsmaß für das kleinere Zeitfenster höher als für das größere. Zudem ist in manchen Fällen erkennbar, dass die Punktwolke für die zeitlich differenzierte Untersuchung (rote Punkte) sich schärfer darstellt und einen Zusammenhang besser beschreibt als im nicht differenzierten Fall (zusätzlich kleine schwarze Punkte, Bild 6.15 rechts unten). Verbesserungen kommen mitunter auch durch anders gewählte Abschnitte zustande, da die abgedeckten Datenbereiche sich nicht entsprechen (Bild 6.16 rechts oben). Bild 6.15 oben zeigt, dass bei Näherungsfunktionen zur Schätzung von Linkgeschwindigkeiten durch deren geringe Datenverfügbarkeit der Signifikanztest bei zeitlicher Differenzierung eine Schwelle darstellt, die oft zum Verwerfen möglicher Zusammenhänge für die Schätzung von Verkehrskenngrößen führt.

Es stellt sich für manche Abschnitte jedoch auch der Fall ein, dass durch die zeitliche Differenzierung das Bestimmtheitsmaß zurückgeht. Ein Grund dafür kann sein, dass im untersuchten Zeitfenster weniger sehr übliche Datenpunkte produziert werden wie z.B. nachts, wenn stundenlang sehr ähnliche Werte protokolliert werden. Diese Punkte liefern bei der Regressionsanalyse sehr stabile Näherungsfunktionen. Entfallen sie durch die zeitliche Differenzierung, entfällt auch ihre stabilisierende Wirkung.

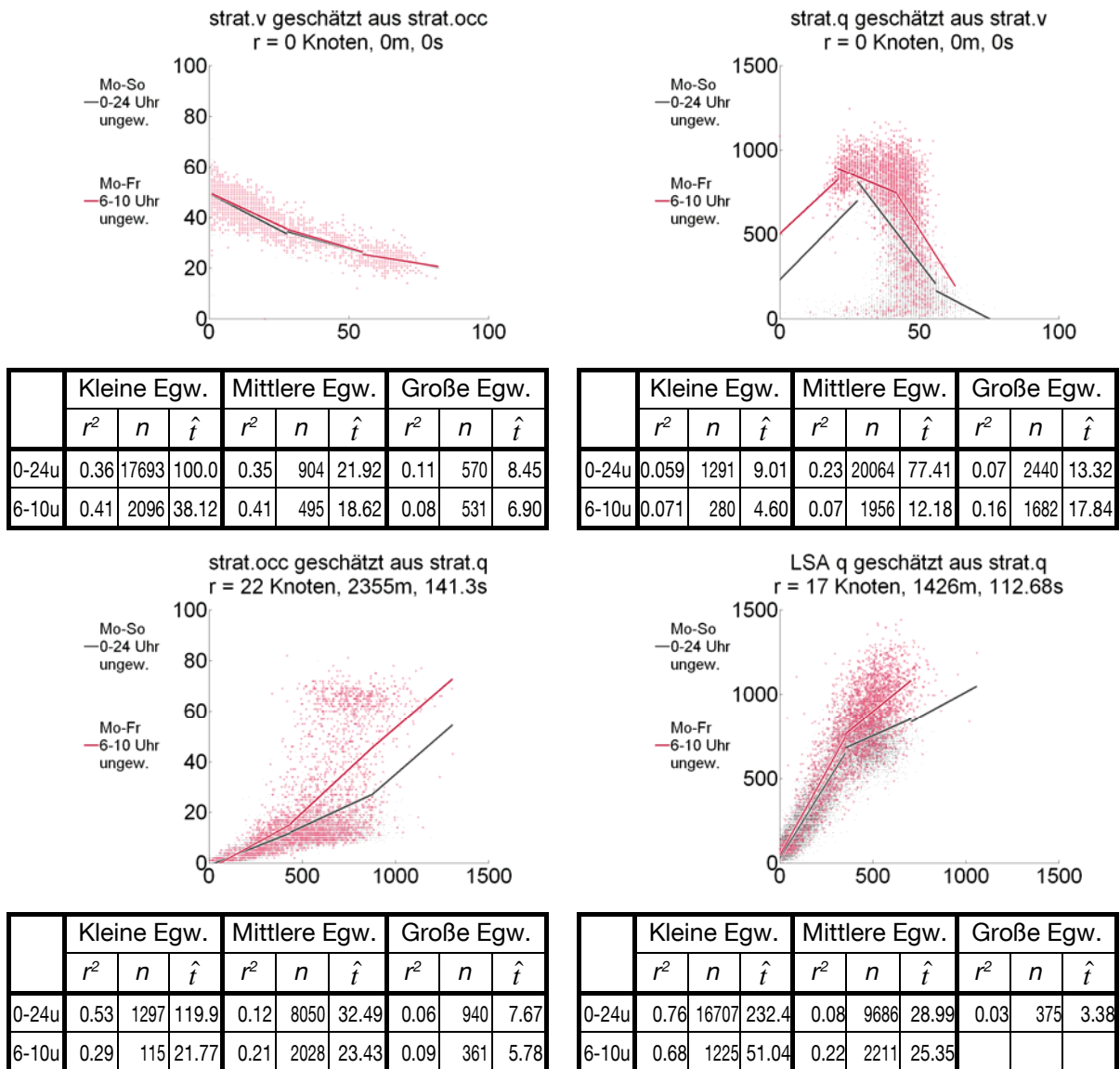
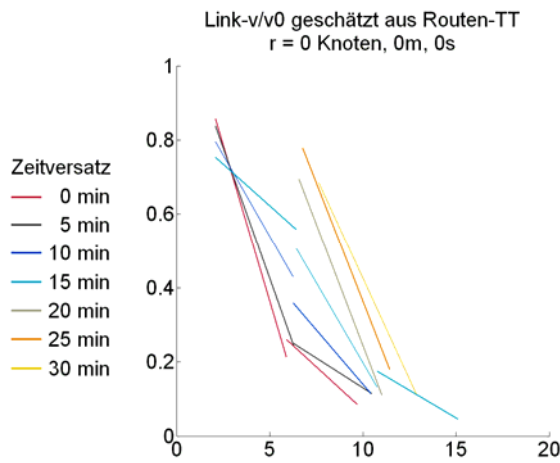


Bild 6.16 Vergleich der Näherungsfunktionen aus der Regressionsanalyse zeitlich undifferenziert und differenziert (Mo-Fr, 6-10 Uhr) – Beispiele für die Schätzung lokaler Daten

Bild 6.17 oben zeigt im Vergleich zu Bild 6.13 oben, dass fast alle Datenpunkte in einem oder mehreren Abschnitten in einem untersuchten Zeitfenster liegen können. So geht bei der Regressionsanalyse der Stichprobenumfang nur in einzelnen Abschnitten zurück. Im Beispiel ist davon lediglich der Abschnitt für kleine Eingangswerte betroffen, die anderen beiden Abschnitte stellen sich fast identisch wie im zeitlich nicht differenzierten Fall dar. Im veränderten Abschnitt führt der reduzierte Stichprobenumfang dazu, dass der Signifikanztest für zeitliche Versätze über 15 min nicht mehr bestanden wird.

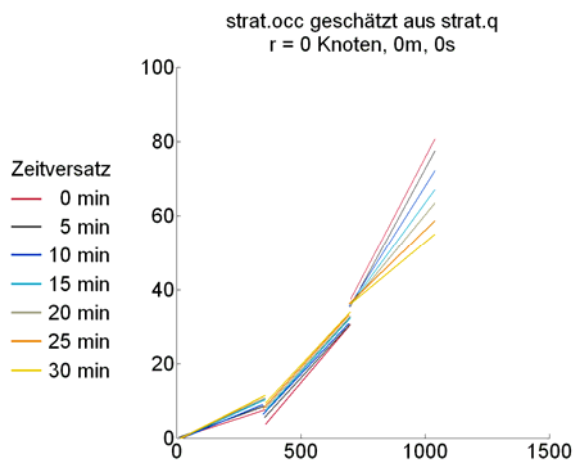
Liegen die entfallenden Punkte jedoch größtenteils in selten besetzten Abschnitten wie im Fall von Bild 6.17 unten im Vergleich zu Bild 6.12 unten, so entfallen signifikante Näherungsfunktionen für Eingangsdaten in selten besetzten Skalenbereichen. Der bei zeitlicher Differenzierung wachsende Stichprobenumfang für mittlere Eingangswerte und 25 min Zeitver-

satz kommt durch die dynamische Abschnittsfestlegung und den geringeren Unterschied zwischen minimalem und maximalem Belegungsgrad zustande.



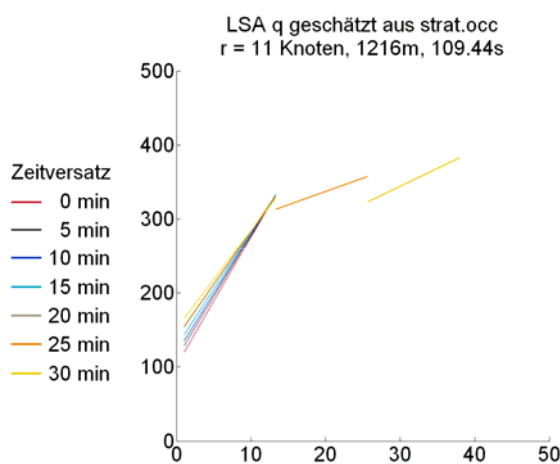
Vgl. Bild 6.13 oben

Δt	Kleine Egw.			Mittlere Egw.			Große Egw.		
	r^2	n	\hat{t}	r^2	n	\hat{t}	r^2	n	\hat{t}
0	0.47	46	6.25	0.26	13	1.96			
5	0.44	48	6.02	0.15	14	1.44			
10	0.22	55	3.85	0.31	15	2.43			
15	0.06	55	1.85	0.39	16	2.98	0.14	13	1.37
20				0.68	18	5.82			
25				0.38	21	3.41			
30				0.28	23	2.88			



Vgl. Bild 6.12 oben

Δt	Kleine Egw.			Mittlere Egw.			Große Egw.		
	r^2	n	\hat{t}	r^2	n	\hat{t}	r^2	n	\hat{t}
0	0.20	618	12.44	0.38	1844	33.88	0.27	990	19.18
5	0.22	611	12.95	0.29	1842	27.50	0.21	989	16.18
10	0.35	600	18.14	0.26	1848	25.35	0.15	997	13.33
15	0.23	625	13.81	0.23	1836	23.55	0.10	977	10.64
20	0.24	628	14.19	0.21	1832	22.03	0.08	975	9.14
25	0.24	611	14.00	0.21	1837	22.30	0.05	993	7.35
30	0.26	634	14.91	0.19	1828	20.45	0.04	969	5.99



Vgl. Bild 6.12 unten

Δt	Kleine Egw.			Mittlere Egw.			Große Egw.		
	r^2	n	\hat{t}	r^2	n	\hat{t}	r^2	n	\hat{t}
0	0.36	2470	37.57						
5	0.35	2470	36.33						
10	0.35	2463	36.44						
15	0.34	2459	35.39						
20	0.32	2456	33.56						
25	0.32	2450	34.21	0.01	305	1.61			
30	0.28	2445	31.12				0.39	8	1.98

Bild 6.17 Beispiele für Ergebnisse der zeitlich differenzierten Regressionsanalyse mit zeitlich versetzten Ein- und Ausgangsdaten

Im Fall der zeitlichen Differenzierung kann die Trivialprognose „Alles bleibt vorerst, wie es ist“ schlechter ausfallen als im undifferenzierten Fall, wie Bild 6.17 Mitte im Vergleich zu Bild 6.12 oben zeigt. Die Schätzungen aus hohen Verkehrsstärken fallen für steigende zeitliche Versätze immer defensiver aus und basieren auf Funktionen mit wesentlich niedrigerem Bestimmtheitsmaß. Auch für kleine Eingangswerte ist das Bestimmtheitsmaß hier geringer. Anders als zuvor sind bei kleinen und mittleren Eingangswerten Unterschiede zwischen den Näherungsgeraden für verschiedene zeitliche Versätze zu erkennen.

So fällt das Fazit für die zeitliche Differenzierung bei der Regressionsanalyse zwiespältig aus: Sie wirkt sich positiv auf die Ergebnisse aus, wenn für das Zeitfenster eine ausreichend große Datenhistorie vorliegt. Jedoch birgt sie auch die Gefahr, dass aufgrund des geringeren Stichprobenumfangs weniger statistisch signifikante Näherungsfunktionen ermittelt werden. Für andere als das hier untersuchte Zeitfenster sollte diese Schwierigkeit eher verstärkt auftreten, da dann für die ohnehin dünn besetzten Skalenbereiche, die Störungen dokumentieren, anteilig noch weniger Messungen zu erwarten sind. Probleme beim Signifikanztest betreffen v.a. Zusammenhänge aus geringverfügbaren Ein- und Ausgangsdaten.

6.5 Rückschlüsse für die Zusammenführung von Einzelschätzwerten

Die verwendete Datenhistorie dokumentiert in aller Regel ungestörte Verkehrsverhältnisse. Störungen stellen eher die Ausnahme dar. So können Zusammenhänge für freien Verkehrsfluss vergleichsweise leicht festgestellt werden, da für sie meist ein ungleich größerer Stichprobenumfang vorliegt. Dieser schlägt sich im Wert der Teststatistik \hat{t} positiv nieder. Das vorgeschlagene Verfahren muss sich aber insbesondere bei der Erkennung ungewöhnlicher Situationen, also von Störungen, bewähren. Da das Bestimmtheitsmaß unabhängig vom Stichprobenumfang ist, wird es im weiteren Verlauf als Kriterium bei der Auswahl der Zusammenhänge verwendet, die zur Schätzwertbildung herangezogen werden.

Abschnittsbildung:

Die abschnittsweise lineare Regressionsanalyse angewendet auf Verkehrsdaten liefert vielversprechende Näherungsfunktionen für die Schätzung von Verkehrskenngrößen mit meist sehr geringen Unstetigkeiten am Übergang zwischen unterschiedlichen Abschnitten. Je nach Nichtlinearität der beobachteten Zusammenhänge bedarf es dazu einer angepassten Anzahl gebildeter Abschnitte. Mit Ausnahme der Funktionen zur Schätzung von Verkehrsstärken hat sich gezeigt, dass drei Abschnitte zur differenzierten Darstellung der Zusammenhänge ausreichen. Da Verkehrsstärken hochverfügbar sind und daher eher zu plausibilisieren sind als zu vervollständigen, werden drei Abschnitte auch hier als ausreichend erachtet. Begründet wird dies mit der Tatsache, dass bei der Überprüfung der Funktionstüchtigkeit eines Detektors nicht ein einzelner Wert, sondern zahlreiche Schätzwerte aus der räumlich-zeitlichen Umgebung herangezogen würden.

Für die abschnittsweise Beobachtung der Zusammenhänge spricht, dass sich für die verschiedenen Abschnitte oft recht unterschiedliche Ergebnisse hinsichtlich Qualität und Signifikanz der Näherungsfunktionen einstellen. So wird bei der Schätzung von Verkehrskenngrößen eine diesbezüglich differenzierte Auswahl der verwendeten Zusammenhänge ermöglicht.

Schätzungen aus lokalen Daten:

Für große Verkehrsstärken liegen mehr Näherungsgeraden vor als für kleine lokale Geschwindigkeiten oder hohe Belegungsgrade als Eingangsgröße, die allesamt bei hohem Verkehrsaufkommen erhoben werden. Dies resultiert aus der Tatsache, dass der erstgenannte Fall im Gegensatz zu den anderen beiden nicht zwingend eine Störungssituation dokumentiert und demzufolge häufiger auftritt.

Auch wenn die Anzahl der Näherungsfunktionen aus Verkehrsstärken an LSA-Detektoren für steigende Verkehrsstärken stärker zurückgeht als an strategischen Detektoren, so gibt es dennoch keinen Grund anzunehmen, dass die ermittelten statistisch signifikanten Näherungsfunktionen aus Verkehrsstärken an LSA-Detektoren für die Schätzung verkehrlicher Kennwerte weniger geeignet sind.

Schätzungen aus Abschnittdaten:

Reisezeiten, die durch Fahrzeugwiedererkennung ermittelt werden, wurden für vergleichbare Ansätze bisher nicht eingesetzt. Mit den gewonnenen Erkenntnissen besteht Grund zur Annahme, dass sie eine aussagekräftige Eingangsgröße insbesondere zur Schätzung von Linkgeschwindigkeiten sein können. Zudem interessiert die Fragestellung, ob die Reisezeiten aus anderen vorliegenden Daten zuverlässig geschätzt werden können, da die Sensorik zur Reisezeitmessung aufwendig und teuer ist. Manche Beispiele zeigen diesbezüglich vielversprechende Eigenschaften.

Schätzungen aus Linkgeschwindigkeiten, abgeleitet aus Positionsmeldungen:

Linkgeschwindigkeiten fungieren beim vorgeschlagenen Ansatz aufgrund ihrer geringen Verfügbarkeit primär als Schätzgrößen. Werden sie als Eingangsgrößen eingesetzt, ist darauf zu achten, dass die Schätzung von Linkgeschwindigkeiten aus sich selbst unterbunden wird. Die aus den Positionsdaten ermittelten Werte sind aufgrund ihrer reduzierten Repräsentativität nur begrenzt als Referenzwerte bei der Bewertung geeignet.

Schätzungen mit zeitversetzten Ein- und Ausgangsgrößen:

Die Ergebnisse aus der zeitversetzten Regressionsanalyse können einen Teil zur möglichst vollständigen Verkehrslageschätzung in einem Netz beitragen, da sie die Datenbasis für die Schätzung insbesondere aus wenig verfügbaren Daten maßgeblich erweitern.

Die Autoregression liefert interessante Ergebnisse. Sie wird bei der Verkehrslageschätzung aus hoch verfügbaren Daten jedoch kaum Wirkungen zeigen, da hier in aller Regel aktuelle Werte vorliegen und daher die Schätzung aus sich selbst entfällt. Auch zur Plausibilisierung ist sie nur bedingt geeignet, weil eine solche eher anhand der Daten aus anderen Sensoren erfolgen sollte. Die Erkenntnisse liefern aber Anhaltspunkte dafür, mit welchem zeitlichen Horizont und mit welcher Zuverlässigkeit aus aktuellen Daten auf die Situation in naher Zukunft geschlossen werden darf. Häufig wird dabei ein Verharren auf dem gegenwärtigen Niveau bei abnehmender Zuverlässigkeit für wachsende zeitliche Versätze prognostiziert. Im Fall von Störungen wird mit wachsendem Versatz zudem eine zunehmende Entspannung der Verkehrssituation erwartet.

Effekte von gewichteter und zeitlich differenzierter Regressionsanalyse:

Die gewichtete lineare Regression liefert keine erkennbaren Vorteile gegenüber der einfachen linearen Regression. Daher wird sie in dieser Arbeit nicht weiter behandelt.

Die Untersuchungen zur zeitlich differenzierten Regressionsanalyse legen die Vermutung nahe, dass dabei bessere Schätzergebnisse entstehen können, auch wenn durch die damit einhergehende Reduzierung der verwendeten Datenbasis weniger Zusammenhänge ihre Signifikanz nachweisen können. Es treten auch Fälle ein, in denen die zeitliche Differenzierung zu geschärften Punktwolken, höheren Bestimmtheitsmaßen und bei nach wie vor großen Stichprobenumfängen zu besseren Ergebnissen beim Signifikanztest führt. So werden Ergebnisse aus der zeitlich differenzierten Regressionsanalyse weiter untersucht.

7. Zusammenführung der Einzelschätzungen und Bewertung der Schätzergebnisse

7.1 Konzept zur Bewertung der Schätzergebnisse

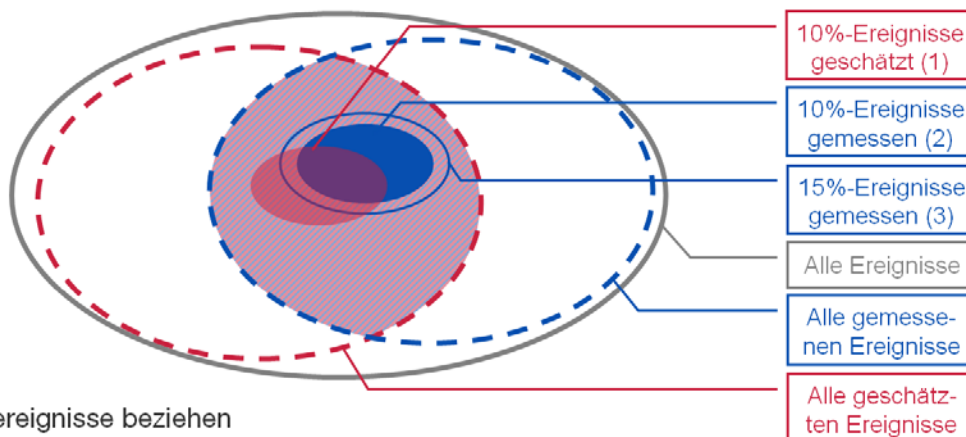
Zur Festlegung geeigneter Kriterien, die bei der Zusammenführung von Einzelschätzwerten berücksichtigt werden, und zur Bewertung der Verfahrensergebnisse werden Referenzdaten benötigt, denen die Schätzwerte gegenübergestellt werden können. Diese Referenzdaten sollen die räumlich-zeitliche Verkehrssituation wahrheitsgetreu abbilden. Hierfür können die im Bewertungsnetz erhobenen Daten herangezogen werden, deren Qualität in Kapitel 5 beschrieben wird. Die Daten aus dem Zeitraum 10.12.2008 bis 15.3.2009 werden als Basis für die Regressionsanalyse verwendet, die letzten sieben Tage aus diesem Zeitraum werden bei der Festlegung der Zusammenführungskriterien eingesetzt. Als Referenzdaten bei der Bewertung zusammengeführter Schätzergebnisse werden Messdaten aus der Woche 16. bis 22.3.2009 (Montag bis Sonntag) herangezogen. Jedoch beschreiben auch die Messdaten die verkehrliche Wahrheit nicht exakt. Daher würde selbst ein idealer Schätzer Abweichungen von diesen Referenzdaten aufweisen. Aufgrund der Vielzahl angestellter Vergleiche zwischen Schätz- und Messwerten können aber auch so ermittelte Maßzahlen stichhaltige Anhaltspunkte für die Ableitung von Zusammenführungskriterien oder für die Beurteilung von Schätzergebnissen liefern.

Typische Fehlermaße zur Bewertung eines Schätzers sind der durchschnittliche Betrag des Fehlers oder der durchschnittliche quadrierte Fehler über alle Schätzwerte. Wie zahlreiche andere Fehlermaße bieten auch diese Maßzahlen die Möglichkeit, dass für ein Schätzverfahren mit grundsätzlich durchschnittlich großen Fehlern und für ein meist sehr exaktes Schätzverfahren mit sehr wenigen sehr großen Fehlern ähnliche Werte ermittelt werden können. Zudem können derartige Maßzahlen die Verteilung der Fehler über den Wertebereich nicht widerspiegeln, was je nach Charakteristik des untersuchten Schätzers maßgebliche Details unkenntlich machen kann.

Im Falle der Schätzung von Verkehrskenngrößen ist ein solches Charakteristikum, dass in den allermeisten Zeitschritten, in denen Schätzungen anzustellen sind, ungestörte Verkehrsverhältnisse vorliegen. Deren Schätzung ist vergleichsweise einfach. So gelingt sie bereits durch schlichte Aussagen wie „immer ungestörter Verkehrsablauf“ sehr gut, was einer zutreffenden Schätzung in den allermeisten Fällen gleichkommt. Nur vergleichsweise selten liegen besondere, gestörte Verkehrsverhältnisse vor, die sich beispielsweise in einem Rückgang der Geschwindigkeit niederschlagen. So entscheidet sich die Qualität eines Schätzers von Verkehrskenngrößen daran, wie häufig er solche besonderen Ereignisse und wie selten er übliche Verhältnisse als besonders erkennt. Daher wird im weiteren Verlauf die Erkennungsrate von Perzentilereignissen als Bewertungskriterium verwendet. So umfassen beispielsweise die 10%-Ereignisse einer Messreihe lokaler Geschwindigkeiten die 10% aller

Messungen mit den geringsten Geschwindigkeiten. Liefert der Schätzer bei diesen Messungen die niedrigsten 10% seiner Schätzwerte zu dieser Messreihe, so beträgt die Erkennungsrate des 10%-Ereignisses 100%. Liegt jedoch nur die Hälfte seiner 10% niedrigsten Schätzwerte zeitgleich zu den 10% geringsten Messwerten vor, so geht die Erkennungsrate auf 50% zurück.

Zur differenzierten Bewertung eines Schätzers werden verschiedene Perzentilereignisse eingesetzt. Da zur Bewertung die Messdaten einer Woche in Fünfminutenintervallen herangezogen werden, liegen bei einer Datenverfügbarkeit von 100% 2016 Referenzwerte vor. So gibt es zur Überprüfung der Erkennungsrate von 1%-Ereignissen maximal 20 Referenzdatensätze, die einer aufsummierten Zeit von 100 min entsprechen, täglich durchschnittlich also knapp 15 min oder drei Schätzwerten. Ein 1%-Ereignis tritt also auch absolut so selten ein, dass es mit der verwendeten Vergleichsdatenbasis gerade noch sinnvoll untersucht werden kann. Bei zurückgehender Datenverfügbarkeit ist die Anzahl der Referenzdatensätze entsprechend abzumindern. Zur Bewertung der Schätzungen aus der zeitlich nicht differenzierten Regressionsanalyse werden daher 1%-, 2%-, 5%-, 10%- und 20%-Ereignisse untersucht. Als weichere Variante der Detektionsraten von Perzentilereignissen wird die Erkennungsrate eines Perzentilereignisses als ein wahrscheinlicheres Perzentilereignis eingesetzt. Sie beschreibt zum Beispiel den Anteil der 5%-Ereignisse, die als 10%-Ereignis erkannt werden (Nomenklatur im weiteren Verlauf 5(10)%-Ereignis). So werden 1(3)%-, 2(5)%-, 5(10)%-, 10(15)%- und 20(30)%-Ereignisse untersucht. Bild 7.1 veranschaulicht Beispiele für die Erkennungsrate von Perzentilereignissen.



Perzentilereignisse beziehen sich stets auf die Schnittmenge gemessener und geschätzter Ereignisse

Erkennungsrate 10%-Ereignis: Anteil der Fläche (1), der über Fläche (2) liegt

Erkennungsrate 10(15)%-Ereignis: Anteil der Fläche (1), der über Fläche (3) liegt

Bild 7.1 Erkennungsrate von Perzentilereignissen einer einzelnen Kenngröße

Für den zeitlich differenzierten Fall ist die Anzahl der Referenzdatensätze während der Woche mit 240 deutlich geringer, da nur Datensätze aus den Zeitfenstern Montag bis Freitag, 6 bis 10 Uhr verwendet werden können. Daher und aufgrund der Tatsache, dass in diesen Zeitbereichen anteilig weitaus mehr besondere Verkehrsereignisse zu erwarten sind, werden die Perzentile bei der Bewertung hier mit 5% (zwölf Referenzdatensätze bei 100% Datenverfügbarkeit), 10%, 15%, 25% und 50% etwas höher gewählt. Bei der weicheren Auswertung kommen 5(10)%-, 10(15)%-, 15(25)%-, 25(35)%- und 50(65)%-Ereignisse zum Einsatz.

Zusätzlich wird zur Beurteilung der Schätzungen aus aktuellen Werten und aus den ermittelten Schätzfunktionen neben der augenscheinlichen Gegenüberstellung von Mess- und ermittelten Ersatzwerten die Wurzel der Summe der quadrierten Fehler ermittelt (RMSE für Root Mean Squared Error). Dieses Fehlermaß berücksichtigt Über- und Unterschätzungen in gleichem Maße, größere Fehler gehen überproportional ein. Die größtmöglichen Fehler treten bei der Schätzung besonderer Verkehrsverhältnisse bei tatsächlich nicht besonderen Verhältnissen oder umgekehrt auf. Durch das Ziehen der Wurzel weist der RMSE dieselbe Einheit wie die untersuchte Kenngröße auf. Mit der Maßzahl können insbesondere die Wirkungen der Kalibrierung untersucht werden.

Aufgrund der geringen Verfügbarkeit und der geringen Repräsentativität der ermittelten Linkgeschwindigkeiten aus Positionsmeldungen ist der Einsatz von RMSE und Detektionsrate von Perzentilereignissen zur Bewertung der Schätzung von Linkgeschwindigkeiten jedoch nur begrenzt sinnvoll. Daher werden zur Bewertung der geschätzten linkbezogenen Kenngrößen die Summen der geschätzten Linkreisezeiten über alle Links einer Route, auf der eine gemessene Reisezeit durch Fahrzeugwiedererkennung als Referenzwert vorliegt, mit den gemessenen Routenreisezeiten verglichen. Auch die Verläufe einzelner geschätzter Linkgeschwindigkeiten können über die Zeit in Augenschein genommen und interpretiert werden. Zudem werden räumlich-zeitliche verkehrliche Muster analysiert, wie sie sich aus den Schätzungen ergeben.

Zum Vergleich mit der entwickelten Methode werden drei weitere Verfahren verwendet:

- Ganglinienverfahren: Es wird für alle 288 Zeitschritte über 24 Stunden eines Tages und für jede Kenngröße je ein Wert ermittelt. Dazu werden alle Messwerte zu diesem Zeitschritt aus allen Tagen beziehungsweise über alle Montage bis Freitage der verwendeten Historie (ca. 90 bzw. 65 Tage) herangezogen und der Median ermittelt. Liegen weniger als 50 Messwerte vor, so wird der zeitliche Suchraum so lange symmetrisch erweitert, bis dieser Stichprobenumfang erreicht ist.
- Trivialprognose: Der zuletzt gemessene Wert einer Kenngröße wird als Schätzwert für den gegenwärtigen Zustand herangezogen.
- Trivial+: Ganglinienverfahren und Trivialprognose werden kombiniert. Üblicherweise wird die Trivialprognose verwendet. Liegt die letzte Messung eines Kennwerts je-

doch weiter als 15 min zurück, so werden Daten aus dem Ganglinienverfahren eingesetzt.

Die Kennwerte zur Bewertung der Schätzverfahren werden für ausgewählte, repräsentative Kenngrößen aus dem Bewertungsnetz dargestellt und diskutiert. Auf eine Mittelung über alle Kenngrößen im Netz wird aus zwei Gründen verzichtet:

- Nicht überall im Netz ereignen sich während der betrachteten Zeiträume Störungen. Dies hat zur Folge, dass sich hier Kenngrößen wie Geschwindigkeiten, Reisezeiten oder Belegungsgrade nie maßgeblich verändern. Die Erkennungsrate von Perzentilereignissen ist für solche Kenngrößen kein aussagekräftiges Kriterium. Zudem ist die Schätzung solcher sich kaum verändernder Kenngrößen vergleichsweise trivial.
- Die Bewertung für die zeitlich differenzierte Anwendung der entwickelten Methode erfolgt nur für das Zeitfenster Montag bis Freitag, 6 bis 10 Uhr. Es ist zu erwarten, dass sich die Ergebnisse hier im Vergleich zum zeitlich nicht differenzierten Fall vor allem auf stadteinwärts führenden Streckenzügen verbessern. Eine Mittelung über alle Kenngrößen im Netz würde folglich die Qualitätsunterschiede zwischen den Schätzergebnissen beider Varianten verwässern.

Abkürzungen in Tabellen und Grafiken bei der Bewertung:

n	Anzahl der bei der Zusammenführung berücksichtigten Einzelschätzungen
0-24/7	Montag bis Sonntag, 0-24 Uhr
6-10/5	Montag bis Freitag, 6-10 Uhr
mAR	mit Autoregression
oAR	ohne Autoregression

7.2 Zusammenführung der Einzelschätzungen und Kalibrierung

Die entwickelte Methode kann zahlreiche einzelne Schätzungen einer einzelnen Kenngröße zu einem Zeitschritt liefern. Diese sind Resultat ebenso zahlreicher Näherungsfunktionen, die bessere oder schlechtere Zusammenhänge zwischen Kenngrößen beschreiben. Es ist zielführend, bei der Zusammenführung der Einzelschätzungen die Schätzergebnisse zu berücksichtigen, die aus starken Zusammenhängen resultieren, deren Zusammenhänge also die höchsten Bestimmtheitsmaße oder Beträge des Korrelationskoeffizienten aufweisen. Die Anzahl der berücksichtigten Einzelschätzungen ist nachfolgend festzulegen. Bestimmtheitsmaß und Betrag des Korrelationskoeffizienten können zudem bei einer Mittelung der ausgewählten Schätzwerte als Gewicht eingesetzt werden, um zu gewährleisten, dass einzelne sehr zuverlässige Schätzungen besonders starken Eingang ins Endergebnis finden. Der Wert der Teststatistik ist hierfür ungeeignet, da er neben anderem auch vom Stichprobenumfang abhängt und so vor allem im Fall besonderer Verkehrsverhältnisse zum Verwerfen von Einzelschätzungen aus guten Zusammenhängen führen kann.

Es muss damit gerechnet werden, dass zusammengeführte Schätzwerte zunächst einer recht defensiven Schätzung gleichkommen können, also die besonderen Ereignisse zwar erkannt, ihre Auswirkungen aber unterschätzt werden. Erfolgt dies systematisch, so kann dieser Umstand durch eine Kalibrierung anhand von Schätzungen für die verwendete Datenhistorie behoben werden. Da sich hierbei nicht über die komplette Skala ein linearer Zusammenhang einstellen muss, wird die Kalibrierung abschnittsweise differenziert gemäß Kapitel 4 durchgeführt. Im Zuge der Kalibrierung kann beobachtet werden, welche Schätzfunktionen letztlich Eingang in vermittelte Schätzwerte finden und welche nicht. Dies ermöglicht die Beantwortung folgender Fragen:

- Welche Datenarten werden bevorzugt als Eingangsgröße verwendet?
- Differiert diese Auswahl für verschiedene Datenarten als Ausgangsgröße?
- Über welche räumlichen Distanzen zwischen den Erhebungsorten von Ein- und Ausgangsgröße erfolgen die verwendeten Einzelschätzungen?
- Welcher Anteil der ermittelten Schätzfunktionen wird letztlich überhaupt verwendet?
- Welchen Anteil machen Schätzungen aus Messwerten vorangegangener Zeitschritte aus?

Für eine Kalibrierung ist es wünschenswert, dass für die Schätzung einer Kenngröße in mehreren Zeitschritten meist dieselben Näherungsfunktionen verwendet werden. Werden gering verfügbare Daten wie Linkgeschwindigkeiten aus Positionsmeldungen als Eingangsgrößen eingesetzt, so gilt dies nur mehr bedingt, da sie nur in einzelnen Zeitschritten vorliegen und den vermittelten Schätzwert deutlich beeinflussen können, in zahlreichen anderen aber nicht. Daher und aufgrund ihrer fragwürdigen Repräsentativität werden Linkgeschwindigkeiten als Eingangsgröße nur dann verwendet, wenn keine infrastrukturseitig erhobenen Kenngrößen für die Schätzung vorliegen. Für die Schätzung aus Linkgeschwindigkeiten wird keine Kalibrierung vorgenommen.

Weiterhin wird im Fall der Verwendung von Linkgeschwindigkeiten als Eingangsgröße pro zeitlichem Versatz nur eine Einzelschätzung zugelassen. So wird verhindert, dass ein Fahrzeug, das zwischen zwei gemeldeten Positionen mehrere Links befahren hat, mit mehreren Einzelschätzungen Eingang in die vermittelte Schätzung findet. Schätzungen ausschließlich aus Linkgeschwindigkeiten sind dort zu erwarten, wo infrastrukturseitige Detektionseinrichtungen komplett fehlen. Im Bewertungsnetz gemäß Bild 6.1 betrifft dies insbesondere den Streckenzug „Leopoldstraße“.

Je nach Anwendungsfall kann das Zulassen der Autoregression, also Schätzungen von Kenngrößen aus ihren Messwerten der vorangegangenen Zeitschritte, sinnvoll sein oder nicht. Zur Untersuchung der Wirkungen werden daher Ergebnisse mit und ohne Autoregression betrachtet.

n	RMSE	Erkennungsrate Perzentilereignis [%]					Erkennungsrate Perzentilereignis (weich) [%]				
		1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
1	0.78	46	50	60	83	82	69	65	86	95	93
2	0.86	31	50	62	84	84	62	73	88	94	92
3	0.93	31	42	66	85	83	62	73	89	94	91
4	0.98	31	38	63	86	82	62	69	89	96	92
5	1.01	31	35	66	85	84	62	73	89	97	93
6	1.04	31	38	63	85	84	69	69	86	96	93
7	1.07	31	38	66	85	84	54	69	88	96	92
8	1.09	31	35	63	84	84	62	69	89	95	92
9	1.11	38	35	62	85	82	54	65	89	96	92
10	1.13	38	35	63	85	79	54	65	89	95	92
15	1.16	38	38	63	84	80	54	62	94	96	92
20	1.19	31	38	62	85	80	46	58	94	95	92

Tab. 7.1 Erkennungsrate von Perzentilereignissen der Reisezeit auf einer Route, Autoregression zugelassen; mit dem Bestimmtheitsmaß gewichtete Mittelung über die n besten Einzelschätzungen

n	RMSE	Erkennungsrate Perzentilereignis [%]					Erkennungsrate Perzentilereignis (weich) [%]				
		1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
1	81.01	21	21	32	43	63	37	39	54	60	85
2	68.27	26	32	52	57	71	47	63	70	73	88
3	70.66	21	26	46	55	68	42	53	70	72	88
4	71.1	26	26	47	54	71	42	58	68	71	87
5	71.77	26	29	47	53	70	37	58	69	70	87
6	73.65	26	29	41	49	68	42	55	61	69	87
7	75.11	26	32	38	51	67	42	50	60	66	87
8	75.88	32	29	39	48	67	42	53	60	65	86
9	76.54	32	29	34	47	65	47	50	58	66	86
10	74.11	32	29	41	52	67	47	53	63	68	87
15	76.11	21	26	40	52	66	47	55	63	68	88
20	78.16	26	26	38	49	65	47	55	59	67	85

Tab. 7.2 Erkennungsrate von Perzentilereignissen der Verkehrsstärke an einem strategischen Detektor, Autoregression nicht zugelassen; mit dem Bestimmtheitsmaß gewichtete Mittelung über die n besten Einzelschätzungen

Tab. 7.1 und Tab. 7.2 sollen nicht zur Bewertung der Schätzergebnisse herangezogen werden. Sie sind Beispiele für zahlreiche Tabellen, auf deren Grundlage die Anzahl der bei der Zusammenführung berücksichtigten Einzelschätzungen festgelegt wird. Hierbei wurden auch Tabellen für den Fall berücksichtigt, dass nur Daten in die Schätzung eingehen, die nicht am Link der Schätzgröße gesammelt werden. Diese Tabellen zeigen, dass sehr seltene Ereignisse bei der Berücksichtigung sehr weniger Einzelschätzungen am besten erkannt werden. Auf der anderen Seite werden weniger seltene Ereignisse oft häufiger richtig geschätzt, wenn die Anzahl der berücksichtigten Einzelschätzungen etwas höher liegt. Dabei nimmt das Fehlermaß für mehr berücksichtigte Einzelschätzungen zu. Es wird festgelegt, dass für die Berücksichtigung von drei und fünf Einzelschätzungen weitere Bewertungen vorgenommen werden.

Die Kalibrierung der zusammengeführten Schätzungen erfolgt durch Gegenüberstellung der historischen Messwerte und ihrer Schätzwerte aus den ermittelten Näherungsfunktionen. Im Fall eines idealen Schätzers liegen die Schätzwerte bei einer grafischen Darstellung auf der Mittelgeraden. Systematische Fehler lassen sich an einer konstanten oder proportionalen Abweichung von dieser Mittelgeraden erkennen, während sich zufällige Fehler in der vertikalen Verteilung um diese konstante oder proportionale Abweichung niederschlagen. Treten systematische Fehler auf, so kann durch die Verwendung einer Korrekturfunktion aus der Kalibrierung eine Verbesserung der Schätzung erzielt werden.

Die Kalibrierungsbeispiele gemäß Bild 7.2 bestätigen im Grundsatz, dass im unkalibrierten Fall insbesondere seltene Verkehrsverhältnisse häufig defensiv geschätzt, also Geschwindigkeiten über- sowie Belegungsgrade und Reisezeiten unterschätzt werden. Die Funktionen werden durch abschnittsweise Regressionsanalyse für drei Abschnitte bei einem Überlappungsgrad von 0,13 bei mindestens fünf Punkten je Abschnitt ermittelt.

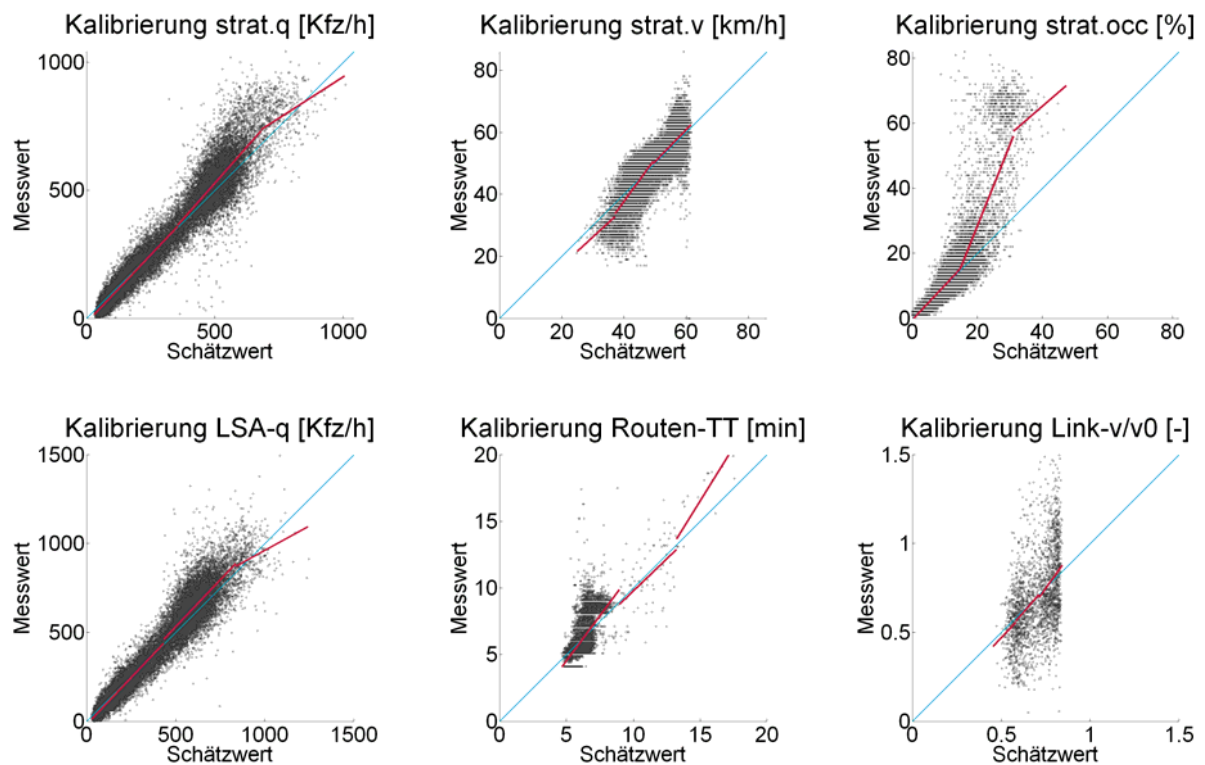


Bild 7.2 Beispiele für die Kalibrierung der Schätzung (rot) verschiedener Kenngrößenarten (blau – idealer unkalibrierter Schätzer, schwarz – Gegenüberstellung einzelner Schätz- und Messwerte)

Die Schätzung von Verkehrsstärken scheint auch schon im unkalibrierten Fall recht gut zu gelingen. Ursachen hierfür könnten sein, dass hohe Verkehrsstärken allein noch kein Störungsindikator sind und dass meist Eingangsdaten von benachbarten Fahrstreifen eine hochwertige Schätzung ermöglichen.

Das gezeigte und weitere Beispiele für die Kalibrierung von linkbezogenen Geschwindigkeiten aus Positionsmeldungen zeigen, dass die Punktwolken hier vergleichsweise stark streuen. Dies ist einmal mehr Folge der geringeren Repräsentativität dieser aus Messungen abgeleiteten Kennwerte.

Im weiteren Verlauf werden unkalibrierte und kalibrierte Schätzergebnisse separat untersucht, um die Wirkung der Kalibrierung auf die Qualität der Verfahrensergebnisse quantifizieren zu können.

7.3 Bewertung geschätzter lokaler Kenngrößen

Die Bewertung der Schätzergebnisse für lokale Kenngrößen wird nach folgenden Ausgangsdatenarten gegliedert:

- Strategisch erhobene Verkehrsstärke
- Strategisch erhobene lokale mittlere Geschwindigkeit
- Strategisch erhobener lokaler Belegungsgrad
- Mit LSA-Detektoren erhobene Verkehrsstärke

Schätzung strategisch erhobener Verkehrsstärken

Die Schätzung der strategisch erhobenen Verkehrsstärke wird am Beispiel eines ausgewählten Detektors vorgenommen. In Tab. 7.3 f. sind zunächst die Ganglinienschätzungen mit den Schätzergebnissen des entwickelten Verfahrens ohne Autoregression sowie die Trivialprognosen mit den Schätzergebnissen mit Autoregression zu vergleichen. Nachfolgend wird ein Vergleich zwischen den Resultaten mit und ohne Autoregression angestellt sowie der Einfluss der Anzahl berücksichtigter Einzelschätzungen und der Kalibrierung diskutiert. Bei der Bewertung der Schätzung weiterer Kenngrößenarten wird weitgehend analog vorgegangen.

Aufgrund der hohen Verfügbarkeit lokaler Daten unterscheiden sich die Ergebnisse für die beiden Varianten der Trivialprognose häufig kaum. Daher werden die Bewertungskenngrößen in den Tabellen für beide gemeinsam angegeben.

Bild 7.3 unterscheidet sich von späteren Abbildungen für andere geschätzte Kenngrößenarten dadurch, dass die Messwerte von Zeitschritt zu Zeitschritt vergleichsweise stark schwanken. Auf die Darstellung der zusammengeführten Schätzergebnisse ohne Autoregression wird verzichtet, da sie den Ergebnissen mit Autoregression sehr ähnlich sind und die Darstellungen somit sehr unübersichtlich werden.

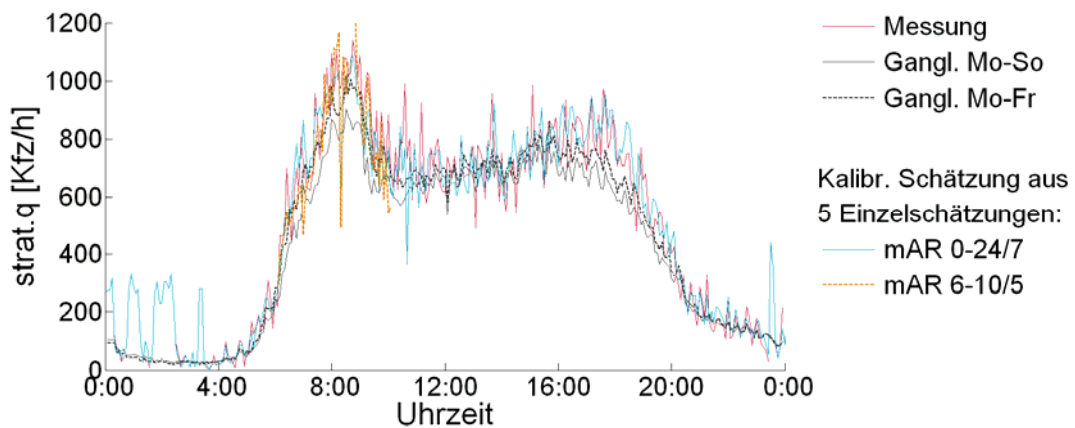


Bild 7.3 Grafischer Vergleich der Verkehrsstärkemessung mit ihrer Schätzung; Dienstag, 17.3.2009

Gemäß Tab. 7.3 ergeben sich für das neue Verfahren ohne Autoregression weit höhere Erkennungsraten von Perzentilereignissen als für das Ganglinienverfahren. Dies gilt im Grundsatz auch für Tab. 7.4, jedoch mit einzelnen Ausnahmen für die Perzentilereignisse 10%, 15%, 25%, 10(15)% und 15(25)% bei zeitlich differenzierter Regressionsanalyse. Auch gelingt für diese Ereignisse überraschenderweise die Schätzung aus der zeitlich nicht differenzierten Regressionsanalyse geringfügig besser als aus der differenzierten. Als Ursache wird hierfür das Schwanken des Messwerts in Betracht gezogen, mit dem ein Vorteil unempfindlicher Verfahren einhergeht. Dafür spricht auch, dass in Tab. 7.4 für das Ganglinienverfahren fast durchgängig höhere Erkennungsraten von Perzentilereignissen aufgeführt sind als für die Trivialverfahren. Die starken Schwankungen während der Hauptverkehrszeit schlagen sich besonders deutlich in erhöhten Fehlermaßen für die Referenzverfahren im Vergleich zu Tab. 7.3 nieder.

Da die Trivialaussage „Alles bleibt so, wie es vorhin war“ für diese Kenngröße offensichtlich keine besonders guten Schätzungen liefert, verwundert es wenig, dass die Miteinbeziehung der Autoregression in beiden Tabellen keine ersichtliche Verbesserung bringt.

Tab. 7.4 zeigt, dass die Erkennungsraten der Perzentilereignisse durch die Kalibrierung kaum beeinflusst werden. Dies konnte erwartet werden, da die Kalibrierung die Reihenfolge der Schätzungen nicht verändert, lediglich an den Übergängen zwischen den Kalibrierungsabschnitten kann es zu Verschiebungen kommen. Bereits die unkalibrierten Schätzergebnisse des neuen Verfahrens weisen geringere Fehlermaße auf als die Vergleichsverfahren. Im kalibrierten Fall gelingt eine weitere Verbesserung. Das Fehlermaß wird deutlich verringert mit Ausnahme der Ergebnisse unter Verwendung der Näherungsfunktionen aus der zeitlich differenzierten Regressionsanalyse. Ursache hierfür ist, dass die Kalibrierung für die zeitlich nicht differenzierte Historie vorgenommen wurde und so auch für die zeitlich differenzierte Schätzung eingesetzt wird. Eine separate Ermittlung der Verbesserungsfunktionen für den zeitlich differenzierten Fall sollte diesen Umstand beheben.

Analysierte Datenhistorie	Kalibriert?	Autoregr.?	n = 3 oder 5?	Verfüg. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
Mo-So	Ganglinie			100	181.96	16	32	55	51	55	47	81	72	62	70
Mo-Fr	Ganglinie			100	99.32	23	54	66	62	66	69	85	80	75	81
	Trivial/Trivial+			100	103.86/103.29	37	41	52	56/57	66	58	65	80	69/70	85
0-24/7	nein	mAR	5	100	82.60	42	59	66	68	74	89	86	83	80	90
		mAR	3	100	82.97	53	59	67	67	74	89	92	85	80	90
		oAR	5	100	82.63	42	59	66	68	74	89	86	84	80	90
		oAR	3	100	82.97	53	59	67	67	74	89	92	85	80	90
	ja	mAR	5	100	75.86	42	59	65	68	74	89	86	83	80	90
		mAR	3	100	75.95	53	54	67	67	74	89	92	85	80	90
		oAR	5	100	75.87	42	59	65	68	74	89	86	84	80	90
		oAR	3	100	75.95	53	54	67	67	74	89	92	85	80	90

Tab. 7.3 Schätzung strategisch erhobener Verkehrsstärken im Zeitraum 0-24/7 – Bewertung

Analysierte Datenhistorie	Kalibriert?	Autoregr.?	n = 3 oder 5?	Verfüg. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						5%	10%	15%	25%	50%	5(10)%	10(15)%	15(25)%	25(35)%	50(65)%
Mo-So	Ganglinie			100	199.72	17	29	53	77	81	33	50	86	83	94
Mo-Fr	Ganglinie			100	131.53	25	50	67	72	83	58	67	81	83	95
	Trivial/Trivial+			100	159.84/159.61	17	42	44	62	78	42	46	69	85	88
0-24/7	nein	mAR	5	100	117.37	42	42	67	77	85	58	71	86	88	93
		mAR	3	100	121.02	50	54	67	75	85	67	75	89	88	92
		oAR	5	100	117.52	42	42	67	77	85	58	71	86	88	93
		oAR	3	100	121.02	50	54	67	75	85	67	75	89	88	92
	ja	mAR	5	100	105.93	42	42	67	77	85	58	71	86	87	93
		mAR	3	100	106.41	50	50	61	75	85	58	71	89	90	92
		oAR	5	100	105.99	42	42	67	77	85	58	71	86	87	93
		oAR	3	100	106.41	50	50	61	75	85	58	71	89	90	92
6-10/5	nein	mAR	5	100	109.67	33	54	61	75	86	67	79	83	92	96
		mAR	3	100	114.82	50	54	56	75	88	67	71	86	90	97
		oAR	5	100	109.67	33	54	61	75	86	67	79	83	92	96
		oAR	3	100	114.82	50	54	56	75	88	67	71	86	90	97
	ja	mAR	5	94	125.38	45	52	59	68	83	73	65	82	91	93
		mAR	3	98	121.47	25	43	54	72	88	58	61	86	91	97
		oAR	5	94	125.12	45	52	59	68	83	73	65	82	91	94
		oAR	3	98	121.47	25	43	54	72	88	58	61	86	91	97

Tab. 7.4 Schätzung strategisch erhobener Verkehrsstärken im Zeitraum 6-10/5 – Bewertung

Für die Berücksichtigung von drei oder fünf Einzelschätzungen bei der Zusammenführung ergeben sich hinsichtlich der Erkennungsraten von Perzentilereignissen nur sehr geringe Unterschiede. Das Fehlermaß RMSE ist für die Schätzung aus fünf Einzelschätzungen jedoch fast durchgängig geringer und damit besser als für den anderen untersuchten Fall.

Bild 7.3 veranschaulicht die Vorteile des entwickelten Ansatzes gegenüber dem Ganglinienverfahren insbesondere für hohe Verkehrsstärken. Auf die Darstellung der Trivialschätzung wird verzichtet, da sie zeitlich gering versetzt die Messwerte wiedergeben würde. Insbeson-

dere nachts gemessene geringe Verkehrsstärken werden offensichtlich gelegentlich überschätzt. Diese Tatsache wirkt sich negativ auf das ermittelte Fehlermaß aus, hätte aber bei einer praktischen Anwendung maximal sehr geringe Auswirkungen, da auch die geschätzten Verkehrsstärken auf einem sehr geringen Niveau liegen. Da in diesem Zeitfenster kein Referenzwert vorliegt, ist als Ursache dafür wahrscheinlich, dass auch zahlreiche andere Messwerte aus der räumlichen Nähe hier nicht verfügbar waren.

Die höheren Fehlermaße in Tab. 7.4 im Vergleich zu Tab. 7.3 resultieren aus der Tatsache, dass im Zeitfenster Montag bis Freitag, 6 bis 10 Uhr vergleichsweise viele verkehrliche Ereignisse und damit implizierte Schwankungen der Kennwerte auftreten. Die reduzierte Verfügbarkeit für kalibrierte Schätzungen bei Nutzung der Schätzfunktionen aus der zeitlich differenzierten Regressionsanalyse kann durch die Seltenheit vergleichbarer Ereignisse in der Historie und dem damit verbundenen Ausbleiben einer Kalibrierungsfunktion für diesen Skalenbereich verursacht werden.

Schätzung strategisch erhobener mittlerer Geschwindigkeiten

Das neue Verfahren ohne Autoregression liefert bei der Schätzung mittlerer lokaler Geschwindigkeiten durchgehend bessere Ergebnisse als das Ganglinienverfahren und fast immer bessere Ergebnisse als die Trivialverfahren (Tab. 7.5). Dies betrifft sowohl die Erkennungsraten von Perzentilereignissen als auch das Fehlermaß RMSE. Durch Zulassen der Autoregression verbessern sich die Ergebnisse weiter.

Für den Fall der Verwendung von fünf statt drei Einzelschätzungen ergeben sich im Allgemeinen höhere Raten und damit bessere Ergebnisse. Die diesbezüglichen Unterschiede beim untersuchten Fehlermaß sind recht gering.

Für die untersuchte Kenngröße verbessert die zeitliche Differenzierung bei der Regressionsanalyse die Schätzung erheblich: Die Detektionsraten der Perzentilereignisse verbessern sich drastisch und liegen auf sehr hohem Niveau, ebenso geht das Fehlermaß RMSE deutlich zurück. Vergleichsweise schwach wirkt sich die Kalibrierung auf das Fehlermaß aus.

Tab. 7.6 dokumentiert durchgehend Vorteile der neuen Methode gegenüber den jeweiligen Vergleichsverfahren. Für die zeitlich differenzierte Regressionsanalyse ergeben sich geringere Fehlermaße und außer für sehr seltene Ereignisse höhere Erkennungsraten als für den undifferenzierten Fall.

Analysierte Datenhistorie	Kalibriert?	Autoregr.?	n= 3 oder 5?	Verfüg. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
Mo-So	Ganglinie			100	5.04	13	14	22	37	46	13	14	33	42	66
Mo-Fr	Ganglinie			100	4.84	0	14	28	39	52	0	24	40	48	69
	Trivial/Trivial+			100	5.06/4.95	27	45	40	51/52	60/61	60	66	60/62	64	76/77
0-24/7	nein	mAR	5	100	4.00	53	66	64	72	76	73	93	85	84	87
		mAR	3	100	4.03	27	48	62	69	74	47	83	81	84	86
		oAR	5	100	4.11	40	59	58	68	75	47	86	78	82	86
		oAR	3	100	4.04	27	52	60	69	74	47	83	81	84	86
	ja	mAR	5	100	3.82	53	66	64	72	76	73	93	85	84	87
		mAR	3	100	3.92	27	48	60	69	74	47	79	81	84	85
		oAR	5	100	3.91	40	55	59	68	75	53	86	77	82	86
		oAR	3	100	3.90	33	52	59	68	74	53	83	81	84	86

Tab. 7.5 Schätzung strategisch erhobener mittlerer Geschwindigkeiten im Zeitraum 0-24/7 – Bewertung

Analysierte Datenhistorie	Kalibriert?	Autoregr.?	n= 3 oder 5?	Verfüg. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						5%	10%	15%	25%	50%	5(10)%	10(15)%	15(25)%	25(35)%	50(65)%
Mo-So	Ganglinie			100	7.31	17	21	44	57	84	25	38	53	73	98
Mo-Fr	Ganglinie			100	6.49	25	25	47	58	84	33	38	56	73	97
	Trivial/Trivial+			100	5.78/5.77	50	50	58	75	83	83	58	89	87	95
0-24/7	nein	mAR	5	100	5.44	75	75	72	83	88	92	83	94	92	98
		mAR	3	100	5.54	42	58	64	83	88	58	67	89	92	98
		oAR	5	100	6.3	42	42	56	72	86	50	54	75	83	95
		oAR	3	100	6.22	50	46	50	72	87	58	54	75	82	97
	ja	mAR	5	100	5.02	75	75	72	83	88	92	83	94	92	98
		mAR	3	100	5.23	42	58	64	83	88	58	67	86	92	97
		oAR	5	100	5.77	50	50	56	73	85	58	67	75	83	94
		oAR	3	100	5.76	58	50	56	73	87	67	67	78	82	97
6-10/5	nein	mAR	5	100	4.3	67	75	75	83	90	100	88	100	95	98
		mAR	3	100	3.84	58	71	81	88	89	100	96	97	97	98
		oAR	5	100	4.3	67	75	75	83	90	100	88	100	95	98
		oAR	3	100	3.84	58	71	81	88	89	100	96	97	97	98
	ja	mAR	5	100	4.25	58	71	75	83	90	100	88	100	95	98
		mAR	3	100	4.02	58	71	81	88	89	100	96	97	97	98
		oAR	5	100	4.53	67	75	75	83	90	100	88	100	95	98
		oAR	3	100	4.27	58	71	81	88	90	100	96	97	97	98

Tab. 7.6 Schätzung strategisch erhobener mittlerer Geschwindigkeiten im Zeitraum 6-10/5 – Bewertung

In Bild 7.4 wird ersichtlich, dass das entwickelte Verfahren gegenüber der Ganglinienmethode Vorteile insbesondere bei markanten und damit bei den interessanten verkehrlichen Ereignissen hat: Im Fall eines starken Rückgangs der Geschwindigkeiten gelangen sehr gute Schätzungen, sicher begünstigt durch eingehende Kenngrößen, die am selben Querschnitt erhoben werden. Inwiefern sich die Ergebnisse ändern, wenn die Daten vom selben Quer-

schnitt nicht als Eingangsgrößen verwendet werden, wird neben anderem in Kapitel 7.5 beleuchtet.

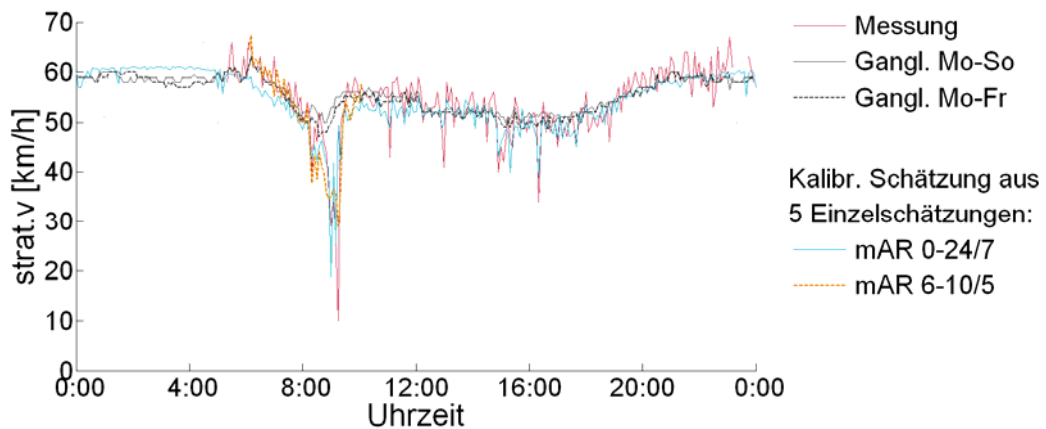


Bild 7.4 Grafischer Vergleich der Messung der mittleren Geschwindigkeit mit ihrer Schätzung; Mittwoch, 18.3.2009

Schätzung strategisch erhobener Belegungsgrade

Die entwickelte Methode ohne Autoregression liefert bei der Schätzung strategisch erhobener Belegungsgrade für die untersuchte Kenngröße gemäß Tab. 7.7 für sämtliche Erkennungsraten von Perzentilereignissen bessere Ergebnisse als das Ganglinienverfahren. Dies gilt auch für den Fall der Schätzung im Zeitraum Montag bis Freitag, 6-10 Uhr aus der zeitlich nicht differenzierten Regressionsanalyse (Tab. 7.8). Besondere Erwähnung verdienen die deutlich höheren Erkennungsraten der Perzentilereignisse 5% und 10% sowie der weichen Perzentilereignisse 5(10)%, 10(15)% und 15(25)% in Tab. 7.8 unter Verwendung der zeitlich differenzierten Historie (6-10/5).

Mit Ausnahme der seltensten Perzentilereignisse liefert das entwickelte Verfahren unter Verwendung der Autoregression höhere Erkennungsraten als das Trivialverfahren. Im Falle der zeitlich differenzierten Regressionsanalyse betrifft dies auch die seltensten Ereignisse.

Erwartungsgemäß gelingen Schätzungen, die die Autoregression zulassen, besser als jene, die solche Zusammenhänge ausschließen. Der Unterschied wächst für seltener werdende Ereignisse.

Das Fehlermaß wird durch die Kalibrierung gemäß Tab. 7.8 deutlich verringert. Wie bei der Schätzung der Verkehrsstärke liefern die unkalibrierten Schätzergebnisse des neuen Verfahrens etwas geringere Fehlermaße als die jeweiligen Vergleichsverfahren. Tendenziell liefert die Zusammenführung von drei statt fünf Einzelschätzungen die besseren Schätzungen. Dies wird dadurch begünstigt, dass lokale Kenngrößen aus solchen Kenngrößen geschätzt werden können, die nahe dem Querschnitt der Schätzgröße gesammelt werden. Dies führt

zur Beobachtung einiger sehr starker Zusammenhänge. Das Hinzuziehen weiterer, schwächerer Zusammenhänge beeinflusst dann die zusammengeführte Schätzung negativ.

Analyzierte Datenhistorie	Kalibriert?	Autoregr.?	n= 3 oder 5?	Verfügb. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
Mo-So	Ganglinie			100	5.63	0	10	38	48	55	0	24	58	59	72
Mo-Fr	Ganglinie			100	5.63	0	14	42	55	64	9	24	58	66	76
	Trivial/Trivial+			100	4.36/4.39	47	38	52	59	72	60	69	66	75	85
0-24/7	nein	mAR	5	100	4.38	20	38	67	73	83	47	72	82	85	94
		mAR	3	100	4.22	27	38	68	73	82	53	76	85	84	95
		oAR	5	100	4.58	7	38	51	65	79	27	55	68	76	89
		oAR	3	100	4.42	20	41	53	69	80	40	59	74	80	90
	ja	mAR	5	97	3.95	20	38	67	72	83	47	72	82	85	93
		mAR	3	99	3.92	27	38	68	73	82	53	76	85	84	95
		oAR	5	97	4.25	7	41	51	64	78	27	55	67	75	89
		oAR	3	99	4.13	20	41	56	69	80	33	62	74	80	90

Tab. 7.7 Schätzung strategisch erhobener Belegungsgrade im Zeitraum 0-24/7 – Bewertung

Analyzierte Datenhistorie	Kalibriert?	Autoregr.?	n= 3 oder 5?	Verfügb. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						5%	10%	15%	25%	50%	5(10)%	10(15)%	15(25)%	25(35)%	50(65)%
Mo-So	Ganglinie			100	10.61	0	21	39	57	84	0	29	53	73	93
Mo-Fr	Ganglinie			100	9.76	0	17	44	67	88	8	33	64	72	97
	Trivial/Trivial+			100	7.57/7.56	50	46	56	68	87	50	63	69	83	96
0-24/7	nein	mAR	5	100	9.25	25	46	67	75	90	25	67	78	83	98
		mAR	3	100	8.97	25	54	67	80	91	50	71	83	88	99
		oAR	5	100	9.57	8	50	50	63	88	33	54	64	80	95
		oAR	3	100	9.28	25	50	56	73	88	42	63	75	83	95
	ja	mAR	5	100	8.26	25	46	67	75	90	25	67	78	83	98
		mAR	3	100	8.12	25	54	67	80	91	33	71	83	88	99
		oAR	5	100	8.72	8	54	56	63	88	33	54	64	80	95
		oAR	3	100	8.49	25	54	61	73	88	42	63	75	83	95
6-10/5	nein	mAR	5	100	7.74	33	67	72	78	92	58	83	92	95	99
		mAR	3	100	7.16	42	63	78	78	91	75	83	89	93	100
		oAR	5	100	7.74	33	67	72	78	92	58	83	92	95	99
		oAR	3	100	7.16	42	63	78	78	91	75	83	89	93	100
	ja	mAR	5	100	6.79	33	67	72	78	92	58	79	89	95	98
		mAR	3	100	6.77	33	58	75	82	91	58	83	89	93	100
		oAR	5	100	6.78	33	67	72	78	92	58	79	89	95	98
		oAR	3	100	6.68	33	63	75	82	91	67	83	89	93	100

Tab. 7.8 Schätzung strategisch erhobener Belegungsgrade im Zeitraum 6-10/5 – Bewertung

Wie erwartet gelingen im untersuchten Beispiel die besten Schätzungen auf der Grundlage der zeitlich differenzierten Regressionsanalyse unter Verwendung von Einzelschätzungen aus der Autoregression und von Kalibrierungsfunktionen.

Bild 7.5 zeigt die Vorteile des entwickelten Ansatzes gegenüber dem Ganglinienverfahren insbesondere für sehr hohe Belegungsgrade und somit für sehr seltene Ereignisse.

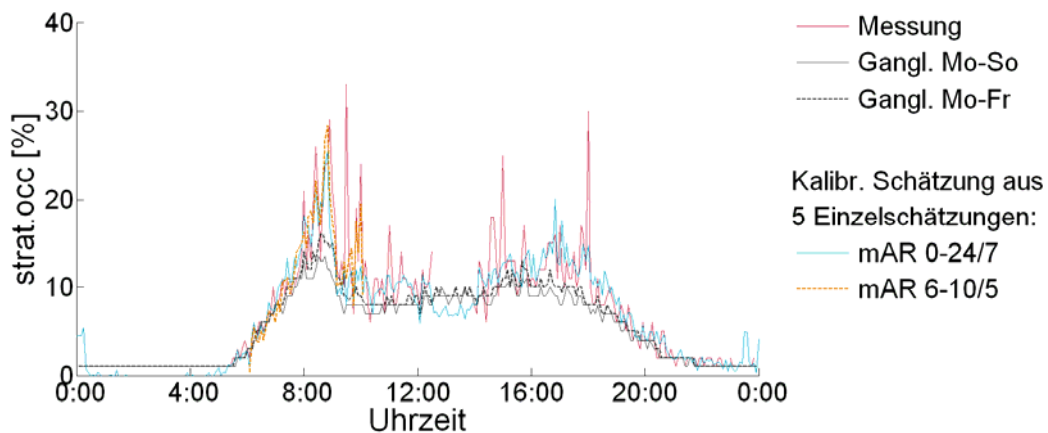


Bild 7.5 Grafischer Vergleich der Belegungsgradmessung mit ihrer Schätzung durch unterschiedliche Verfahren; Donnerstag, 19.3.2009

Schätzung von Verkehrsstärken an LSA-Detektoren

Das ausgewählte Beispiel zur Bewertung der geschätzten Verkehrsstärken an LSA-Detektoren bestätigt im Wesentlichen die bisherigen Beobachtungen. Die Ergebnisse für die Berücksichtigung von drei Einzelschätzungen werden nicht angezeigt; sie liefern unwesentlich geringere Erkennungsraten und etwas höhere Werte für den RMSE.

Außer für das 1%-Ereignis ergeben sich für das Verfahren ohne Autoregression durchgehend höhere Erkennungsraten und geringere RMSE als für das Ganglinienverfahren. Mit Autoregression sind die Ergebnisse durchgängig besser als für die Trivialverfahren. Hier sind die Unterschiede tendenziell für die sehr seltenen Ereignisse am größten (Tab. 7.9 f.). Es stellen sich jedoch entgegen den Erwartungen für den Zeitraum 6-10/5 bei Berücksichtigung der Autoregression sogar geringfügig geringere Raten als für den Fall ohne Autoregression ein (Tab. 7.10). Grund hierfür ist der lebhafte Verkehr in diesem Zeitfenster mit schnell und stark schwankenden Kenngrößen, deren Schätzung sich verschlechtert, wenn eine für Autoregression typische Einzelschätzung „Alles bleibt, wie es ist“ einfließt. Dies lässt sich auch am höheren Fehlermaß RMSE für die Trivialprognose im Vergleich zur Ganglinie (Mo-Fr) erkennen. Es ist zu vermuten, dass die Schätzung aus einem Messwert vom Nachbarfahrstreifen besser gelingt als aus dem eigenen Messwert des vorangegangenen Zeitschritts.

Das Fehlermaß RMSE steigt im zeitlich differenzierten Fall bei der Kalibrierung entgegen den Erwartungen an. Ursache hierfür ist wie bei der Schätzung strategisch erhobener Verkehrsstärken die Verwendung der Kalibrierungsfunktionen aus dem zeitlich undifferenzierten Fall auch zur Anpassung der zeitlich differenzierten Schätzung. Um hier die beabsichtigte Verbesserung erzielen zu können, müsste die Kalibrierung ebenso zeitlich differenziert erfolgen.

Analysierte Datenhistorie	Kalibriert?	Autoregr.?	n= 3 oder 5?	Verfügb. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
Mo-So	Ganglinie			100	157.64	42	30	42	44	53	42	62	59	56	73
Mo-Fr	Ganglinie			100	97.38	31	27	57	58	66	46	69	75	69	81
	Trivial/Trivial+			100	109.09/108.45	16	35	46	46	62	32	59	63	65/64	83/82
0-24/7	nein	mAR	5	100	73.51	42	49	65	69	76	79	84	89	83	90
		oAR	5	100	78.17	32	43	62	66	73	68	86	88	82	90
	ja	mAR	5	100	69.18	42	49	65	69	76	79	84	89	83	90
		oAR	5	100	73.57	37	41	62	67	73	63	84	88	82	90

Tab. 7.9 Schätzung von Verkehrsstärken an LSA-Detektoren im Zeitraum 0-24/7 – Bewertung

Analysierte Datenhistorie	Kalibriert?	Autoregr.?	n= 3 oder 5?	Verfügb. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						5%	10%	15%	25%	50%	5(10)%	10(15)%	15(25)%	25(35)%	50(65)%
Mo-So	Ganglinie			100	203.7	45	43	41	63	84	55	43	71	79	95
Mo-Fr	Ganglinie			100	146.21	27	26	38	68	82	45	35	74	82	94
	Trivial/Trivial+			100	168.32/168.13	9	30	41	61	79	18	52	62	74	91
0-24/7	nein	mAR	5	100	119.58	36	48	59	77	86	55	70	88	93	96
		oAR	5	100	123.79	45	48	56	75	84	55	65	88	91	95
	ja	mAR	5	100	110.52	36	48	59	77	86	55	70	85	93	96
		oAR	5	100	115.15	55	43	53	77	84	55	61	88	91	95
6-10/5	nein	mAR	5	100	111.03	45	48	59	72	87	64	65	79	93	95
		oAR	5	100	116.13	45	48	62	74	86	64	65	82	93	95
	ja	mAR	5	94	126.92	27	48	56	74	86	45	57	84	94	95
		oAR	5	94	132.60	27	52	59	74	84	45	62	81	94	93

Tab. 7.10 Schätzung von Verkehrsstärken an LSA-Detektoren im Zeitraum 6-10/5 – Bewertung

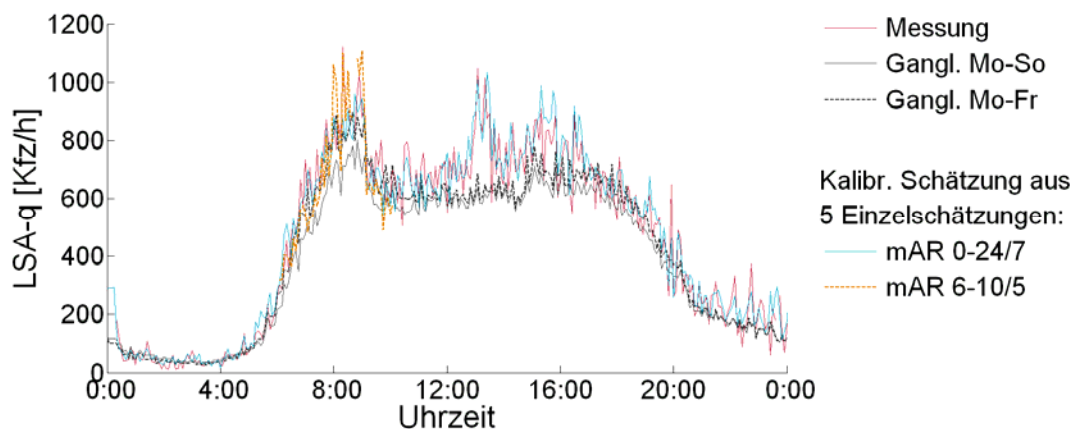


Bild 7.6 Grafischer Vergleich der Verkehrsstärkemessung an einem LSA-Detektor mit ihrer Schätzung; Freitag, 20.3.2009

In Bild 7.6 sind Vorteile für die sehr hohen Werte sowohl in der Vormittagsspitze als auch für die nachmittägliche Hauptverkehrszeit zu erkennen. Im Gegensatz zum eingesetzten Ganglinienverfahren kann die entwickelte Methode hier auf die freitags frühere und häufig auch stärkere Abendspitze bezüglich der Verkehrsnachfrage reagieren.

7.4 Bewertung geschätzter Abschnittsdaten

Gemäß Tab. 7.11 f. gelingt die Schätzung von Routenreisezeiten mit dem neuen Verfahren besser als mit dem Ganglinienverfahren. Dafür sprechen die überwiegend höheren Erkennungsraten sowie die geringeren Fehlermaße vor allem für den kalibrierten Fall. Die Verbesserungen durch die Kalibrierung werden durch wenig stetige Übergänge zwischen den Abschnitten der Kalibrierungsfunktion verursacht. Ähnliche Ergebnisse liefert der Vergleich der Regressionsmethode mit den Trivialverfahren, jedoch ist deren geringes Fehlermaß für die Schätzung dieser Kenngröße unerreicht. Dieser Unterschied zu den bisherigen Beobachtungen kann darin begründet liegen, dass die Routenreisezeit weniger stark schwankt als lokal erhobene Daten.

Analysierte Datenhistorie	Kalibriert?	Autoregr.?	n= 3 oder 5?	Verfügb. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
Mo-So	Ganglinie			90	1.52	10	24	58	73	73	30	62	94	85	83
Mo-Fr	Ganglinie			88	1.40	13	24	57	72	81	25	59	93	92	93
	Trivial/Trivial+			100	0.38/0.40	70	81	92	93	85	100	100	100	100	92/93
0-24/7	nein	mAR	5	99	0.90	70	57	94	92	80	90	100	100	99	94
		oAR	5	95	1.79	20	10	46	60	60	30	45	74	78	67
	ja	mAR	5	98	0.49	70	57	94	92	80	90	100	100	99	94
		oAR	5	92	0.84	10	32	77	73	58	90	100	92	75	66

Tab. 7.11 Schätzung von Routenreisezeiten im Zeitraum 0-24/7 – Bewertung

Analysierte Datenhistorie	Kalibriert?	Autoregr.?	n= 3 oder 5?	Verfügb. [%]	RMSE	Erkennungsrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
						5%	10%	15%	25%	50%	5(10)%	10(15)%	15(25)%	25(35)%	50(65)%
Mo-So	Ganglinie			100	3.16	18	36	48	69	82	27	45	76	85	92
Mo-Fr	Ganglinie			100	2.71	9	36	52	64	83	18	59	73	89	91
	Trivial/Trivial+			100	0.77	82	77	82	93	94	100	91	100	100	99
0-24/7	nein	mAR	5	100	1.88	73	59	79	95	95	91	82	100	100	100
		oAR	5	100	3.61	18	18	27	46	73	27	27	48	69	87
	ja	mAR	5	100	0.93	73	59	79	95	95	91	82	100	100	100
		oAR	5	85	1.35	0	28	63	80	88	44	61	100	89	91
6-10/5	nein	mAR	5	100	1.69	73	77	79	91	91	100	86	100	100	98
		oAR	5	100	2.87	9	55	53	57	76	36	59	69	72	90
	ja	mAR	5	100	1.26	82	77	82	91	91	100	95	100	100	98
		oAR	5	100	2.59	27	59	66	69	78	45	73	78	85	92

Tab. 7.12 Schätzung von Routenreisezeiten im Zeitraum 6-10/5 – Bewertung

Tab. 7.12 dokumentiert bei den Erkennungsraten durchgehende Vorteile für die zeitlich differenzierte Regressionsanalyse als Grundlage für die Schätzfunktionen, die Fehlermaße wiederum sind für den anderen untersuchten Fall geringer und damit besser.

Werden fünf Einzelschätzungen berücksichtigt, so liegen die Erkennungsraten von Perzentilereignissen fast immer etwas höher als für drei berücksichtigte Einzelschätzungen, die zusammengeführt wiederum etwas geringere Fehlermaße RMSE aufweisen. Die Ergebnisse für die Berücksichtigung dreier Einzelschätzungen werden in den Tabellen nicht aufgeführt.

Bild 7.7 zeigt wie die vorherigen Bilder Vorteile des neuen Schätzverfahrens vor allem für besondere Ereignisse. Zu den erhöhten Schätzwerten gegen 15 Uhr liegen keine Referenzmessungen vor. Da die Schätzwerte für andere Zeitbereiche ohne Referenzwerte wie z.B. nachts auf geringerem Niveau liegen, ist es denkbar, dass es sich hier um eine richtige Schätzung handelt. Es liegen aber keine Werkzeuge zur Überprüfung dieser Erwägung vor.

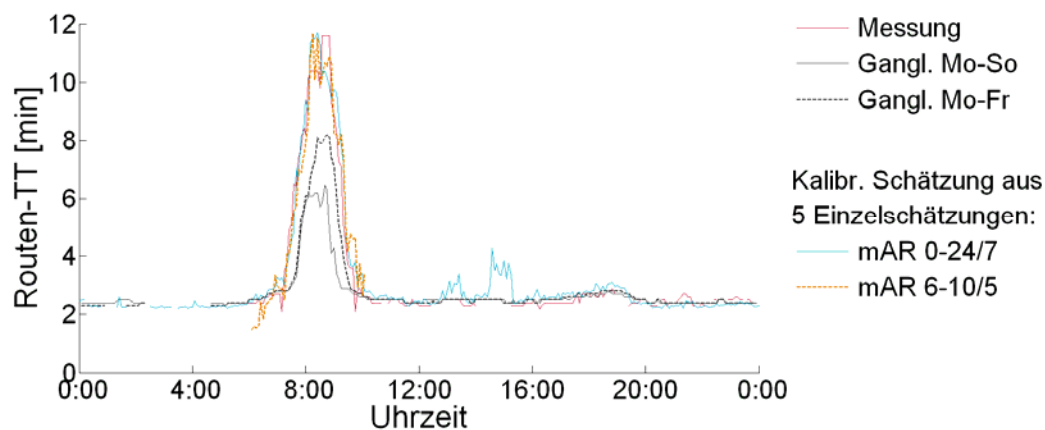


Bild 7.7 Grafischer Vergleich der Reisezeitmessung auf einer Route mit ihrer Schätzung; Dienstag, 17.3.2009, Route 1-2

7.5 Bewertung geschätzter Linkgeschwindigkeiten aus Positionsmeldungen

Aussagekräftige Referenzmessungen zur direkten Bewertung der geschätzten linkbezogenen Geschwindigkeiten bzw. Reisezeiten (in der Folge werden beide Begriffe verwendet) liegen nicht vor, da die Werte, die aus Positionsmeldungen abgeleitet werden, einzeln nicht als repräsentativ betrachtet werden können (Linkreisezeiten aus erhobenen Positionsmeldungen sind in Bild 7.8 ff. unter „Messung“ verzeichnet). Zur Beurteilung der Qualität der Schätzwerte werden daher folgende Schritte gewählt:

- Diskussion der zeitlichen Verläufe der Schätzwerte
- Qualitativer Abgleich der Schätzwerte mit vorhandenen Werten aus Positionsmeldungen

- Betrachtung und Diskussion der räumlich-zeitlichen verkehrlichen Muster, die sich aus den geschätzten Linkgeschwindigkeiten ergeben
- Vergleich erhobener Routenreisezeiten mit den Summen der geschätzten Linkreisezeiten über die Links, die auf der Route liegen

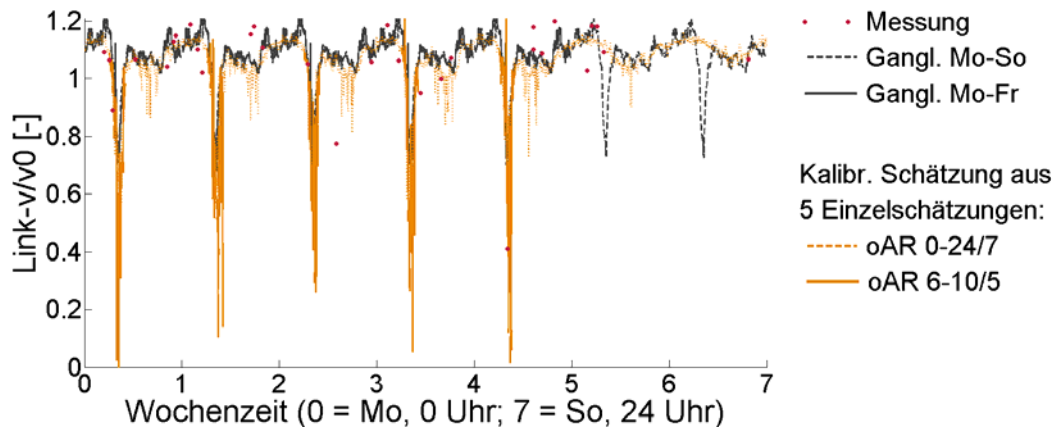


Bild 7.8 Grafischer Vergleich von linkbezogenen Geschwindigkeitsschätzungen an einem Link am Mittleren Ring, Höhe Olympiapark; Montag, 16.3.2009 bis Sonntag, 22.3.2009

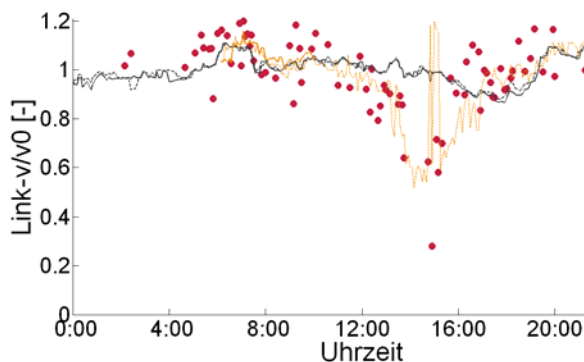


Bild 7.9 Grafischer Vergleich von linkbezogenen Geschwindigkeitsschätzungen an einem Link auf der A9 Richtung Nürnberg; Freitag, 20.3.2009 (Legende siehe Bild 7.8)

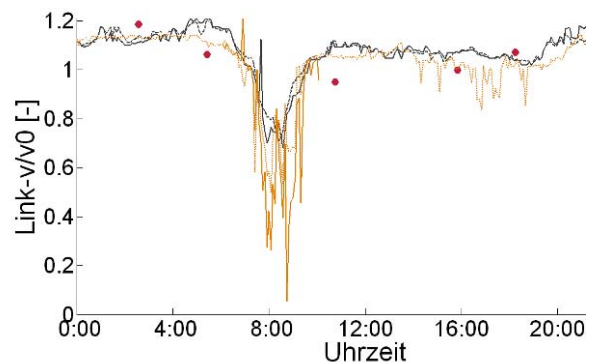


Bild 7.10 Grafischer Vergleich von linkbezogenen Geschwindigkeitsschätzungen an einem Link am Mittleren Ring, Höhe Olympiapark; Donnerstag, 19.3.2009 (Legende siehe Bild 7.8)

In Bild 7.8 ff. fehlen Graphen für die Schätzung mit Autoregression, da sich für Linkgeschwindigkeiten aufgrund der geringen Datenverfügbarkeit und der vergleichsweise großen zufälligen Fehler bei der Autoregression so gut wie keine signifikanten Zusammenhänge finden lassen. Somit sind die Schätzergebnisse für die Fälle mit oder ohne Autoregression in der Regel dieselben. Die Graphen verlaufen dann deckungsgleich.

Bild 7.8 zeigt einen plausiblen Verlauf der geschätzten linkbezogenen Geschwindigkeit für den gewählten Link über den Zeitraum der Bewertungswoche. Hier wie auch in Bild 7.9 ist insbesondere für die üblichen, also hohen Geschwindigkeiten erkennbar, dass die erhobe-

nen Werte merklich auch kurzfristig um die Schätzung schwanken. Dies ist ein Indiz für ihre fragwürdige Repräsentativität und stützt die These, dass hier die Schätzung die Verkehrslage besser widerspiegeln könnte als der einzelne erhobene Wert. Anders als die Ganglinie verläuft die Schätzung nicht jeden Tag identisch. Besonders starke Geschwindigkeitseinbrüche jeweils in der morgendlichen Hauptverkehrszeit am Montag und am Freitag sind plausibel, ebenso wie die deutlichsten nachmittäglichen Rückgänge freitags, die zudem – ebenfalls plausibel – früher erfolgen als an anderen Werktagen. Dass die Ganglinie über alle Tage (Mo-So) am Wochenende Nachteile gegenüber dem Schätzverfahren hat, war zu erwarten und soll nicht weiter in die Diskussion eingehen.

Es wird ersichtlich, dass sämtliche unüblichen Erhebungsereignisse (Werte kleiner als 1,0 in Bild 7.8 und kleiner als 0,8 in Bild 7.9) recht gut mit den schätzungsgemäßen Geschwindigkeitseinbrüchen übereinstimmen. Die Abweichungen der Messwerte von der Ganglinie am frühen Nachmittag gemäß Bild 7.9 sind für einen Freitag üblich. Sie sind auch in den Schätzergebnisse gut erkennbar.

Auch Bild 7.10 zeigt einen plausiblen Verlauf der Schätzung im Laufe des ausgewählten Tages. Zur morgendlichen Hauptverkehrszeit wird ein deutlicher Geschwindigkeitseinbruch geschätzt, unter Verwendung der Näherungsfunktionen aus der zeitlich differenzierten Regressionsanalyse drastischer als für den zeitlich nicht differenzierten Fall. Weitere betrachtete Diagramme zeigen ähnliche Muster, jedoch liegen oft keinerlei Messwerte vor. So spricht bisher alles für plausible Schätzungen der Linkreisezeiten.

Bild 7.11 f. zeigt Geschwindigkeitseinbrüche insbesondere südlich des Autobahnkreuzes zur morgendlichen Hauptverkehrszeit, am deutlichsten montags, wenn sich Wochenendpendler zu den Tagespendlern addieren. Eine weitere Störung am Samstagabend ist erkennbar. Ihre Ursache lässt sich im Ende eines Fußballspiels in der Allianz-Arena und der daraus resultierenden verstärkten Verkehrsnachfrage vermuten (FC Bayern München – Karlsruher SC am 21.3.2009, Abpiff gegen 17:15 Uhr; die Allianz-Arena liegt in den Darstellungen unmittelbar vor dem Autobahnkreuz). Aufgrund nicht vorhandener sonstiger Daten werden bei der Schätzung dieser Kenngrößen ausschließlich Routenreisezeiten als Eingangsgrößen verwendet. Weiter stromabwärts ist langfristig eine Baustelle mit Fahrstreifenreduktion und Fahrbahnverengung eingerichtet, die als eine Ursache der Geschwindigkeitseinbrüche gelten kann. Die geschätzten Störungsmuster sind in diesem Zusammenhang plausibel. Es zeigt sich in Bild 7.11, dass für manche räumlich-zeitliche Ausschnitte mangels statistisch signifikanter Näherungsfunktionen oder auch mangels Eingangsgrößen keine Schätzwerte generiert werden können. Diese Ausschnitte sind im vorliegenden Fall jedoch meist nachts und am Sonntag und somit in Zeitfenstern unkritischer Verkehrszustände auszumachen.

Gemäß Bild 7.12 werden die Geschwindigkeitsrückgänge unter Verwendung der Näherungsfunktionen aus zeitlich differenzierter und zeitlich nicht differenzierter Regressionsanalyse ähnlich ermittelt, für den zeitliche differenzierten Fall sind sie etwas drastischer.

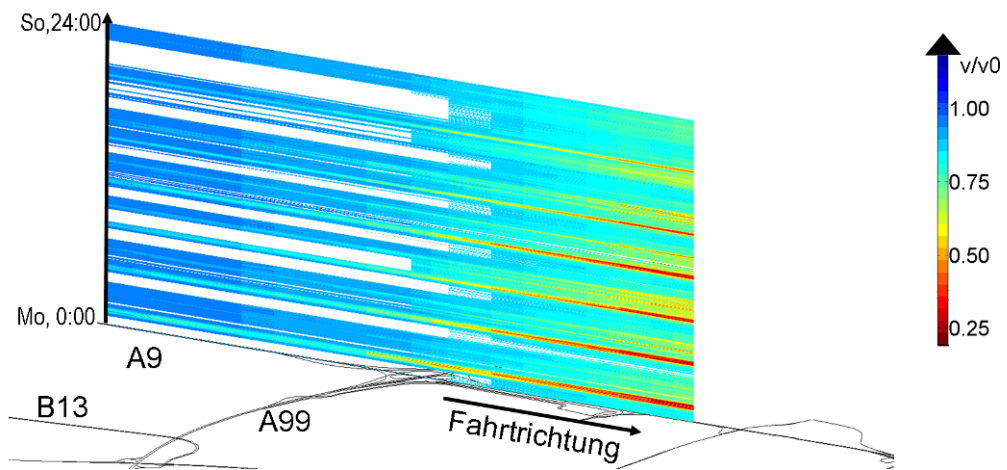


Bild 7.11 Schätzung des linkbezogenen Quotienten aus geschätzter und freier Geschwindigkeit in 5-min-Intervallen von Montag, 16. (unten) bis Sonntag, 22.3.2009 (oben) auf der Route 1-2

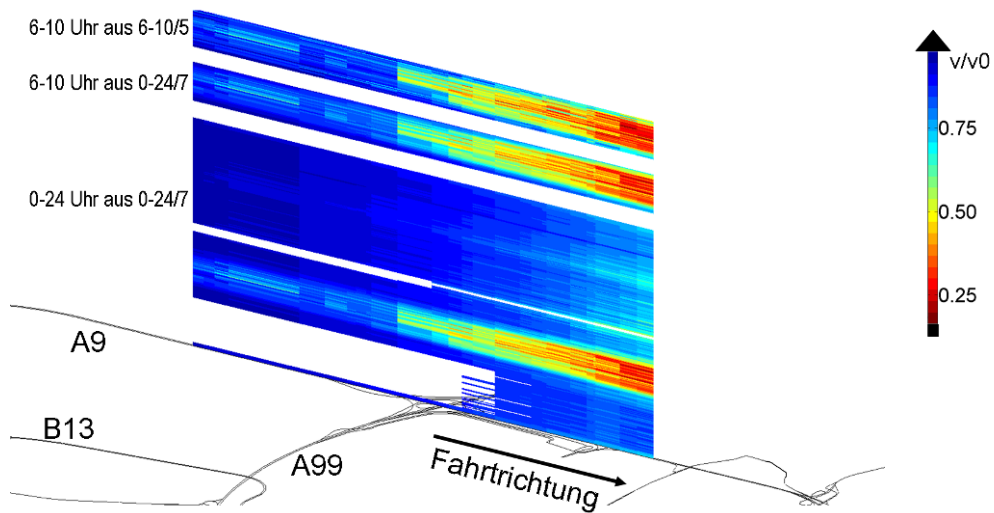


Bild 7.12 Wie Bild 7.11, jedoch nur für Montag, 16.3.2009, zusätzlich mit der Schätzung aus der zeitlich differenzierten Regressionsanalyse für 6 bis 10 Uhr

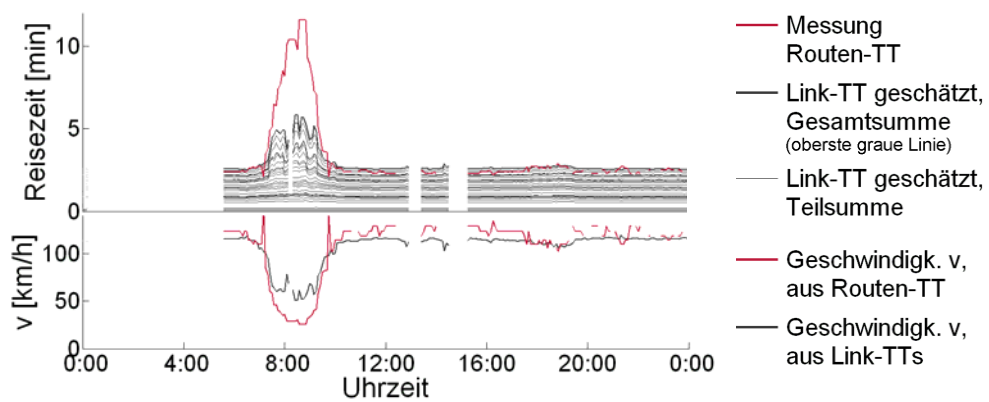


Bild 7.13 Vergleich der gemessenen Routenreisezeit und der Summe der dort geschätzten Linkreisezeiten mit schrittweise gebildeten Teilsommen, Dienstag, 17.3.2009, Route 1-2

Nach Bild 7.13 spiegeln die geschätzten Linkreisezeiten die gemessenen Routenreisezeiten qualitativ gut wider, jedoch unterschätzen sie den Referenzwert für wachsende Zeitverluste. Dies legt in Verbindung mit den Erkenntnissen aus Bild 7.12 den Schluss nahe, dass die zeitlich differenzierte Regressionsanalyse zu besseren Ergebnissen führt. Die Unterschätzung der Linkreisezeiten könnte auch durch einen weiteren Kalibrierungsschritt zumindest reduziert werden.

Die fehlenden Schätzungen nachmittags in Bild 7.13 resultieren aus der Tatsache, dass keine Eingangsgrößen verfügbar sind. Es wird ersichtlich, dass die Schätzung erst einige Zeitschritte nach dem Ausbleiben der gemessenen Routenreisezeit abbricht. Dies ist Folge der zeitlich versetzten Regressionsanalyse, die auch die Schätzung der Linkreisezeiten aus älteren Eingangsgrößen ermöglicht.

Da auf manchen Routen in einigen Zeitschritten keine Fahrzeuge wiedererkannt wurden und folglich auch keine Reisezeiten vorliegen, weist der Verlauf des roten Graphen in Bild 7.13 und Bild 7.15 f. wiederholt Lücken auf.

Für Routenreisezeiten von zwei bis drei Minuten liegt die Summe der geschätzten Linkreisezeiten oft etwas über dem Messwert. Da aber der Messwert auf dem linken Fahrstreifen erhoben wird und sich die Schätzwerte auf alle Fahrstreifen beziehen, muss hier nicht ein Wert zwingend besser oder schlechter sein als der andere.

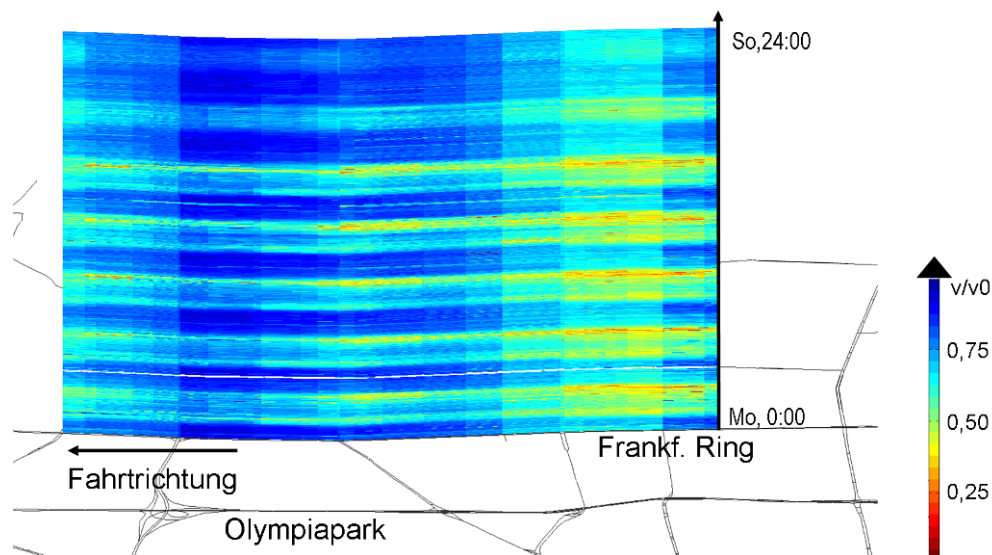


Bild 7.14 Schätzung des linkbezogenen Quotienten aus geschätzter und freier Geschwindigkeit in 5-min-Intervallen von Montag, 16. (unten) bis Sonntag, 22.3.2009 (oben) auf der Route 5-4

Nachdem sich die bisherigen Ausführungen auf einen Autobahnabschnitt bezogen, stellen Bild 7.14 f. Schätzergebnisse für einen innerörtlichen Abschnitt mit zahlreichen LSA-Detektoren dar. Auch die hier geschätzten verkehrlichen Muster sind für eine überwiegend stadtauswärts führende Route plausibel. Offensichtlich stellen sich insbesondere abends

und immer wieder an ähnlichen Stellen des untersuchten Abschnitts größere Reisezeitverluste ein. Bild 7.15 zeigt für den Donnerstag der Bewertungswoche, dass sich die Summe der Linkreisezeiten recht gut der gemessenen Routenreisezeit annähert, jedoch bleibt auch hier eine gewisse Differenz. So kann festgehalten werden, dass die Linkreisezeiten tendenziell defensiv geschätzt werden, dass also mit einer Überschätzung eines solchen Wertes üblicherweise nicht gerechnet werden muss. Stellenweise ist durch breiter werdende Schichten zwischen den Teilsummenschriften erkennbar, dass auf manchen Links größere Reisezeitverluste auftreten als auf anderen.

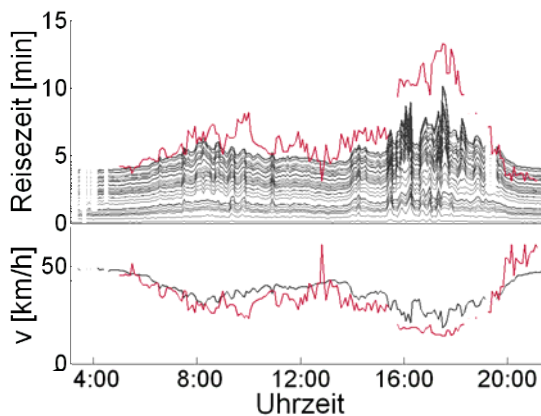


Bild 7.15 Wie Bild 7.13, jedoch für Donnerstag, 19.3.2009, Route 5-4

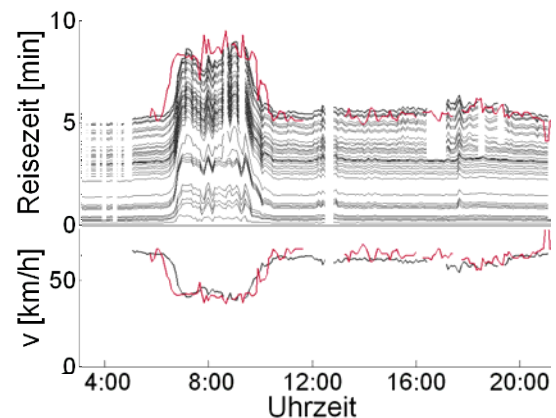


Bild 7.16 Wie Bild 7.13, jedoch für Montag, 16.3.2009, Route 2-3

Bild 7.16 zeigt für eine dritte Route, die teilweise auf der Autobahn und teilweise im städtischen Netz verläuft, eine sehr gute Übereinstimmung der aufsummierten geschätzten Linkreisezeiten mit der gemessenen Routenreisezeit im Tagesverlauf. Neben der Routenreisezeit stehen auf dieser Route auch lokale Daten aus strategischen Detektoren als Eingangsgrößen für die Schätzung zur Verfügung.

So deuten sämtliche angestellte Untersuchungen auf qualitativ plausible und quantitativ vernünftige, tendenziell defensive Schätzungen von Linkreisezeiten hin. Ihre zeitlichen Verläufe beschreiben erwartete Muster. Der Vorteil der entwickelten Methode gegenüber Ganglinienverfahren, die als statistische Alternative zur Verfügung stehen, ist die Möglichkeit, auf aktuelle, vom Üblichen abweichende infrastrukturseitig erhobene Messwerte zu reagieren. Dieser Vorteil wird in den diskutierten Abbildungen ersichtlich. Die zeitliche Differenzierung bei der Regressionsanalyse scheint zu etwas besseren Ergebnissen zu führen. Dabei ist jedoch zu berücksichtigen, dass durch den dann zurückgehenden Stichprobenumfang und der damit einhergehenden geringeren Signifikanz der historisch beobachteten Zusammenhänge größere Lücken in den räumlich-zeitlichen Verkehrslageschätzungen entstehen können.

7.6 Weitergehende Sensitivitätsanalysen

Mit den zusammengeführten Schätzungen aus bis zu fünf Einzelschätzungen ohne Autoregression, die für den Kalibrierungszeitraum errechnet wurden, werden weitere Untersuchungen angestellt, um die Wirkungszusammenhänge im Verfahrensablauf greifbarer zu machen. 26.807.003 Einzelschätzungen wurden berücksichtigt. Dabei wurden 41.338 der 550.763 Schätzfunktionen genutzt, 9.160 davon mindestens 100 Mal.

Weitere 446.243 Näherungsfunktionen mit Linkgeschwindigkeiten als Eingangsgröße und 7.830 Funktionen aus der Autoregression werden nicht betrachtet. So werden weniger als 2% der Näherungsfunktionen mindestens 100 Mal zur Schätzung herangezogen, über 92% aller Näherungsfunktionen kommen nie zum Einsatz.

Tab. 7.13 stellt die Anteile verfügbarer und genutzter Schätzfunktionen differenziert nach Datenarten von Ein- und Ausgangsgrößen und nach zeitlichen Versätzen gegenüber. Diese Anteile summieren sich für jede geschätzte Datenart zu 100%. Linkgeschwindigkeiten aus Positionsmeldungen sind als Eingangsgröße ausgeschlossen. Die Tabelle erlaubt folgende Rückschlüsse:

- Bei der Schätzung lokaler Daten spielen Eingangsgrößen mit einem Alter von mehr als fünf Minuten eine vernachlässigbar kleine Rolle.
- Routenreisezeiten werden zur Schätzung lokaler Kenngrößen kaum eingesetzt.
- Summiert über alle zeitlichen Versätze werden alle Datenarten aus derselben Datenart überproportional häufig geschätzt.
- Strategisch erhobene mittlere lokale Geschwindigkeiten als Eingangsgröße spielen bei der Schätzung von Verkehrsstärken an LSA-Detektoren und von Routenreisezeiten so gut wie keine Rolle. Ihr Anteil an der Schätzung von Linkgeschwindigkeiten liegt ebenfalls weit unter dem Anteil verfügbarer Schätzfunktionen.
- Strategisch erhobene Belegungsgrade werden bei der Schätzung aller Datenarten, ausgenommen Verkehrsstärken an LSA-Detektoren, überproportional häufig als Eingangsgröße eingesetzt. Am deutlichsten kann die Beobachtung für die Schätzung von Routenreisezeiten gemacht werden. Der Unterschied ist hier im Vergleich zur Schätzung aus strategisch erhobenen Verkehrsstärken drastisch.
- Der Anteil genutzter Funktionen zur Schätzung von Routenreisezeiten und Linkgeschwindigkeiten aus Verkehrsstärken an LSA-Detektoren ist zwar geringer als der Anteil verfügbarer Funktionen, jedoch steuern sie einen wesentlichen Anteil zu den zusammengeführten Schätzungen bei.
- Routenreisezeiten werden bei der Schätzung von Linkgeschwindigkeiten weit überproportional als Eingangsgröße eingesetzt.
- Bei der Schätzung von Linkgeschwindigkeiten und mit Abstrichen von Routenreisezeiten werden auch größere Anteile aus zeitversetzten Eingangsdaten geschätzt. Am

deutlichsten ist dies für die Schätzung von Linkgeschwindigkeiten aus Verkehrsstärken an LSA-Detektoren zu erkennen.

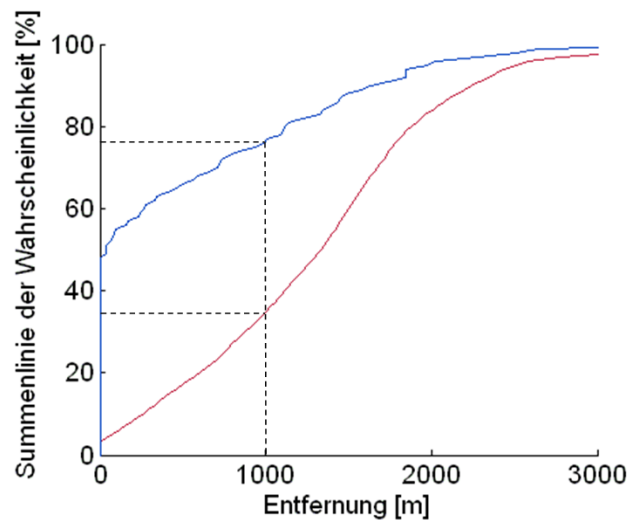


Bild 7.17 Summenlinien der Distanzverteilung zwischen Ein- und Ausgangsgröße für verfügbare (rot) und genutzte (blau) Schätzfunktionen

Bild 7.17 zeigt, dass weniger als 5% der Schätzfunktionen eine Entfernung von 0 zwischen den Erhebungsorten von Ein- und Ausgangsgröße aufweisen. Dies ist der Fall, wenn beide Kenngrößen am selben Link erhoben werden, oder wenn eine Kenngröße eine Routenreisezeit ist und der Erhebungsort der anderen Kenngröße auf dieser Route liegt. Diese Schätzfunktionen liefern jedoch fast 50% der berücksichtigten Einzelschätzungen. Wie erwartet spielen bei den weiteren Einzelschätzungen Näherungsfunktionen mit wachsenden Distanzen also eine abnehmende Rolle. So werden in mehr als 75% der Fälle Schätzfunktionen mit einer Distanz von bis zu 1000m zwischen den Erhebungsorten von Ein- und Ausgangsgröße verwendet, während der Anteil der ermittelten Schätzfunktionen unter 35% liegt (Beispiel gemäß den gestrichelten Linien in Bild 7.17). Für die anderen beiden Distanzmaße freie Reisezeit und Anzahl der Knotenpunkte zwischen den Erhebungsorten ergeben sich sehr ähnliche Verhältnisse.

Resultierend aus den vorangegangenen Beobachtungen werden folgende Fragestellungen beispielhaft beleuchtet:

- Wie verändert sich die Schätzung lokaler Kenngrößen, wenn nur Schätzungen aus Kenngrößen zulässig sind, die an einem anderen Link erhoben werden?
- Wie verändert sich die Schätzung von Linkgeschwindigkeiten, wenn ausschließlich lokale Daten als Eingangsgrößen verwendet werden?
- Wie verändert sich die Schätzung von Linkgeschwindigkeiten, wenn ausschließlich Routenreisezeiten als Eingangsgrößen verwendet werden?

geschätzt erklärend		strat.q		strat.v		strat.occ		LSA q		Routen-TT		Link-v/v0	
		verf.	gen.	verf.	gen.	verf.	gen.	verf.	gen.	verf.	gen.	verf.	gen.
strat.q	0	2.8	18.7	2.7	5.7	2.8	13.9	2.9	4.9	2.4	0.2	2.6	0.9
	5	2.8	4.4	2.7	1.7	2.7	4.4	2.9	1.6	2.5	0.3	2.6	0.6
	10	2.8	1.7	2.7	1.1	2.7	1.8	2.9	0.5	2.4	0.3	2.6	0.7
	15	2.8	1.0	2.6	1.4	2.7	2.4	2.9	0.6	2.5	0.0	2.5	0.8
	20	2.8	0.9	2.7	1.3	2.7	1.6	2.9	0.4	2.5	1.3	2.5	1.3
	25	2.8	0.5	2.6	1.1	2.7	1.5	2.9	0.4	2.4	1.2	2.5	1.0
	30	2.8	0.8	2.6	1.1	2.7	2.8	2.9	0.5	2.5	0.2	2.6	0.5
	Σ	19.5	27.9	18.6	13.5	19.0	28.4	20.4	8.9	17.1	3.4	17.9	5.8
strat.v	0	2.5	5.4	2.6	20.0	2.5	9.5	2.5	0.0	2.0	0.2	1.8	0.7
	5	2.6	1.7	2.5	2.3	2.5	1.9	2.5	0.0	2.0	0.3	1.8	0.4
	10	2.5	0.7	2.5	0.6	2.5	0.9	2.5	0.0	2.0	0.1	1.8	0.4
	15	2.5	0.4	2.5	0.3	2.4	0.3	2.6	0.0	2.0	0.2	1.8	0.3
	20	2.5	0.2	2.5	0.3	2.4	0.2	2.5	0.0	2.0	0.1	1.8	0.3
	25	2.5	0.2	2.5	0.6	2.4	0.3	2.5	0.0	1.9	0.1	1.9	0.3
	30	2.5	0.1	2.4	0.2	2.4	0.1	2.5	0.0	1.9	0.0	1.8	0.6
	Σ	17.7	8.8	17.6	24.3	17.1	13.1	17.7	0.0	13.7	1.0	12.8	2.9
strat.occ	0	1.9	14.3	2.1	28.8	2.0	25.5	1.9	2.9	1.7	8.4	1.9	2.2
	5	1.9	3.2	2.0	6.0	1.9	1.9	1.8	1.1	1.7	7.0	1.9	2.5
	10	1.8	0.9	2.0	1.8	1.9	0.5	1.8	0.1	1.7	4.1	1.8	3.0
	15	1.8	0.3	2.0	1.0	1.9	0.2	1.8	0.0	1.7	6.7	1.9	1.8
	20	1.8	0.2	1.9	2.8	1.9	0.6	1.8	0.1	1.7	0.9	1.8	2.3
	25	1.7	0.2	1.9	2.0	1.8	1.2	1.8	0.2	1.6	1.9	1.9	1.8
	30	1.7	0.2	1.9	1.3	1.7	0.2	1.8	0.3	1.7	12.7	1.8	2.2
	Σ	12.5	19.3	13.9	43.6	13.2	30.1	12.7	4.6	11.7	41.5	13.0	15.7
LSA q	0	6.2	35.0	6.2	10.9	6.2	20.4	6.1	63.1	7.0	8.4	7.1	4.6
	5	6.2	7.5	6.1	0.9	6.2	4.7	6.1	12.0	7.1	0.7	7.1	4.4
	10	6.2	0.6	6.2	0.2	6.2	1.0	6.0	3.4	7.3	4.6	7.0	3.7
	15	6.1	0.4	6.1	0.6	6.2	0.5	6.1	1.8	7.2	7.3	7.0	5.4
	20	6.2	0.2	6.1	0.6	6.3	0.3	6.0	1.2	7.2	2.5	7.0	7.8
	25	6.1	0.1	6.1	1.4	6.2	0.8	6.0	2.1	7.3	2.6	7.0	8.0
	30	6.1	0.1	6.1	1.0	6.3	0.4	6.0	2.8	7.3	2.5	7.0	6.1
	Σ	43.1	43.9	42.9	15.6	43.6	28.1	42.4	86.3	50.4	28.6	49.2	40.0
Routen-TT	0	1.0	0.0	1.0	1.5	1.0	0.3	1.0	0.1	1.0	5.8	1.0	9.5
	5	1.0	0.0	1.0	0.3	1.0	0.0	1.0	0.0	1.0	3.4	1.0	9.9
	10	1.0	0.0	1.1	0.5	1.0	0.0	1.0	0.0	1.0	7.2	1.0	5.3
	15	1.1	0.0	1.0	0.1	1.0	0.0	1.0	0.0	1.0	0.6	1.0	5.0
	20	1.1	0.0	1.0	0.3	1.0	0.0	1.0	0.0	1.0	3.3	1.0	1.4
	25	1.0	0.0	1.0	0.1	1.0	0.0	1.0	0.0	1.0	0.5	1.0	1.8
	30	0.9	0.0	1.0	0.3	1.0	0.0	0.9	0.0	1.0	4.7	1.0	2.7
	Σ	7.1	0.0	7.1	3.1	7.0	0.3	6.8	0.1	7.1	25.5	7.1	35.6
Absolut (Mio)		0.02	4.64	0.02	3.64	0.02	3.64	0.13	2.9	0.03	0.37	0.32	1.31

Tab. 7.13 Anteile [%] von Datenarten als Eingangsgrößen an verfügbaren (verf.) und genutzten (gen.) Schätzfunktion bei der Schätzung anderer Datenarten, differenziert nach zeitlichen Versätzen

7.6.1 Schätzung lokaler Kenngrößen ohne Eingangsgrößen vom selben Link

Bild 7.18 und Tab. 7.14 sind mit Bild 7.4 und Tab. 7.5 zu vergleichen und erlauben eine Einschätzung des Einflusses von Eingangsdaten vom selben Link auf die Schätzergebnisse von lokalen mittleren Geschwindigkeiten an strategischen Detektoren. Der Vergleich der Abbildungen zeigt, dass sich für übliche Verkehrszustände mit recht hohen Geschwindigkeiten kaum erkennbare Unterschiede ergeben. Bei den seltenen Ereignissen zeigen sich jedoch deutliche Nachteile für den Fall ohne Eingangsgrößen vom selben Link. Ein Grund hierfür mag sein, dass die Kalibrierung nicht gesondert für den Fall mit Eingangsdatenbeschränkung vorgenommen wurde, was zu einem stärkeren Absinken des Schätzwerts führen könnte. Aus diesem Grund wird in Tab. 7.14 auf die Darstellung der Ergebnisse aus der kalibrierten Schätzung verzichtet. Die Kalibrierung wirkt sich aber nicht auf die Erkennungsraten der Perzentilereignisse aus. Ihr deutlicher Rückgang im Vergleich zu Tab. 7.5 zeigt, dass Eingangsdaten vom selben Link auf die Qualität der Schätzung lokaler mittlerer Geschwindigkeiten maßgeblichen positiven Einfluss haben. Dennoch liegen die Raten auch ohne Eingangsdaten vom selben Link deutlich über denen für die Ganglinienverfahren und ungefähr auf dem Niveau der Trivialprognose.

Durch die vorgenommene Eingangsdatenbeschränkung entfällt die Autoregression.

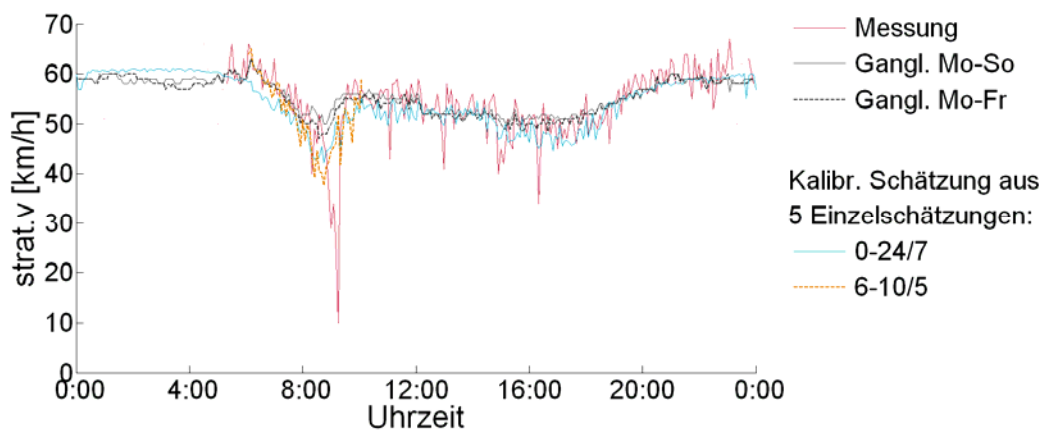


Bild 7.18 Schätzung lokaler Geschwindigkeiten ohne Eingangsgrößen vom selben Link (vgl. Bild 7.4)

Analysierte Datenhistorie	Kalibriert?	n = 3 oder 5?	Verfügb.[%]	RMSE	Erkennungsraten Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
					1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
0-24/7	nein	5	100	4.35	27	34	42	59	69	33	66	59	71	81
		3	100	4.31	27	38	45	60	71	33	62	62	73	83

Tab. 7.14 Schätzung lokaler Geschwindigkeiten ohne Eingangsgrößen vom selben Link im Zeitraum 0-24/7 (vgl. Tab. 7.5)

Für die Schätzung von Verkehrsstärken am ausgewählten LSA-Detektor ergeben sich durch den Ausschluss von Eingangsdaten vom selben Link sogar Verbesserungen (Tab. 7.15). Die Bewertungskenngrößen erreichen das Niveau wie bei der Schätzung mit Autoregression. Offensichtlich wirkte sich die Berücksichtigung anderer Sensoren positiv aus. Für den Untersuchungszeitraum konnten dafür keine tieferen Ursachen ermittelt werden.

Analysierte Datenhistorie	Kalibriert?	n= 3 oder 5?	Verfüg. [%]	RMSE	Erkennungrate Perzentilereignis [%]					Erk.-rate Perzentilereignis (weich) [%]				
					1%	2%	5%	10%	20%	1(3)%	2(5)%	5(10)%	10(15)%	20(30)%
0-24/7	nein	5	100	73.52	42	49	65	69	76	79	84	89	83	90

Tab. 7.15 Schätzung der Verkehrsstärke an einem LSA-Detektor ohne Eingangsgrößen vom selben Link (vgl. Tab. 7.9)

7.6.2 Schätzung von Linkgeschwindigkeiten ausschließlich aus lokalen Daten oder ausschließlich aus Routenreisezeiten

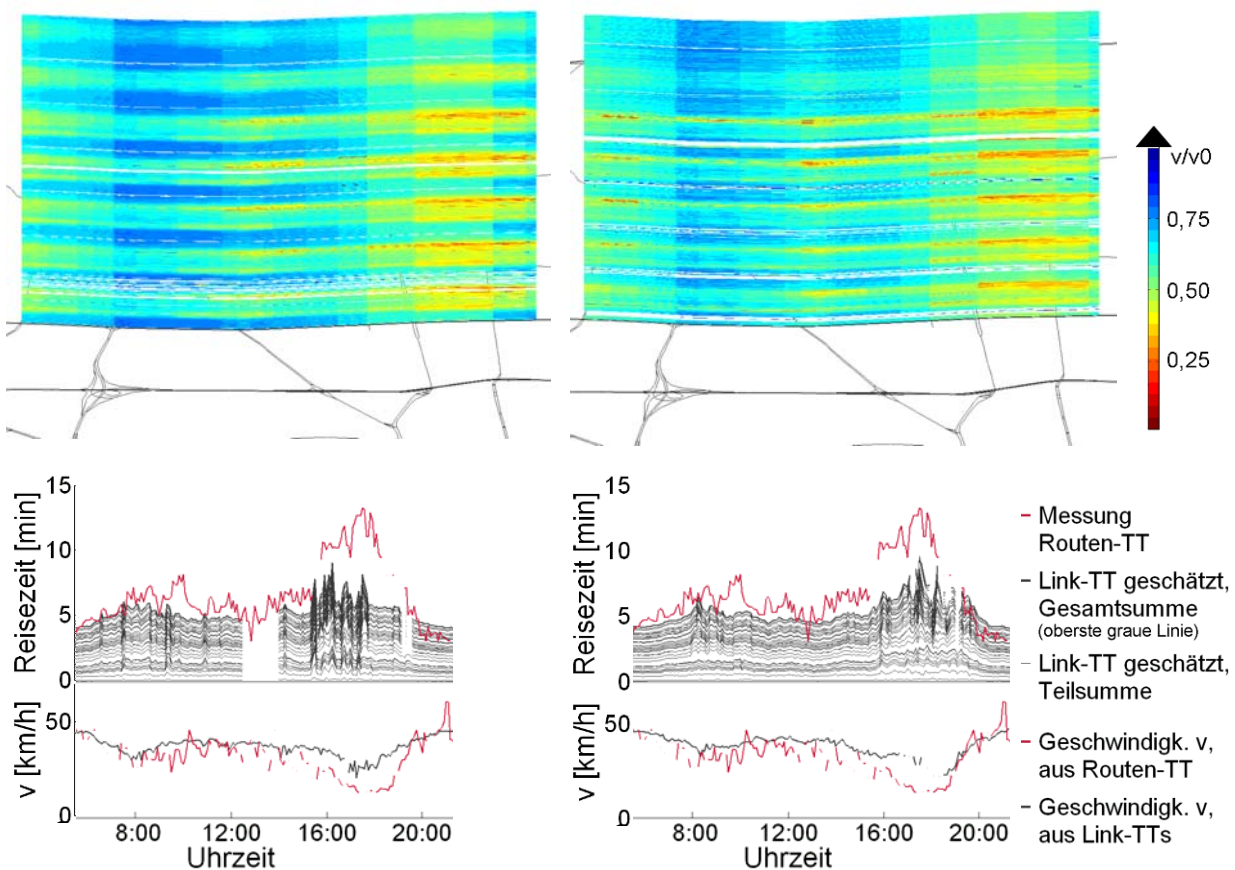


Bild 7.19 Schätzung von Linkgeschwindigkeiten ausschließlich aus lokalen Daten (links) und ausschließlich aus Routenreisezeiten (rechts; vgl. Bild 7.14 f.)

Am Beispiel der Route 5-4 wird untersucht, inwiefern sich lokale und abschnittsbezogene Eingangsdaten auf die Schätzung von Linkgeschwindigkeiten auf einer reisezeitdetektierten

Route auswirken (Bild 7.19). Die Konturplots zeigen, dass in allen Fällen Reisezeitverluste in plausibler Art und Weise räumlich-zeitlich aufgeteilt werden: Die geschätzten linkbezogenen Geschwindigkeiten sinken insbesondere vor einschlägigen Knotenpunkten in der abendlichen Hauptverkehrszeit ab. Tendenziell lässt sich feststellen, dass Geschwindigkeitsrückgänge aus lokalen Daten räumlich abgegrenzter geschätzt werden als aus Routenreisezeiten. Der Vergleich der Ganglinien in Bild 7.19 unten zeigt, dass steigende Routenreisezeiten für beide untersuchte Fälle sich auch in erhöhten geschätzten Linkreisezeiten niederschlagen. Dabei ist festzustellen, dass geschätzt aus lokalen Daten die Übereinstimmung mit den Messwerten sowohl qualitativ als auch quantitativ schlechter ausfällt als geschätzt aus Routenreisezeiten. Der defensive Charakter der Linkreisezeitenschätzung bleibt für die untersuchten Fälle durchgängig erhalten. Das Ausbleiben der Schätzwerte aus lokalen Daten am Nachmittag ist darauf zurückzuführen, dass für dieses Zeitfenster keine lokalen Daten vorliegen.

Die Schätzung aus sämtlichen verfügbaren Daten liefert augenscheinlich jedoch bessere Ergebnisse (Bild 7.15). Dies kann als Anhaltspunkt dafür gelten, dass die Nutzung sämtlicher verfügbarer Daten zu verbesserten Schätzergebnissen führt und dass das Bestimmtheitsmaß ein geeignetes Werkzeug für die Auswahl von Einzelschätzungen darstellt.

8. Resümee

Überblick über das neue Schätzverfahren

Eine neue statistische Methode zur Schätzung verkehrlicher Kenngrößen wurde unter Berücksichtigung der Anforderungen in Kapitel 3.3 entwickelt und erprobt (Überblick in Bild 8.1). Schätzfunktionen werden durch eine abschnittsweise Regressionsanalyse unter Verwendung historischer Daten ermittelt. Die Abschnitte werden überlappend gewählt, um Unstetigkeiten an den Übergängen zu minimieren. Unter Verwendung statistischer Kenngrößen und unter Berücksichtigung verfügbarer Messwerte als Eingangsdaten werden einzelne Näherungsfunktionen zur Schätzung einer Kenngröße ausgewählt und darauffolgend ihre Schätzergebnisse zusammengeführt. Systematische Schätzfehler werden durch eine Kalibrierung aufgefangen. Die kalibrierten zusammengeführten Schätzergebnisse können zur Vervollständigung fehlender und zur Plausibilisierung gemessener Werte herangezogen werden. Erprobung und Bewertung des Verfahrens erfolgten mit lokalen Daten aus strategischen und LSA-Detektoren, mit infrastrukturseitig gemessenen Routenreisezeiten und mit Linkgeschwindigkeiten, die aus Positionsmeldungen abgeleitet werden.

Nachfolgende zusammenfassende Aussagen zeigen, dass die Anforderungen an die Methode gemäß Kapitel 3.3 erfüllt werden:

Sämtliche verkehrlich relevanten Daten können mit der Methode verarbeitet werden. Damit ist sie auch auf künftige, heute noch nicht genutzte Messgrößen anwendbar. Das Verfahren stellt keinerlei Anforderungen an die Verfügbarkeit einzelner Eingangsdaten. Die Berücksichtigung von Güteindikatoren wird durch eine gewichtete Regressionsanalyse ermöglicht. Im Fall des Einfließens von fahrzeuggenerierten Daten wie beispielsweise von Positionsmeldungen kann netzweit und linkfein die aktuelle Verkehrslage geschätzt werden. Der regressionsanalytische Ansatz stellt dabei die umfassende Nutzung von Erfahrungswerten sicher, die aus einer Datenhistorie gewonnen werden können. Durch die vorherige Berechnung sämtlicher Näherungsfunktionen ermöglicht die vorgestellte Methode eine sehr schnelle Ermittlung der Schätzgrößen. Somit begünstigt das Verfahren auch die Verkehrslageschätzung in großen Netzen. Sollte dennoch die Größe eines Netzes zu Problemen hinsichtlich des Zeitbedarfs der Verkehrslageschätzung führen, steht der Bildung von Netzausschnitten zur Datenverarbeitung mit mehreren Rechnern verfahrensseitig nichts im Wege.

Überblick: Abschnittsweise Regressionsanalyse zur Schätzung von Verkehrskenngrößen

Offline-Telle: blau
Online-Telle: orange

- 1) Es werden nur Zusammenhänge zwischen Kenngrößen untersucht, deren Erhebungsorte ein Kriterium der räumlichen Nähe erfüllen.
- 2) Es werden nur Näherungsfunktionen verzeichnet, die den t-Test mit einem Signifikanzniveau von 90% bestehen.
- 3) Zusammenführung durch gewichtete Mittelung der Einzelschätzungen aus den Näherungsfunktionen mit dem höchsten Bestimmtheitsmaß.

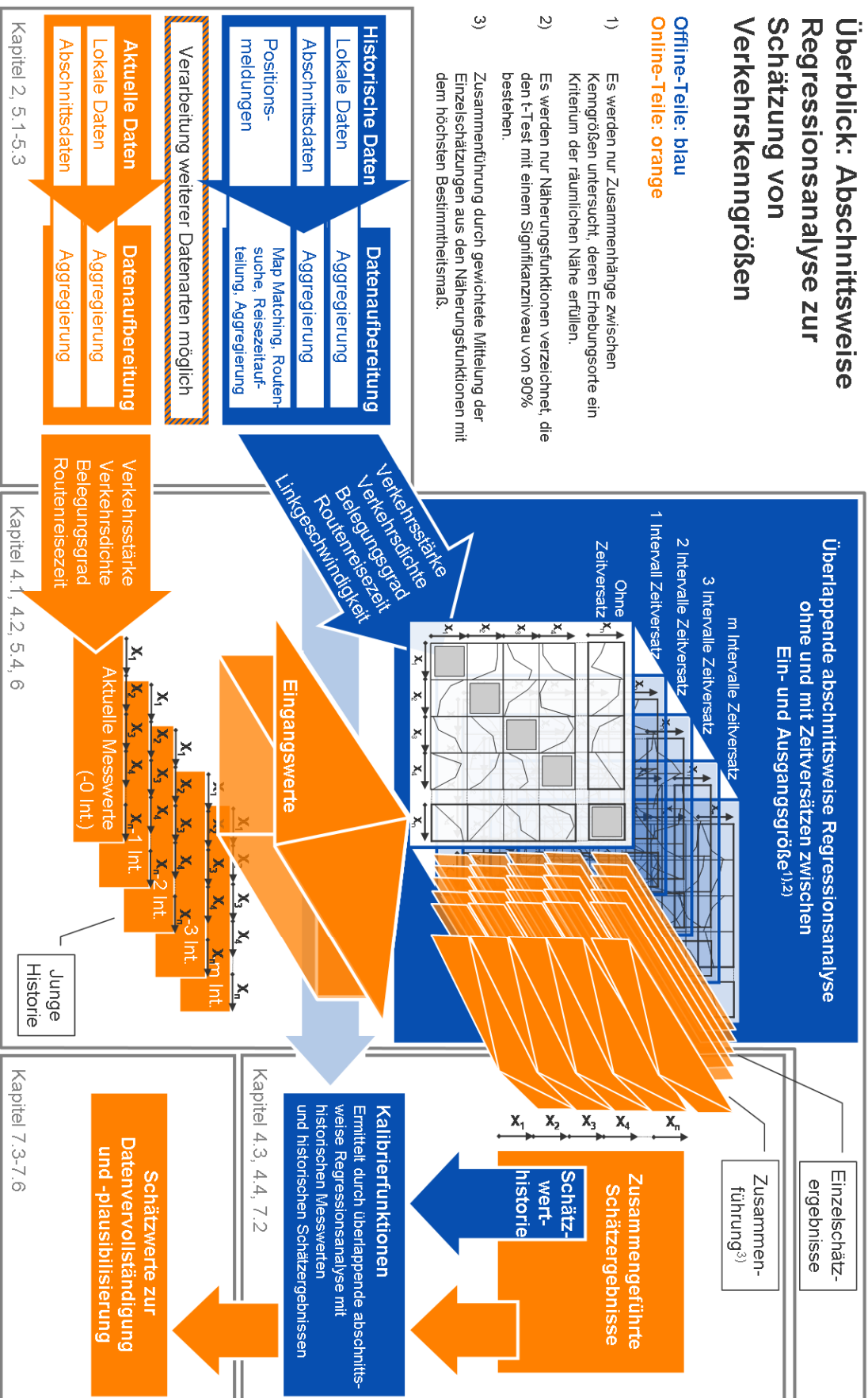


Bild 8.1 Überblick über das Verfahren zur Schätzung von Verkehrskenngrößen durch abschnittsweise Regressionsanalyse

Justierung und Bewertung des Verfahrens

Die Bildung von drei Abschnitten bei der Suche nach Näherungsfunktionen erweist sich als geeignet. Sie reicht aus, um die meisten Zusammenhänge plausibel nachzubilden. Bei der Bildung einer größeren Abschnittszahl wächst die Gefahr, dass aufgrund des zurückgehenden abschnittsbezogenen Stichprobenumfangs weniger statistisch signifikante Näherungsfunktionen ermittelt werden können. Lediglich zur Schätzung von Verkehrsstärken wären wegen des annähernd parabolischen Zusammenhangs mit Geschwindigkeiten und Verkehrsdichten oder Belegungsgraden mehr Abschnitte wünschenswert. Da aber Verkehrsstärken hoch verfügbar sind, sind sie durch das beschriebene Verfahren eher zu plausibilisieren als zu vervollständigen. Für diesen Zweck sind auch die Ergebnisse bei der Bildung von drei Abschnitten dienlich.

Die beobachteten Zusammenhänge und die abgeleiteten Näherungsfunktionen bestätigen im Grundsatz Erwartungen, die auf das Fundamentaldiagramm zurückgehen. Vorüberlegungen, die zur Regressionsanalyse mit zeitlich versetzten Ein- und Ausgangsgrößen angestellt wurden, bestätigen sich ebenfalls: Mit wachsendem Zeitversatz werden die Näherungsfunktionen defensiver, schätzen also unkritischere Verkehrszustände, und die Qualität der Zusammenhänge nimmt ab. Dies schlägt sich in einem Rückgang des Bestimmtheitsmaßes nieder. Aus manchen Messwerten kann insbesondere im Störfall bereits wenige Minuten nach der Erhebung nicht mehr statistisch fundiert auf die aktuelle Situation geschlossen werden.

Die gewichtete lineare Regression als Alternative zur einfachen linearen Regression wurde zur Berücksichtigung von Messwerten unterschiedlicher Güte erprobt. Sie führte zu keinen nennenswerten Veränderungen bei den ermittelten Näherungsfunktionen. Vieles spricht dafür, dass sich Messwerte geringerer Güte in diffuseren Punktwolken, schwächeren Zusammenhängen und damit auch im Fall der einfachen Regression in geringeren Bestimmtheitsmaßen niederschlagen. Damit disqualifizieren sich Näherungsfunktionen aus solchen Messwerten häufig bei der Auswahl verwendeter Zusammenhänge. Die ursprünglich für die gewichtete lineare Regression angedachte Aufgabe wird also zumindest zu großen Teilen durch die bereits vorliegenden statistischen Auswertungen abgedeckt. Somit bietet die gewichtete lineare Regression im vorgestellten Verfahren keine nennenswerten Vorteile gegenüber der linearen Einfachregression.

Zur Bewertung zusammengeführter Schätzergebnisse, also der letztendlichen Schätzung von Verkehrskenngrößen, für infrastrukturseitig erhobene Kenngrößen werden neben dem Fehlermaß „Wurzel aus dem mittleren quadrierten Fehler“ die Erkennungsraten von Perzentilereignissen herangezogen. Als Referenz werden Messwerte eingesetzt, die während einer Woche aufgezeichnet wurden. Die Bewertungsgrößen werden für verschiedene Varianten der neuen Methode sowie für Ganglinienverfahren und die Trivialprognose „Alles bleibt, wie es gerade noch war“ ermittelt. Ihre Gegenüberstellung zeigt, dass das regressionsbasierte

Verfahren fast durchgängig höhere Erkennungsraten und meist geringere Fehlermaße aufweist als das jeweilige Vergleichsverfahren. Die Kalibrierung der Schätzergebnisse reduziert das Fehlermaß weiter. Die Zusammenführung von bis zu fünf Einzelschätzungen erweist sich als geeignet. Das Zulassen der Autoregression, also der Schätzung aus dem letzten Messwert für die zu schätzende Größe, wirkt sich positiv auf die Qualität der Schätzungen aus. Da infrastrukturseitig erhobene Daten aber vergleichsweise hoch verfügbar sind, sind sie durch das vorgestellte Verfahren eher zu plausibilisieren als zu vervollständigen. Hierbei ist die Verwendung des zuvor gemessenen Werts als Eingangsgröße jedoch zu hinterfragen.

Aufgrund fehlender geeigneter Referenzwerte wird die Bewertung geschätzter Linkgeschwindigkeiten qualitativ und häufig aggregiert über Routen vorgenommen, auf denen infrastrukturseitig Reisezeiten gemessen werden. Plausible zeitliche Verläufe geschätzter Linkreisezeiten können beobachtet werden, auch nach Abgleich mit vereinzelt verfügbaren Werten aus Positionsdaten. Steigende Reisezeiten auf infrastrukturseitig untersuchten Routen schlagen sich auch in höheren geschätzten Reisezeiten für manche Links auf diesen Routen nieder. Dabei ergeben sich plausible räumlich-zeitliche Störungsmuster. Die Summe aller geschätzten Linkreisezeiten über eine reisezeitdetektierte Route legt den Schluss nahe, dass die Schätzung defensive Werte liefert, also die tatsächlichen Linkreisezeiten eher unterschätzt.

Durch die Schätzung von Linkgeschwindigkeiten gelingt die netzweite linkfeine Verkehrslageschätzung weitgehend. Nur selten werden für einen Link keine Schätzwerte ausgegeben. Dies ist der Fall, wenn entweder keine geeigneten Eingangsgrößen oder keine statistisch signifikanten Schätzfunktionen zu verfügbaren Eingangsgrößen vorliegen. Da diese Fälle ausschließlich nachts beobachtet wurden, betreffen diese Lücken in der räumlich-zeitlichen Verkehrslage Bereiche, in denen üblicherweise keine Störungen des Verkehrsablaufs auftreten. So bietet das regressionsbasierte Verfahren eine Lösung für die stark begrenzte räumlich-zeitliche Verfügbarkeit von linkbezogenen Geschwindigkeiten aus Positionsmeldungen.

Die zeitlich differenzierte Regressionsanalyse wurde für das Zeitfenster Montag bis Freitag, 6 bis 10 Uhr erprobt. Es ergeben sich hierbei allenfalls leicht erhöhte Detektionsraten der Perzentilereignisse. Während die Zusammenhänge durch die zeitliche Differenzierung geschärft werden, geht gleichzeitig der Stichprobenumfang zurück, so dass Argumente für eine wachsende Signifikanz solchen für eine geringere entgegenstehen. Liegen historische Daten in einem Umfang vor, dass die Signifikanz beobachteter Zusammenhänge aufgrund des Stichprobenumfangs nicht mehr in Frage steht, so kann mit besseren Schätzergebnissen bei zeitlich differenzierter Regressionsanalyse gerechnet werden. Mit den vorliegenden Daten war dies nicht durchgängig nachzuweisen.

Aufgrund der geringen Verfügbarkeit und der fragwürdigen Repräsentativität einzelner aus Positionsmeldungen ermittelter Linkreisezeiten wurden Linkinformationen als Eingangsgrößen für das entwickelte Schätzverfahren ausgeschlossen.

Unter den lokalen Daten erweist sich der Belegungsgrad als Datenart, die als Eingangsgröße zur Schätzung aller anderen Datenarten überproportional häufig eingesetzt wird. Offensichtlich liefert er stärkere Zusammenhänge als die Verkehrsstärke, die kein eindeutiger Störungsindikator ist, und als die lokale mittlere Geschwindigkeit, die häufig erst beim Eintreten einer Störung am Querschnitt absinkt und zuvor auf konstant hohem Niveau verharrt.

Aufgrund der zahlreichen LSA-Detektoren im Bewertungsnetz werden auch hier erhobene Verkehrsstärken häufig als Eingangsgröße ausgewählt. Sprechen die statistischen Kenngrößen der Schätzfunktionen dafür, sollten auch diese Daten zum Verfahren zugelassen werden. Sollte aus diesen Detektoren auch der Belegungsgrad verfügbar sein, könnte sein Beitrag zum Schätzergebnis mutmaßlich groß sein. Im vorliegenden Netz war er nicht verfügbar.

Infrastrukturseitig erhobene Reisezeiten stellen für die Schätzung linkbezogener Geschwindigkeiten eine wichtige Eingangsgröße dar. Dies wird unter anderem daran erkennbar, dass sie weit überproportional zum erhobenen Datenumfang und zur Anzahl ermittelter Näherungsfunktionen als Eingangsgröße ausgewählt werden. Aufgrund ihrer räumlichen Aussagekraft tragen sie also einen großen Anteil zur netzweiten linkfeinen Verkehrslageschätzung mit dem beschriebenen Verfahren bei.

Ansatzpunkte für weitere Arbeiten

Das neue Verfahren zur Schätzung von Verkehrskenngrößen kann Grundlage für weitere Untersuchungen sein und für bestehende Verfahren einen Mehrwert leisten. So könnten in einem ersten Schritt zur Verbesserung der Verkehrslageschätzung die geschätzten Linkreisezeiten so kalibriert werden, dass ihre Summe über eine reisezeitdetektierte Route mit der gemessenen Reisezeit übereinstimmt. Es ist zu vermuten, dass eine exponentielle Anpassung, die also höhere Zuschläge für Links mit höheren geschätzten Reisezeitverlusten vorsieht, ein besseres Ergebnis liefert als das gleichmäßige Verteilen des Reisezeitunterschieds über die gesamte Route. Neben diesen beiden wären auch weitere diesbezügliche Ansätze zu sichten und zu bewerten.

Darüber hinaus könnte die neue Methode mit weiteren Datenarten wie beispielsweise mit Reisezeiten aus der Fahrzeugwiedererkennung durch die Analyse von Signaturen über Induktivschleifen oder mit Umfelddaten erprobt werden. Es kann erwartet werden, dass sich für Schneefall oder Straßenglätte gute Zusammenhänge zwischen Umfelddaten und Geschwindigkeiten einstellen würden. Ob diese Umfelddaten durch infrastruktur- oder fahrzeugseitige Sensorik erfasst werden, spielt zunächst keine Rolle. So ist auch die Erprobung mit weiteren Informationen denkbar, die fahrzeugseitig erhoben werden. Vielversprechend erscheinen hier beispielsweise Fahrzeugabstandsdaten, da sie die Ableitung von Aussagen bezüglich der Verkehrsdichte erlauben. Diese Kenngrößenart stellt kontinuierlich und netzweit geschätzt eine aussagekräftige Entscheidungsgrundlage für Maßnahmen des Ver-

kehrmanagements dar. Außerdem erscheint die Integration von Reisezeitinformationen straßengebundener öffentlicher Verkehrsmittel besonders aussichtsreich, da solche Daten einerseits als Eingangsgrößen in regelmäßigen Zeitschritten vorliegen, andererseits als Schätzgrößen die Fahrzeugankunftsprognose verbessern und so im Rahmen künftiger dynamischer ÖV-Routenempfehlungen Verwendung finden können.

Ein Netz, das an Knotenpunkten für verschiedene Abbiegebeziehungen verschiedene Links vorsieht, ermöglicht insbesondere bei der Offline-Auswertung von Positionsmeldungen die Ableitung höherwertiger Informationen. Diese wiederum führen zu einer feiner aufgelösten Verkehrslageschätzung, da dann für jede einzelne Abbiegebeziehung Schätzungen der linkbezogenen Geschwindigkeiten angestellt werden können.

Eine umfangreichere Datenhistorie wäre insbesondere zur Beobachtung und Schätzung sehr seltener Ereignisse hilfreich. Sie würde zudem verbesserte Untersuchungen zur Anwendung der neuen Methode differenziert nach Tagestypen und Zeitbereichen ermöglichen. Eine Online-Anwendung würde zu einer ständig wachsenden Historie führen, so dass sich hier mit dem längeren Fortbestand erweiterte Möglichkeiten böten.

Hochverfügbare linkbezogene Geschwindigkeiten aus Positionsmeldungen für einen kleinen Netzausschnitt oder nur für einen Streckenzug eventuell aus gezielt durchgeführten Befahrungen könnten dafür herangezogen werden, die Fähigkeit der Schätzmethode zur Verfolgung der Ausbreitung von Verkehrszuständen zu überprüfen. Nachfolgend wäre ein diesbezüglicher Vergleich mit modellbasierten Ansätzen von Interesse, die Verkehrsströme propagieren.

Die Übertragung von Erfahrungen, die in einem Netzausschnitt gesammelt wurden, auf einen anderen Ausschnitt im selben oder auch in einem anderen Netz könnte einen wichtigen Beitrag zur Vergrößerung des Erfahrungsschatzes und damit zu einer Quasi-Erhöhung des Stichprobenumfangs leisten. Hierzu bedürfte es eines Verfahrens, das die Bewertung der Ähnlichkeit zwischen Netzausschnitten unter Beachtung geometrischer und verkehrlicher Bedingungen ermöglicht.

Werden die Ergebnisse der netzweiten linkfeinen Verkehrslageschätzung, die durch die neue Methode gewonnen werden, archiviert, so kann dieses Archiv als Historie für das Verfahren zur musterbasierten Störungsausbreitungsprognose gemäß Kapitel 3.2.2 eingesetzt werden. Auch für das Prognoseverfahren wäre die zuvor erwähnte Übertragung von Erfahrungen von Netzausschnitt zu Netzausschnitt eine Bereicherung.

Die Schätzungen des vorgestellten Verfahrens können mit Schätzungen anderer Methoden gekoppelt werden. Dies kann zu verbesserten Ergebnissen und zu einer höheren räumlich-zeitlichen Verfügbarkeit von Schätzwerten führen. Als weitere statistische Methoden können beispielsweise Ganglinienverfahren, deren Ergebnisse üblicherweise eine hohe Verfügbar-

keit aufweisen, oder instanzbasierte Ansätze, die mehrdimensionale Muster berücksichtigen, einen Mehrwert bieten. Darüber hinaus können statistisch ermittelte Schätzgrößen mit Ansätzen zusammengeführt werden, die auf ein Verkehrsflussmodell zurückgreifen. Werden solche Modelle, ob makroskopisch oder mikroskopisch, mit Schätzungen aus der neuen Methode angereichert und kalibriert, so ist eine verbesserte Wirkungsabschätzung verschiedener Verkehrssteuerungsszenarien möglich. Die Schätzung von Verkehrskenngrößen aus aktuellen Daten und aus der abschnittswisen Regressionsanalyse kann auf diesem Wege zur verbesserten Auswahl von Maßnahmen des Verkehrsmanagements beitragen.

Literaturverzeichnis

- AMONES [2010]: Anwendung und Analyse modellbasierter Netzsteuerungsverfahren in städtischen Straßennetzen. Interner Bericht zum BMVBS-Projekt im Rahmen der Förderinitiative Mobilität 21. Universität Stuttgart, Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik; Technische Universität Darmstadt, Fachgebiet Verkehrsplanung und Verkehrstechnik; Technische Universität München, Lehrstuhl für Verkehrstechnik; Technische Universität Braunschweig, Institut für Verkehr und Stadtbauwesen
- BÁRDOSSY, A. [1996]: Statistik für Bauingenieure. Vorlesungsumdruck, Institut für Wasserbau, Lehrstuhl für Wasserbau und Wasserwirtschaft, Universität Stuttgart
- BELL, M.; SHIELD, C.; BUSCH, F.; KRUSE, G. [1997]: Stochastic User Equilibrium Path Flow Estimator. Transportation Research, Part C, Vol. 5(1997), Nr. 3/4
- BERMANN, E.; BUSCH, F. [2006]: Online Traffic State Estimation for the City of Graz. Proceeding of the ITS World Congress 2006, October 8-12, 2006, London, Great Britain
- BERNARD, M.; HACKNEY, J.; AXHAUSEN, K. [2006]: Correlation of link travel speeds. 6th Swiss Transport Research Conference, Monte Verica/Ascona, 15.-17. März 2006
- BMVBW (Bundesministerium für Verkehr, Bau und Wohnungswesen) [2002]: Technische Lieferbedingungen für Streckenstationen, Ausgabe 2002. Bundesanstalt für Straßenwesen, Bergisch Gladbach. Verlag für neue Wissenschaft GmbH, Bremerhaven
- BÖHNKE, P. [1980]: Ein Verfahren zur aktuellen Analyse und Prognose der Verkehrsabläufe innerhalb von Straßenabschnitten aus lokal gemessenen Fahrzeugfolgen. Dissertation, Aachen
- BÖKER, G. [2000]: Zustandsschätzung im Schnellstraßenverkehr auf der Basis von Einzelfahrzeugdaten. at – Automatisierungstechnik 8(2000), Oldenbourg Verlag
- BORGELT, C.; KLAWONN, F.; KRUSE, R.; NAUCK, D. [2003]: Neuro-Fuzzy-Systeme. 3. Auflage. Vieweg Verlag, Wiesbaden
- BRAKATSOULAS, S.; PFOSER, D.; SALAS, R.; WENK, C. [2005]: On Map Matching Vehicle Tracking Data. Proceedings of the 31st VLDB Conference, Trondheim, Norwegen
- BRILON W., SCHNABEL W. [2003]: Bewertung der Qualität des Verkehrsablaufs auf Hauptverkehrsstraßen. Straßenverkehrstechnik 1.2003, Kirschbaumverlag, Bonn
- BÜHLER, M. [2006]: GPS / Galileo. Fachseminar in Verteilte Systeme, Institute for Pervasive Computing, Distributed Systems Group, ETH Zürich
- BUSCH, F.; DINKEL, A.; LEONHARDT, A.; ZIEGLER, J. [2006]: Benchmarking für Verkehrsdatenerfassungs- und Steuerungssysteme. Herausgegeben vom Bundesministerium für Verkehr, Bau und Stadtentwicklung, Heft 949, Bonn, ISBN 3-86509-575-5
- BUSCH, F.; GHIO, A. [1994]: Automatic Incident Detection on Motorways by Fuzzy Logic. Proc. Traffic and Transport Solutions, Amsterdam, 1994

- BUSCH, F.; GHIO, A.; CREMER, M.; HENNINGER, T. [1994]: Dynamische Schätzung des Verkehrszustands auf Autobahnen – ein modellbasierter Filteransatz. Sonderveröffentlichung der Siemens AG, München
- COIFMAN, B. [1998a]: Vehicle Reidentification and Travel Time Measurement in Real-Time on Freeways Using the Existing Loop Detector Infrastructure. Transportation Research Record No. 1643, Transportation Research Board, Washington, D.C., S. 181-191
- COIFMAN, B. [1998b]: A New Algorithm for Vehicle Reidentification and Travel Time Measurement on Freeways. Applications of Advanced Technologies in Transportation, ASCE, Newport Beach, 1998, S. 167-174
- COIFMAN, B.; CASSIDY, M. [2001]: Vehicle Reidentification and Travel Time Measurement on Congested Freeways. Transportation Research. Part A: Policy and Practice Vol. 36 No. 10, S. 899-917
- CREMER, M. [1979]: Der Verkehrsfluß auf Schnellstraßen. Springer Verlag Berlin, Heidelberg, New York
- DAGANZO, C. [1993]: The Cell Transmission Model, Part I: A Simple Dynamic Representation of Highway Traffic.
- DAGANZO, C. [1994]: The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory. Transportation Research Part B, Vol. 28, Nr. 4
- DAGANZO, C. [1995]: The Cell Transmission Model, Part II: Network Traffic. Transportation Research Part B, Vol. 29, Nr. 2
- EHMKE, J.; MEISEL, S. [2008]: Charakterisierung des städtischen Straßenverkehrs mit Floating Car Data und Data Mining. Straßenverkehrstechnik 10/2008, Kirschbaumverlag, Bonn
- EID, J. [2001]: Entwicklung und Analyse eines Verkehrsinformationsalgorithmus zur grafischen Anzeige des Verkehrszustands im Rahmen der MOBINET-Maßnahme Ring-Info. Diplomarbeit am Fachgebiet für Verkehrstechnik und Verkehrsplanung, Technische Universität München
- FGSV (Forschungsgesellschaft für Straßen- und Verkehrswesen) [2000]: Begriffsbestimmungen – Teil Verkehrsplanung, Straßenentwurf und Straßenbetrieb. Ausgabe 2000. FGSV Verlag GmbH, Köln
- FGSV (Forschungsgesellschaft für Straßen- und Verkehrswesen) [2003]: Hinweise zur Datenvervollständigung und Datenaufbereitung in verkehrstechnischen Anwendungen. FGSV Verlag GmbH, Köln
- FGSV (Forschungsgesellschaft für Straßen- und Verkehrswesen) [2005]: Hinweise zum Fundamentaldiagramm – Grundlagen und Anwendungen. Ausgabe 2005. FGSV Verlag GmbH, Köln
- FRIEDRICH, B.; ALMASRI, E. [2005]: Modellbasierte Optimierung der Versatzzzeiten mit dem Cell Transmission Model. Heureka, Karlsruhe, 2. und 3. März 2005

- 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 (IIS), Orlando, Florida, USA
- FRIEDRICH, M.; VORTISCH, P. [2005]: Verfahren zur dynamischen Verkehrsumlegung – Ein methodischer Überblick. Straßenverkehrstechnik 03/2008, Kirschbaumverlag, Bonn
- FRÖHLICH, S. [2006]: Untersuchungen zur effizienten Generierung, Archivierung und Fortschreibung von Ganglinien für die Nutzung in operativen Verkehrsmanagementsystemen. Diplomarbeit an der Professur für Verkehrsleitsysteme und –prozessautomatisierung, Technische Universität Dresden
- GAZIS, D. C.; HERMAN, R.; ROTHERY, R. W. [1961]: Nonlinear Follow The Leader Models of Traffic Flow. Operations Research, Vol. 9
- GIPPS, P. [1981]: A Behavioural Car-Following Model for Computer Simulation. Transportation Research Part B, Vol. 15B
- GÖRZ, G.; ROLLINGER, C.; SCHNEEBERGER, J. (Hrsg.) [2000]: Handbuch der künstlichen Intelligenz. 3. Auflage. Oldenbourg Verlag, München
- GRÄNICHER, H. [1996]: Messung beendet – Was nun? Einführung und Nachschlagewerk für die Planung und Auswertung von Messungen. – 2., überarb. Aufl. – Zürich: vdf, Hochschulverlag AG an der ETH Zürich, ISBN 3-7281-2258-0; Stuttgart: Teubner, ISBN 3-519-13659-7
- GRÜBER, B.; RÖHR, T. [2007]: Reisezeitmessungen setzen sich durch. Straßenverkehrstechnik 05/2007, Kirschbaumverlag, Bonn
- GRÜLLER, R. [1997]: NAVSTAR-GPS, das globale Positionierungssystem. Schweizer Eisenbahn-Revue 5/1997, Minirex AG, Luzern
- GÜRTLER, S. [2005]: Mikroskopische Simulation des Verkehrsablaufs in Ballungsräumen mit zellularen Automaten. Diplomarbeit, Lehrstuhl für Verkehrstechnik, Technische Universität München
- HENNINGER, T. [2001]: Ein Verfahren zur gekoppelten Schätzung von Kantenbelastungen, Abbiegequoten und Störungen in Stadtstraßennetzen. Dissertation, Arbeitsgruppe Automatisierungstechnik, Technische Universität Hamburg-Harburg
- HERZ, R.; SCHLICHTER, H.; SIEGENER, W. [1992]: Angewandte Statistik für Verkehrs- und Regionalplaner. 2., neubearbeitete und erweiterte Auflage 1992. Werner-Verlag, Düsseldorf, ISBN 3-8041-1971-9
- HOOPS, M.; BUSCH, F.; GLOGER, R.; KELLER, H. [1999]: NEMO, Netzmodellierung zur Erstellung eines Verkehrslageberichts in einem Strategie- und Service-Zentrum am Beispiel des Projektes Munich-COMFORT. Endbericht im Rahmen des Projektes LLAMD/Munich-COMFORT

- JANKO, J. [1994]: Probleme der Reisezeitprognose in einem Leitsystem für den Straßenverkehr. Dissertation, TU Berlin
- KATES, R.; BOBINGER, R.; STOICA, I.; MÜCK, J. [2005]: Fusion von Verkehrsprognosen für die Praxisanwendung. Straßenverkehrstechnik 06/2005, Kirschbaumverlag, Bonn
- KIM, Y. [2002]: Online Traffic Flow Model Applying Dynamic Flow-Density Relations. Dissertation am Fachgebiet für Verkehrstechnik und Verkehrsplanung, Technische Universität München
- LEONHARDT, A. [2008]: Ein Instanzbasiertes Lernverfahren zur Prognose von Verkehrskenngrößen unter Nutzung räumlich-zeitlicher Muster. Dissertation am Lehrstuhl für Verkehrstechnik, Technische Universität München
- LEONHARDT, A. [2010]: Ein Instanzbasiertes Lernverfahren zur Prognose von Verkehrskenngrößen unter Nutzung räumlich-zeitlicher Muster. Straßenverkehrstechnik 02/2010, Kirschbaumverlag, Bonn
- LESTER, P. [2005]: A* Pathfinding for Beginners. www.gamedev.net
- LIGHTHILL, M.; WHITHAM, G. [1955]: On kinematic waves – II. A theory of traffic flow on long crowded roads. Proceeding of the Royal Society, London, Series A, Mathematical and Physical Sciences, Vol. 229, No. 1178, Mai 1955
- LINAUER, M.; DIN, K.; ASAMER, J.; NOWOTNY, B.; NORMAN, N.; HEININGER, M.; HEININGER, D.; PAUKERL, A.; TUREK, T.; STEINER, P. [2006]: Intelligente Infrastruktur; FUSION – Floating Car Daten und Sensordaten intelligent fusionieren. Endbericht. Bundesministerium für Verkehr, Innovation und Technologie, Wien
- LISTL, G.; DAMMANN, W. [2009]: Untersuchungen zum Einsatz von Taxi-Floating Car Data im Ballungsraum Rhein-Main. Straßenverkehrstechnik 03/2009, Kirschbaumverlag, Bonn
- LOGI, F.; ULLRICH, M. [2001]: The Dynamic Network Model DINO: Documentation. Arbeitspapier, Fachgebiet Verkehrstechnik und Verkehrsplanung, Technische Universität München
- MAIER, F. [2004]: Einfluss der Streckenauslastung auf die Fahrzeiten unterschiedlicher Kfz-Fahrergruppen. Diplomarbeit, Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik, Uni-versität Stuttgart
- MAIER, F.; BRAUN, R.; BUSCH, F.; MATHIAS, P. [2008]: Musterbasierte Kurzfristprognose der innerörtlichen Störungsausbreitung und automatische Maßnahmenempfehlung. Straßenverkehrstechnik 09/2008, Kirschbaumverlag, Bonn
- MAIER, F.; ROTH, C. [2008]: Verfahren zur Wiedererkennung von Fahrzeugen anhand der Signaturen über Induktivschleifen. Straßenverkehrstechnik 02/2008, Kirschbaumverlag, Bonn
- MATHIAS, P. [1999]: Statische und dynamische Verkehrslegung mit Rekurrenten Neuronalen Netzen. Dissertation an der RWTH Aachen

- MATSCHKE, I.; FRIEDRICH, B.; HEINIG, K. [2004]: Data Fusion Technique in the Context of Traffic State Estimation. Proceeding of the Triennial Symposium on Transportation Analysis TRISTAN V, June 13 - 18, 2004, Le Gosier - Guadeloupe, French West Indies
- MESCHENDÖRFER J. [2004]: Fahrzeugklassifizierung und Fahrzeugwiedererkennung anhand von Mikrodaten lokaler Detektoren. Diplomarbeit am Lehrstuhl für Verkehrstechnik, Technische Universität München, München
- MEIER, J.; WEHLAN, H. [2000]: AZTEK – Automatische zeitlich-räumliche Störungserkennung. Technischer Bericht, Institut für Systemdynamik und Regelungstechnik, Universität Stuttgart
- MIN, W.; WYNTER, L.; AMEMIYA, Y. [2007]: Road Traffic Prediction with Spatio-Temporal Correlations. IBM Research Report RC24275 (W0706-018)
- NAGEL, K.; SCHRECKENBERG, M. [1992]: A cellular automaton model for freeway traffic. Journal de Physique I 2, Les Ulis Cedex A
- OFFERMANN, F. [2001]: Ein Neuro-Fuzzy-Modell zur Reisegeschwindigkeitsabschätzung auf Richtungsfahrbahnen basierend auf einer Fusion lokaler und fahrzeuggenerierter Verkehrsdaten. Dissertation am Lehrstuhl für Straßenwesen, Erd- und Tunnelbau, Rheinisch-Westfälische Technische Hochschule Aachen
- OH, C., RITCHIE, S. G. [2002]: Anonymous Vehicle Tracking for Real-Time Traffic Surveillance. Technical report UCI-ITS-TS-WP-02-16, University of California, Irvine
- OH, S., RITCHIE, S. G., OH, C. [2001]: Real-Time traffic Measurement from Single Loop Inductive Signatures. 81st Annual TRB Meeting, Washington, D.C.
- PFANNERSTILL, E. [1983]: Ein Verfahren zur Merkmalsextraktion aus Linienmustern zur Zuordnung von Fahrzeugkollektive. Dissertation, Aachen
- PINKOFSKY, L. [2006]: Typisierung von Ganglinien der Verkehrsstärke und ihre Eignung zur Modellierung der Verkehrsnachfrage. Dissertation, Institut für Verkehr und Stadtbauwesen, Technische Universität Braunschweig
- SAAKE, G.; Sattler, K.-U.; Keim, D. [2000]: Datenbank- und Visualisierungstechnologie in der Informationsfusion. Simulation und Visualisierung 2000, 11. Märztagung an der Universität Magdeburg
- SANDERS, P.; SCHULTES, D. [2005]: Genauer Hochleistungs-Routenplaner. Heidelberger Innovationsforum, 30.11.2005
- SCHIMANDL, F.; ZHANG, M.; MUSTAFA, M.; MENG, L. [2009]: Real time application for traffic state estimation based on large sets of floating car data. mobil.TUM International Scientific Conference on Mobility and Transport – ITS for larger Cities, München, 12.-13. Mai 2009
- SCHENDZIELORZ, T. [2003]: Entwicklung und Anwendung eines Bewertungsverfahrens für DINO. Diplomarbeit am Lehrstuhl für Verkehrstechnik, Technische Universität München

- SCHNABEL, W.; LOHSE, D. [1997]: Grundlagen der Straßenverkehrstechnik und der Verkehrsplanung – Band 1. Verlag für Bauwesen, Berlin
- SCHÖBER, M. [2009]: Robuste Verkehrszustandsschätzung und Störungserkennung auf Schnellstraßen. Dissertation an der Universität Stuttgart
- SPANGLER, M. [2006]: Systembeschreibung Echtzeitermittlung von Reisezeiten. Arbeitspapier, Lehrstuhl für Verkehrstechnik, Technische Universität München
- SPANGLER, M. [2009]: Reisezeitbasierte Verfahren für die Verkehrszustandsanalyse von städtischen Hauptverkehrsstraßen. Dissertation am Lehrstuhl für Verkehrstechnik, Technische Universität München
- SEYDEL, I. [2008]: Möglichkeiten der Nutzung multivariater Analysemethoden zur Untersuchung von Verkehrsdaten. Diplomarbeit an der Professur für Verkehrsleitsysteme und -prozessautomatisierung, Technische Universität Dresden
- STAHEL, A. [1995]: Statistische Datenanalyse – Eine Einführung für Naturwissenschaftler. Verlag Vieweg, ISBN 3-528-06653-9
- SUN, C., RITCHIE, S.G., TSAI, K., JAYAKRISHNAN, R. [1999]: Use of vehicle signature analysis and lexicographic optimization for vehicle reidentification on freeways. Technical report UCI-ITS-WP-98-17, University of California, Irvine
- THOMAS, N.E. [1998]: Multi-state and multi-sensor incident detection systems for arterial streets. Transportation Research Part C, Vol. 6 (1998), Nr. 5/6
- TRB [2000]: Highway Capacity Manual, National Research Council, Washington, D.C.
- VAN ZUYLEN, H. J.; WILLUMSEN, L. G. [1980]: The Most Likely Trip Matrix Estimation from Traffic Counts. Transportation Research Part B, Vol. 14(1980), Nr. 3
- VORTISCH, P. [2006]: Modellunterstützte Messwertpropagierung zur Verkehrslageschätzung in Stadtstraßennetzen. Dissertation am Institut für Verkehrswesen, Universität Karlsruhe
- VORTISCH, P.; ORTGIESE, M. [2005]: Vom Stau zur Verkehrsinformation – Datenfusion als Teil eines Gesamtprozesses. at – Automatisierungstechnik 53 (2005), Oldenbourg Verlag
- WELCH, G.; BISHOP, G. [2004]: An Introduction to the Kalman Filter. http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf, eingesehen am 31. August 2009
- WAGNER, P. [2007]: Überblick FCD-Projekte des DLR: Erfahrungen, Potenzial und Limiten. <http://www.itsswitzerland.com>, eingesehen am 09. Oktober 2009
- WIEDEMANN, R. [1974]: Simulation des Straßenverkehrs. Schriftenreihe des Instituts für Verkehrswesen, Universität Karlsruhe
- ZADEH, L. [1965]: Fuzzy Sets. Information and Control, Vol. 8, Nr. 3
- ZIMMERMANN, H.-J. [1999]: Intelligente Technologien und Data Mining zur Unterstützung des Verkehrsmanagements. Tagungsbericht bei der HEUREKA 1999, Karlsruhe

Abkürzungsverzeichnis

abs. (5)	Absolute Häufigkeit der Näherungsfunktionen bei Bildung von fünf Abschnitten
abs. (3)	Absolute Häufigkeit der Näherungsfunktionen bei Bildung von drei Abschnitten
DSRC	Dedicated Short Range Communication
Egw.	Eingangswert
f.	und der/die/das Folgende
ff.	und die Folgenden
FCD	Floating Car Data
GPRS	General Packet Radio Service
GPS	Global Positioning System
gr.	Abschnitt für große Eingangswerte, also rechter Abschnitt
Hz	Hertz (Einheit bei der Frequenzmessung)
IR	Infrarot
Link-v/v0	Quotient aus linkbezogener Geschwindigkeit aus Positionsmeldungen und freier Geschwindigkeit gemäß Netzinformationen [-]
Lkw	Lastkraftwagen
LOS	Level of Service
LSA	Lichtsignalanlage
LSA q	mit einem LSA-Detektor erhobene Verkehrsstärke [Fz/h]
k	Verkehrsdichte [Fz/km]
kl.	Abschnitt für kleine Eingangswerte, also linker Abschnitt bei grafischer Darstellung
m	Meter (Längeneinheit)
mAR	mit Autoregression
mi.	Abschnitt für mittlere Eingangswerte, also mittlerer Abschnitt
Mio.	Million(en)
n	Stichprobenumfang oder Anzahl der bei der Zusammenführung berücksichtigten Einzelschätzungen
Nb.	Anzahl der Beziehungen, die nach der Suchraumbeschränkung gemäß Kapitel 5.4 untersucht wurden („Nb.“ für Nachbarschaft)

oAR	ohne Autoregression
ÖPNV	Öffentlicher Personennahverkehr
Pkw	Personenkraftwagen
q	Verkehrsstärke [Fz/h]
r	Radius oder Widerstand (resistance)
r^2	Bestimmtheitsmaß
rel. (3)	Relative Häufigkeit bezogen auf alle untersuchten Beziehungen bzw. sämtliche Näherungsfunktionen bei Bildung von drei Abschnitten für diesen zeitlichen Versatz
s	Sekunde
Radar	Radio Detection And Ranging
RFID	Radio Frequency Identification
Routen-TT	Reisezeit [min] durch Fahrzeugwiedererkennung
RMSE	Wurzel der Summe quadrierter Fehler (Root Mean Squared Error)
strat. q	mit strategischem Detektor erhobene Verkehrsstärke [Fz/h]
strat. v	mit strategischem Detektor erhobene mittlere lokale Geschwindigkeit [km/h]
strat. occ	mit strategischem Detektor erhobener Belegungsgrad [%]
\hat{t}	Wert der t-Statistik
Δt	Zeitversatz
TLS	Technische Lieferbedingungen für Streckenstationen
v.a.	vor allem
0-24u	Verwendung aller Daten (7 Tage/Woche, 24 Stunden/Tag) ungewichtete lineare Regression
0-24g	wie 0-24u, jedoch gewichtete lineare Regression
6-10u	Verwendung der Daten zwischen Montag und Freitag, jeweils 6-10 Uhr, ungewichtete lineare Regression
0-24/7	Montag bis Sonntag, 0-24 Uhr
6-10/5	Montag bis Freitag, 6-10 Uhr

Abbildungsverzeichnis

Bild 1.1	Grundsätzlicher Ablauf der Schätzung von Verkehrskenngrößen bei Verwendung eines regressionsbasierten Ansatzes.....	3
Bild 2.1	Fahrzeugsignaturen als Rohdaten (oben) und nach der Normierung	10
Bild 2.2	Zwei einander zugeordnete Pkw-Signaturen und ihr Steigungsverlauf	11
Bild 2.3	Zwei einander zugeordnete Lkw-Signaturen und ihr Steigungsverlauf	11
Bild 2.4	Ableitung von linkbezogenen Kenngrößen aus Flottenpositionsdaten – Überblick.....	14
Bild 2.5	Gerasterte digitale Karte mit räumlich indizierten Links.....	17
Bild 2.6	Abgleich einer Position mit dem Netz basierend auf Spatial Indexing	17
Bild 2.7	Positionen mit Polygonzug einer Route – über Spatial Indexing ausgewählte eventuell zugehörige Links, nicht verworfene Links durchgezogen eingezeichnet.....	18
Bild 2.8	Abstandsbestimmung Position-Link und genaue Verortung (1).....	19
Bild 2.9	Abstandsbestimmung Position-Link und genaue Verortung (2).....	19
Bild 2.10	Vergleich Dijkstra-Algorithmus – A*-Algorithmus	23
Bild 2.11	Ergebnis der Routensuche ohne und mit optimierter Teilroutenkombination	24
Bild 3.1	Überblick über das Verfahren zur Störungsausbreitungsprognose	35
Bild 3.2	Störungsmuster, Zeit von unten nach oben fortschreitend.....	36
Bild 3.3	Zusammenstellung der gesichteten Verfahren zur Verkehrslageschätzung	40
Bild 4.1	Konzeptioneller Ansatz zur Schätzung verkehrlicher Kenngrößen	43
Bild 4.2	Messwertmatrix für das Netz gemäß Bild 4.3 für fünf Zeitschritte mit Schätzung fehlender Daten unter Verwendung beobachteter Zusammenhänge	45
Bild 4.3	Beispielnetz.....	45
Bild 4.4	Diagonale Beobachtung von Zusammenhängen zwischen Verkehrskenngrößen und Autoregression.....	45
Bild 4.5	Näherungsgerade aus einfacher linearer Regression.....	49
Bild 4.6	Näherungsgerade aus gewichteter linearer Regression	49
Bild 4.7	Einfache lineare Regression für Stützstellen aus einer kubischen Funktion.....	50
Bild 4.8	Einfache lineare Regression für Stützstellen aus einer q-k-Punktvolke.....	50
Bild 4.9	Abschnittsweise lineare Regression – optimiert nach dem Bestimmtheitsmaß	51
Bild 4.10	Abschnittsweise lineare Regression mit einer konstanten Anzahl von Messpunkten	52
Bild 4.11	Abschnittsweise lineare Regression mit äquidistanten Abschnitten.....	52
Bild 4.12	Lineare Regression für äquidistante Abschnitte und Überlappungen von 0,2 (oben) und 0,5 (unten)	53
Bild 4.13	Zusammenhänge zwischen zahlreichen Kenngrößen und Schätzung von Kenngrößen basierend auf aktuellen Messwerten	54
Bild 5.1	München Nord – ursprüngliches Netz.....	57
Bild 5.2	Verfügbarkeit von Linkgeschwindigkeiten aus Positionsmeldungen	58
Bild 5.3	Netz zur Verkehrslageschätzung	58
Bild 5.4	Quotienten v/v_0 aus linkbezogenen Geschwindigkeiten (netzweit aus Positionsdaten ermittelt) und linkbezogenen freien Geschwindigkeiten über die Zeit	59
Bild 5.5	Lage der lokalen Detektoren im Netzausschnitt.....	60
Bild 5.6	Netzabdeckung der Routen mit infrastrukturseitiger Reisezeitmessung für den vorliegenden Netzausschnitt	61

Bild 5.7	Summenlinie der Datenverfügbarkeit für infrastrukturseitig gemessene Reisezeiten in 5-min-Intervallen über 73 Routen.....	61
Bild 5.8	Suchraum nach Zusammenhängen zu einem untersuchten Link	64
Bild 6.1	Bewertungsnetz mit Querschnitten zur Fahrzeugwiedererkennung und Benennung von Streckenzügen	67
Bild 6.2	Erläuterung zum Verständnis und zur Verknüpfung von Tab. 6.1, Tab. 6.2 und Tab. 6.3 .	69
Bild 6.3	Beispiele für Näherungsfunktionen mit Verkehrsstärken aus strategischen Detektoren als erklärende Kenngröße, jeweils für drei (schwarz) und für fünf (blau) Abschnitte, ohne Zeitversatz; Distanz r zwischen den Erhebungsorten in unterschiedlichen Maßeinheiten: Anzahl der dazwischen liegenden Knotenpunkte, Fahrentfernung [m] und freie Reisezeit [s].....	75
Bild 6.4	Beispiele für Näherungsfunktionen mit lokalen mittleren Geschwindigkeiten aus strategischen Detektoren als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz.....	77
Bild 6.5	Beispiele für Näherungsfunktionen mit Belegungsgraden aus strategischen Detektoren als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz.....	79
Bild 6.6	Beispiele für Näherungsfunktionen mit Verkehrsstärken aus LSA-Detektoren als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz.....	80
Bild 6.7	Näherungsfunktionen für strategisch erhobene lokale Kenngrößen mit den Messwerten eines strategischen Detektors auf dem benachbarten Fahrstreifen als Eingangsgröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte	82
Bild 6.8	Beispiele für Näherungsfunktionen mit Routenreisezeiten als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz.....	84
Bild 6.9	Beispiele für Näherungsfunktionen mit Linkgeschwindigkeiten als erklärende Kenngröße, jeweils für drei (schwarz) und fünf (blau) Abschnitte, ohne Zeitversatz.....	85
Bild 6.10	Näherungsfunktion für Linkgeschwindigkeiten aus Linkgeschwindigkeiten bei nahegelegenen Links	87
Bild 6.11	Mäßiger Zusammenhang zwischen lokaler und linkbezogener Geschwindigkeit.....	87
Bild 6.12	Beispiele für Näherungsfunktionen mit verschiedenen Zeitversätzen aus lokalen Daten samt statistischen Kennwerten.....	89
Bild 6.13	Beispiele für Näherungsfunktionen mit verschiedenen Zeitversätzen aus Abschnittdaten samt statistischen Kennwerten.....	91
Bild 6.14	Beispiele für Ergebnisse der Autoregression mit verschiedenen Zeitversätzen samt statistischen Kennwerten.....	92
Bild 6.15	Vergleich der Näherungsfunktionen aus der Regressionsanalyse zeitlich undifferenziert (ungewichtet und gewichtet) sowie für Mo-Fr, 6-10 Uhr (ungewichtet) – Beispiele für die Schätzung von Linkgeschwindigkeiten und Routenreisezeiten	93
Bild 6.16	Vergleich der Näherungsfunktionen aus der Regressionsanalyse zeitlich undifferenziert und differenziert (Mo-Fr, 6-10 Uhr) – Beispiele für die Schätzung lokaler Daten.....	96
Bild 6.17	Beispiele für Ergebnisse der zeitlich differenzierten Regressionsanalyse mit zeitlich versetzten Ein- und Ausgangsdaten.....	97
Bild 7.1	Erkennungsrate von Perzentilereignissen einer einzelnen Kenngröße.....	102

Bild 7.2	Beispiele für die Kalibrierung der Schätzung (rot) verschiedener Kenngrößenarten (blau – idealer unkalibrierter Schätzer, schwarz – Gegenüberstellung einzelner Schätz- und Messwerte).....	107
Bild 7.3	Grafischer Vergleich der Verkehrsstärkemessung mit ihrer Schätzung; Dienstag, 17.3.2009	109
Bild 7.4	Grafischer Vergleich der Messung der mittleren Geschwindigkeit mit ihrer Schätzung; Mittwoch, 18.3.2009	113
Bild 7.5	Grafischer Vergleich der Belegungsgradmessung mit ihrer Schätzung durch unterschiedliche Verfahren; Donnerstag, 19.3.2009	115
Bild 7.6	Grafischer Vergleich der Verkehrsstärkemessung an einem LSA-Detektor mit ihrer Schätzung; Freitag, 20.3.2009.....	116
Bild 7.7	Grafischer Vergleich der Reisezeitmessung auf einer Route mit ihrer Schätzung; Dienstag, 17.3.2009, Route 1-2	118
Bild 7.8	Grafischer Vergleich von linkbezogenen Geschwindigkeitsschätzungen an einem Link am Mittleren Ring, Höhe Olympiapark; Montag, 16.3.2009 bis Sonntag, 22.3.2009	119
Bild 7.9	Grafischer Vergleich von linkbezogenen Geschwindigkeitsschätzungen an einem Link auf der A9 Richtung Nürnberg; Freitag, 20.3.2009 (Legende siehe Bild 7.8).....	119
Bild 7.10	Grafischer Vergleich von linkbezogenen Geschwindigkeitsschätzungen an einem Link am Mittleren Ring, Höhe Olympiapark; Donnerstag, 19.3.2009 (Legende siehe Bild 7.8).....	119
Bild 7.11	Schätzung des linkbezogenen Quotienten aus geschätzter und freier Geschwindigkeit in 5-min-Intervallen von Montag, 16. (unten) bis Sonntag, 22.3.2009 (oben) auf der Route 1-2	121
Bild 7.12	Wie Bild 7.11, jedoch nur für Montag, 16.3.2009, zusätzlich mit der Schätzung aus der zeitlich differenzierten Regressionsanalyse für 6 bis 10 Uhr.....	121
Bild 7.13	Vergleich der gemessenen Routenreisezeit und der Summe der dort geschätzten Linkreisezeiten mit schrittweise gebildeten Teilsummen, Dienstag, 17.3.2009, Route 1-2	121
Bild 7.14	Schätzung des linkbezogenen Quotienten aus geschätzter und freier Geschwindigkeit in 5-min-Intervallen von Montag, 16. (unten) bis Sonntag, 22.3.2009 (oben) auf der Route 5-4	122
Bild 7.15	Wie Bild 7.13, jedoch für Donnerstag, 19.3.2009, Route 5-4	123
Bild 7.16	Wie Bild 7.13, jedoch für Montag, 16.3.2009, Route 2-3.....	123
Bild 7.17	Summenlinien der Distanzverteilung zwischen Ein- und Ausgangsgröße für verfügbare (rot) und genutzte (blau) Schätzfunktionen.....	125
Bild 7.18	Schätzung lokaler Geschwindigkeiten ohne Eingangsgrößen vom selben Link (vgl. Bild 7.4)	127
Bild 7.19	Schätzung von Linkgeschwindigkeiten ausschließlich aus lokalen Daten (links) und ausschließlich aus Routenreisezeiten (rechts; vgl. Bild 7.14 f.).....	128
Bild 8.1	Überblick über das Verfahren zur Schätzung von Verkehrskenngrößen durch abschnittsweise Regressionsanalyse.....	132

Tabellenverzeichnis

Tab. 6.1	Absolute Häufigkeit von Näherungsfunktionen bei drei und fünf gebildeten Abschnitten sowie relative Häufigkeit bei drei Abschnitten für verschiedene Arten von erklärenden und geschätzten Kenngrößen und für unterschiedliche zeitliche Versätze.....	71
Tab. 6.2	Aufteilung der absoluten Häufigkeiten gemäß Tab. 6.1 auf die Abschnitte bei Bildung von drei Abschnitten.....	72
Tab. 6.3	Mittleres Bestimmtheitsmaß der Näherungsgeraden gemäß Tab. 6.2	73
Tab. 6.4	Anzahl der Näherungsgeraden aus dem Zeitfenster werktags 6-10 Uhr relativ zur Anzahl aus der zeitlich nicht differenzierten Regressionsanalyse gemittelt über alle untersuchten zeitlichen Versätze [%].....	95
Tab. 7.1	Erkennungsrate von Perzentilereignissen der Reisezeit auf einer Route, Autoregression zugelassen; mit dem Bestimmtheitsmaß gewichtete Mittelung über die n besten Einzelschätzungen	106
Tab. 7.2	Erkennungsrate von Perzentilereignissen der Verkehrsstärke an einem strategischen Detektor, Autoregression nicht zugelassen; mit dem Bestimmtheitsmaß gewichtete Mittelung über die n besten Einzelschätzungen.....	106
Tab. 7.3	Schätzung strategisch erhobener Verkehrsstärken im Zeitraum 0-24/7 – Bewertung	110
Tab. 7.4	Schätzung strategisch erhobener Verkehrsstärken im Zeitraum 6-10/5 – Bewertung	110
Tab. 7.5	Schätzung strategisch erhobener mittlerer Geschwindigkeiten im Zeitraum 0-24/7 – Bewertung.....	112
Tab. 7.6	Schätzung strategisch erhobener mittlerer Geschwindigkeiten im Zeitraum 6-10/5 – Bewertung.....	112
Tab. 7.7	Schätzung strategisch erhobener Belegungsgrade im Zeitraum 0-24/7 – Bewertung....	114
Tab. 7.8	Schätzung strategisch erhobener Belegungsgrade im Zeitraum 6-10/5 – Bewertung....	114
Tab. 7.9	Schätzung von Verkehrsstärken an LSA-Detektoren im Zeitraum 0-24/7 – Bewertung..	116
Tab. 7.10	Schätzung von Verkehrsstärken an LSA-Detektoren im Zeitraum 6-10/5 – Bewertung..	116
Tab. 7.11	Schätzung von Routenreisezeiten im Zeitraum 0-24/7 – Bewertung.....	117
Tab. 7.12	Schätzung von Routenreisezeiten im Zeitraum 6-10/5 – Bewertung.....	117
Tab. 7.13	Anteile [%] von Datenarten als Eingangsgrößen an verfügbaren (verf.) und genutzten (gen.) Schätzfunktion bei der Schätzung anderer Datenarten, differenziert nach zeitlichen Versätzen	126
Tab. 7.14	Schätzung lokaler Geschwindigkeiten ohne Eingangsgrößen vom selben Link im Zeitraum 0-24/7 (vgl. Tab. 7.5).....	127
Tab. 7.15	Schätzung der Verkehrsstärke an einem LSA-Detektor ohne Eingangsgrößen vom selben Link (vgl. Tab. 7.9).....	128

Anhang 1: Vorabveröffentlichungen

MAIER, F.; ROTH, C. [2008]: Verfahren zur Wiedererkennung von Fahrzeugen anhand der Signaturen über Induktivschleifen. Straßenverkehrstechnik 02/2008, Kirschbaumverlag, Bonn

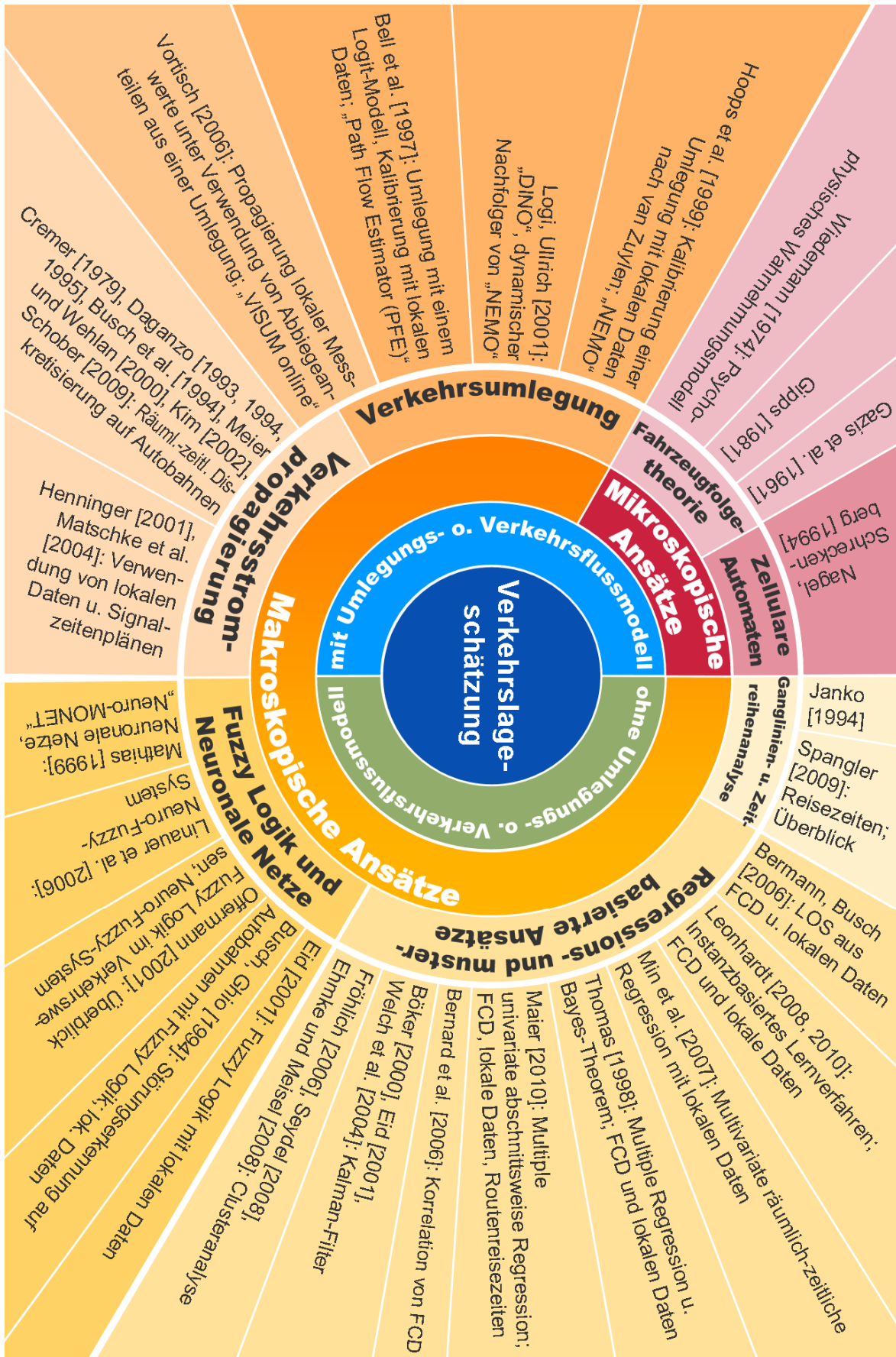
BUSCH, F.; MAIER, F. [2008]: Musterbasierte Störungserkennung und -ausbreitungsprognose. Fachtagung "Verkehrsmanagement und Verkehrstechnologien", Halle/Saale, 20./21.Mai 2008

MAIER, F.; BRAUN, R.; BUSCH, F.; MATHIAS, P. [2008]: Pattern-based short term prediction of urban congestion propagation and automatic response. TEC magazine 06/2008

MAIER, F.; BRAUN, R.; BUSCH, F.; MATHIAS, P. [2008]: Musterbasierte Kurzfristprognose der innerörtlichen Störungsausbreitung und automatische Maßnahmenempfehlung. Straßenverkehrstechnik 09/2008, Kirschbaumverlag, Bonn

Anhang 2: Überblick über Verfahren zur Verkehrslageschätzung

Wie Bild 3.3, ergänzt um das vorgestellte neue Verfahren (nächste Seite).



Schriftenreihe des Lehrstuhls für Verkehrstechnik

Univ.-Prof. Dr.-Ing. Fritz Busch

ISSN 1612-9431

Heft 1	Steinhoff, Christiane: Online Bewertung der Akzeptanz und der Wirksamkeit präventiver Maßnahmen durch Streckenbeeinflussungsanlagen auf Autobahnen	2003
Heft 2	Hoops, Marcus: Methodik zur Qualitätssicherung von aggregierten Daten eines Messsystems im Betrieb von verkehrstechnischen Anlagen	2008
Heft 3	Zhang, Ying: Vergleich der Verkehrsentwicklung in deutschen Großstädten und Shanghai sowie Herleitung von Handlungserfordernissen	2003
Heft 4	Nocera, Silvio: Steuerung des Sektorbetriebs bei nachfrageabhängiger ÖPNV-Bedienung	2004
Heft 5	Underberg, Robert: Bereitstellung und Nutzung von Messwerten des Verkehrsablaufs im ÖPNV im ländlichen Raum	2004
Heft 6	Vukanovic, Svetlana: Intelligent link control framework with empirical objective function: INCA	2009
Heft 7	Belzner, Heidrun: Erkennung von Verkehrsstörungen unter Verwendung erweiterter fahrzeuggenerierter Daten	2009
Heft 8	Lange, Ulrich: Wirkung komplexer Maßnahmen auf die Verkehrsmittelwahl von Umlandpendlern	2009
Heft 9	Leonhardt, Axel: Ein Instanzbasiertes Lernverfahren zur Prognose von Verkehrskenngrößen unter Nutzung Räumlich-Zeitlicher Verkehrsmuster	2009
Heft 10	Braun, Robert: Ein echtzeitfähiger Evolutionärer Algorithmus zur netzweiten Optimierung der Lichtsignalsteuerung	2009
Heft 11	Spangler, Matthias: Reisezeitbasierte Verfahren für die Verkehrszustandsanalyse von städtischen Hauptverkehrsstraßen	2009
Heft 12	Kloth, Holger: Aufgaben der Landkreise bei der Planung und Erfolgskontrolle des Öffentlichen Personennahverkehrs	2010
Heft 13	Maier, Friedrich: Abschnittsweise Regressionsanalyse zur Schätzung von Verkehrskenngrößen	2010

