Graz University of Technology

Master Thesis

Traffic Flow Reconstruction on Motorways by Data Fusion

conducted at the
Institute for Highway Engineering and Transport Planning
Graz University of Technology, Austria

in co-operation with
Signal Processing and Speech Communications Laboratory
Graz University of Technology, Austria

by
Erwin Nindl

Supervisors:
Univ.-Prof. Dr.-Ing. Martin Fellendorf

Supervising Assistants:
Dipl.-Ing. Paul Meissner
Dipl.-Ing. Robert Neuhold
Dipl.-Ing. Dipl.-Ing. Dr.techn. Thomas Reiter

Assessors/Examiners:
Assoc.Prof. Dipl.-Ing. Dr.mont. Franz Pernkopf

Graz, December 10, 2013

Abstract

The primary purpose of this Master's thesis was to investigate methodologies for traffic state estimation on Austrian motorways based on various data sources. This work was motivated to provide insight into traffic state estimation methods performing data-fusion of multiple sources.

This thesis is organised in two major parts: (1) The theoretical part, including an introduction to the terminology of traffic engineering, a comprehensive literature review of the latest traffic flow models, and data-fusion concepts. (2) The practical part, dealing with the implementation and validation of algorithms based on the idea of interpolation with respect to the characteristic shock waves present in traffic flow, i.e. the *Generalised Adaptive Smoothing Method (GASM)* and the *Extended Generalised Treiber-Helbing Filter (EGTF)*. Within the thesis, important preprocessing steps for the data, which are crucial for the subsequent reconstruction methods, were identified and discussed in detail. Furthermore, suitable methods (e.g. based on extracted measures like travel times and evaluation using Microsimulation data as ground truth) for the evaluation of the experiments were discussed.

The conducted experiments revealed that the GASM and the related EGTF were well suited for velocity field reconstruction on Austrian motorways. Findings based on the experiments led to the conclusion that the EGTF showed superior performance in most cases. One of the most important outcomes was that a preceding bias correction of velocities obtained from stationary sensors is essential for satisfactory results of traffic state estimation. In addition the conducted experiments highlighted certain points for potential improvement.

Kurzfassung

Die vorliegende Masterarbeit liefert die Untersuchung von auf unterschiedlichen Datenquellen basierenden Verfahren, die der Rekonstruktion der Verkehrslage auf österreichischen Autobahnen dienen. Darüber hinaus wird darauf abgezielt, sich einer Datenfusion der verschiedenen Quellenlagen zu bedienen.

Der Aufbau der folgenden Arbeit gliedert sich in zwei Hauptbereiche: (1) Einem Theorieteil, der Definitionen relevanter Begriffe aus dem Verkehrswesen, eine umfassende Literaturrecherche geeigneter Verkehrsflussmodelle sowie eine Übersicht zu modernen Datenfusionskonzepten bietet. (2) Einem praktischen Teil zum Zwecke der Umsetzung und Validierung der beiden Algorithmen – der *Generalised Adaptive Smoothing Method (GASM)* und dem *Extended Generalised Treiber-Helbing Filter (EGTF)* – basierend auf der Idee der Interpolation unter spezieller Berücksichtigung auf die im Verkehrsfluss vorhandenen Schockwellen. In diesem Rahmen wurden die wichtigsten, für die daraus folgenden Rekonstruktionsmethoden grundlegenden Schritte der Datenvorverarbeitung im Detail betrachtet. Weiters wurden Möglichkeiten der Bewertung der gewonnenen experimentellen Ergebnisse diskutiert wie beispielsweise durch abgeleitete Größen von Reisezeiten oder direkt durch Durchführung einer Mikrosimulation.

Die durchgeführten Experimente belegen, dass sowohl GASM als auch der EGTF zur Verkehrslagerekonstruktion auf österreichischen Autobahnen geeignet sind, wobei aus den durch die Experimente gewonnenen Resultaten die Schlussfolgerung gezogen wird, dass der EGTF in den meisten Fällen konkretere Ergebnisse liefert. Die Tatsache, dass eine Korrektur der systematischen Abweichungen von gemittelten Geschwindigkeiten als wesentlich für den Erhalt zufriedenstellender Ergebnisse der Verkehrslagenrekonstruktion gilt, kann als eine der bedeutendsten Erkenntnisse dieser Arbeit angesehen werden. Darüber hinaus können durch die in dieser Masterarbeit durchgeführten Experimente mögliche zukünftige Ansatzpunkte zur Verbesserung der genannten Methoden geliefert werden.

Problem- und Aufgabenstellung

Problemstellung

Auf österreichischen Autobahnen und Schnellstraßen ist aus Gründen der Verkehrssteuerung, Statistik und Mautverrechnung Sensorik zur automatisierten Erhebung von Verkehrsdaten installiert. Generell eignen sich diese Daten ebenso zur Schätzung der aktuellen Verkehrslage. Aus der rekonstruierten Verkehrslage können unterschiedliche Ereignisse wie Unfälle, Staus oder andere Vorfälle detektiert werden. Besonders in der Nähe von Ballungsräumen ist eine Vielzahl ortsfester Sensoren entlang des Straßenverlaufs installiert (*Euler'sche Datenquellen*). Zwischen den ortsfesten Sensoren liegen in der Regel keine Messdaten vor, wodurch die Verkehrslage anhand der zur Verfügung stehenden Daten rekonstruiert werden muss.

Neben den klassischen Detektorquerschnittsdaten stehen vermehrt hochaufgelöste Trajektorien einzelner Verkehrsteilnehmer, sogenannte Floating-Car-Daten, zur Verfügung. Diese Daten aus der Sicht einzelner Fahrzeuge werden unter anderem auch *Lagrange'sche Daten* genannt.

In dieser Arbeit sollen verschiedene Rekonstruktionsverfahren auf Basis gegebener Detektorquerschnittsdaten verglichen werden. Weiters soll mittels geeigneter Algorithmen eine Schätzung des Verkehrszustandes basierend auf Euler'schen und Lagrange'schen Datenquellen evaluiert und implementiert werden. Schließlich sollen die entwickelten Verfahren auf ihre Robustheit hinsichtlich des Fehlens einzelner Datengruppen sowie der Messfehler der Datenquellen evaluiert werden.

Aufgabenstellung

Die folgende Liste enthält wesentliche Bearbeitungspunkte der Masterarbeit; Änderungen mit fortschreitendem Erkenntnisstand während der Bearbeitung sind möglich:

1. Recherche, Bewertung und Anwendung bestehender Verfahren zur Verkehrszustandsschätzung aus Euler'schen Datenquellen.

2. Recherche möglicher Verfahren zur Verkehrsschätzung aus Datenquellen verschiedener Kategorien.

3. Entwicklung einer Algorithmik zugeschnitten auf das hochrangige österreichische Straßennetz und Implementierung in MATLAB.

4. Validierung der entwickelten Algorithmen anhand von Simulationsdaten sowie an realen Daten aus dem österreichischen Verkehrsnetz.

5. Untersuchung der Robustheit der entwickelten Algorithmen, insbesondere die Auswirkung des Fehlens einer Datengruppe bzw. einer schlechten lokalen Detektionsdichte auf die Schätzung.

Ein PC mit Berechungssoftware sowie allgemeiner Bürosoftware stehen dem Diplomanden an beiden Instituten zur Verfügung. Der Diplomand verpflichtet sich, die Software sowie die bereitgestellten Daten ausschließlich zur Anfertigung der Masterarbeit zu nutzen. Die simulierten sowie realen Verkehrsdaten werden dem Diplomanden vom Institut für Straßen und Verkehrswesen zur Verfügung gestellt.

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

date	(signature)

Contents

1 Introduction **15**

2 Traffic Flow Theory **17**
2.1 Traffic Data . . . 17
2.1.1 Microscopic Data . . . 18
2.1.2 Aggregated or Macroscopic Data . . . 18
2.2 The Fundamental Relation of Traffic Flow . . . 19
2.3 Elementary Patterns of Congested Traffic . . . 21
2.4 Summary . . . 22

3 Dynamic Traffic Flow Models **23**
3.1 Categorisation of Traffic Flow Models . . . 23
3.2 Continous Macroscopic Models . . . 25
3.2.1 The First-Order Conservation Equation . . . 26
3.2.2 First Order Model or Lighthill, Whitham and Richards Model . . . 27
3.2.3 First Order Model based on Cumulative Flows / Hamilton-Jacobi Equation 29
3.3 Discrete Macroscopic Models . . . 31
3.4 Modelling in Lagrangian Coordinates . . . 32
3.5 Summary . . . 33

4 Traffic State Estimation **35**
4.1 Newtonian Relaxation or Nudging Method . . . 36
4.2 Kalman Filter based Estimation . . . 37
4.3 Variational Data Assimilation . . . 38
4.4 Heuristic Methods . . . 39
4.5 Adaptive Smoothing Method Variants . . . 40
4.5.1 Adaptive Smoothing Method (ASM) . . . 40
4.5.2 Generalised Adaptive Smoothing Method (GASM) . . . 44
4.5.3 Extended generalised Treiber-Helbing Filter (EGTF) . . . 45
4.6 Classification of Models and State Estimation Methods . . . 46
4.7 Summary . . . 47

5 Data Sources **49**
5.1 Test-Site Austrian A4 Motorway . . . 49
5.2 Stationary Detectors . . . 52
5.3 Floating Cars . . . 56
5.3.1 GPS based Position Estimation . . . 56
5.3.2 GPS based Velocity Estimation . . . 59
5.4 Travel Time Data . . . 62
5.4.1 Total Travel Times from Virtual Trajectories . . . 63
5.5 Microsimulation . . . 65
5.6 Detailed Record of Available Data . . . 67
5.7 Summary . . . 67

6 Implementation **69**
6.1 Traffic-Data Container . . . 69
6.2 Microsimulation Framework . . . 70
6.3 Treiber-Helbing Filter . . . 73
6.4 Summary . . . 76

7 Experimental Evaluation and Findings **77**
7.1 Sensitivity Analysis of GPS-based Velocity Estimation . . . 77
7.2 Calibration . . . 79
7.2.1 Model-Based Calibration on Stationary Data Only . . . 79
7.2.2 Calibration by Reconstruction and Subsequent Validation . . . 82
7.3 Behaviour of ASM-based Methods . . . 84
7.3.1 Stationary Input-Data Only . . . 84
7.3.2 Floating Car Data . . . 88
7.3.3 Reconstruction with Combined Data-Sources – Data Fusion . . . 91
7.3.4 ASM Performance under the Presence of Noise . . . 94
7.4 Summary . . . 97

8 Conclusion **99**

Bibliography **103**

Appendix A WGS84 Datum Resolution **113**

Appendix B Calibration **115**

Appendix C Global Performance Measures **123**

Appendix D Software Documentation **125**

List of Abbreviations

AIT	Austrian Institute of Technology
ANPR	Automatic number plate recognition
ASFINAG	Autobahnen- und Schnellstraßen-Finanzierungs-Aktiengesellschaft
ASM	Adaptive Smoothing Method
BEV	Bundesamt für Eich- und Vermessungswesen
CFL	Courant–Friedrichs–Lewy
CTM	Cell Transmission Model
EGTF	Extended Generalised Treiber-Helbing filter
EKF	Extended Kalman filter
ELWR	Extended LWR
EnKF	Ensemble Kalman filter
FC	floating car
FFT	fast Fourier transform
FIR	finite-impulse-response
GASM	Generalised Adaptive Smoothing method
GPS	Global Positioning System
HJ	Hamilton-Jacobi
ISV	Institute for Highway Engineering and Transport Planning
KF	Kalman filter
LWR	Lighthill Whitham Richards
MA	moving average
PCHIP	Piecewise Cubic Hermite Interpolating Polynomial
PDE	partial differential equation
pdf	probability density function
PF	Particle filter
PRNG	pseudo-random number generator
SPS	Standard Position Service
TLM	tangent linear model
TT	travel time
UnKF	Unscented Kalman filter
VISSIM	**V**erkehr **I**n **S**tädten **SI**mulations**M**odell
w.r.t.	with respect to

1 Introduction

The Austrian motorway system is equipped with point sensors performing automatic data-collection for various reasons, such as traffic management, statistics, and toll charging. In addition, that data can be applied to various problems, such as the detection of congestion, accidents etc. In general, detailed knowledge about the evolution of the traffic state over time forms the basis for planning any expensive enhancement of the traffic infrastructure. Especially the motorway system close to urban areas is equipped with a dense network of non-moving, i.e. *stationary*, traffic sensors. In accordance to the theory of fluid dynamics, data from such stationary sensors is referenced to as *Eulerian data*. Normally, no information about the traffic stream is available between those stationary traffic sensors. Thus, the traffic status between the individual sensors has to be reconstructed.

In addition to classic stationary sensors more and more data from inside the traffic stream has become available recently. In accordance to the theory of fluid dynamics, data from such moving sensors is referenced to as *Lagrangian data*. This so-called *floating-car data* may be collected by classical odometers, GPS enabled devices or indirectly from evaluation of mobile-phone roaming data. The high resolution trajectories obtained from single vehicles can be used to increase the performance of the aforementioned traffic status estimation methods.

Problem Description and Requirements

It is the objective of this thesis to evaluate different methodologies for traffic state reconstruction on the basis of data originating from both sensor categories, i.e. Eulerian and Lagrangian data-sources. Based on this two data-sources, it is a goal to improve quality and reliability of traffic-status estimation. Finally, the robustness of the methods with respect to noisy or missing data is evaluated.

The following list includes the key points of this thesis:

1. Research and evaluation of existing methodologies for traffic state estimation based on data from stationary detectors.

2. Research of possible methodologies for traffic state estimation based on heterogeneous data sources, such as stationary detectors or GPS equipped vehicles.

3. Development of traffic state estimation algorithms suitable for the Austrian motorway system and the respective implementation in MATLAB.

4. Validation of those developed algorithmic methods based on both, data from microsimulation and on real-world data from the Austrian traffic network.

5. Evaluation of the robustness of developed algorithms, especially the impact of measurement errors, the impact of the spatial and temporal data density, or the absence of individual data points on the algorithm performance.

The real-world data and the microsimulation model are provided by the Institute for Highway Engineering and Transport Planning (ISV). Both, data and microsimulation model originate from research projects conducted by the ISV.

According to the aforementioned key points this thesis is structured among seven further chapters. First, we give an introduction to the fundamentals of traffic engineering, and the modelling of traffic dynamics in Chapters 2 and 3. Second, we introduce different concepts of data-fusion and state-estimation in Chapter 4. Third, we describe the test-site and its respective data-sources in Chapter 5, followed by a discussion of implementation issues in Chapter 6. The conducted experiments and its results are listed in Chapter 7. Finally, the main findings and an outlook on future work are stated in Chapter 8.

2 Traffic Flow Theory

This chapter gives a short introduction to traffic flow theory. Beginning with the definition of categories of traffic data sources, we continue with an enumeration of traffic data belonging to a single vehicle (Sec. 2.1.1). Next, we introduce aggregated or macroscopic data (Sec. 2.1.2), as well as the fundamental relation of traffic flow and its related states (Sec. 2.2).

2.1 Traffic Data

There are several ways to classify data from traffic sensors. One way is to characterise the data of traffic sensors along two dimensions. The first dimension is related to the spatio-temporal characteristics, i.e. if the data represents local traffic quantities (e.g. time headway, flow) or if it represents quantities over space (e.g. journey speed, travel time). The second dimension is related to the degree of aggregation, where the data represents aggregated or averaged quantities (e.g. aggregated flows, averaged speeds). Tab. 2.1 provides an overview those two dimensions of traffic data as well as examples, respectively.

	Single vehicle	Aggregated
Local	Vehicle passage, -speed and -length	Traffic flow, time and harmonic mean speed
Spatial	Vehicle travel time, journey speed, trajectory samples	Space mean speed, mean travel time, mean journey speed

Table 2.1: Classification of data from traffic sensors with respect to their spatio-temporal characteristics and degree of aggregation including some examples.

Another way of classification comes from the field of fluid dynamics: A distinction is made between measurements taken from inside of the vehicle stream (i.e. GPS trajectories of single vehicles) and measurements taken from outside of the vehicle stream (i.e. cross section sensors). These measurements are referred to as *Lagrangian* and *Eulerian* measurements, respectively.

2.1.1 Microscopic Data

Microscopic data or *single vehicle data* are quantities attributed to either one single vehicle or to a single vehicle and its preceding vehicle. The following quantities can be measured by observing of one single vehicle [1]:

v_α velocity of vehicle α,
x_α vehicle position,
d_α front to front distance,
t_α time headway, and
T_α rear to front time headway.

Those quantities can be used to describe the state of a vehicle at a fixed point in time. A set of ordered microscopic data including at least the position and time $\{t_i, x_\alpha(t_i)\}_{i=1}^{N}$, with $t_i < t_{i+1}$, is referred to as *vehicle trajectory*.

2.1.2 Aggregated or Macroscopic Data

Especially with traditional cross section sensors, microscopic data is aggregated respectively over fixed time intervals Δt or a road section S. Aggregation time intervals may vary, common values are between 20 s and 15 min. A common value for Δt in Austria is 60 s[1]. We give a summary of the most common used aggregated quantities in traffic engineering [1, 2] below:

Flow q. The flow q, sometimes also called flux, is the count of vehicles ΔN which pass a road-cross-section at location x during a time interval Δt:

$$q = \frac{\Delta N}{\Delta t} \qquad \text{and} \qquad [q] = \text{Veh}\,\text{s}^{-1}. \tag{2.1}$$

Time mean speed V_t. The time mean speed V_t is the average speed of vehicles $\alpha = 0 \ldots N-1$ passing one location x during a given time interval Δt:

$$V_t = \langle v_\alpha \rangle = \frac{1}{N} \sum_{\alpha=\alpha_0}^{\alpha_{0+N-1}} v_\alpha \qquad \text{and} \qquad [V_t] = \text{m}\,\text{s}^{-1}. \tag{2.2}$$

Space mean speed V_s. The space mean speed V_s is the average speed of all vehicles $\alpha = 0 \ldots N-1$ on a road section S and a specific time t

$$V_s = \frac{1}{\langle \frac{1}{v_\alpha} \rangle} = N \left(\sum_{\alpha=\alpha_0}^{\alpha_{0+\Delta N-1}} \frac{1}{v_\alpha} \right)^{-1} \qquad \text{and} \qquad [V_s] = \text{m}\,\text{s}^{-1}. \tag{2.3}$$

The time mean speed is never less than space mean speed, the relation is given by [3, 4]

$$\text{E}\{V_t\} = \text{E}\{V_s\} + \frac{\sigma_s^2}{\text{E}\{V_s\}}, \tag{2.4}$$

where $\text{E}\{\cdot\}$ is the expectation operator and σ_s^2 denotes the variance of V_s. Methods to estimate the space mean speed V_s from a given time mean speed V_t are summarised in Sec. 5.2.

[1] All stationary sensors in Austria are operated by a single organisation: the semi-governmental highway operator ASFINAG.

Station speed. The station speed is the aggregation of per-lane speeds to a speed per cross section. It is generally defined as the weighted average of lane speeds [5]

$$v(t) = \frac{\sum_{l=1}^{L} (q_l(t) \cdot v_l(t))}{\sum_{j=1}^{L} q_j(t)}, \tag{2.5}$$

where L is the number of lanes, $q_l(t)$ and $v_l(t)$ are the flow and speed at lane l of a detector station, respectively.

Density k. The density k (or concentration), sometimes also denoted as ρ, within a road section S at a given time t is the inverse of the average spacing of the $\alpha = 0 \dots N-1$ vehicles:

$$k = \frac{N}{\sum\limits_{\alpha} d_\alpha} = \frac{q}{V_s} \qquad \text{and} \qquad [k] = \text{Veh}\,\text{m}^{-1}. \tag{2.6}$$

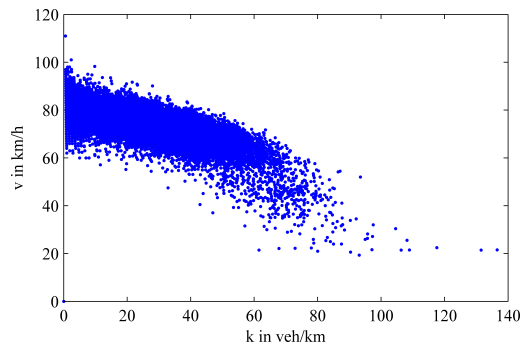
These primary quantities of traffic flow are sufficient to model the complex behaviour of the traffic stream. Before giving an overview of that complex behaviour in Chapter 3, an introduction into the *fundamental relation of traffic flow* will be provided.

2.2 The Fundamental Relation of Traffic Flow

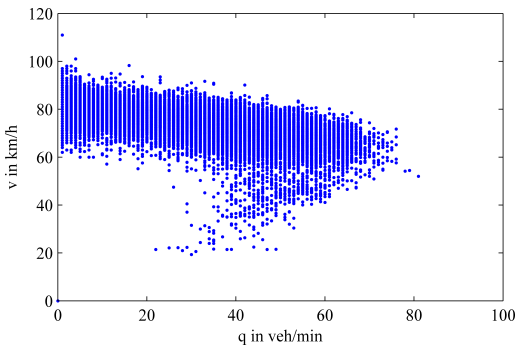
The relationship between the three fundamental variables of traffic flow, namely the mean speed v, flow q, and density k is called the *fundamental relation of traffic flow*

$$q = k \cdot v. \tag{2.7}$$

Please note that v is the space mean speed V_s. This relation can be observed when the traffic stream is stationary, i.e. the flow rates are constant along space and time. When the traffic is not stationary, a different behaviour may be observed. Fig. 2.1 shows exemplary data of one-minute aggregated speed-flow and speed-density combinations obtained by a stationary traffic sensor. An evident relationship between k and v and as a consequence between v and q is visible to the naked eye.



(a) Density-speed relation



(b) Flow-speed relation

Figure 2.1: Scatter plots of one-minute aggregated data from June 04, 2012 until June 27, 2012 at kilometre $x = 230\,\text{m}$ of the Austrian motorway A4, driving direction east. A higher aggregation-interval may yield a different distribution of the data.

A natural way to model the density-speed or the flow-speed relation is to fit a parametric

function to the scattered data. These fitted parametric functions are the so called *equilibrium equations*:

$$v(x,t) = \mathrm{V_e}(k(x,t)) \tag{2.8}$$
$$q(x,t) = \mathrm{Q_e}(k(x,t)), \tag{2.9}$$

and describe a statistical relationship between the three fundamental traffic quantities in addition to eq. (2.7). The equilibrium equations can be transformed by using the fundamental relation in eq. (2.7). A number of different parametric functions, alias models, were suggested in the past. In [6], Regler provides an overview of different equilibrium equations that have been proposed. The simplest model is to fit a straight line to the density-speed relation.

All three relations can be combined to the so called *fundamental diagram* as illustrated in Fig. 2.2. The fundamental diagram or the equilibrium equations, respectively, are a basic building block for almost all state of the art traffic modelling and estimation methods as described in Chapters 3 and 4. The three extrema in the fundamental diagram are related to special states of

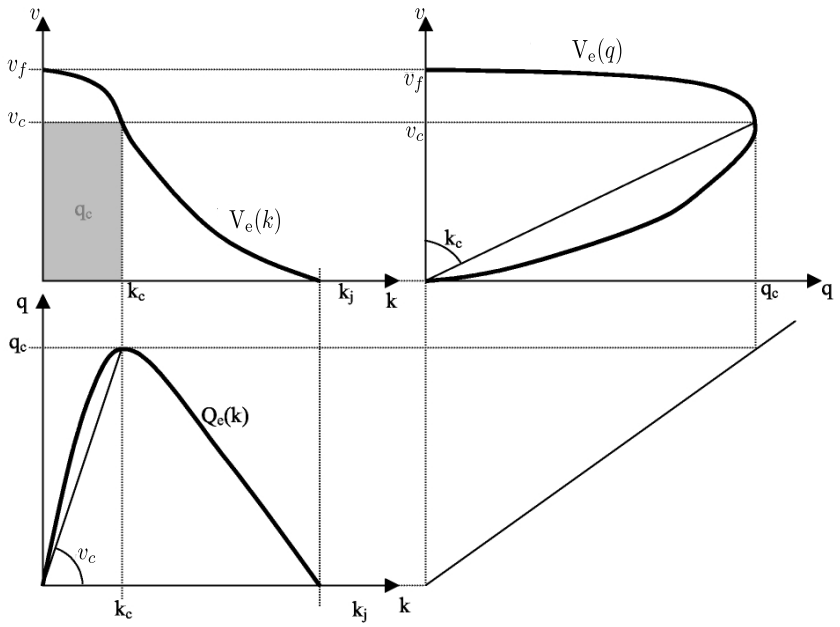


Figure 2.2: The fundamental diagrams of traffic flow, including free flow speed v_f; density k_c, flow q_c and speed v_c at capacity, and the jam density k_j (Fig. adapted from [7]).

traffic flow. These special states are referred to as *free flow*, *synchronised flow* and *wide moving jam* [8]:

Free flow. At free flow, every driver can choose his velocity freely. Traffic is characterised by highest reachable speeds v_f, densities are comparably low.

Capacity. At capacity, flow rates reach their maximum q_c. Traffic at capacity is normally in synchronised flow with the density k_c. Beyond densities of k_c, both flow and speed decrease. When crossing the capacity limit a step-like decrease of capacity can be observed. This phenomenon, not shown in Fig. 2.2, is called *capacity drop* [9].

Jam. Near the jam point, both flow and speed approach zero. This point is characterised by the greatest possible density, the so called jam density k_j.

While free flow conditions might provide the highest comfort for the individual driver, a utilisation of a motorway near capacity offers economic and environmental advantages. Since the demand of capacity can vary over a wide range over time it is rather difficult to design a road with sufficient capacity for every situation, and consequently traffic jams can occur. Below we describe different elementary patters of congested traffic that can be observed.

2.3 Elementary Patterns of Congested Traffic

Congested traffic can be seen as the traffic that has a highest impact on individuals, society, and the economy. Congestion can occur in various patterns with different levels of complexity. Even complex traffic breakdowns are composed of elementary patterns [10, 11]. Those patterns differ in spatial and temporal extent, and the homogeneity of the congested field. All of those six patterns are listed below:

- *Pinned Localized Cluster (PLC)*: a region with congested speeds at a fixed location over a longer period of time.
- *Moving Localized Cluster (MLC)* propagate upstream with the characteristic speed v_c.
- *Oscillating Congested Traffic (OCT)* is just characterized by oscillating speeds in the congested range, i.e. unstable traffic flows.
- *Triggered Stop and Go Pattern (TSG)* may be viewed as special case of OCT but with high velocity traffic flows between the upstream propagating jams.
- *Homogeneous Synchronized Traffic (HST)* synchronised traffic between the states of free flow and congestion over a spatially extended area.
- *Homogeneous Congested Traffic (HCT)* congested speeds without oscillations over a spatially extended area.

An illustration of the described congestion patterns is shown in Fig. 2.3. It is obvious, that these congestion patterns differ in homogeneity and in spatial-temporal extend. Accordingly, these patterns have a different impact on individual vehicles, e.g. the total travel time. Stationary sensors and probe vehicles are distributed over a road stretch sparsely, and therefore can gather data from these patterns only partly. As a consequence spatio-temporal traffic estimation methods, as described later in this work, may perform differently on traffic data originating from these six different elementary patterns.

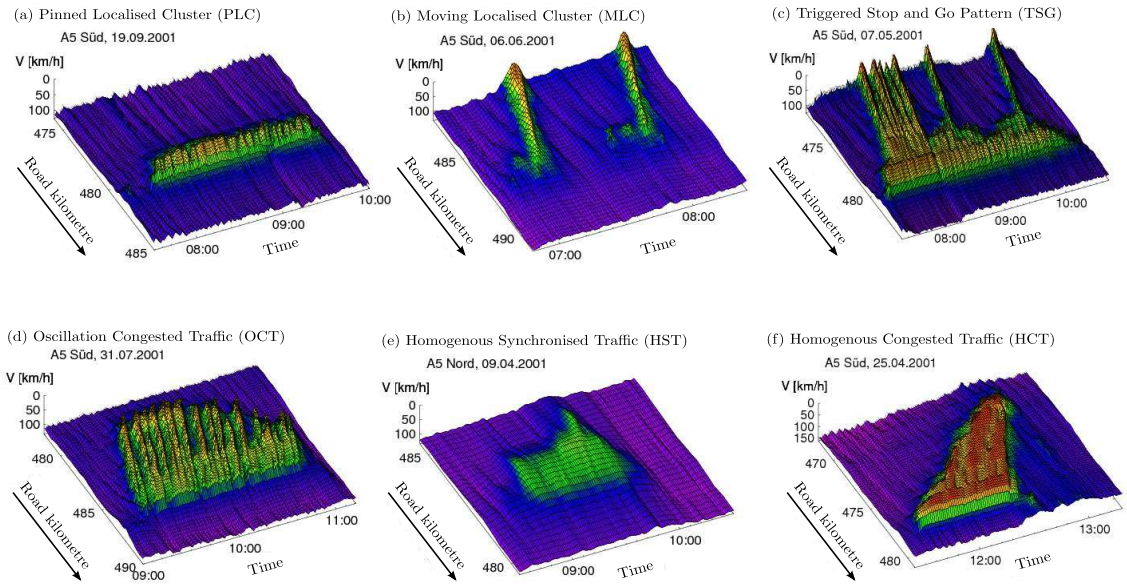


Figure 2.3: Illustration of the elementary patterns of congested traffic. Even complex traffic breakdowns are composed of these six elementary congestion patterns (Fig. from [11]).

2.4 Summary

In this chapter we introduced the basic traffic data, including base-data belonging to single vehicle as well as data belonging to the overall traffic stream. We mentioned fundamental relation of traffic flow, the relation between the fundamental quantities of traffic flow prevailing under stationary traffic conditions. Furthermore, we stated the equilibrium equations, and their combination with the fundamental relation to the so called fundamental diagram. Finally, the elementary patterns of congested traffic have been described. In the next chapter we give an overview about the modelling of the complex dynamics of traffic flow.

3 Dynamic Traffic Flow Models

One major question in traffic engineering is the dynamic change of traffic state over space and time. Unfortunately, the fundamental relation in Sec. 2.2 does not provide a mechanism to analyse how such traffic states evolve. In this chapter, we introduce dynamic models for describing the dynamics of traffic flow.

Research on the modelling of traffic flow started in the 1950s. Lighthill and Whitham presented a model for vehicular flow [12] based on the analogy of the traffic flow with the flow of a fluid. Since the publication of their traffic model a lot of different traffic models have been proposed. Depending on the research objective, different types of mathematical models are used. In [13], Hoogendoorn and Bovy give an overview of the multitude of proposed models and state variables in the period between the years 1955 and 2001.

In the next chapter, we will first introduce different categories of traffic flow models (Sec. 3.1), followed the first-order conservation principle (Sec. 3.2.1) and the resulting macroscopic model (Sec. 3.2.2) as well as an alternative formulation based on cumulative flows (Sec. 3.2.3) and its respective solution methods in continuous domain. Finally, mention a popular scheme for the solution in the discrete domain (Sec. 3.3).

3.1 Categorisation of Traffic Flow Models

Traffic flow models may be categorised on the basis of different criteria. Those criteria are roughly counted among one of the following points:

- Detail or scale of the process model,
- mathematical, and
- conceptional criteria.

The aforementioned criteria are elaborated in more detail in the following.

Scale of the process model. The modelling of traffic flow can be performed on different scales. The subsequent enumeration ranges from coarse to a fine modelling scale:

Picoscopic: The same facts as for microscopic models apply here. Further, those models contain a description of vehicle sub-units, i.e. the shape of the vehicle, the engine, or the

brakes. Such models may be useful for cooperative vehicle systems and automotive safety technologies.

Microscopic: Every individual vehicle is modelled including its detailed behaviour. This includes driver behaviour, vehicle class, acceleration, slowdown and lane-change of single vehicles. Such models should be used if one wants to investigate the interaction between individual vehicles or heterogeneous vehicle classes, or the influence of single vehicles on the traffic stream.

Mesoscopic: Almost the same aspects as for the macroscopic models can be applied here. Only the concentration of vehicles is described with statistical distributions. In addition, such models include parameters for the behaviour individual vehicles, e.g. drivers expect density changes when adjusting their speed.

Macroscopic: The traffic stream is modelled like a liquid flowing through a pipe. The models incorporate local aggregated traffic quantities, such as speed, flow, and density. Consequently, no information of individual vehicles can be obtained from the model, i.e. lane-change behaviour. Macroscopic models are computationally efficient, even with large road networks. It is easy to incorporate various heterogeneous data sources.

Fig. 3.1 gives an overview of the mentioned modelling scales, underlying physical principles and associated state variables.

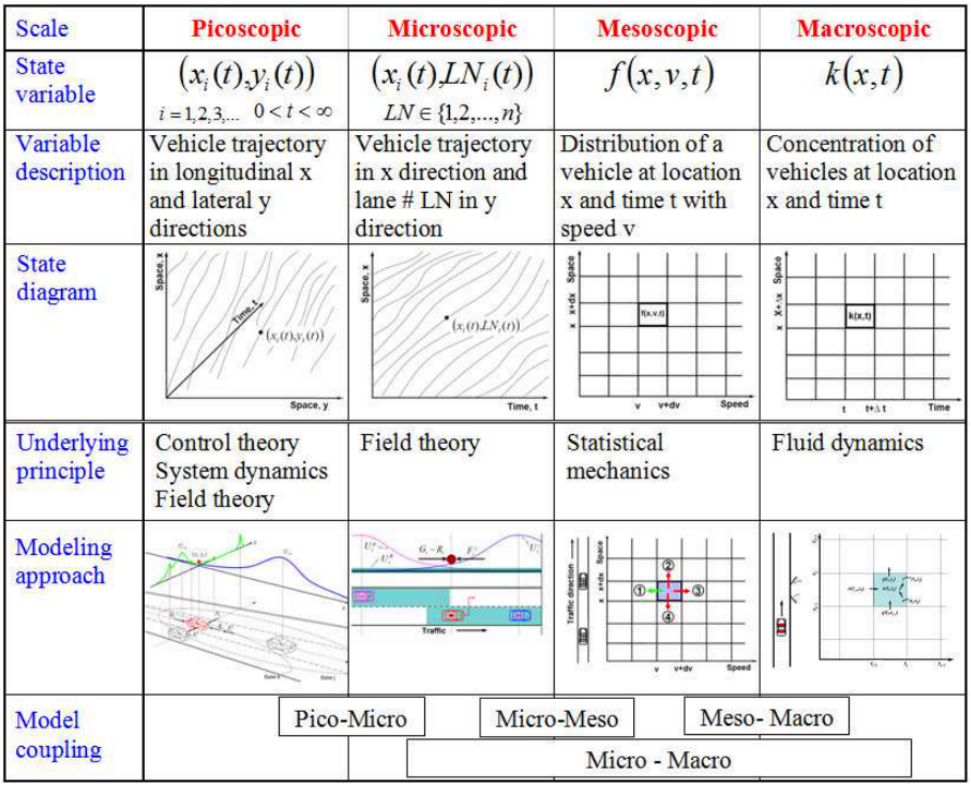


Figure 3.1: Modelling scales of traffic-flow models (Figure from [14]).

Usually the scale of application correlates with the scale of the process model. Models may be used to describe the dynamics of a single roadway stretch, a mixed vehicle class roadway junction, a suburban arterial road network, a single highway or even a high-level transport network.

Recently, several compositional modelling approaches have been proposed [14–16]. They combine the advantages of the different scales of the process model. The dynamics of a whole road network is described with a macroscopic model for example. Additionally some points of special

interest, e.g. highway intersections, are described with a microscopic model. Both models share their process state using data-fusion methods.

Mathematical criteria. The modelling of traffic can be classified according to mathematical criteria. Most models of dynamic systems have a differential equation as key element. Models can be classified according to the form of their differential equation.

In fluid dynamics, two conceptionally different approaches of process modelling exist. One is to model the stream with a view from outside the stream. The other is to model the stream from the view from inside the stream. These two different representations of the flow field coordinate system are referred to as *Eulerian* and *Lagrangian* modelling, respectively. On the basis of computational benefits, the modelling in Lagrangian coordinates have recently become more popular (see Sec. 3.4).

Models typically describe the evolution of single-vehicle-quantities or aggregated quantities over space and time. That evolution can be modelled *continuous* or *discrete*. Mixed models have been proposed as well, i.e. one state variable is continuous, the other is discrete. This setting is so called *semi-discrete*.

Models can be built with exclusively exact relationships or with the influence of some random processes. These different approaches are named as *deterministic* and *stochastic* representation of the process, respectively.

Conceptional criteria. For problem solution two conceptionally different approaches exist.

Heuristics: Heuristics are based on empirical knowledge. Relationships between different quantities are modelled by fitting parametric functions to empirical data. In general, these models do not rely on a physical law.

First principles modelling: In a first principle model traffic stream quantities are calculated directly from established laws of physics. In contrast to Heuristics these models are based on no assumptions such as empirical or fitted parameters.

Even though there is a wide variety of criteria for classifying traffic models, the above mentioned criteria are the most important ones for our work. Our research objective is to examine different methods for the fusion of heterogeneous data from motorways, acquired by static cross-section sensors and floating cars. The behaviour of single vehicles is not of interest; We solely want to reconstruct the macroscopic state of the motorway. In recognition of our requirements and the the different categories of models described above, the use of a macroscopic traffic model seems to be obvious. In the following section we provide more information about these macroscopic models.

3.2 Continous Macroscopic Models

As mentioned above, macroscopic models are used to describe the dynamics of traffic flow on the level of aggregated quantities such as flow, speed, and density. The traffic stream is modelled as a one dimensional, compressible fluid. The dynamics of the stream can be characterised with a partial difference equation. This section focuses on the derivation of first-order dynamic equations for traffic flow, followed by an introduction of the strongly related family of *LWR (Lighthill Whitham Richards) models*, and its simplifications with a triangular fundamental diagram. Finally we discuss possible solution methods in the continuous domain.

3.2.1 The First-Order Conservation Equation

The physical modelling of a compressible fluid is based on conservation laws. Some basic quantities of traffic flow are locally conserved: They can neither be created or destroyed, nor are they able to "jump" from one position to an other. The number of conserved quantities is related to the order of the resulting model, i.e. mass conservation results in first-order models, mass & momentum conservation results in second-order models. A multitude of first-order and second order models has been proposed [13], but in the following we focus on first-order models only. The use of second order models has been criticised in the past [17, 18], as they may produce unrealistic results, i.e. reversed traffic flow. For further information on second-order models we refer to [1, 2, 13].

Mass conservation implies that the mass of vehicles is conserved, i.e. all vehicles that enter a road section will leave the road section again and no vehicles will be stored or created. This criterion is fulfilled in general for high rank roads. The conservation equation, also called continuity equation, can be easily derived by investigating a small volume of roadway (Fig. 3.2) as a traffic continuum. Because of its fundamental nature, we summarise the derivation of the

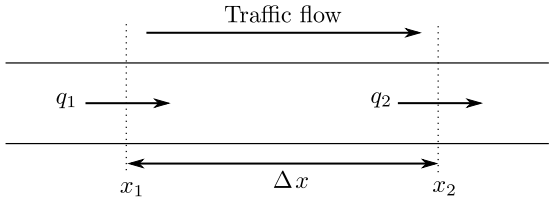


Figure 3.2: Small volume of roadway traffic to illustrate the mass-conservation equation.

conservation equation from [19]: The flow rates at the section boundary are

$$q_1 = \frac{N_1}{\Delta t} \qquad \text{and} \qquad q_2 = \frac{N_2}{\Delta t}, \tag{3.1}$$

the change of the vehicles in the section is

$$\Delta N = N_2 - N_1 = (q_2 - q_1)\Delta t = \Delta q \Delta t. \tag{3.2}$$

Assume, that the traffic densities in the section at t_1 and t_2 are k_1 and k_2, respectively. There are $M = k\Delta x$ vehicles in the section. Thus, the change of vehicles in the section can be expressed as

$$\Delta M = k_1 \Delta x - k_2 \Delta x = (k_1 - k2)\Delta x = -\Delta k \Delta x. \tag{3.3}$$

The number of vehicles in a section is conserved during the same time interval, therefore $\Delta N = \Delta M$, i.e.

$$\Delta q \Delta t = -\Delta k \Delta x \tag{3.4}$$

$$\Delta q \Delta t + \Delta k \Delta x = 0 \tag{3.5}$$

$$\frac{\Delta q}{\Delta x} + \frac{\Delta k}{\Delta t} = 0. \tag{3.6}$$

Let $\Delta x \to 0$ and $\Delta t \to 0$, then the above difference equation becomes a *partial differential equation* (PDE) :

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \qquad \text{or} \qquad \frac{\partial k}{\partial t} + \frac{\partial (kV)}{\partial x} = 0. \tag{3.7}$$

This equation models the flow field from the point of view of stationary detectors. This is referred to as model in *Eulerian* formulation. Note that eq. (3.7) holds only for a homogeneous road section, if there are changes in the profile, sinks or sources, additional terms have to be added [1].

A model from the point of a moving observer is called *Lagrangian* formulation. Equation (3.7) in Lagrangian formulation can be written as

$$\frac{\mathrm{d}k}{\mathrm{d}t} = \frac{\partial k}{\partial t} + V\frac{\partial k}{\partial x} = -k\frac{\partial V}{\partial x}, \tag{3.8}$$

which leads to a very intuitive interpretation of the solution: If the vehicles further ahead reduce their speed, the density will rise, and vice versa. Equation (3.8) can be used to solve any differentiable flow field.

This essential difference between the Eulerian and Lagrangian formulation has a large impact on the solution of the continuity equation. In the Eulerian formulation, information can propagate in both spatial directions of the flow field. When using numerical solution methods one has to consider information from both directions: (1) the maximum admissible flow in driving direction, and (2) the provided vehicle flow from upstream direction. On the other hand, with the Lagrangian formulation, the solution only depends on the traffic state in front of the observer, the traffic state in the rear has no influence on the solution. Although the Lagrangian formulation is less intuitive than the Eulerian formulation, a solution in the Lagrangian formulation can done with less computational effort.

3.2.2 First Order Model or Lighthill, Whitham and Richards Model

Up to this point eq. (3.7) is under-determined. Lighthill, Whitham and Richards formulated the problem completely using the fundamental relation of traffic flow. The combination of eqs. (2.7), (2.9) and (3.7) gives the well known *LWR (Lighthill Whitham Richards) model*, which can be written as [12]

$$\frac{\partial k}{\partial t} + c(k)\frac{\partial k}{\partial x} = 0, \tag{3.9}$$

with the *characteristic speed* given by the gradient of the equilibrium equation

$$c(k) = \frac{\mathrm{d}Q_e(k)}{\mathrm{d}k}. \tag{3.10}$$

The characteristic speed represents the velocity at which continuous density fluctuations propagate through the traffic stream depending on the instantaneous local density[2]. Equation (3.9) belongs to the class of non linear wave equations, which implies another characteristic of the traffic stream: In situations when $\frac{\partial k}{\partial x}$ becomes infinite, the local density follows a discontinuous change resulting in so-called *shock waves*. These discontinuities travel with the shock wave speed through the flow field (Fig. 3.3). For sufficiently large changes in density, the speed of a shock front $x_{12}(t)$ can be expressed by[3]

$$c_{12} = \frac{\mathrm{d}x_{12}(t)}{\mathrm{d}t} = \frac{q_2 - q_1}{k_2 - k_1}. \tag{3.11}$$

The LWR model is sufficient to describe traffic flow behaviour for a range of different traffic situations. Various different fundamental diagrams have been proposed for the application of

[2] Also known as *dispersion*.
[3] Also known as the *Rankine–Hugoniot jump relation* in physics.

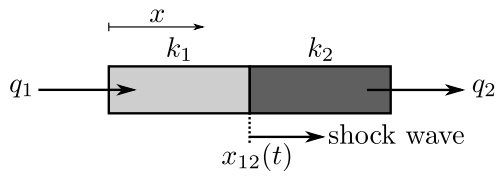


Figure 3.3: Illustration of a shock wave: discontinuities in density k and flow q are travelling along the road with the speed $c_{12} = \frac{\mathrm{d}x_{12}(t)}{\mathrm{d}t}$ (Fig. from [1]).

the LWR model. In general, the solution of the hyperbolic non linear wave equation is difficult. As a consequence, multiple simplifications have been proposed with the one described in the following being of great importance.

Triangular Fundamental Diagram. In [20–22], Newell discusses different drawbacks of the LWR model. One point concerns the selection of the particular equilibrium equation: (1) in general the LWR model is not able to describe the traffic flow in detail, and (2) especially for queueing applications the exact shape of the equilibrium equation is not important as long as it reflects the three characteristic states of traffic flow, i.e. free flow, capacity, and wide-moving jam. As a consequence Newell proposes the simplest possible flow-density relation, defined by

$$Q_e(k) = \begin{cases} c_{\text{free}} \cdot k & k \leq \frac{q_{\max}}{c_{\text{free}}} \quad \text{(free flow)} \\ \frac{(k_j - k) \cdot q_{\max} \cdot c_{\text{free}}}{k_{\text{j}} c_{\text{free}} - q_{\max}} & \text{otherwise} \quad \text{(congested flow)}, \end{cases} \tag{3.12}$$

where overall equation is defined by the free-flow velocity c_{free}, the capacity $q_{\max}$, and the jam-density k_j. This simple flow-density relationship results in the two characteristic velocities

$$c(k) = \frac{\partial}{\partial k} Q_e(k) = \begin{cases} c_{\text{free}} & k \leq \frac{q_{\max}}{v_0} \quad \text{(free flow)} \\ c_{\text{cong}} & \text{otherwise} \quad \text{(congested flow)}. \end{cases} \tag{3.13}$$

Traffic flow is simplified to two dominant regimes, *free flow* and *congested flow*, respectively as illustrated in Fig. 3.4. Thus, shock waves can only traverse with two constant speeds c_{free}

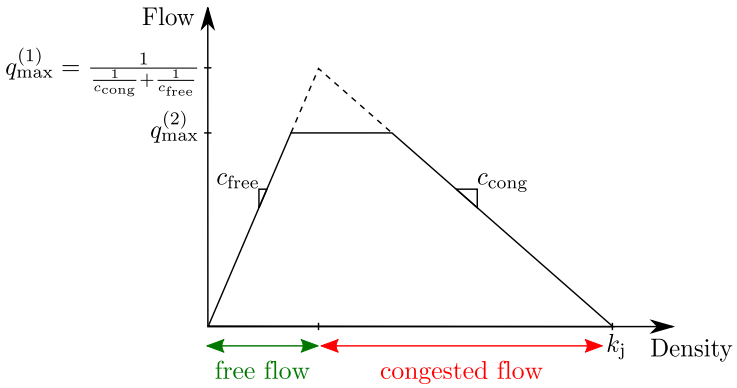


Figure 3.4: Triangular and trapezoidal flow-density diagram. Both diagrams share the constant shock wave speeds c_{free} and c_{cong}, but differ in the maximum flow $q^{(1)}_{\max}$ and $q^{(2)}_{\max}$, for the triangular and the trapezoidal fundamental diagram, respectively.

and c_{cong}. Furthermore the constant propagation speeds prevent dispersion. Even if the model seems to be simple, it has enough descriptive power to sufficiently describe the traffic of high-rank roadways.

Solution methods. The analytic solution of an LWR model requires well-funded knowledge in the field of analysis and differential equations and can be summarised as an initial value problem of the non linear hyperbolic equation, known as *Riemann problem*. This problem can be solved by the application of weak conditions, which results in an infinite number of possible solutions. Weak solutions cover those with discontinuities, as well as continuous solutions. Although there is an infinite number of possible solutions, there is only one with physical plausibility, the so-called *entropy solution*. That plausibility can be ensured by restricting the possible shock wave speeds (see eq. (3.11)) to a maximum rate. For further information about the entropy solution we refer to [23, 24].

Unfortunately, the first proposed entropy solution does not hold for bounded domains; The stated method is only possible for infinite long highways with no on-ramps or off-ramps. As a consequence, when solving the LWR PDE in continuous formulation and numerical schemes, they might provide meaningless solutions. In [25], Strub and Bayen introduce the weak formulation of boundary conditions, for which the initial-boundary value problem of the LWR PDE is well posed.

The presence of shock waves, and therefore discontinuities in the traffic flow (see Fig. 3.3), makes the direct solution of eq. (3.9) difficult. In order to obtain the right solution it is necessary to follow the path wave front $x_{12}(t)$ (i.e. wave-front tracking), which may be computationally elaborative. A possible remedy is to formulate eq. (3.9) on the basis of cumulative flows as described in the following section.

3.2.3 First Order Model based on Cumulative Flows / Hamilton-Jacobi Equation

In [20–22], Newell presents an alternative formulation of the LWR model based on the *cumulative flow* $N(x,t)$, which represents the cumulative number of vehicles passing a location x by time t starting from the passage of a reference vehicle. Accordingly, the fundamental variable can be written as

$$k(x,t) = \frac{-\partial N(x,t)}{\partial x} \qquad \text{and} \qquad q(x,t) = \frac{\partial N(x,t)}{\partial t}, \tag{3.14}$$

and the identity

$$\frac{\partial^2 N(x,t)}{\partial x \partial t} = \frac{\partial^2 N(x,t)}{\partial t \partial x} \tag{3.15}$$

is equivalent to eq. (3.9). The dynamic equation formulated with cumulative flows is referenced to as *Hamilton-Jacobi (HJ) PDE*.

One main advantage of formulating the dynamic equation on the basis of $N(x,t)$ is that it guarantees the conservation of vehicles. A shock is represented by a discontinuity in the first derivative of $N(x,t)$, but in opposition to eq. (3.9) the conservation equation is valid across the shock wave front $x_{12}(t)$. Thus, for a constant flow-density relation and known boundary values $N(x_0,t_0)$ and $q(x_0,t_0)$ it is possible to calculate $N(x,t)$ at all points along the characteristic curve through (x_0,t_0). For multi-valued solutions, Newell proposes to take the minimum-surface of $N(x,t)$ as the plausible solution. This obtained solution agrees with the weak entropy solution mentioned earlier.

Solution based on Variational Formulation

In [26,27], Daganzo completes Newell's method, where the minimum-surface solution is ensured by a *variational formulation*. The solution of the kinematic wave problem is represented by a set of continuum least cost paths in space and time. Daganzo's method minimises the cost to

reach a point, based on the vehicle count as a cost metric. Based on the dual formulation of his approach he is able to prove a relation between microscopic and macroscopic traffic flow models [28]. Their method is able to incorporate time-dependent restrictions: For example signalling or moving restrictions such as slow buses or in general time varying fundamental diagrams can be incorporated straightforwardly by introduction of a space or time dependent cost. In this case, the solution of the variational formulation can be obtained by dynamic programming and is not exact in general but for many cases.

In general, the variational theory allows elegant solutions for global optimisation problems [29]. It finds a widespread application in mathematics and engineering, where the optimal solution to a mathematical problem is the minimum of its value, which may be written as

$$\mathbf{x}_{\text{opt}} = \arg\min_{\mathbf{x}} \mathcal{J}(\mathbf{x}), \tag{3.16}$$

where $\mathcal{J}(\mathbf{x})$ is the problem formulation, i.e. the cost-function, which has to be minimised. The minimisation can be realised numerically using gradient descent, or if certain conditions are fulfilled, in explicit formulation. Besides the solution of the dynamic equations, variational methods have been applied to optimal control problems in the field of traffic engineering. We mention further examples in Sec. 4.3.

Lax-Hopf Algorithm

In [30, 31], Claudel and Bayen propose a method for solving the HJ PDE semi-analytically for arbitrary concave fundamental diagrams. Their method, based on *viability theory*, is able to incorporate piecewise-constant initial and boundary conditions, as well as floating car data formulated as internal conditions (see Fig. 3.5). The computational principle of the algorithm

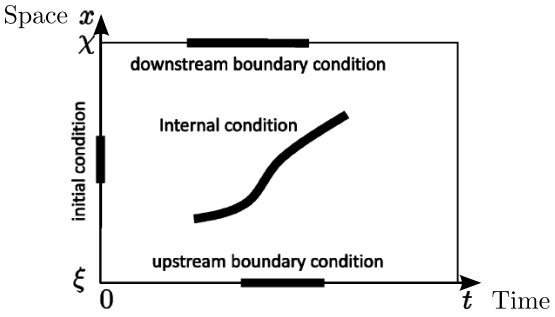


Figure 3.5: Illustration of possible conditions, i.e. sensor data, for the the solution of the Lax-Hopf algorithm (Fig. from [30]).

is based on the minimisation of closed form partial solutions. The solutions can be obtained at a computational cost proportional to the count of initial and boundary conditions.

In [32], Mazaré et al. discuss the method in more detail. Further, they point out the differences to the variational method by Daganzo [27] and propose a faster algorithm for triangular fundamental diagrams. An implementation for MATLAB is available for download on the *Mobile Millennium* project-website [33]. However, it does not support the incorporation of internal boundary conditions.

The main advantage over other solution methods is the available closed-form solution which provides exact results for general concave fundamental diagrams. In general, closed form solutions can be assumed to be many times faster compared to iterative or numerical solution schemes. It is possible to incorporate measurements from heterogeneous data sources, but the method ignores the different measurement errors. Further disadvantages are the inability to use time-dependent fundamental diagrams, and no exact results for road-networks.

The solution methods of both formulations of the dynamic equation require a strong attention on the valid formulation of the dynamic equations and its respective boundary conditions. A simpler approach is to discretise the differential equation followed by a numerical solution technique. All methods based on this principle are usually summarised under the term *discrete modelling*. In Sec. 3.3 we give an overview about the basic principles of discrete models.

3.3 Discrete Macroscopic Models

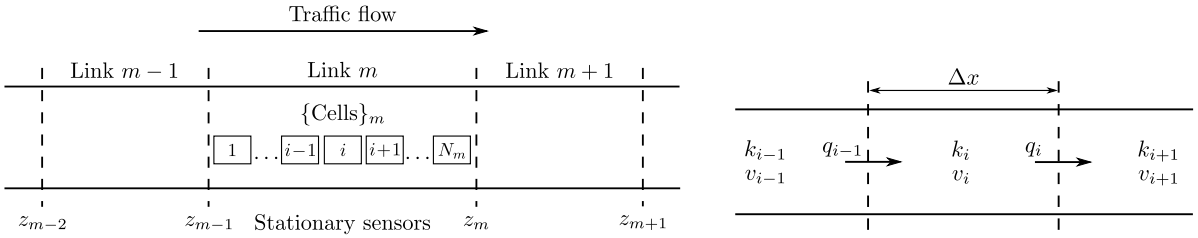
As already previously mentioned, the solution of the dynamic equation in continuous form is a demanding problem. In this section we describe principles of discretisation of the dynamic equation and the related solution methods. All of these methods discretise the derivatives in eq. (3.7) followed by a numerical integration. On basis of the Taylor expansion it is possible to analyse the error residual introduced by the finite differences. Below we summarise the well known cell transmission model based on first order discretisation scheme followed by a summary of higher order discretisation schemes.

First Order Discretisation Schemes

In [34, 35], Daganzo describes the *Cell Transmission Model (CTM)*, namely the first approach to solve the LWR model with a triangular fundamental diagram based on *Godunov's method*. *Godunov's method* or *Godunov's scheme* is a numerical scheme for solving partial difference equations [36], based on the idea that the solution domain is divided into finite volumes, i.e. cells, where the Riemann problem is solved at each inter-cell boundary. Daganzo's approach is divided into two steps:

1. Discretisation and solution with the *first-order up-winded finite volume scheme*, and
2. incorporation of weak boundary conditions with the *minimum supply-demand scheme*.

First-order up-winded finite volume scheme. The main idea of the upwind scheme is to divide a road stretch into *cells* with constant length Δx and a *piece-wise* constant distribution of speed v_i and density k_i as illustrated in Fig. 3.6. The dynamic equation is approximated by a



(a) Links (b) Cells

Figure 3.6: A discrete traffic flow model composed of motorway links with stationary sensors on the boundaries as illustrated in Fig. (a). A link is composed of a sequence of cells with equal length and associated quantities of traffic flow as illustrated in Fig. (b).

first-order forward difference and evolved over time with uniform steps $t \to t + \Delta t$, which yields following state-space equation

$$\begin{cases} \frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \\ q = kv \\ q(k) = Q_e(k) \end{cases} \implies k_i^{t+\Delta t} = k_i^t + r\left(q_{i+1}^t - q_i^t\right) \qquad i = 1 \dots N_m \tag{3.17}$$

with the step-size

$$r = a\frac{\Delta t}{\Delta x}. \tag{3.18}$$

The step-size r must satisfy the *Courant–Friedrichs–Lewy (CFL) condition*

$$|r| = \left| a\frac{\Delta t}{\Delta x} \right| \leq 1 \qquad \text{with} \qquad a = c_{\max}, \tag{3.19}$$

where $c_{\max}$ represents the largest velocity at which information can propagate through the system. In other words, the physical domain must be contained in the numerical domain and information has to pass every cell without skipping an intermediate cell. By implication this yields the practical restriction: both, spatial and temporal resolution have to be selected with respect to the each other.

Minimum supply-demand scheme. Moreover, information can travel in both directions, which involves that the flow between two cells is limited to the minimum of the supply and the demand beside upstream and downstream the cell boundary, respectively. Loosely speaking, the supply is given by the number of vehicles leaving cell $i-1$ and the demand is given by the the amount of empty space in cell i.

The Godunov scheme is just first-order accurate, which implies that the error, given by the remainder of the Tailor expansion, is bounded linearly by $\mathcal{O}(r)$. This may introduce high numerical diffusion, especially at positions with strong variations of the state variable.

Higher Order Discretisation Schemes

For second-order models, Godunov's method results in considerable errors. In opposition, the *MacCormack method* is a discretisation scheme based on second order differences [1]. The algorithm comprises of two steps: A predictor step which is the same as with Godunov's method followed by a corrector step. Unlike Godunov's method, the MacCormack method does not introduce diffusion errors.

Summary

In short, the Godunov scheme is suitable for the discretisation of first-order dynamic equations. Since Daganzo proposed the CTM, solution methods in discrete domain gained increasing popularity. For higher order equations, it is necessary to approximate the derivative by higher order discrete differences. Other numeric Riemann-solvers can be found in mathematical literature, but up to now they don not have a broad usage in traffic engineering.

3.4 Modelling in Lagrangian Coordinates

In the theory of fluid dynamics there are two fundamental different ways of formulating the dynamic equation, either with *Eulerian* or with *Lagrangian* coordinates. Traditionally, the kinematic wave model in traffic engineering is formulated in Eulerian coordinates, i.e. from the view of a stationary observer. Recently, formulations of the kinematic wave model in Lagrangian coordinates are receiving increasing attention as they simplify the respective solution methods [37–40].

Information in the traffic stream, i.e. waves, never travels faster than the traffic itself. In the Lagrangian formulation, the coordinate system travels with the same velocity as the traffic. As a consequence, information can only propagate in the opposite direction of the traffic stream. Or formulated from the view of the driver, it is only necessary to respond to the traffic conditions in front of the vehicle. This fact leads to various simplified solution methods. E.g. for the Godunov scheme, the minimum supply-demand step can be omitted and the method reduces to an upwind scheme. Further examples of models based on Lagrangian formulation can be found in Sec. 4.6.

3.5 Summary

This chapter covered a short overview of possible modelling principles for the description of traffic dynamics. Further, we have put the focus on macroscopic traffic flow models, as they are suitable for the incorporation of data from heterogeneous sources. We explained the law of vehicle conservation (mass conservation), which can be seen as the theoretical foundation for all first-order macroscopic traffic-flow models. Subsequently, we introduced the continuous, first-order LWR flow model, as well as the Hamilton-Jacobi equation, an alternative first-order formulation based on cumulative flows. For both formulations we outlined possible solution methods in the continuous domain. Finally, we mentioned a discretisation strategy using the example of the CTM and the underlying discretisation scheme after Godunov.

Although we discussed various forms and respective solution methods for the dynamic equation, there still remains an open issue: The mentioned modelling principles do not reflect the real world exactly, and all measurements are comprised with certain errors. The following chapter deals with feasible approaches on how to incorporate imperfect observations to imperfect process models as good as possible.

4 Traffic State Estimation

The term *state estimation* refers to the estimation of a state of a physical process based on given observations and assumptions about the process. There are different views on the estimation problem, considering control theory, estimation theory, variational analysis, and probability theory. Fig. 4.1 illustrates an exemplary state estimation problem of traffic engineering: For given floating car observations $\{v(x_j(t_i), t_i)\}$ a spatio-temporal velocity field $v(x,t)$ should be estimated. An alternative view of state estimation is as a *data assimilation* problem [42]: how

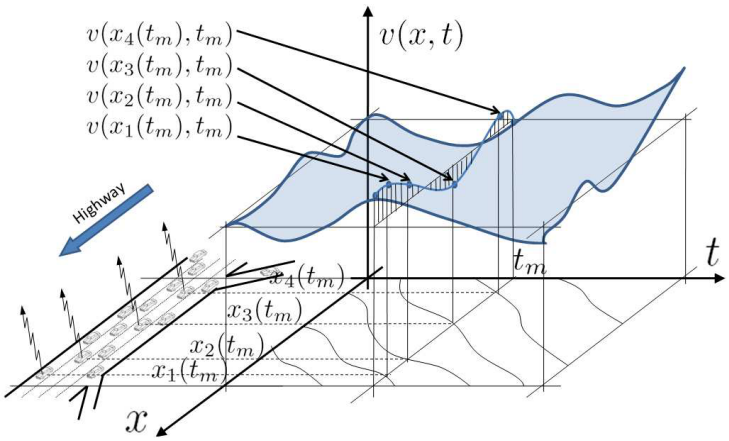


Figure 4.1: Illustration of the distributed velocity field $v(x,t)$ to be reconstructed from Lagrangian samples (Fig. from [41]).

to *optimally* incorporate observations of a process into a given physical model. All process observations are estimates up to a certain point. Measured data from sensors is comprised of additional noise. Some equations in the model are approximated, e.g. the equilibrium equations, and so introduce additional noise. Moreover, the process model does not cover all factors, which introduces additional process noise.

Data assimilation can be achieved with different approaches. The most widely used techniques in engineering and science are based on

- heuristics,
- Newtonian relaxation,
- Kalman filter variants, and

- variational methods / optimal interpolation.

All these concepts require a transformation of the measured quantity $\mathbf{y}_k$ to the quantity representing the process state $\mathbf{x}_k$ by inversion of the measurement equation with the general form

$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{v}_k) + \vartheta_k, \tag{4.1}$$

where ϑ_k denotes the measurement noise. The equation $h(\cdot)$ states the relation between the state variable and the observation, which may include additional time-varying behaviour denoted by $\mathbf{v}_k$. In some cases, the reformulation of the process equation with an alternative state variable can lead to considerable simplifications of the observation equation, e.g. the reformulation of the CTM based on velocities as proposed in [43].

Although our experiments are conducted solely with heuristic methods, we give an outline of all aforementioned data assimilation concepts below. Most of those concepts are well-known in the field of electrical engineering and computer science, nonetheless we also mention methods from the field of meteorology.

4.1 Newtonian Relaxation or Nudging Method

The *Newtonian relaxation* or *nudging method* basically consists of adding a source- or feedback-term to the dynamic state equation. The nudging method for a spatio-temporal field may be written as

$$\begin{cases} f(z,x,t) = \lambda(x,t) \cdot (z - z_{\text{obs}}) & \text{dynamic equation} \\ z(x_0,t_0) = z_0 & \text{initial condition} \end{cases} \tag{4.2}$$

where x and t are the spatial and the temporal variables, and z and z_{obs} denote the state and the observed variable, respectively. The term $\lambda(x,t)$ is called nudging factor or nudging gain. The feedback-term is proportional to the difference between the observed state variable z_{obs} and the equivalent quantity z computed by the integral of the dynamic differential equation. As a consequence, the feedback term forces the dynamic equation $f(z,x,t)$ to fit z_{obs}.

In [44], Herrera and Bayen use the nudging method for assimilation of data from floating cars and stationary loop detectors. To be specific, floating car data is incorporated by nudging, whereas data from stationary detectors is incorporated by weak boundary conditions (see Sec. 3.2.2). They propose to add a nudging term to the dynamic equation (3.9) denoted by

$$\frac{\partial k}{\partial t} + \frac{\partial q(k)}{\partial x} = -\sum_{j=1}^{J} \lambda\left(x - x_j(t), t\right) \cdot \left[k(x_j(t),t) - k_{\text{obs}}(x_j(t),t)\right], \tag{4.3}$$

where $j = 1 \dots J$ is the vehicle index of the floating cars. Local densities from mobile sensors k_{obs} are estimated from speed observations and transformed to the state variable by inversion of eq. (2.8). A possible nudging factor is mentioned with

$$\lambda(x,t) = \begin{cases} \frac{1}{T_a} \exp\left(-\frac{x^2}{X_{\text{nudge}}^2}\right) \exp\left(-\frac{t-t_{\text{obs}}}{T_d}\right) & \text{if} \quad x \le \alpha X_{\text{nudge}} \quad \text{and} \quad t > t_{\text{obs}} \\ 0 & \text{otherwise,} \end{cases} \tag{4.4}$$

where T_a determines the strength of the nudging factor, T_d and X_{nudge} reflect how the effect of a observation decreases over space and time, respectively. The factor α, with $\alpha > 0$, reflects the area of influence of the measurements.

For implementation purposes, Herrera and Bayen propose a discretisation of eq. (4.3) based on the first-order Godunov scheme (as described in Sec. 3.3). They validate their method on reference-data under simplified conditions. The simplified conditions comprise of (1) the fact that the observed roadway section includes no on- and off-ramps, and (2) only traffic from the two outer left lanes was considered. In summary, Herrera and Bayen are able to reconstruct the spatio-temporal density distribution with satisfactory results, although they note, their conclusions are preliminary and more analysis needs to be done, with respect to their simplifications.

We were not able to find more publications about the application of the nudging method in highway data assimilation problems.

4.2 Kalman Filter based Estimation

The Kalman filter (KF) and its variants are one of the most important and widely-used state estimation methods in engineering. The Kalman filter requires a system-description as discrete model in state-space formulation, where the system state at time $k+1$ is evolved from the prior state at time k which may be written as

$$\begin{cases} \mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{w}_k, \mathbf{u}_k) + \eta_k & \text{state-space equation} \\ \mathbf{y}_k = h(\mathbf{x}_k, \mathbf{v}_k) + \xi_k & \text{observation equation,} \end{cases} \tag{4.5}$$

where $\mathbf{x}_k$ is the internal state variable, $\mathbf{u}_k$ is usually denoted as the control vector and can be used to model external requirements, e.g. traffic demand on the network boundaries, η_k and ξ_k denote the process- and the observation-noise, respectively. Moreover, $\mathbf{w}_k$ and $\mathbf{v}_k$ denote possible time-variation of the process and the observation, e.g. varying capacity or kinematic wave speed, respectively.

On the basis of eq. (4.5), Kalman filtering consists of an iterative two-step *prediction-correction* scheme (Fig. 4.2) [45]. In the prediction step, the time-update based on the state-space equation

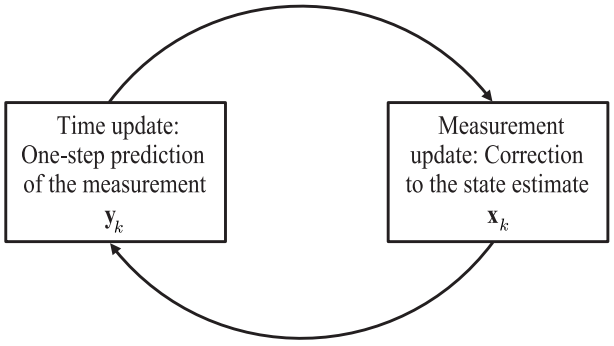


Figure 4.2: Illustration Kalman filter two-step predictor-corrector scheme.

is taken; in the correction step, the measurement update is taken in order to correct the estimate of the current state. The classic Kalman filter is limited to *linear* state-space models and observation equations. Furthermore, it is important that the noise is white Gaussian noise as well as all variances are known. If the model perfectly matches the real world, the state estimate is optimal in terms of the least squares error. In the case of non-linear process and/or observation equations, different extensions of the Kalman filter are available: the Extended Kalman filter (EKF), the Particle filter (PF), the Unscented Kalman filter (UnKF), the Ensemble Kalman filter (EnKF), and many more. Depending on the specific form of eq. (4.5) one has to select

the appropriate variant. In [46], Chen gives an overview of the wide field of Bayesian filtering, including most of the mentioned Kalman filter variants.

Recently, Kalman filtering techniques have become rather popular in the field of traffic engineering. Based on the well known discrete macroscopic and mesoscopic traffic models, i.e. CTM, METANET, and its modifications, Kalman filtering techniques are widely used for estimation of traffic state $\mathbf{x}_k$, as well as for estimation the alteration of road conditions $\mathbf{w}_k$. We refer to Sec. 4.6 for an overview of proposed combinations of discrete traffic models combined with Kalman filtering methods.

4.3 Variational Data Assimilation

In this section we want to introduce data assimilation based on *variational calculus*. As already stated in eq. (3.16), the calculus of variation allows to find a global optimal solution w.r.t. a cost-function $\mathcal{J}(\mathbf{x})$. In contrast to Kalman filtering, variational methods integrate the observations in a non-sequential way. The minimisation based on variational theory was already introduced in Sec. 3.2 in conjunction with the solution of continuous models and optimal control problems. First, we introduce variational data-assimilation techniques in general, followed by an example state estimation example from traffic engineering.

Variational data assimilation techniques, i.e. 3D-Var and 4D-Var, are widely used in the field of meteorology. Similar to the Kalman filtering method, it combines predicted states from a forecast model with given observations. Below we describe the 4D-Var technique for combining data from different sources in a statistically optimal way [47]. The cost-function for 4D-Var can be written as

$$\mathcal{J}(\mathbf{x}) = \mathcal{J}_b(\mathbf{x}) + \mathcal{J}_{\text{obs}}(\mathbf{x}) \tag{4.6}$$

$$\mathcal{J}(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \sum_{i=1}^{N} (\mathbf{y}_i - h_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1} (\mathbf{y}_i - h_i(\mathbf{x}_i)) , \tag{4.7}$$

where $\mathcal{J}_b$ and $\mathcal{J}_{\text{obs}}$ are the forecast (background) and the observation cost functions, respectively. The matrices $\mathbf{B}$ and $\mathbf{R}$ denote the background and the observation error covariance matrices, respectively. If the sequence of model states $\mathbf{x}_i$ is a solution of the model equation, the minimum for $\mathcal{J}(\mathbf{x})$ represents the optimal fit of the model to the data. In general this yields a non-linear constrained optimisation problem, but under the assumptions of

1. a causal process model, and
2. the process model can be linearised around the optimal state x_{opt}, i.e. the process model can be transformed to a *tangent linear model (TLM)* without considerable errors,

the minimisation problem simplifies to a quadratic cost function. The respective minimum can be found with derivation followed by gradient descent. The solution of the 4D-Var method is strongly related to the ones obtained by Kalman filtering methods [47, 48].

Variational data assimilation is not very popular in the field of traffic engineering. By now we neither found scientific articles nor conference papers regarding this topic, but nonetheless, Volpi applies in her master's thesis [49] the 4D-Var to a data-fusion problem based on the second-order model by Payne and Whitham. She is able to achieve results with artificial input-data. However, Volpi mentions some drawbacks and the need for further research. We are not aware of any further publication using a variational data-assimilation approach.

Although not directly related to variational data-assimilation, we want to mention a method based on the theory of optimal control, as it also relies on variational minimisation techniques. In [50, 51], Jacquet et al. introduce the *Extended LWR model (ELWR)*, an extension of the standard LWR model which includes ramp-flows using so-called saturation functions and cor-

responding ramp-flow-rates. Based on the ELWR they are able to formulate a linearised first variation which can be solved by gradient minimisation of the cost-function

$$\min_y \mathcal{J}(k,u) = \mathcal{J}_{\text{obs}}(k) + \mathcal{J}_{\text{bar}}(u), \tag{4.8}$$

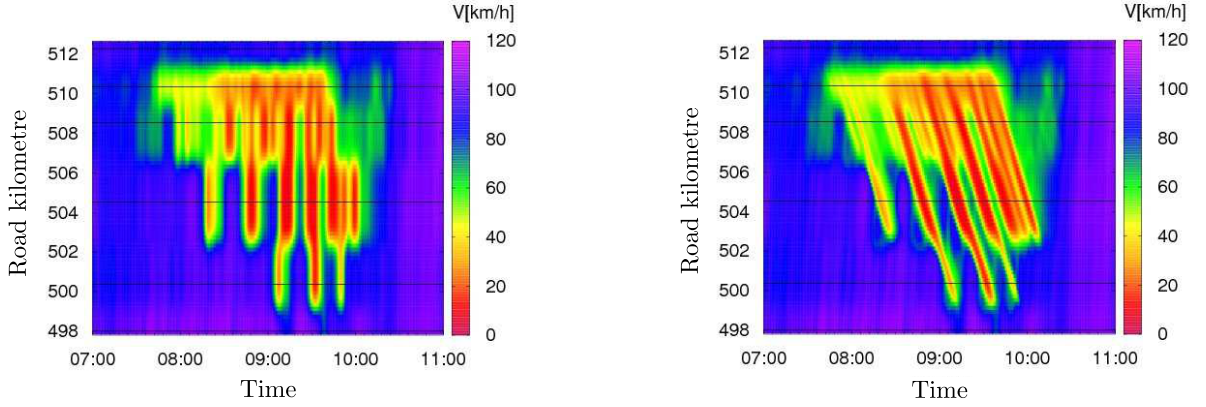
based on the observation k and the ramp rates, i.e. barrier conditions, given by the variable u. The authors give an example for optimal ramp-metering, and for initial traffic density estimation. One of the main advantages of their method in contrast to the Godunov scheme is that the discretisation is done at the optimisation step.

4.4 Heuristic Methods

The principles behind heuristics have been already introduced in Sec. 3.1. A first, impracticable approach to reconstruct a spatio-temporal field of speed, density, or flow would be smoothed interpolation or higher-order regression of the scattered sensor data.

A similar procedure is conceivable for the data assimilation of time-variant fields: Local interpolation methods can be applied for *successive correction* of data obtained from a forecast model and observed values. For example, the *Cressman correction* [52] adapts a forecast at a specific grid-point by linear combination of the residuals between predicted and observed quantities within an influence radius around the grid-point. The residuals inside the radius are weighted by their distance.

All these classical interpolation methods ignore the existence of characteristic shock waves in traffic flow, and weight the observations *isotropic* in each direction of the spatio-temporal plane. In [53], Treiber and Helbing present a spatio-temporal interpolation method, the *Adaptive Smoothing Method* (ASM), based on *anisotropic* interpolation. Their method exploits the fact, that perturbations propagate with *nearly constant speed* upstream and downstream under congestion and free flow conditions as already mentioned in Sec. 3.2.2. The transition between free-flow and congested traffic is evaluated by a non-linear velocity filter. Reconstructed flow fields obtained by isotropic smoothing and the ASM are compared in Fig. 4.3.



(a) Isotropic interpolation

(b) Anisotropic interpolation

Figure 4.3: Comparison of flow field reconstruction with isotropic and anisotropic smoothed interpolation. In contrast to isotropic interpolation illustrated in (a), the adaptive smoothing method considers the characteristic speeds of traffic flow as illustrated in (b). The horizontal lines represent the stationary detectors (Figs. from [11]).

Even though the ASM is based on simplifying assumptions when compared to continuous

flow models, it is possible to achieve good results for traffic state reconstruction from empirical data [11, 54]. Since we have conducted several experiments with different ASM variants, a detailed description of the algorithm-family follows in the section below.

4.5 Adaptive Smoothing Method Variants

The *Adaptive Smoothing Method* and its variants are of great importance for our work. Our implemented data-fusion algorithms are mostly based on this family of algorithms. In [53], Treiber and Helbing propose the ASM for reconstruction of spatio-temporal traffic variables $z(x,t)$, such as flow q, speed V, and density k, based on *stationary detector data*. While the ASM was first suggested for homogeneous data only, some extensions for heterogeneous data-sources have been proposed later [1, 42, 54].

All ASM-variants assume the following partly heuristic properties of traffic flow [53]: Based on the assumption of a bilinear fundamental diagram (see Fig. 3.4) and depending on the particular traffic state, perturbations propagate with constant speeds, namely c_{free} and c_{cong} under free-flow and under congested traffic state, respectively. Even in traffic states with "synchronised" congested traffic flow, and the consequential high traffic densities or low velocities, perturbations propagate with velocities near c_{cong}. Further, the method should smooth high-frequency fluctuations of the input-data in both spatial and temporal dimension.

In the sections below we describe all ASM variants found in literature with emphasis on the particular differences, as well as the different usable data-sources.

4.5.1 Adaptive Smoothing Method (ASM)

Initially, the ASM was proposed for homogeneous input-data from stationary cross-section sensors only. Typical input-data z_i is obtained by stationary cross-section sensors which can include the average velocity, the vehicle flow, the occupancy, or derived quantities such as the traffic density k. For given ordered input-data

$$d_i = \{x_i, t_i, z_i \,|\, x_i \leq x_{i+1}, t_i \leq t_{i+1}\} \qquad i = 1 \ldots N, \tag{4.9}$$

a spatio-temporal field $z(x,t)$ in the range $x \in [x_1, x_N]$ and $t \in [t_1, t_N]$ can be calculated. Overall the algorithm consists of two main parts:

1. The input-data is filtered by a spatio-temporal low-pass w.r.t. the two characteristic propagation velocities; Missing values are reconstructed in this step.
2. A non-linear filter superimposes both velocity fields by linear combination with an adaptive weighting factor.

The characteristic fields are calculated by two-dimensional smoothed interpolation which can be written as

$$z_{\text{free}}(x,t) = \frac{1}{\mathcal{N}_{\text{free}}(x,t)} \sum_{i \in A_{\text{free}}(x,t)} \phi_0 \left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{free}}} \right) \cdot z_i, \qquad \text{and} \tag{4.10}$$

$$z_{\text{cong}}(x,t) = \frac{1}{\mathcal{N}_{\text{cong}}(x,t)} \sum_{i \in A_{\text{cong}}(x,t)} \phi_0 \left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{cong}}} \right) \cdot z_i, \tag{4.11}$$

where ϕ_0 is a low-pass interpolation kernel, $\mathcal{N}_{\text{free}}$ and $\mathcal{N}_{\text{cong}}$ are variable normalisation factors, and $A_{\text{free}}(x,t)$ and $A_{\text{cong}}(x,t)$ define the area where $\phi_0 \not\approx 0$. As ϕ_0 should have a low-pass

characteristic, Treiber and Helbing propose an exponential kernel denoted by

$$\phi_0(x,t) = \exp\left\{-\left(\frac{|x|}{\sigma} + \frac{|t|}{\tau}\right)\right\}, \tag{4.12}$$

with the smoothing-width parameters σ and τ. Instead of an exponential function, the use of other low-pass kernels, such as a bivariate Gaussian kernel, would be possible. This kernel is skewed towards the characteristic velocities respectively in eqs. (4.10) and (4.11) which can be written as

$$\phi_{\text{free}}(x,t) \equiv \phi_0\left(x, t - \frac{x}{c_{\text{free}}}\right) \qquad \text{and} \tag{4.13}$$

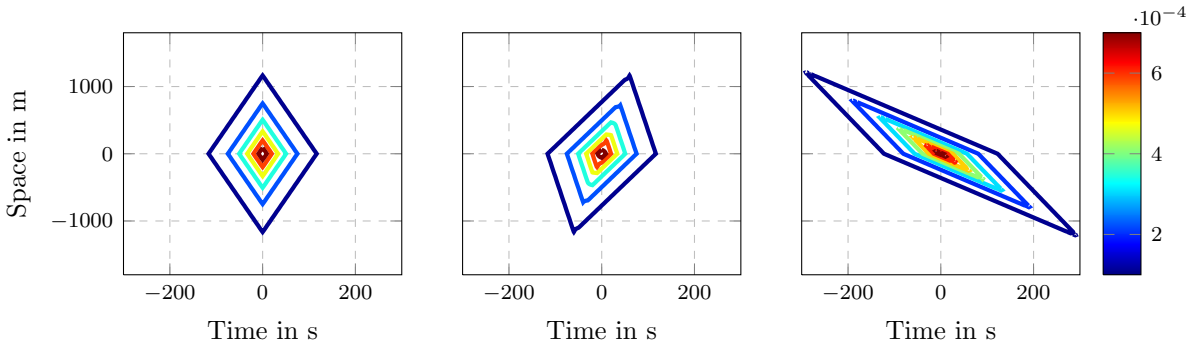
$$\phi_{\text{cong}}(x,t) \equiv \phi_0\left(x, t - \frac{x}{c_{\text{cong}}}\right). \tag{4.14}$$

Those interpolation kernels define the weight of a single data-point (x_i, t_i, z_i) in the estimation of $z(x,t)$ with

$$\beta_{\text{free},i}(x,t) = \phi_{\text{free}}(x_i - x, t_i - t), \qquad \text{and} \tag{4.15}$$

$$\beta_{\text{cong},i}(x,t) = \phi_{\text{cong}}(x_i - x, t_i - t) \tag{4.16}$$

which is determined by the distance between the point (x,t) and the data-points (x_i, t_i) as illustrated in Fig. 4.4. The factors $\mathcal{N}$ normalise the sum in eqs. (4.10) and (4.11) with respect



(a) Isotropic kernel ϕ_0 (b) Free-flow: ϕ_{free} (c) Congested-flow: ϕ_{cong}

Figure 4.4: Illustration of the spatio-temporal low-pass interpolation kernels. The isotropic kernel in (a) gets skewed by the characteristic velocities c_{free} or c_{cong} as illustrated in (b) and (c).

to the current number of data-points in $A(x,t)$ denoted by

$$\mathcal{N}_{\text{free}}(x,t) = \sum_{i \in A_{\text{free}}(x,t)} \beta_{\text{free},i}(x,t) \qquad \text{and} \tag{4.17}$$

$$\mathcal{N}_{\text{cong}}(x,t) = \sum_{i \in A_{\text{cong}}(x,t)} \beta_{\text{cong},i}(x,t). \tag{4.18}$$

Finally, the quantity z can be reconstructed at a spatio-temporal point (x,t) as follows

$$z(x,t) = w(x,t)\, z_{\text{cong}}(x,t) + [1 - w(x,t)]\, z_{\text{free}}(x,t). \tag{4.19}$$

The weight factor $w(x,t)$ controls the superposition of the fields under free flow and congested conditions. The factor should be $w \approx 0$ at high speeds and $w \approx 1$ at low speeds. Thus, one can

use different weighting functions bounded between $[0, 1]$. Any function of quantities of traffic flow, i.e. density/occupancy, speed or flow, which is able to discriminate between free-flowing and congested traffic may be possible. For a reconstruction of a velocity field $V(x,t)$, Treiber et al. recommend in [53] the sigmoid function

$$w(x,t) = W(V_{\text{free}}(x,t), V_{\text{cong}}(x,t)) = \frac{1}{2}\left[1 + \tanh\left(\frac{V_{\text{thr}} - \min(V_{\text{free}}(x,t), V_{\text{cong}}(x,t))}{\Delta V}\right)\right], \quad (4.20)$$

where $V_{\text{free}}(x,t)$ and $V_{\text{cong}}(x,t)$ correspond to the aforementioned $z_{\text{free}}(x,t)$ and $z_{\text{cong}}(x,t)$, respectively. Further, V_{thr} denotes the threshold velocity and ΔV denotes the transition width between free-flow and congested traffic, respectively. The function $w(x,t)$ is illustrated in Fig. 4.5. One can interpret the min-function in eq. (4.20) intuitively: the field with higher density and

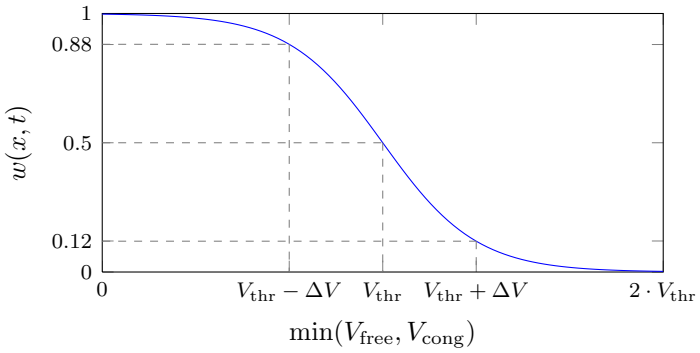
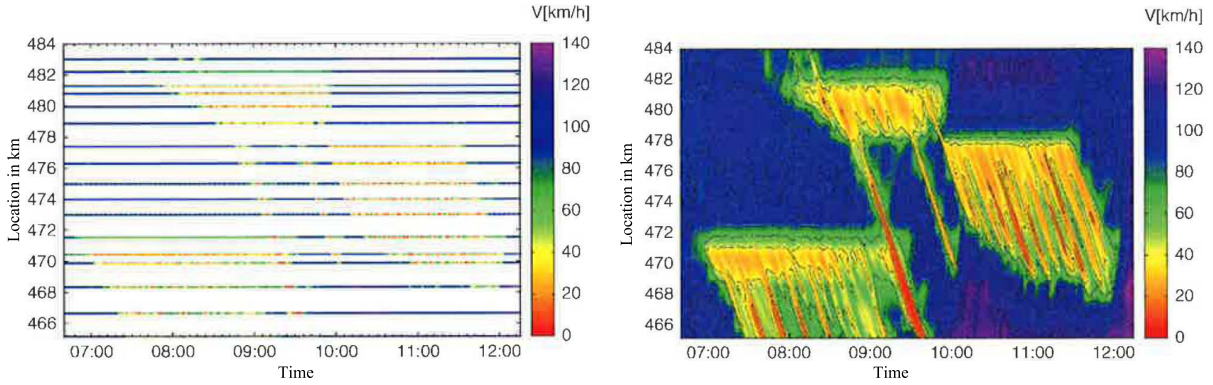


Figure 4.5: The non-linear weighting-factor $w(x,t)$ for combination of the characteristic velocity fields.

lower speed dominates the decision about a particular traffic state.

Input-data from stationary cross-section detectors (indicated by the horizontal lines), and a reconstructed velocity field based on the ASM are shown in Fig. 4.6



(a) Sensor data

(b) Reconstruction

Figure 4.6: Reconstruction of a velocity field $V(x,t)$ from stationary detector-data (indicated by horizontal lines) based on the adaptive smoothing method (Figs. from [1]).

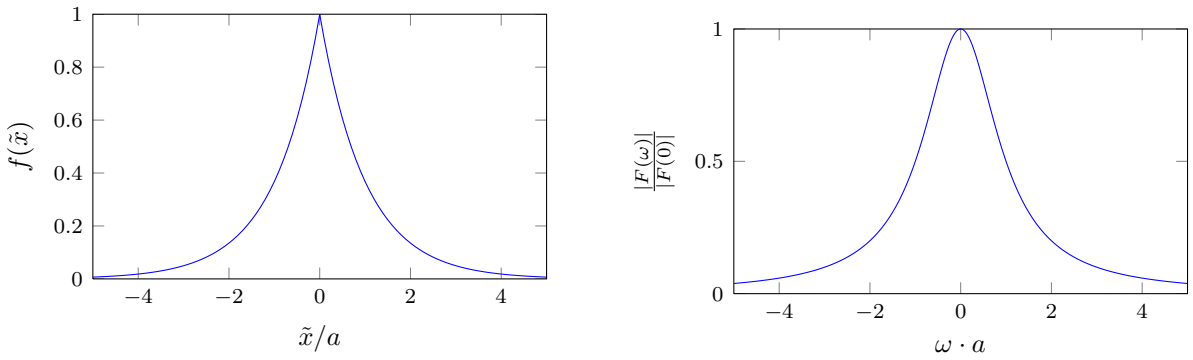
Frequency Characteristics of the Interpolation Kernel. Every cross-section through the zero point of the interpolation-kernel results in an one-dimensional exponential kernel

$$f(\tilde{x}) = \mathrm{e}^{-\frac{|\tilde{x}|}{a}}, \quad (4.21)$$

with kernel-width parameter a, depending on the respective angle of the cross-section in the x/t plane. The Fourier transform of $f(\tilde{x})$ yields the Lorentzian function [55]

$$\mathcal{F}_{\tilde{x}}[f(\tilde{x})] = F(\omega) = \frac{2/a}{1/a^2 + \omega^2}. \tag{4.22}$$

Both representations of $f(\tilde{x})$, i.e. in spatial/time-domain, and in frequency domain, are illustrated in Fig. 4.7.



(a) Time domain

(b) Fourier transform

Figure 4.7: Illustration of the two-sided exponential kernel in spatial/time-domain (a) and the respective Fourier transform in (b).

The 6 dB cut-off frequency ω_c is defined by the relation

$$\frac{1}{2}|F(0)| \stackrel{!}{=} |F(\omega_c)| \tag{4.23}$$

which yields

$$\omega_c(a) = \pm\frac{1}{a} \qquad \text{and} \qquad f_c(a) = \pm\frac{1}{2\pi a}, \tag{4.24}$$

or a respective minimum (cut-off) wave-length of

$$L_c(a) = \pm 2\pi a. \tag{4.25}$$

This cut-off wavelength is important, when considering the spatio-temporal interpolation as sampling problem. Especially for data from stationary detectors it is easy to specify the *average* sampling intervals. This fact is reflected in the recommended algorithm parameters σ and τ as described in the next paragraph.

Calibration. The selection of the algorithm parameters c_{free}, c_{cong}, V_{thr}, and ΔV including a sensitivity analysis is discussed in [53]. Treiber and Helbing provide no specific calibration procedure justified by the heuristic nature of the algorithm. Instead, they give recommendations for the selection of the parameters based on empirical studies. Tab. 4.1 provides a summary of recommended parameters to use for German motorways. A predefined set of parameters seems to be inappropriate for every motorway, because the capacity and the characteristic velocities will change. However, the algorithm is not very sensitive against the variation of the parameters in Tab. 4.1 [53, 54].

One possible method of calibration can be the optimisation of the parameters based on repeated random sub-sampling validation. In his doctoral thesis, Schreiter proposes the calibration

Parameter	Description	Value	SI
σ	Spatial smoothing width	avg. $\Delta x_i/2$	
τ	Temporal smoothing width	avg. $\Delta t_i/2$	
c_{free}	Characteristic velocity under free-flow conditions	$70\,\text{km}\,\text{h}^{-1}$	$19.4\,\text{m}\,\text{s}^{-1}$
c_{cong}	Characteristic velocity under congested conditions	$-15\,\text{km}\,\text{h}^{-1}$	$-4.2\,\text{m}\,\text{s}^{-1}$
V_{thr}	Threshold between congested and free-flow traffic	$60\,\text{km}\,\text{h}^{-1}$	$16.7\,\text{m}\,\text{s}^{-1}$
ΔV	Transition width between congested and free-flow traffic	$20\,\text{km}\,\text{h}^{-1}$	$5.6\,\text{m}\,\text{s}^{-1}$

Table 4.1: Proposed parameters of the adaptive smoothing method (ASM) to use for German motorways by [53].

of characteristic speeds by detecting lines in the spatio-temporal plane using the Hough transform [56]. In general, any method for calibration of the flow-density fundamental diagram may be used to calibrate the traffic dependent parameters.

The initial proposal of the ASM is restricted to stationary input-data only. Usually, this input-data is aligned on a spatio-temporal grid regularly (see Fig. 4.6). As a consequence, the sums in eqs. (4.10), (4.11), (4.17) and (4.18) can be implemented efficiently as described in Chapter 6.

4.5.2 Generalised Adaptive Smoothing Method (GASM)

In [1,54], Treiber et al. propose the *Generalised ASM* (GASM), an extension of the ASM, which is also able to estimate a spatio-temporal field $z(x,t)$ based on heterogeneous input-data. A first approach for heterogeneous input-data would be the reconstruction with the ASM. This ignores the different measurement uncertainties of each data-source and would lead to a sub-optimal solution. The GASM, on the other hand, introduces a data-source specific weight-factor which is based on following requirements: (1) the input-data is free of systematic errors, (2) the error variances θ_j^2 of each data-source are known, and (3) the errors of the different data-sources are uncorrelated.

Basically, flow fields for the different classes of input-data, i.e. $z_{\text{cong}}^{(j)}(x,t)$ and $z_{\text{free}}^{(j)}(x,t)$, are calculated first, followed by linear combination weighted with respect to their expected error variance, which can be written as

$$z(x,t) = \sum_j \alpha^{(j)} \cdot \left\{ w^{(j)}(x,t) z_{\text{cong}}^{(j)}(x,t) + \left[1 - w^{(j)}(x,t)\right] z_{\text{free}}^{(j)}(x,t) \right\}, \tag{4.26}$$

where $\alpha^{(j)}$ denotes the weighting factor for a given data-source with index j. The combination of the different data-sources should minimise the overall error-variance of the the reconstruction $\theta(\{\alpha^{(j)}\})$. This yields to an optimisation problem given by

$$\{\alpha^{(j)}\} = \underset{\{\alpha^{(j)}\}}{\arg\min}\, \theta(\{\alpha^{(j)}\}) = \min \sum_j \left(\alpha^{(j)}\right)^2 \theta_j \qquad \text{subject to} \qquad \sum_j \alpha^{(j)} = 1. \tag{4.27}$$

Treiber et al. recommend the same parameters for the GASM as for the ASM in the previous section. The weights $\alpha^{(j)}$ are decoupled of the spatio-temporal weights $\beta_i(x,t)$ in eqs. (4.10) and (4.11) and as a consequence ignore the sampling density of the different data-sources. The EGTF, another extension of the ASM, solves this issue. The full algorithm is described in the section below.

4.5.3 Extended generalised Treiber-Helbing Filter (EGTF)

In [42], van Lint and Hoogendoorn propose the *Extended Generalised Treiber-Helbing Filter (EGTF)*. Based on the GASM, referenced as *Generalised Treiber-Helbing Filter (GTF)* in their work, the EGTF is also able to reconstruct a spatio-temporal field $z(x,t)$ from heterogeneous data-sources. In contrast to the GASM, their method considers the spatio-temporal density of the different data-sources.

The EGTF consists of several different steps. First, for each data-source j, the input-data $\{z_i\}^{(j)}$ is smoothed at the points $\{x_i,t_i\}^{(j)}$ with the ASM as described in Sec. 4.5.1. Hereinafter, $z_i^{(j)}(x,t)$ denotes the smoothed data-point for a particular data-source j. Additionally, the weighting factor $w^{(j)}(x,t)$ from eq. (4.20) is reused later. For the fusion of multiple data-sources into a common field $z(x,t)$ they propose the linear combination

$$z(x,t) = \frac{\sum\limits_j \alpha^{(j)}(x,t) \sum\limits_{i\in A^{(j)}(x,t)} \phi_i^{(j)}(x,t) z^{(j)}(x,t)}{\sum\limits_j \alpha^{(j)}(x,t) \sum\limits_{i\in A^{(j)}(x,t)} \phi_i^{(j)}(x,t)}. \tag{4.28}$$

This equation contains two data-source specific weight factors $\alpha^{(j)}(x,t)$ and $\phi_i^{(j)}(x,t)$.

The factor $\alpha^{(j)}(x,t)$ represents the reliability of the input data at (x,t) and is based on a priori estimates of the measurement uncertainty of data-source j. The value of $\alpha^{(j)}(x,t)$ is based on the following assumptions:

1. $\alpha^{(j)}$ is inverse proportional to the standard deviation $\Theta^{(j)}$ of a particular data-source j.
2. $\Theta^{(j)}$ is constant in free-flow as well as under congested traffic conditions, with a smaller value under congested conditions. Accordingly, $\Theta_0^{(j)}$ denotes the standard deviation of the measurement error under congestion ($v(x,t) \ll V_{\text{thr}}$) and $\left[1+\mu^{(j)}\right]\Theta_0^{(j)}$ under free-flow conditions ($v(x,t) \gg V_{\text{thr}}$), where the parameter $\mu^{(j)}$ represents the increased error standard deviation and has to be selected before.

The particular traffic state is represented by $w^{(j)}(x,t)$, so the first weighting factor can be written as

$$\alpha^{(j)}(x,t) = \frac{1}{\Theta_0^{(j)}\left[1+\mu^{(j)}(1-w^{(j)}(t,x))\right]}. \tag{4.29}$$

The second factor in eq. (4.28), namely $\phi_i^{(j)}(x,t)$, considers the fact that data from different sources will result in different traffic flow conditions. Further, it covers the fact that one data-source will include more data around a particular estimation point (x,t). Using eqs. (4.15), (4.16) and (4.20) the authors propose the factor

$$\phi_i^{(j)}(x,t) = w^{(j)}(x,t)\cdot\beta_{\text{cong},i}^{(j)}(x,t) + \left[1-w^{(j)}(x,t)\right]\cdot\beta_{\text{free},i}^{(j)}(x,t). \tag{4.30}$$

For optimal results, $\Theta_0^{(j)}$ and $\mu^{(j)}$ must be estimated for each data-source. This requires refined knowledge of the particular sensors, or the acquisition of reference measurements. If both methods are not possible for a particular data-source, an educated guess may deliver sufficient reconstruction results.

The authors use the EGTF to fuse velocity data obtained from stationary detectors and floating cars with travel time data obtained by an *automatic numberplate recognition (ANPR)* system. In order to obtain densely sampled velocities from the ANPR system, the travel times were first converted to virtual trajectories. Furthermore, they evaluate not only different spac-

ings of stationary detectors and the ANPR system, but also different percentages of floating cars.

4.6 Classification of Models and State Estimation Methods

This section offers a compact overview of traffic flow models, solution methods or respective data-fusion methods found in literature. The overview is structured along three categories: (a) solution methods based on a discrete dynamic equation, (b) solution methods based on a continuous dynamic equation, although it may be solved numerically later, and (c) heuristic methods. Especially the presented methods for continuous dynamic equations may be able to incorporate heterogeneous input-data (HI) as boundary conditions, but ignore the measurement errors. A comprehensive overview is provided by Tab. 4.2.

(a) Discrete dynamic equation

Model	**proposed by**	**State estimation method**	**HI**
Cell Transmission Model (CTM)	Daganzo [34]	PF by Sau et al. [57]	
Switching Mode Model (SMM)	Muñoz et al. [58]	Mixture KF by Sun et al. [59, 60]	
		Nudging Method by Herrera and Bayen [44]	✓
		KF by Herrera and Bayen [44]	✓
METANET	Papageorgiou et al. [61]	EKF by Wang and Papageorgiou [62]	
		Comparison of EKF and UKF by Heygi et al. [63]	
		Parallelised PF (PPF) by Heygi et al. [64]	
		EKF by Abdi [65]	
		EKF trained ANN by Abdi et al. [66]	
		Road-networks with PPF by Mihaylova et al. [67]	
		Road-networks with PGSPF by Mihaylova et al. [67]	
Hybrid Stochastic Model (HSM)	Boel and Mihaylova [68]	direct by Boel and Mihaylova [68, 69]	
		UKF by Mihaylova et al. [70, 71]	
		PF by Mihaylova et al. [71]	
Discrete Payne-Whitham Model	Payne and Whitham [72]	4D-Var data-fusion by Volpi [49]	
CTM for velocity (CTM-v)	Work et al. [43, 73]	EnKF by Work et al. [41, 43, 73]	✓
Lagrangian CTM	Wageningen-Kessels et al. [38, 74]	EKF by Yuan et al. [39, 75]	✓

(b) Continuous dynamic equation

Model	proposed by	Solution method	HI
Extended LWR (ELWR)	Jacquet et al. [50]	With linearised first variation and solution with gradient descent by Jacquet et al. [50, 51]	
PDE in Hamilton-Jacobi (HJ) form	Daganzo [26, 27]	Variational formulation and solution via dynamic programming by Daganzo [27]	
		Lax-Hopf based closed form solution by Claudel and Bayen [31]	✓
		Accelerated algorithm for triangular fundamental diagrams by Mazaré et al. [32]	✓
Lagrangian HJ PDE	Han et al. [40]	Closed form solution by Han et al. [40]	✓

(c) Heuristic methods

Method	proposed by	HI
Adaptive Smoothing Method (ASM)	Treiber et al. [53]	
Generalized Adaptive Smoothing Method (GASM)	Treiber et al. [54]	✓
Extended Generalized Treiber-Helbing Filter (EGTF)	van Lint et al. [42]	✓

*Table 4.2: Overview of recently proposed mesoscopic and macroscopic traffic models, related solution methods and data-fusion methods respectively. The column **HI** indicates the capability of processing heterogeneous input data.*

4.7 Summary

This chapter provided a conceptual overview of modern data-fusion methodologies used in the area of traffic engineering. This overview was structured among four fundamentally different approaches for solving the problem of data-fusion. In combination with the overview of traffic flow modelling in Chapter 3, this chapter may serve as a starting point for future studies in this particular research area.

Since we implemented algorithms based on the idea of anisotropic interpolation, an extensive description of those methods was included. This description should serve as a basis for the comprehension of our further work, including requirements regarding the used data sources, the description of particular implementation issues, and the evaluation of conducted experiments.

5 Data Sources

Let us remember our research objective, namely the traffic state estimation by fusion of input-data originating from various sources. Although methods for fusion of heterogeneous data have been presented in the previous chapter, the different data is in fact recorded in different domains and must be transformed to an appropriate physical domain before fusion. In particular, the methodologies based on anisotropic interpolation (see Sec. 4.5) require input data in form of eq. (4.9) independent from the specific sensor type. Therefore, the following data preprocessing steps have become necessary: Stationary sensors must be associated with a particular road kilometre. GPS-based floating car data is recorded in a two-dimensional spatial domain and has to be (1) associated with a particular road and (2) transformed to the one-dimensional spatial domain. If not measured directly, velocities must also be estimated from trajectories. Furthermore, the specification of the error for all data-sources is an important prerequisite for all mentioned data-fusion algorithms.

First, we describe our test-site, the Austrian motorway A4 (Sec. 5.1) and its respective data sources (Sec. 5.2 to 5.4). Second, give information about the used microsimulation software and its respective model (Sec. 5.7).

5.1 Test-Site Austrian A4 Motorway

We were able to perform experiments on real-world data from the Austrian A4 Ostautobahn motorway between junction Prater and Vienna International Airport as well as a detailed microsimulation model representing the same area. The real-world data originates from different partners in the framework of projects carried out by the ISV. The geometry of our test-site is illustrated in Fig. 5.1, including the positions of ramps and stationary detectors, plus the section for travel-time measurements.

Road Geometry

Information about the road geometry acts as a key element in processing data obtained from floating cars, such as matching to a given driving direction, estimating the road kilometre or if required the floating car velocity.

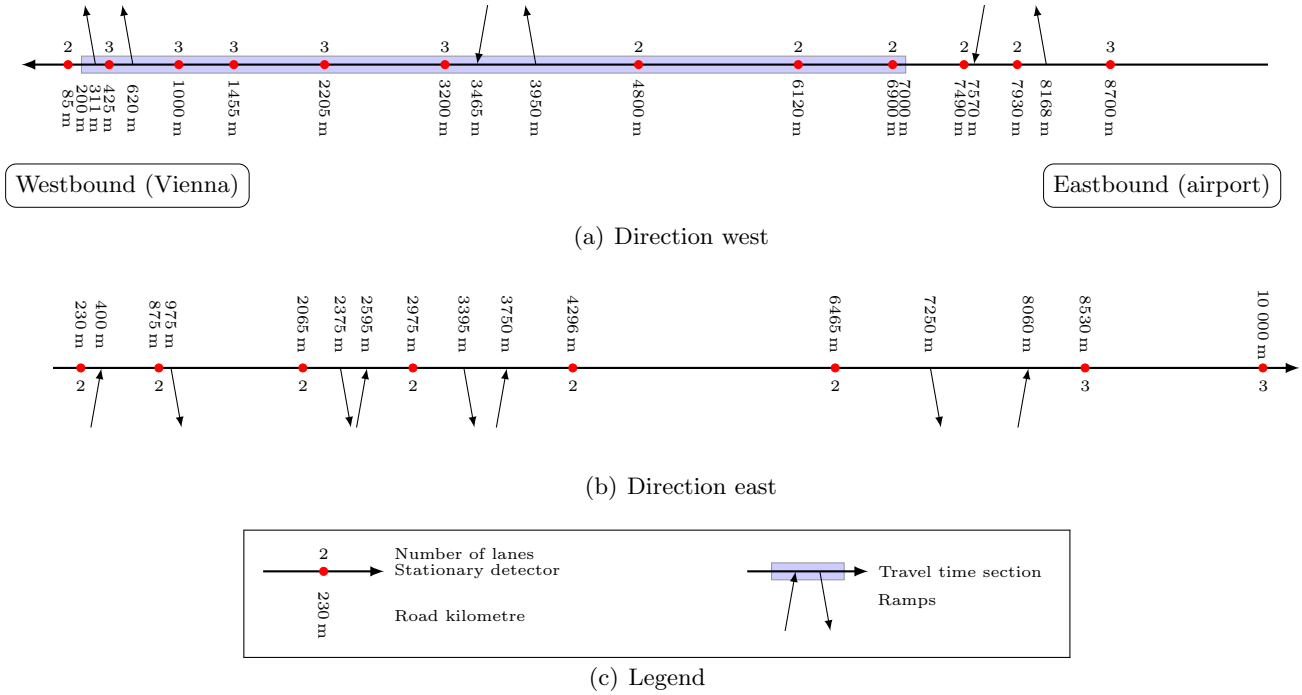


Figure 5.1: Illustration of the test-site A4 Ostautobahn between Wien-Mitte (road kilometre 0) and Vienna International Airport (approx. road kilometre 10). The figure includes ramps, stationary detectors, and the section for travel-time measurement with respective road kilometres.

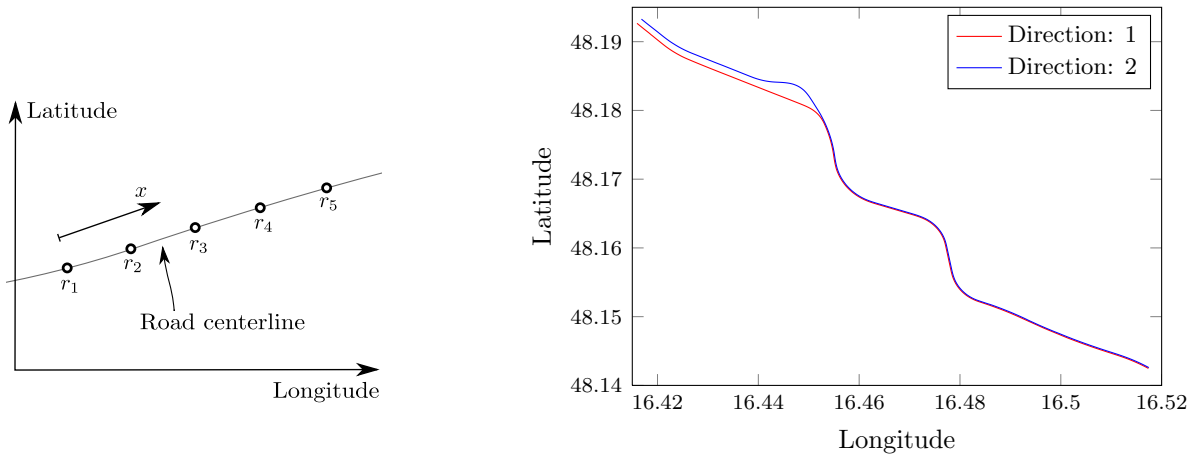
The geometry of the road is modelled as a geospatial poly-line without junctions. The poly-line is basically a list of data-points composed of four entities

$$r_i = \{lat_i, lon_i, x_i, h_i\}\,, \qquad i = 1 \ldots M \tag{5.1}$$

where $\{lat_i, lon_i\}$ is the geospatial datum in the WGS84 reference coordinate system[4], x_i is the road kilometre, and h_i is the heading direction[5]. No lane information is included in the model. The coordinates of the model are going along the centreline of the real road surface. The model data for the Austrian A4 motorway is available at ISV, having a road kilometre sampling distance of $\Delta x = x_{i+1} - x_i = 5\,\text{m}$. The model is illustrated in Fig. 5.2. To achieve a lower sampling distance of e.g. $\Delta x = 1\,\text{m}$, the model is interpolated with a one-dimensional spline interpolation for both geospatial coordinates. Data for new motorway stretches can be obtained from different data-sources, e.g. *Graphenintegrations-Plattfom Austria (GIP)*[6], *OpenStreetMap*[7], or other geographic information providers.

An overview of all available data sources is provided in Tab. 5.1, including the type of the measurement, the data provider, and a brief description of the data characteristics. Due to the fact that the data comes from third-party providers, only limited information besides the short description in Tab. 5.1 is available. For example, we have no detailed information about the used sensors, the measuring arrangement, or the dimension of expected measurement errors. In order to specify characteristics of the data, e.g. expected measurement errors, we discuss the aforementioned data sources on the basis of general information, e.g. data-sheets and scientific literature. The following sections provide a review about stationary-detectors (Sec. 5.2), followed by floating cars (Sec. 5.3) and the travel time measurements based on automatic number plate

[4] `http://spatialreference.org/ref/epsg/4326/`
[5] 1: In the same direction as the road kilometres; 2: In opposite direction of the road kilometres.
[6] `http://www.gip.gv.at`
[7] `http://www.openstreetmap.org/`



(a) Model

(b) Geographic model extent

Figure 5.2: The model of the motorway is shown in (a). The model combines a geospatial datum lon/lat with a road kilometre x and a heading direction h in a single data-point $r_i = (lat_i, lon_i, x_i, h_i)$. The geographic extent of the model data for the motorway A4 is shown in (b).

recognition (ANPR) for validation purposes (Sec. 5.4). The final section (Sec. 5.5) describes the available microsimulation tool by which every aforementioned data source can be simulated.

Type	**Project / Provider**	**Description**
Cross sectional data	REFEREE[8], ASFINAG	One-minute aggregated measurements of flow and velocity, divided per lane and for two vehicle classes.
Floating car data	REFEREE / AIT	WGS84 coordinates with timestamps from GPS equipped taxis, having an average sampling interval of $\Delta T \approx 20 - 36\,\mathrm{s}$.
	IMPAKT[9]	WGS84 coordinates with timestamps and velocity information from GPS equipped probe vehicles, having an average sampling interval of $\Delta T = 1\,\mathrm{s}$.
Travel times	REFEREE	Five-minute aggregated travel time measurements obtained by an ANPR system solely on the first lane obtained between the road kilometres $X_{\mathrm{BC}} = 7 - 0.2\,\mathrm{km}$, driving direction west.
Microsimulation model	IMPAKT / TU Graz	Microsimulation model for PTV-VISSIM, calibrated on traffic conditions similar to workdays between 17:00 and 18:00 on the basis of stationary detector data.

Table 5.1: Overview of available data for this thesis originally obtained by research partners in the framework of conducted research projects of the ISV.

[8] *Referenzierung und Evaluierung von verkehrstechnischen Effekten* is a project partially financed by BmVIT under contract FFG-Nr. 831730 by the project partners AIT, ITS Vienna Region, SLR Engineering OG, EBE Solutions and TU Graz.

[9] *Indikatoren zur multikriteriellen Analyse kooperativer Transportsysteme* is a project partially financed by BmVIT under contract FFG-Nr. 831747 by the project partners rk communication mobility e.U., TU Graz, TraffiCon GmbH and Technische Universität München.

5.2 Stationary Detectors

Stationary detectors perform measurements on the traffic stream at one single point of a road, or on a small section of a road. In terms of data as defined in eq. (4.9), the component $x_i = const.$ for every individual sensor.

Classic stationary sensors are inductive single- or double loop detectors [11]. In addition, a wider range of sensors became available recently, including microwave radar, ultrasonic- and laser-ranging, passive infra-red sensors for single point measurements and automatic number plate recognition for section based measurements. In combination with an information processing system a velocity sensor may be used to perform vehicle tracking, counting and classification [76]. Furthermore, a fusion of different sensors may be used to perform simultaneous velocity measurements, vehicle counting, and classification [77].

Measurement Accuracy

Weather and road conditions can vary in a wide range over time. Those factors may have an impact on the accuracy of a stationary traffic sensor. Manufacturers of traffic sensors are obligated to specify a guaranteed range of measurement uncertainty for their sensors, although the true distributions of errors are not published[10]. Tab. 5.2 provides a summary of the uncertainties for single-vehicle measurements of different commercial traffic sensors. One can see that the uncertainties lie between $\pm 1\,\mathrm{km\,h^{-1}}$ and $\pm 5\,\mathrm{km\,h^{-1}}$ for velocities below $100\,\mathrm{km\,h^{-1}}$ and between $\pm 1\,\%$ and $\pm 5\,\%$ for velocities above $100\,\mathrm{km\,h^{-1}}$, respectively. However, the manufacturers do not specify the relative frequency of cases where the error specifications may be expected to be valid. Under the assumption that (1) the specified uncertainty ranges are valid in 99.7 % of all cases, and (2) the measurement errors are Gaussian distributed with zero mean, the ranges of uncertainty can be interpreted as the $\pm 3\sigma$ region around the true value.

Name	**Manufacturer**	**Speed**		**Counting**
		$\leq 100\,\mathrm{km\,h^{-1}}$	$> 100\,\mathrm{km\,h^{-1}}$	
TDC1	Swarco	$\pm 5\,\mathrm{km\,h^{-1}}$	$\pm 5\,\%$	$\pm 3\,\%$
TDC3, TDC4	Swarco	$\pm 3\,\mathrm{km\,h^{-1}}$	$\pm 3\,\%$	$\pm 3\,\%$
UMRR 29	Smart Microwave Sensors	$\pm 1\,\mathrm{km\,h^{-1}}$	$\pm 1\,\%$	n.a.
MultaStat CRM NG	Multanova AG	$\pm 2\,\mathrm{km\,h^{-1}}$	$\pm 2\,\%$	n.a.

Table 5.2: Uncertainties of different available commercial stationary traffic sensors [77–81].

Since no further information about the sensors on the motorway A4 is available, there is no information on the real uncertainties. Moreover, real world data-sets, as provided by ASFINAG, are aggregated over fixed time intervals of 1 min and contain missing values of approximately 3-5 %[11]. Those missing values are interpolated based on daily traffic track records.

The measured velocity $\tilde{v}_\alpha$ of an individual vehicle α can be written as

$$\tilde{v}_\alpha = v_\alpha + \eta_{\mathrm{v},\alpha}, \tag{5.2}$$

where v_α denotes the true velocity and $\eta_{\mathrm{v},\alpha}$ the zero-mean Gaussian measurement noise. An aggregated data-point v_i has to be interpreted as the sum of the average vehicle velocity $\overline{v}_{\alpha,i}$ and the average measurement noise $\overline{\eta}_i$ determined from all vehicles $\{\alpha\}_i$ passed in the respective

[10] For instance, a request to Jenoptik GmbH in Germany regarding information about the real uncertainty of their traffic sensors was rejected at March 26, 2013 with reference to company secrets.

[11] Source: Prof. Fellendorf at March 11, 2013

aggregation interval i:

$$v_i = \overline{v}_{\alpha,i} + \overline{\eta}_i = \frac{1}{q_i} \sum_{\alpha \in \{\alpha\}_i} v_\alpha + \eta_{\mathrm{v},\alpha}, \tag{5.3}$$

where both quantities are assumed to be Gaussian and independently distributed for simplification[12]. The averaging of multiple velocities yields a lower variance when compared to the variance of single samples. For aggregated quantities by the arithmetic mean, the reduced variance is given by

$$\mathrm{Var}(v_i) = \sigma^2_{v_i} = \frac{\sigma^2_{v_\alpha} + \sigma^2_{\eta_\mathrm{v}}}{q_i} \tag{5.4}$$

where q_i denotes the momentary vehicle flow in the given aggregation interval i. That should be considered when adding measurement noise to aggregated, synthetically generated detector data.

Increasing Time-Aggregation Intervals

The individual vehicle velocities are aggregated to one-minute intervals by our data provider. For certain analyses, it may be necessary to increase the time-aggregation interval, e.g. from one minute to five or 15 minutes. The increase of the aggregation interval can be seen as illustrated in Fig. 5.3. This process is composed of two steps, a moving average (MA) step, followed

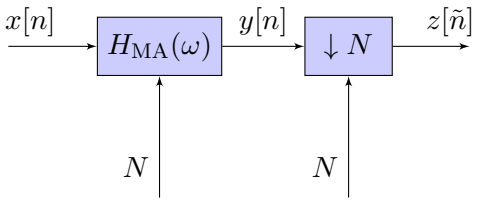


Figure 5.3: Illustration of a system for increasing the aggregation time-interval of a quantity x. The system is composed of two steps: x is filtered by a moving-average filter with order N, followed a down-sampling step with the same rate N.

by a down sampling step. Both operations must be performed with the same parameter N, determined by the respective aggregation intervals T_{aggr} of the signals $x[n]$ and $z[\tilde{n}]$ as following

$$N = \frac{T_{\mathrm{aggr},z}}{T_{\mathrm{aggr},x}}. \tag{5.5}$$

The moving average can be realised as finite-impulse-response (FIR) filter, denoted by

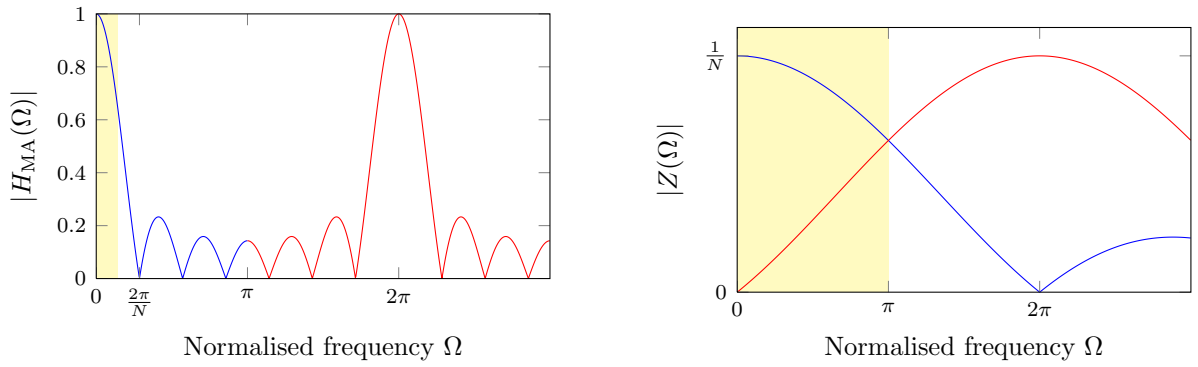
$$h_{\mathrm{MA}}[n] = \begin{cases} \frac{1}{N} & 0 \le n < N \\ 0 & \text{otherwise,} \end{cases} \tag{5.6}$$

or by its respective discrete-time Fourier transform

$$H_{\mathrm{MA}}(\omega) = \frac{1}{N} \frac{e^{-j\omega N/2}}{e^{-j\omega/2}} \frac{\sin\left(\frac{\omega N}{2}\right)}{\sin\left(\frac{\omega}{2}\right)}. \tag{5.7}$$

[12] While this fact is not true for velocities $> 100\,\mathrm{km\,h^{-1}}$ as stated in Tab. 5.2.

The magnitude of eq. (5.7) is illustrated in Fig. 5.4(a). The consecutive down-sampling step removes N samples, and as a consequence, (1) reduces the signal energy and (2) stretches the spectrum by the factor N [82]. The resulting, stretched main-lobes for the overall system are illustrated in Fig. 5.4(b). In order to prevent aliasing in $z[\tilde{n}]$ it is necessary to formulate a



(a) MA filter

(b) Stretched lobes

Figure 5.4: Spectral characteristics of the system system illustrated in Fig. 5.3. In avoidance of aliasing in $z[\tilde{n}]$ it is necessary to limit the maximum frequency in $x[n]$ to the highlighted frequency range. Only the two most important contributions to $Z(\Omega)$ are shown. In principle, all side-lobes from the replicated spectra are present in (b) as well.

criterion for the maximum admissible frequency in $x[n]$ as follows

$$f_{\mathrm{max},x} \leq \frac{1}{2 \cdot T_{\mathrm{aggr},z}}. \tag{5.8}$$

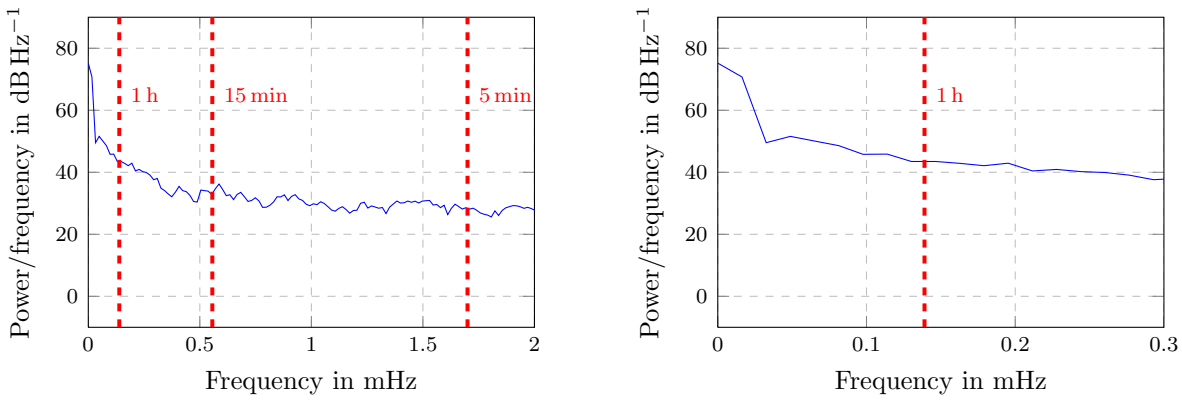
If this criterion is not fulfilled it is mandatory to limit the maximum frequency $f_{\mathrm{max},x}$ by preceding low-pass filtering. Fig. 5.5 shows the power spectral density (PSD) [83] of exemplary one-minute aggregated velocity data obtained by a stationary detector[13]. The dashed vertical lines indicate $f_{\mathrm{max},x}$ for common aggregation time-intervals $T_{\mathrm{aggr},z}$. Depending on the distance between the maximum of the PSD to the components greater than $f_{\mathrm{max},x}$, certain errors may be introduced. The example shows that especially for large changes of the aggregation time-interval, e.g. from one minute to one hour, the criterion of eq. (5.8) might become of importance. In this case, the distance of the aliased components to the maximum is less than 40 dB and thus the introduced errors can be expected to be $> 1\,\%$.

Correction of Time-mean Velocity Bias

As mentioned earlier, stationary sensors aggregate velocities with the time-mean average. The fundamental relation of traffic flow on the other hand, is based on the space-mean velocity (see Sec. 2.2) as well as all mentioned traffic reconstruction methods. Further, the time-mean speed V_t is never less than the space-mean speed V_s, as stated in eq. (2.4). There are several methods to estimate the space-mean speed from a given time-mean speed [3, 84, 85].

Subsequent assumptions can be used to specify ranges of the bias. Under the assumption that V_t and V_s are similarly distributed, they share the same standard deviation ($\sigma_t \approx \sigma_s$). Equation (2.4) can be rearranged to a quadratic equation in the normalised form. The solution

[13] The PSD was obtained by Welch's method using the following parameters: signal length: 5890 samples, window type: Hamming, window length: 1024 samples, overlap: 50 %.



(a) (b) Detail view

Figure 5.5: Power spectral density of one-minute aggregated velocity data from May 17, 2013 15:00 until May 23, 2013 20:00 at road kilometre $x = 4800\,\mathrm{m}$ of the motorway A4, driving direction west (cf. Fig. 7.9). The dashed vertical lines indicate the maximum admissible frequency in $x[n]$ for common aggregation time-intervals $T_{\mathrm{aggr},z}$.

is given by

$$V_s = \mathrm{E}\{V_t\} \cdot \left(\frac{1}{2} + \sqrt{\frac{1}{4} - \left(\frac{\sigma_t}{\mathrm{E}\{V_t\}} \right)^2} \right) \tag{5.9}$$

$$V_s \approx V_t \cdot \left(\frac{1}{2} + \sqrt{\frac{1}{4} - CV_V^2} \right) = V_t \cdot const \tag{5.10}$$

where V_t is the aggregated time mean speed and CV_V is the coefficient of variation

$$CV_V = \frac{\sigma_t}{\mathrm{E}\{V_t\}}. \tag{5.11}$$

A lower CV_V means less fluctuations in the traffic flow: $CV_V < 10\,\%$ for synchronised traffic, $20\,\% < CV_V < 70\,\%$ under free-flow conditions, and $CV_V \approx 100\,\%$ for urban arterial roads in the US [86]. The coefficient of variation can be estimated from existing data; additionally some ranges have been mentioned in literature:

- 15-25 % for US freeways and expressways [87, p. 2-10],
- around 16 % for German motorways [3, p. 57],
- 5-20 % of an example motorway in the Netherlands [1, p. 127].

On this basis, a CV_V of 5-20 % yields without any correction an over estimation of the space-mean speed V_s of 0.3-4.4 %. Under the assumption of a constant, well known CV_V, e.g. 16 %, it is possible to realise a simple correction scheme based on one multiplication. Please note that the simplifying assumption $\sigma_t \approx \sigma_s$ is not valid in general; on the contrary, the relationship between the variances is approximately a bilinear function which leads to correction schemes that show a consideration of the momentary traffic density [85].

5.3 Floating Cars

Data from floating cars (FC) has to be obtained in the form of eq. (4.9), but in contrast to data from stationary detectors

1. data-points are not as regularly distributed over the spatio-temporal plane, and
2. in general, the positions x_i and velocities v_i are not available directly.

In this section we discuss the estimation of the FC position in Sec. 5.3.1 in the one-dimensional road coordinate system, followed by the estimation of the velocity in Sec. 5.3.2.

5.3.1 GPS based Position Estimation

The positions of floating cars can be determined by a mechanical odometer or the satellite-based Global Positioning System (GPS)/Standard Position Service (SPS). An odometer accumulates the travelled distance over time, and therefore comes with the drawback of error accumulation. With GPS/SPS, on the other hand, the position estimation is carried out anew every time step. SPS has the capability to measure time, coordinates in the WGS84 reference coordinate system, in some cases speed, and several performance indicator values, i.e. the DOP-values [88].

Position Accuracy

Basically, GPS/SPS works with a pseudo-range measurement between the GPS-receiver and all visible GPS-satellites. At least four satellites are necessary to estimate the position and time without ambiguity, as indicated in Fig. 5.6. Multiple error sources are known that affect the

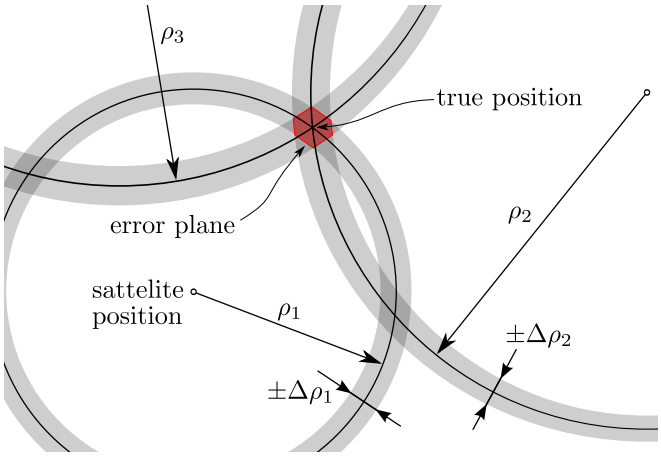


Figure 5.6: Principle of GPS pseudo-range position estimation process. At least three visible satellites are necessary to estimate the position without ambiguity. Uncertainties in the pseudo-range measurements result in an error plane around the true position [88].

pseudo-range measurement [88, 89]. These include errors in the satellite position, clock errors, atmospheric effects, interference effects, relativistic effects, and noise. Accordingly, a model for pseudo-range measurement P can be written as [90]

$$P = \rho + c(dT + dt) + d_{ion} + d_{trop} + e, \tag{5.12}$$

where ρ is the correct signal delay, c is the speed of light in vacuum, dT and dt are the time delays on the receiver and the satellite respectively, d_{ion} and d_{trop} are the time delays in the ionosphere and the troposphere respectively, and e represents the measurement noise and unmodelled multipath signal propagation. In addition, numerical errors occur in the computation of the receiver position.

Due to this wide variety of possible error sources the quality of the GPS-SPS may vary greatly. Commonly used accuracy measures for GPS positioning [88, 91] are summarised in Tab. 5.3. Experiments show that the distribution of horizontal GPS position error can be

Measure		**Definition**
CEP	Circular error probable	The radius of a circle, centred at the antenna position, containing 50 % the points in a scatter plot.
(67 %, 95 %)	Horizontal (67 %, 95 %) accuracy distribution	The radii of two circles, centred at the antenna position, containing 67 % and 95 % of the points in a scatter plot.
rms_1	Root mean square error in one dimension	The square root of the average of the squared error in one dimension, e.g. the north error.
d_{RMS}, rms_2	Root mean square error in two dimensions	The square root of the average of the squared horizontal error.

Table 5.3: Commonly used GPS accuracy measures [88, 91].

assumed as bivariate Gaussian distribution. Furthermore, when assuming uncorrelated errors of the horizontal coordinate estimates, rms_1 represents the standard deviation σ. The χ^2-distribution describes the distribution of the sum of squares of normally distributed random variables. This property might be exploited to convert one accuracy measure to an other [91]. Modern GPS data-loggers specify a CEP of around 3 m [92, 93].

As already mentioned above, spatial coordinates are obtained in the WGS84-based polar coordinate system as longitude and latitude. The spatial resolution on the spherical surface is not directly related to the numerical resolution of the coordinates. This relationship varies with subject to the particular coordinates. Please find a detailed description in Appendix A.

Trajectory Matching

All presented traffic state estimation methods incorporate spatial data based on a one-dimensional road coordinate system. But floating-car trajectories are obtained in the non-Cartesian two-dimensional WGS84 reference coordinate system. Some literature is available to match GPS-trajectories to a given road-network, for example the method [94] by Lou et al., whereby an open source implementation is available[14].

On the basis of the geometric properties of our test-site and the applied approach for flow field reconstruction, it is possible to formulate simplifying requirements for the trajectory matching method: (1) the test-site is a continuous motorway stretch without branches and a characteristic geometry (from north-west to south-east); (2) traffic flow reconstruction is performed on a macroscopic level, and thus, no lane information is required. Accordingly, we propose a simplified method for matching floating-car trajectories based on nearest neighbour search on WGS84 coordinates. Note that the numerical distances in longitude and latitude do not reflect the spatial distances directly. We provide additional information about this fact in Appendix A. For comparing short spatial distances, e.g. in the scale of 100 m, on the basis of WGS84 coordinates it is adequate to normalise every point $X_i = [lon_i, lat_i]^T$ with

$$X_{i,\text{norm}} = \begin{bmatrix} \cos\left(\overline{lat}\right) & 0 \\ 0 & 1 \end{bmatrix} \cdot X_i, \tag{5.13}$$

[14] `https://code.google.com/p/traveltimeanalysis/`

where $\overline{lat}$ is the average latitude over all points. Accordingly, we propose the following scheme for the A4 motorway:

1. For each driving direction, remove trajectory points with a maximum numerical distance in longitude and latitude, e.g. $\varepsilon_{\text{lon,lat}} = 2 \times 10^{-4}$.
2. Select the road segment with the smallest numerical distance.
3. Check the driving direction of the floating-car. Especially the coordinates of the A4 are monotonous decreasing as illustrated in Fig. 5.2(b). This property can be tested with the *Pearson product-momentum correlation coefficient* [95] denoted by

$$r_{x,y} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}}, \tag{5.14}$$

   with $\{x_i\} = \{lon_i\}_{i=1}^{N}$ or $\{x_i\} = \{lat_i\}_{i=1}^{N}$ and $\{y_i\} = 1 \dots N$. This yields rules to discriminate between both driving directions as summarised in Tab. 5.4.
4. Finally, the floating car coordinates (lon_i, lat_i) are converted to the one-dimensional road coordinates x_i by nearest-neighbour search in the road model from eq. (5.1).

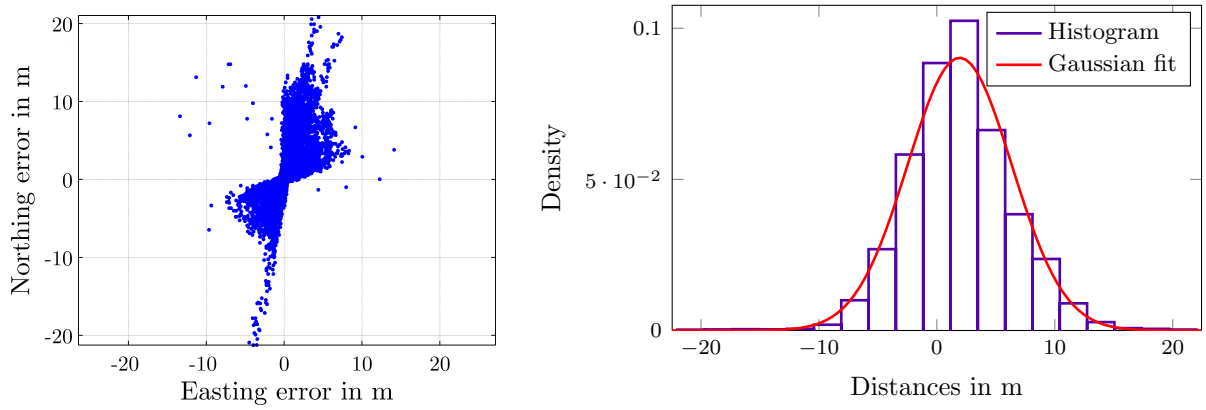
	Driving direction	
	south-east	north-west
r_{lon}	> 0	< 0
r_{lat}	< 0	> 0

Table 5.4: Proposed rules for discriminating the driving directions for the Austrian A4 motorway.

As already mentioned, the GPS position error can be assumed as bivariate Gaussian. Thus, the distances between the GPS points and the matched road points should follow a one-dimensional Gaussian distribution as well. The distances between the assigned road segment and the corresponding high-resolution GPS trajectories[15] obtained for the project IMPAKT are illustrated in Fig. 5.7. The spatial distribution of errors in Fig. 5.7(a) reflects the orientation of the A4 motorway. The distribution of absolute distances is shown in Fig. 5.7(b), where the distances from data-points below the road, i.e. in the south-west quadrant of Fig. 5.7(a), are counted as negative. The positive bias is caused by the unequal distribution of driving directions: only 11 trajectories are heading direction south-east, while 23 trajectories are heading direction north-west.

As stated, the estimated FC position x_i is comprised of a position error, but the algorithms presented in Sec. 4.5 ignore the incorporation of such errors. Nonetheless, position errors contribute indirectly to error of the measured velocities. Because of this, the error variances of estimated velocities may be higher as discussed in the subsequent section.

[15] With a sampling time-interval of $\Delta T = 1\,\text{s}$.



(a) Scatter plot

(b) Distribution of distances

Figure 5.7: Matching distances between real-world GPS points and road model. The bow-tie-shape in (a) reflects the spatial orientation of the road, i.e. from north-west to south-east. The distribution of distances incorporates a slight bias of $\overline{\delta} = 1.9\,\mathrm{m}$ caused by the unequal distribution of driving directions in the dataset.

5.3.2 GPS based Velocity Estimation

The vehicle speed is an important physical quantity for the presented data-fusion algorithms in this work. Depending on the used GPS receiver, speed may be measured directly by the receiver or has to be estimated from recorded track-points later. Direct speed estimation in GPS receivers is usually done by using the Doppler effect. In the following we describe different concepts of GPS based speed estimation methods.

Velocity Estimation based on the Doppler Effect

GPS receivers continuously track the the carrier frequencies of all GPS satellites in range. The difference between those tracked frequencies and the well known carrier frequencies is the so called Doppler-shift. The knowledge of the Doppler-shift might be used to calculate the relative speeds between the satellites and the GPS receiver, and consequently the three-dimensional velocity of the GPS-receiver. For a summary of the computation we refer to [88]. Modern GPS receivers can estimate the velocity with an accuracy of about $\pm 0.1\,\mathrm{m\,s^{-1}}$ [92, 93]. In contrast to position estimation, the speed estimation is more robust against atmospheric effects.

Velocity Estimation based on Coordinates

In some GPS receivers, the velocity estimation is not implemented. Thus, the vehicle velocity between two consecutive points passed at times t_1 and t_2 ($t_1 < t_2$) can be estimated from their coordinates

$$\mathbf{v}_{1,2} = \frac{d\mathbf{u}}{dt} = \frac{\mathbf{u}(t_2) - \mathbf{u}(t_1)}{t_2 - t_1}, \tag{5.15}$$

where $\mathbf{u}(t) = [x(t), y(t)]^T$ denotes the position vector in Cartesian coordinates at given time. Compared to a recorded trajectory from a moving GPS receiver, the geometry of a high rank road stretch follows a smooth course. As a consequence of position errors orthogonal to the driving direction, the measured GPS trajectory would meander around the real vehicle trajectory as illustrated in Fig. 5.8. Because of the greater length of the measured trajectory in comparison to the real trajectory, the velocity would be over-estimated systematically.

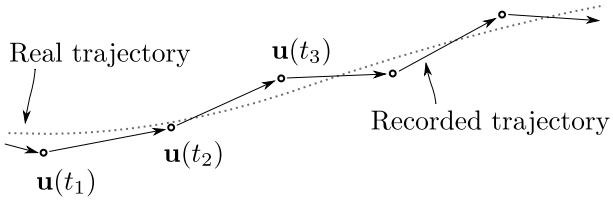


Figure 5.8: Over-estimation of the way travelled caused by coordinate errors orthogonal to the current moving trajectory.

Several methods can be used to improve the speed estimation from GPS trajectories. All of these methodologies use models that correct errors in the estimation process: E.g. a vehicle can only drive on the road surface, therefore the trajectory is limited to a spatial area; or a vehicle follows the laws of mechanics and can only change its speed in certain ranges, therefore one can apply a kinematic vehicle model. These two possibilities are described in the subsequent paragraphs.

Speed Estimation based on Coordinates combined with a Kinematic Vehicle Model

The motion of bodies can be described using the theory of Kinematics. If the forces on a body are completely known, its motion can be described by a differential equation. For the solution of such a differential equation the initial conditions of the kinematic system must be known. Generally, it is difficult to determine the initial state of a kinematic system completely. With Kalman Filtering techniques, on the other hand, one can determine the state of a kinematic system in an iterative manner [96]. These filtering techniques have been introduced in Sec. 4.2.

In order to get a process model, the *kinematic position vector*

$$\mathbf{X}_k = [x_k, y_k, z_k]^T , \tag{5.16}$$

including the three-dimensional coordinates, the *kinematic velocity vector* $\dot{\mathbf{X}}_k = \frac{\partial \mathbf{X}_k}{\partial t}$, and the *kinematic acceleration vector* $\ddot{\mathbf{X}}_k = \frac{\partial^2 \mathbf{X}_k}{\partial t^2}$ can be used. With these three vectors one can build models with rising complexity. One of the most common models applied to land vehicles is the *four-state constant velocity model* [96, 97] given by

$$\mathbf{X}_k = \mathbf{X}_{k-1} + \Delta t \cdot \dot{\mathbf{X}}_{k-1} + \frac{1}{2}\Delta t^2 \cdot \mathbf{w}_{k-1}, \tag{5.17}$$

where $\Delta t = t_k - t_{k-1}$ and the acceleration is included in the noise-term $\mathbf{w}_{k-1}$. The model assumes a vehicle moving at almost constant velocity and almost constant direction. Acceleration, jerks, and uncertainties in the model are represented by a represented with a Brownian motion along each canonical axis. For the Kalman prediction step, the linear kinematic model has to be written in state-space formulation

$$\mathbf{x}_k = \mathbf{T}_{k-1} \cdot \mathbf{x}_{k-1} + \mathbf{B}_{k-1} \cdot \mathbf{w}_{k-1}, \tag{5.18}$$

where T is the state-transition matrix, $\mathbf{B}$ is the noise-input matrix. The state vector contains both, position- and velocity-state variables:

$$\mathbf{x}_k = [x_k, y_k, z_k, \dot{x}_k, \dot{y}_k, \dot{z}_k]^T . \tag{5.19}$$

A detailed description of the algorithm including the derivation of the covariance matrices, and a stability analysis for the four-state constant velocity model is given in [96].

The previously described four-state constant velocity model is simple and provides a good

introduction to the underlying methodology but may lack descriptive power for various situations. Examples of more complex motion models and mode switching models can be found in literature in the respective literature [97–99].

Increased Robustness by Incorporation of a Road Model

A different approach on robust GPS based speed estimation is to project the obtained GPS trajectory on a given road model and compute the velocity from the corresponding road kilometres. This procedure can be summarised as follows:

1. Obtain the consecutive GPS coordinates (lon_1, lat_1) and (lon_2, lat_2).
2. For each GPS coordinate look for the nearest neighbour in the road model to get the corresponding road kilometres $x_{R,j}$ and $x_{R,k}$ for the coordinates (lon_1, lat_1) and (lon_2, lat_2), respectively.
3. Calculate the vehicle velocity with the matched coordinates $x_{R,k}$ and $x_{R,j}$.

The overall procedure is illustrated in Fig. 5.9. The FC velocity is defined by the trajectory

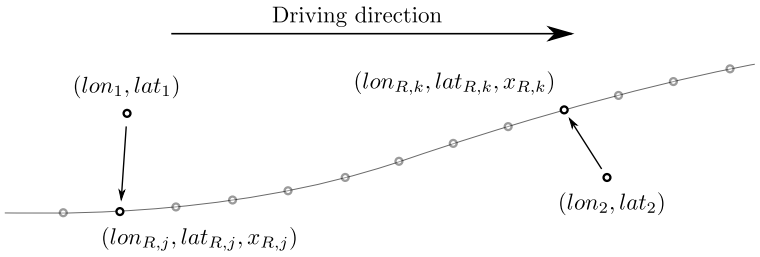


Figure 5.9: Speed estimation based on GPS coordinates (lon_i, lat_i) and a road model $(lon_{R,j}, lat_{R,j}, x_{R,j})$. The GPS coordinates are projected on the road model by nearest neighbour search, followed by a velocity calculation using the road kilometre $x_{R,j}$.

gradient. We propose the use of following two discrete gradients:

Piece-wise Linear Interpolation. Calculation of the velocity with a simple forward- and backward-difference, respectively denoted by

$$v_k = \Delta x_k(t_k) \qquad \text{and} \qquad v_k = \nabla x_k(t_k). \tag{5.20}$$

The velocity is given by the gradient of the preceding linear interpolant. This method is suitable for high-resolution trajectories, comprised of low noise.

Piecewise Cubic Hermite Interpolating Polynomial (PCHIP). For higher sampling time-intervals we propose the weighted harmonic average between the preceding and succeeding slopes as defined for the the shape-preserving PCHIP by Moler [100] for the calculation of the velocities with

$$v_k = \left[\frac{1}{w_1 + w_2} \left(\frac{w_1}{\varphi_{k-1}} + \frac{w_2}{\varphi_k} \right) \right]^{-1}, \tag{5.21}$$

where the slopes are denoted by the forward difference

$$\varphi_k = \Delta x_k(t_k), \tag{5.22}$$

the weights are denoted by

$$w_1 = 2h_k + h_{k-1}, \qquad w_2 = h_k + 2h_{k-1}, \qquad \text{and} \qquad h_k = t_{k+1} - t_k. \tag{5.23}$$

In general, a PCHIP is any piecewise cubic polynomial that interpolates the given data, and has specified derivatives at the interpolation points.

Both interpolation methods, as illustrated in Fig. 5.10, assume the data-points as exact and do not consider position errors. Overall, the uncertainty of the methods is influenced by the quality

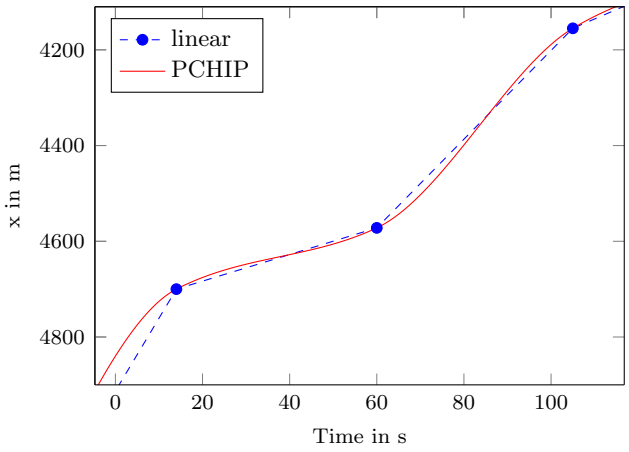


Figure 5.10: Comparison of interpolated FC trajectories with piece-wise linear and PCHIP interpolants.

of the road model, the sampling intervals Δx and ΔT of the road model and the GPS trajectory, respectively and the variance of the GPS coordinates $\sigma_{lon,lat}$. Each of these uncertainties and the nearest neighbour search affect the uncertainty of the estimated speed. For continuous equations, the propagation of uncertainties can be calculated straightforwardly. The nearest neighbour step is a non-differentiable function and breaks this property. Consequently, the propagation of uncertainties is studied with a Monte-Carlo based sensitivity analysis as described in Chapter 7.

5.4 Travel Time Data

As already mentioned, performance evaluation based on global error measures is only possible in areas where reference measurements are available. In general, such reference measurements are expensive to obtain. An alternative for performance-evaluation is to extract some features from the estimated spatio-temporal field that are easier to measure.

The *travel time (TT)* over a given road section can be a meaningful measure for traffic flow. For instance, the ratio between peak travel-time and free travel-time can be seen as an indicator for congestion, as the travel-time is influenced by the overall traffic conditions on that particular section. Travel-times can be measured relatively inexpensive with ANPR systems, i.e. stationary sensors are only required at the beginning and the end of a given road section.

As already mentioned above, the test-site A4 motorway includes a section for travel time measurements. Two ANPR cameras were installed on the outer right lane for travel time measurements on the section $X_{\mathrm{BC}} = 7000 - 200\,\mathrm{m}$. Thus, the ANPR system covers only a fraction of the passing vehicles. In particular, velocities on the outer right lane are generally lower when compared to the average velocity over all lanes, and as a consequence, the obtained travel times may contain a positive bias. The TT data is provided by the AIT aggregated in five minute intervals, but it is possible to request individual vehicle TT if required for future research projects.

In the following we discuss a method for extracting TT from spatio-temporal fields as obtained by traffic state reconstruction algorithms described in Chapter 4.

5.4.1 Total Travel Times from Virtual Trajectories

The reconstruction of travel-times from a spatio-temporal velocity field is similar to the estimation from (very dense) stationary detectors. In [101], Ni and Wang give an overview of travel-time estimation based on spatio-temporal ITS data. In contrast to other algorithms, where the aggregated traffic quantities are assumed to be constant along the cell boundaries, data reconstructed by the ASM-family is only defined point-wise on a spatio-temporal grid (see Fig. 5.11). Therefore, some of their mentioned algorithms are not applicable, e.g. [102]. Under

(a) Section wise (b) Point wise

Figure 5.11: Difference between section-wise and point-wise defined spatio-temporal data. Section-wise data is defined over a spatio-temporal region as shown in (a), whereas point-wise data is solely defined on points as shown in (b)

the assumption of a point-wise defined spatio-temporal velocity field with sufficiently small sampling distances we propose the following algorithm for travel-time estimation which is similar to the method by Ni and Wang [101].

The trajectory of a virtual vehicle $\{t, x(t)\}$ can be expressed by the differential equation

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = V(t, x(t)), \tag{5.24}$$

where $V(t, x)$ represents a spatio-temporal velocity field. In order to obtain the travel-time between two spatial points A and B, one can rearrange eq. (5.24) to

$$\frac{\mathrm{d}t(x)}{\mathrm{d}x} = \frac{1}{V(t(x), x)}. \tag{5.25}$$

By the integration of eq. (5.25) it is possible to calculate the total travel time for the road section $[x_A, x_B]$ and the starting time t_A as following

$$TT(t_A, x_A, x_B) = t_B - t_A = \int_{t=t_A}^{t_B} \mathrm{d}t = \int_{x=x_A}^{x_B} \frac{1}{V(t(x), x)} \mathrm{d}x, \tag{5.26}$$

where it is only possible to solve the integral over the spatial coordinate. In order to calculate the travel times from a discrete spatio-temporal velocity field $\{v_{i,j}\}$ we propose to estimate discrete vehicle trajectories with piece-wise constant speeds $\{t_k, x_k, v_k\}$ as illustrated in Fig. 5.12. Accordingly, the integral in eq. (5.26) becomes

$$\hat{TT}(t_A, x_A, x_B) = \hat{t}_B - t_A = \sum_{k=1}^{K} \delta t_k = \sum_{k=1}^{K} \frac{\delta x_k}{v_k}, \tag{5.27}$$

where δt_k and δx_k are the instantaneous step-widths bounded by the constant sampling intervals of $\{v_{i,j}\}$ with $\delta t_k \leq \Delta t$ and $\delta x_k \leq \Delta x$. The instantaneous sampling-widths are determined by

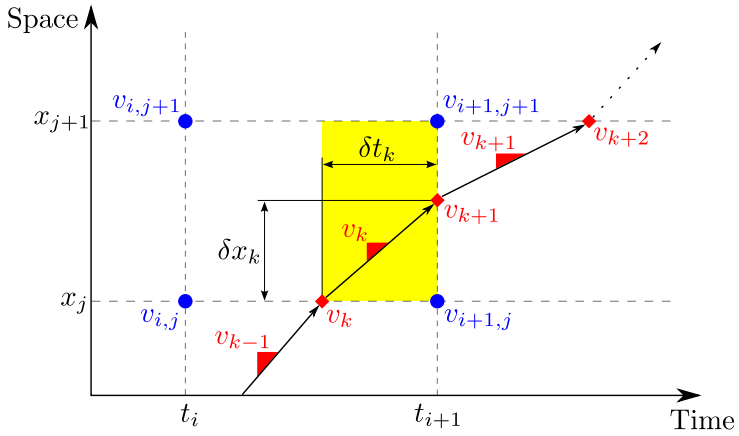


Figure 5.12: Scheme for reconstruction of trajectories of virtual vehicles $\{t_k, x_k, v_k\}$ based on a regular grid of velocities (blue), piece-wise constant velocities (red) and variable step-widths δt_k and δx_k. The residual cell for the time-step k is highlighted in yellow.

the size of the *residual cell* (see Fig. 5.12) and the instantaneous velocity v_k which is interpolated from $\{v_{i,j}\}$. Width $w_{t,k}$ and height $w_{x,k}$ of a residual cell for a time-step k are denoted by

$$w_{t,k} = \left\lceil \frac{t_k}{\Delta t} \right\rceil \Delta t - t_k \qquad \text{and} \tag{5.28}$$

$$w_{x,k} = \left\lceil \frac{x_k}{\Delta x} \right\rceil \Delta x - x_k. \tag{5.29}$$

Finally, the instantaneous step-widths are given by

$$\delta x_k = \begin{cases} v_k \cdot w_{t,k} & v_k < \frac{w_{x,k}}{w_{t,k}} \\ w_{x,k} & \text{otherwise} \end{cases} \tag{5.30}$$

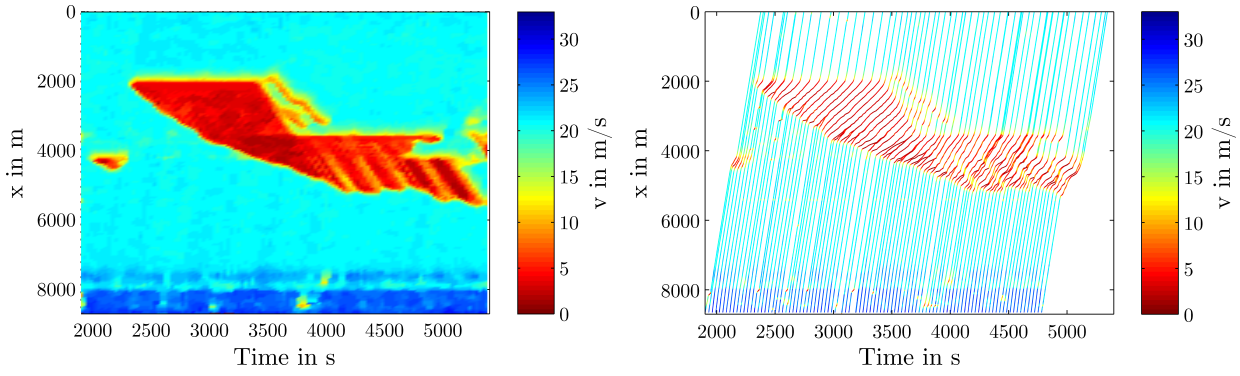
and

$$\delta t_k = \begin{cases} w_{t,k} & v_k < \frac{w_{x,k}}{w_{t,k}} \\ \frac{w_{x,k}}{v_k} & \text{otherwise.} \end{cases} \tag{5.31}$$

With eqs. (5.27) to (5.31) it is possible to calculate virtual vehicle trajectories and corresponding travel times simultaneously.

A simulated spatio-temporal velocity field with its corresponding virtual trajectories is shown in Fig. 5.13 and the extracted travel times are shown in Fig. 5.14. In areas with congestion, some trajectories are attracted to their neighbouring trajectories (see Fig. 5.13(b)). As a consequence the extracted total travel times fluctuate over time. In order to suppress those fluctuations we propose the smoothing with a moving average filter. Further, travel times from ANPR systems are typically aggregated to 5 min intervals. In order to be comparable, the raw TT has to be smoothed and aggregated with the same time interval. Raw travel times and the respective smoothed versions are illustrated in Fig. 5.14. The aggregated block-average can be seen as a down-sampled version of the moving average. Both averaged versions vary slightly due to implementation details[16].

[16] The span of the moving average is rounded to the nearest odd number.



(a) Ground truth velocity field (b) Virtual trajectories

Figure 5.13: A simulated spatio-temporal velocity field in (a) and generated virtual trajectories for travel time estimation in (b).

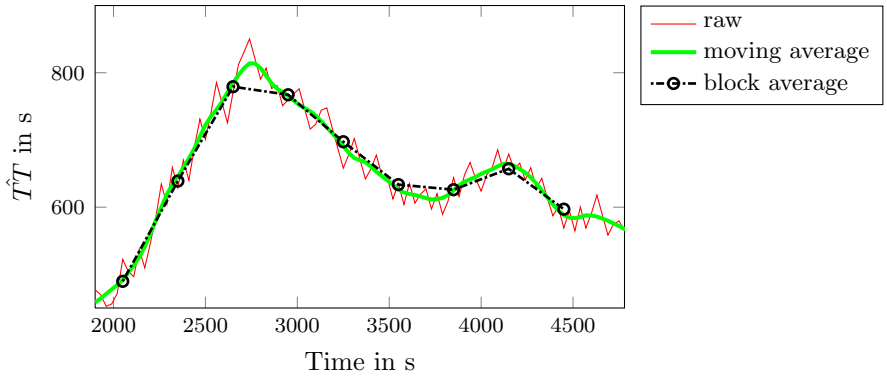


Figure 5.14: Estimated total travel times based on the discrete spatio-temporal velocity field $\{v_{i,j}\}$ from Fig. 5.13.

5.5 Microsimulation

Validation of measurements on a traffic flow presents an enormous challenge for researchers. One possibility to get around this issue is the use of synthetic data generated by microsimulation. Nowadays, microsimulation tools serve as a basic tool in the field of traffic engineering and science. Thus, many different microsimulation tools have evolved up to now.

Although there are many microsimulation tools, we decided to use *VISSIM*[17] for our work, since a license for version 5.40 is available at the ISV. VISSIM [103] is a microscopic multi-modal traffic simulation software by the PTV Planung Transport Verkehr AG[18]. Traffic is modelled by the psychophysical car-following model by Wiedemann. Initially developed for urban traffic, VISSIM later has been calibrated for traffic on motorways. Today, PTV is numbered among the global market leaders on traffic simulation.

As part of the project IMPAKT, the ISV developed a fully calibrated VISSIM model of the Austrian A4 motorway. The model consists of the main road section between road kilometre 0 km (Knoten Prater) and road kilometre close to 10 km (Vienna International Airport) and the respective on- and off-ramps as illustrated in Fig. 5.15. Coordinates in VISSIM are represented in a Cartesian coordinate system with metric coordinates; no geographic coordinate reference systems are supported. The geometry of our VISSIM model is based on the deprecated *MGI*

[17] **V**erkehr **I**n **S**tädten **SI**mulations**M**odell
[18] `http://www.ptvgroup.com`

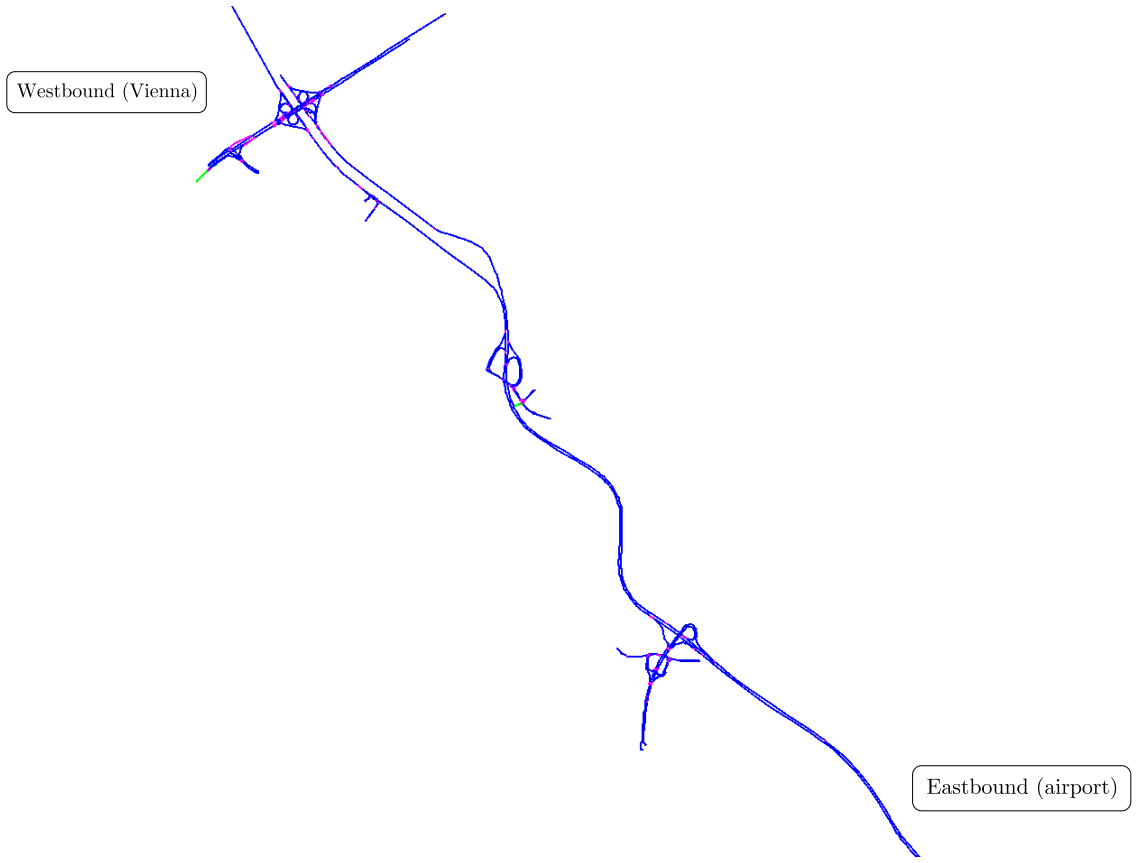


Figure 5.15: Screenshot of the VISSIM model used representing our test-site, the Austrian A4 motorway. The model consists besides the main road also of the respective on- and off-ramps (cf. Fig. 5.1)

(Ferro) / Austria East Zone (EPSG:31283) [104] geodetic datum.

The model can produce realistic traffic conditions when compared to a working day between 17:00 and 18:00. The overall traffic stream is formed by superposition of 82 different traffic streams following static routes. The flow rates of the particular streams were determined in advance with the macroscopic simulation tool PTV-VISUM on the basis of a complete road model of the region and stationary detector data.

A complete simulation run is performed over 10 000 s with a resolution of five steps per second and can be divided into two phases: the preparation phase (e.g. until $t = 1800\,\mathrm{s}$) where the empty road gets filled with individual vehicles, and the main phase (i.e. the remaining time until $t = 7200\,\mathrm{s}$) where the traffic conditions may be assumed as valid. In each case, after simulation time $t = 7200\,\mathrm{s}$ the flows into the model come to a halt and the road empties out again. In order to guarantee reproducibility, VISSIM performs simulations on the basis of a pseudo-random number generator (PRNG), where the behaviour can be influenced by setting an integer random seed variable[19].

It is possible to export different categories of measurements from a VISSIM simulation. Therefore, we are able to generate stationary sensor data, floating car data, as well as ground-truth data for evaluation. Stationary detectors can be added using *Data Collection Points*; floating car data can be exported using the *Vehicle Record*; and as a ground-truth, space- and time-aggregated traffic variables, such as link flow, density and speed, can be exported using *Link Evaluation*. These three types of data all provide the required information to perform a traffic flow reconstruction followed by an evaluation.

[19] Always 12 in the experiments conducted within this thesis.

In order to simulate real-world data it is necessary to (1) transform the data to the appropriate coordinate system and (2) consider the measurement noise introduced by the imperfect sensors as described earlier in this chapter. This is not possible in VISSIM directly and has to be done afterwards. For further implementation details we refer to Sec. 6.2.

5.6 Detailed Record of Available Data

This section offers a compact overview of the whole range of data available, including the detailed attributes, its spatio-temporal extent, and the average data density. Even though only a fraction of the mentioned data was used in our conducted experiments, all data had to be examined with respect to the usability for our thesis. The complete record is provided in Tab. 5.5.

Category	Attributes	Time period	Further information
Stationary detectors	Flow, velocity and headway time; one-minute aggregated and two vehicle classes (codes 34 and 36 in accordance with [105])	01.01.2012 – 27.06.2012	Eastbound: 11 detectors in the section $0.23 - 13.79\,\text{km}$; westbound: 16 detectors in the section $0.85 - 13.50\,\text{km}$
		01.05.2013 – 31.05.2013	Solely westbound: 10 detectors in the section $0.85 - 7.49\,\text{km}$
GPS trajectories	WGS84 coordinates (numerical resolution of four fractional digits) and timestamps obtained from GPS equipped taxis (REFEREE)	01.06.2012 – 28.06.2012	Average trajectory sampling interval $\Delta T = 32\,\text{s}$; 40 trips per hour; limited to driving direction west
		01.03.2013 – 31.05.2013	Average trajectory sampling interval $\Delta T = 22\,\text{s}$; 40 trips per hour; limited to driving direction west
	WGS84 coordinates (numerical resolution of seven fractional digits), timestamps, and respective velocities obtained from probe vehicles (IMPAKT)	15.06.2012 – 21.06.2012	Trajectory sampling interval $\Delta T = 1\,\text{s}$; 92 trips in the respective time-period in the greater Vienna area, but too few trajectories captured on the A4 motorway
Travel times	Five-minute aggregated travel times for five different vehicle classes (codes 6, 7, 8, 10, and 11 in accordance with [105]) on the section $X_{\text{BC}} = 7 - 0.2\,\text{km}$ in driving direction east	05.11.2012 – 05.06.2013	Only those vehicles entering and leaving the section on the outer right lanes were considered
Simulation	Any desired data within the motorway section 0-10 km	7200 s less the preparation time (e.g. 1800 s)	Simulation resolution: five steps per second; random seed: 12

Table 5.5: Detailed record of all datasources available for our thesis.

5.7 Summary

In this chapter we presented our test-site including all associated data-sources. The data was provided with only little additional information by the project partners. Therefore, sources of measurement uncertainties were identified based on a set of common assumptions. We also described transformations of data from one domain to another which are required for particular data sources. Furthermore, we came to the conclusion that some of these transformations are

only allowed under certain conditions, i.a. the change of aggregation time-intervals requires some spectral properties as an important precondition. Finally, we described methods for the estimation of certain data that is not directly available, such as the driving direction, floating car velocities in some cases, or the extraction of travel times from spatio-temporal velocity fields. The subsequent chapter provides further information on implementation issues of the abovementioned preprocessing and estimation algorithms.

6 Implementation

Within our project we implemented software in MATLAB with respect to compatibility with version 2011a or later, including batch-like scripts, simple functions and object-oriented components. The object-oriented parts build the core of the implemented software and are comprised of data-containers for the different data sources as well as implementations of the aforementioned reconstruction algorithms based on the idea of Treiber and Helbing as described in Sec. 4.5. The data-containers hold the data and can carry out important conversions and preprocessing functions. Furthermore, we implemented software for calibration and validation of the specific algorithms. Class-diagrams of all implemented classes are provided in Appendix D.

Subsequently, we describe special aspects of our implementation. These include implementation issues of the traffic-data containers in Sec. 6.1, followed by the microsimulation framework in Sec. 6.2, and the Treiber-Helbing filter variants in Sec. 6.3.

6.1 Traffic-Data Container

With regard to the implementation of data containers, each data category is represented by a separate container-class which provides the appropriate data-fields and data-processing methods. The data from ANPR systems, floating cars, and stationary detectors is made available by the respective providers via CSV-files. Thus, a container-class for a particular data-category is implemented as an abstract class, where the method for parsing a CSV-file with a provider specific format has to be implemented in a derived class.

At the moment, data fields are implemented as separate vectors, where elements with the same index are counted among a common data-point d_i. This fact allows elegant data-filtering on the basis of *logical indexing*[20]. However, the implementation with individual vectors also implies a limitation of the maximum number of data-points given by the available working memory. In order to cope with larger datasets we suggest the use of a file-based database. An example would be *SQLite* and its corresponding MATLAB-binding which is freely available under [106] and [107], respectively.

[20] `https://www.mathworks.com/help/matlab/math/matrix-indexing.html`

6.2 Microsimulation Framework

As already mentioned in Sec. 5.5, both a license of VISSIM and a fully calibrated model of the Austrian A4 motorway from the project IMPAKT are available at the ISV. Subsequently, we describe the realised system for performance evaluation of the implemented data-fusion and traffic flow reconstruction algorithms. The overall evaluation system is illustrated in Fig. 6.1. It is composed of two parts: VISSIM, its respective traffic model, and the output configuration

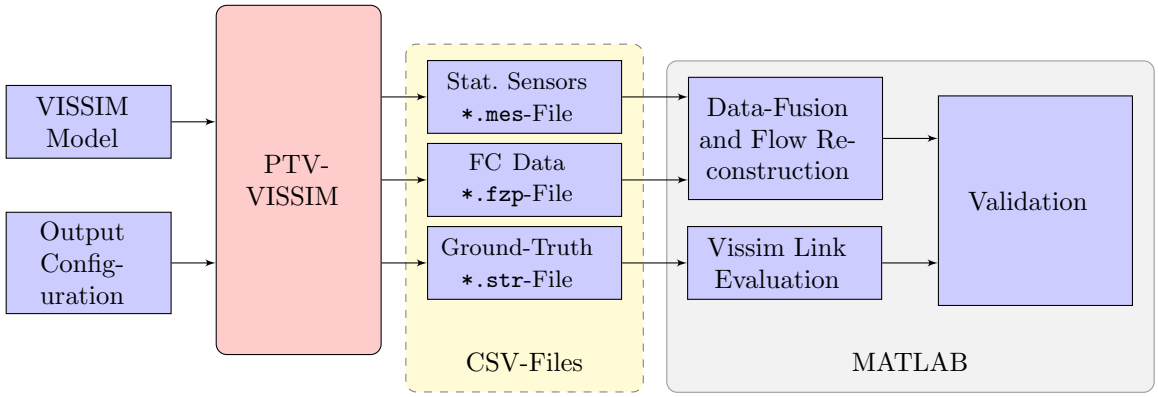


Figure 6.1: Evaluation system based on a microsimulation with PTV-VISSIM. *The data is passed to Matlab using CSV-files: the extensions* *.mes *for stationary detectors,* *.fzp *for floating car data, and* *.str *for the complete link-states.*

on the left; our implemented MATLAB software on the right. Both parts share the traffic-data with CSV-files. Spatio-temporal resolution of the ground-truth data, vehicle composition, the ratio of floating cars, the position of the stationary detectors, and the road conditions have to be configured in the VISSIM model. It is possible to create congested traffic patterns by addition of either

1. *reduced speed areas*,
2. a lower *desired speed decision*, or
3. the reduction of the number of lanes for a short road segment.

The first two possibilities can be configured in the VISSIM graphical interface over a limited period of time, while the third is only possible for the overall simulation run. An alternative is to control VISSIM via the COM programming interface where the model parameters can be modified during the simulation run. Unfortunately, the creation of congestion patterns as mentioned in Fig. 2.3 is not straight-forward and requires a comprehensive trial and error procedure.

In order to be consistent with data from real measurements, coordinates for floating car data and ground-truth data have to be converted initially.

Conversion of Floating-Car Coordinates. The geometry of our VISSIM model is based on the deprecated *MGI (Ferro) / Austria East Zone (EPSG:31283)* [104] geodetic datum. On the other hand, GPS devices store geospatial data as WGS84–based datum in longitude and latitude. Therefore, the simulated vehicle trajectories are transformed to the WGS84 with the widely used Helmert transformation [96]. The Helmert transformation is a three-dimensional similarity transformation of an initial point $X = [x, y, z]^T$ to a transformed point $X' = [x', y', z']^T$ which can be written as

$$X' = C + \mu \cdot R(\omega, \varphi, \kappa) \cdot X, \tag{6.1}$$

where C is the translation vector, μ is a unit-less scale factor, and R is the rotation matrix. R is composed of three orthogonal rotation matrices

$$R = R_x(\omega) \cdot R_y(\varphi) \cdot R_z(\kappa), \tag{6.2}$$

which represent a rotation around each canonical axis, x, y, and z, respectively. Finally, the transformation may be written with seven parameters

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix} + \mu \cdot \begin{bmatrix} 1 & r_z & -r_y \\ -r_z & 1 & r_x \\ r_y & -r_x & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{6.3}$$

and the scale factor $\mu = 1 + m \cdot 10^{-6}$. Normally, the transformation parameters from local geographic reference systems to WGS84 are provided by the national geographic agencies, i.e. the *Bundesamt für Eich- und Vermessungswesen (BEV)*[21] in Austria. The parameters for conversion from EPSG:31283 to WGS84 are summarised in Tab. 6.1. There are already some

c_x	c_y	c_z	m	r_x	r_y	r_z
577.326	90.129	463.919	2.4232	5.137	1.474	5.297

Table 6.1: Helmert parameter for a transformation from the deprecated MGI (Ferro) / Austria East Zone (EPSG:31283) *to the WGS84 reference coordinate system [104].*

libraries that support the transformation between known cartographic projections. In our work we implemented the Helmert transformation using the command-line tool *Cs2cs* from the *Proj.4 Cartographic Projections Library* [108]. In order to call Cs2cs from MATLAB we implemented a wrapper-function[22] working under both, Windows and Unix-based operating systems.

Correction of Systematic Geometric Errors. The geometry of the VISSIM model does not correspond to the real road model in detail. Certain sections of the road, especially in driving direction north-west, have a systematic geometric error as illustrated in Fig. 6.2. We

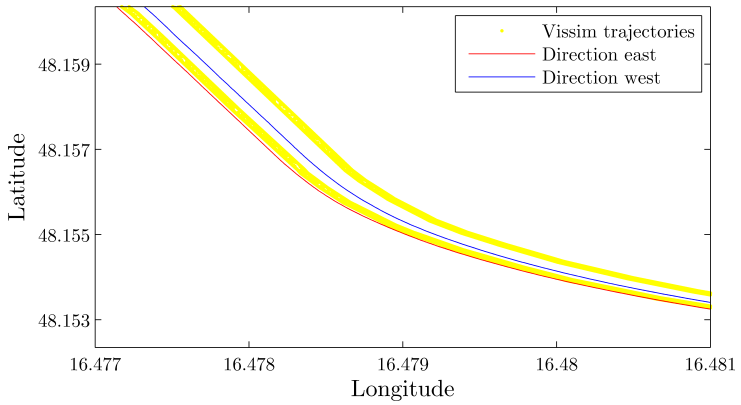


Figure 6.2: Illustration of the systematic mismatch between simulated vehicle trajectories and the road model. Especially the trajectory points in driving direction west do not agree with the real road geometry.

define the systematic error between the vehicle coordinate X_{veh} and the nearest point on the

[21] `https://portal.bev.gv.at/`
[22] `https://www.mathworks.com/matlabcentral/fileexchange/41203-cs2cs-wrapper-for-matlab`

road X_{road} with

$$\Delta X = X_{\text{veh}} - X_{\text{road}}. \tag{6.4}$$

Both components of the systematic error, driving direction north-west are illustrated in Fig. 6.3. These geometric errors may cause problems with certain experiments. In order to correct the

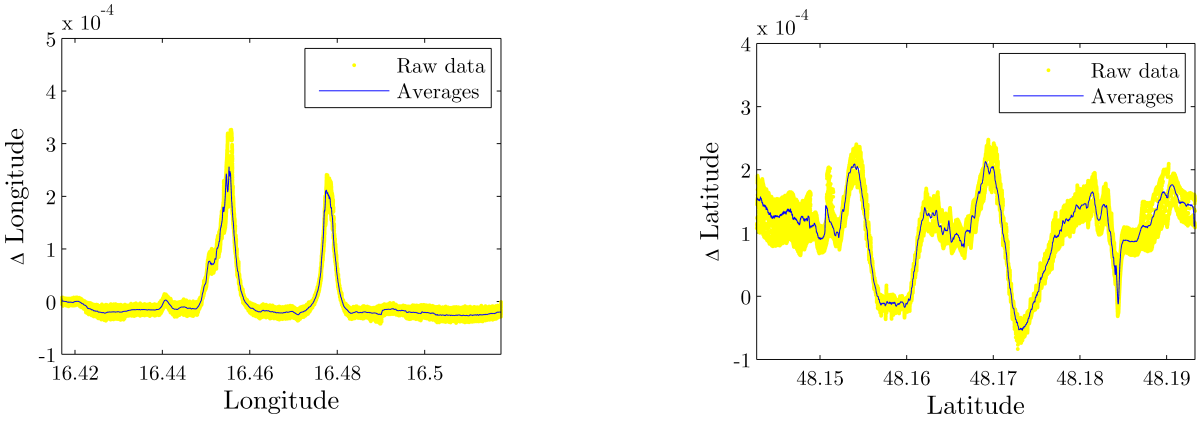
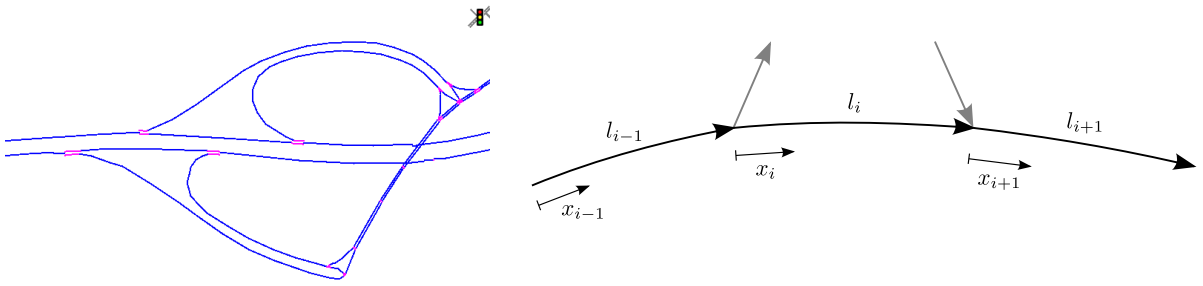


Figure 6.3: Systematic geometric error between simulated trajectories and real road.

systematic errors we propose the correction of the geometric error by the following scheme:

1. Simulate an adequate number of floating car points.
2. Calculate an average error $\overline{\Delta X}$ for each point in the road model.
3. For a new FC trajectory determine the nearest point on the road and subtract $\overline{\Delta X}$ subject to the respective spatial location x.

Conversion of One-Dimensional Road Coordinates. In VISSIM, roads are composed of *links* and *connectors* as illustrated in Fig. 6.4. Every link or connector comes with a local



(a) VISSIM screenshot (b) Scheme of consecutive links

Figure 6.4: A model of a road network in VISSIM is composed of links *and* connectors*, with blue and magenta colour in (a), respectively. Each link comes with its own local one-dimensional coordinate system as illustrated in (b).*

one-dimensional coordinate system. Ground-truth data from Link Evaluation exists initially as space-aggregated traffic variables for each link separately in the local link coordinate system. An analysed motorway section $\{l_i\}_{i=1}^{L}$ may include several consecutive links and connectors as illustrated in Fig. 6.4(b). But ground-truth data is required in an evenly sampled spatio-temporal grid over the whole road section. In order to generate the evenly sampled data we propose the following procedure:

1. For the first link l_1, associate the corresponding road kilometre.
2. For all following links, accumulate the distance information of all intermediate links.
3. Re-sample the spatio-temporal data on an evenly spaced grid. The re-sampling of the overall velocity field realised using MATLAB's `TriScatteredInterp` class[23] which performs a linear interpolation.

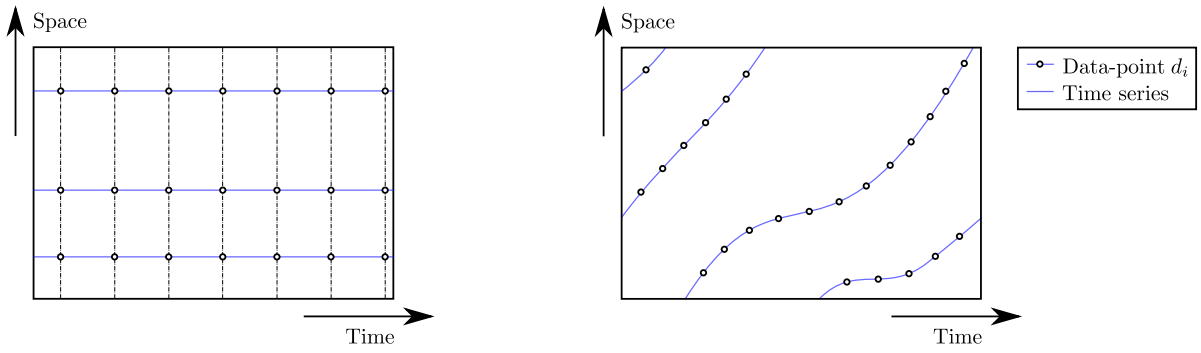
It is a fact, that traffic flow fields should not be interpolated in an isotropic manner since it ignores the characteristic propagation velocities (see Sec. 4.4). However, for sufficient small spatio-temporal sampling distances the propagation velocities are negligible.

6.3 Treiber-Helbing Filter

This section provides information about the implementation of ASM-based algorithms (described in Sec. 4.5) and is structured among two parts: First, we state the requirements of our implementation with regard to the container classes providing the input data; Second, we discuss some aspects of the computational complexity as well as feasible ways of an efficient implementation.

Processable Data Sources

The ASM was initially proposed for stationary detector data only, but was extended to arbitrary spatio-temporal data later. A main difference between data from stationary detectors and data from floating cars is the alignment of the respective time series data: Time series from stationary detectors are aligned horizontally on a spatio temporal grid, whereas time series from floating cars follow the momentary flow velocity in the spatio temporal plane as illustrated in Fig. 6.5.



(a) Stationary detector data (b) Floating car data

Figure 6.5: Comparison of spatio-temporal alignment from different data categories: stationary detector data is lattice-like distributed as shown in (a), whereas floating car trajectories follow the momentary flow velocity as shown in (b).

However, all ASM-based algorithm variants ignore the connection of data via time series and consider every data-point as separate information. Therefore, it is possible to use any data-source as long as: (1) it consists of sufficiently enough data scattered over the spatio-temporal area of reconstruction, and (2) the data-points d_i consist at least of information about road kilometre, time, and velocity (c.f. eq. (4.9)). Accordingly, our implementation is able to perform a reconstruction with data from any container class with the only requirement that the data-fields velocity v, time-stamp *time*, and road kilometre x are declared as public attributes.

[23] `https://www.mathworks.com/help/matlab/ref/triscatteredinterp.html`

Computational Complexity and Efficient Implementation Variants

Although the ASM variants are formulated in continuous domain, reconstruction is performed for implementation purposes point-wise on a predefined spatio-temporal grid

$$(x,t) \in ([x_0^{\mathrm{rec}},\ldots,x_N^{\mathrm{rec}}],[t_0^{\mathrm{rec}},\ldots,t_M^{\mathrm{rec}}]). \tag{6.5}$$

Especially the evaluation of the double-sum in eqs. (4.10), (4.11), (4.17) and (4.18) at every point in (x,t) is computational costly. But the aforementioned equations share a similar structure, which makes implementations with increased efficiency possible. The basic idea behind more efficient implementations is that for the reconstruction of a point $z(x,t)$, the characteristic kernel β_{char} (see eq. (4.16) and eq. (4.13)) decays with increasing distance w.r.t its centre. Thus, it is sufficient to consider only input data z_i in an influence area $A(x,t)$. For $A(x,t)$, we propose the *minimum bounding rectangle (MBR)* of an indicator function $\mathbb{1}_A$ defined by

$$A(x,t) = \mathrm{MBR}\left\{\mathbb{1}_A\left(\beta_{\mathrm{char}}(x,t) \not\approx 0\right)\right\} = \mathrm{MBR}\left\{\mathbb{1}_A\left(\beta_{\mathrm{char}}(x,t) \geq \varepsilon_A\right)\right\}, \tag{6.6}$$

where ε_A defines the minimum weight. The size of $A(x,t)$ depends on the kernel parameter σ and τ, and on the characteristic propagation velocity c_{char} and can be written as

$$\mathrm{area}\left\{A(x,t)\right\} = 4a^2 \cdot \left[\max\left\{\tau, \frac{1}{c_{\mathrm{char}}}\sigma\right\} \times \sigma\right] = const. \tag{6.7}$$

The parameter a is related to the minimum weight ε_A, i.e. for an exponential kernel one can write $a = \log(\varepsilon_A)$. The anisotropic interpolation kernel β_{char} and the size of the corresponding MBR are illustrated in Fig. 6.6. Two possible approaches for an efficient implementation are described below.

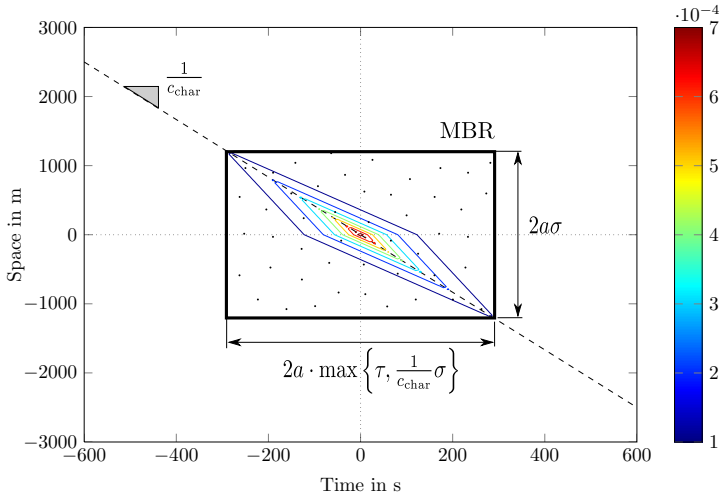


Figure 6.6: Illustration of a two-dimensional anisotropic interpolation kernel β_{char} for a characteristic velocity $c_{\mathrm{char}} < 0$. The respective minimum bounding rectangle (MBR) given by $\beta_{\mathrm{char}} \not\approx 0$.

Direct Implementation

In this section, we describe the direct implementation of the aforementioned double sums. Although slow `for`-loops can be easily avoided via vectorisation, it is necessary to lookup all data-points in range $\{z_i\}_{i\in A(x,t)}$ for a reconstruction of a single point $z(x,t)$. Even though Schreiter et al. discussed the computational complexity of a direct implementation in [109], they ignored the complexity of the lookup step.

At the moment we implemented the lookup with *logical indexing* in MATLAB, which can be seen as a sequential or linear search. The complexity depends linearly on the number n^{in} of input points z_i denoted in big O notation by

$$\text{lookup}(\{z_i\}_{i\in A(x,t)}) \in \mathcal{O}\left(n^{\text{in}}\right) = \mathcal{O}\left(\frac{X^{\text{in}} \cdot T^{\text{in}}}{\Delta x^{\text{in}} \cdot \Delta t^{\text{in}}}\right), \tag{6.8}$$

where X^{in} and T^{in} define the spatio-temporal extend of the input data, and Δx^{in} and Δt^{in} denote the average spacing of the data-points. Thus, the complexity of the overall algorithm can be denoted by

$$ASM_{\text{direct}} \in \mathcal{O}\left(\frac{X^{\text{rec}} \cdot T^{\text{rec}}}{\Delta x^{\text{rec}} \cdot \Delta t^{\text{rec}}}\left[\underbrace{\frac{X^{\text{in}} \cdot T^{\text{in}}}{\Delta x^{\text{in}} \cdot \Delta t^{\text{in}}}}_{\text{lookup}} + \underbrace{\frac{a^2}{\Delta x^{\text{in}} \cdot \Delta t^{\text{in}}}}_{\text{sum}}\right]\right). \tag{6.9}$$

However, the larger of both, which can be either *lookup* or *sum*, will determine the final computational complexity. Especially for large input-datasets the lookup part will dominate. Therefore, we propose the use of special data structures for the input data.

In general, search problems can be accelerated by the use of tree based data structures. While the insertion into tree structures requires more effort, search can be done on average in $\mathcal{O}\left(\log\left(n^{\text{in}}\right)\right)$ time. In [110], Guttman proposed the *R-tree* data structure for indexing multi-dimensional information, such as geographic coordinates, polygons, or even arbitrary higher dimensional data-sets. No native implementations of R-trees are available in MATLAB, but there are freely available third-party libraries. The *GML LidarK Library* [111] for example, provides an open-source framework for processing multidimensional point data such as LIDAR scans. The library also includes a data-structure called *Seg-tree*, a variety of the R-tree. For the SQLite database mentioned in Sec. 6.1 an R-tree module is available as well [112].

Interpolation with Convolution

The second possibility of an efficient implementation is to align the input data z_i on the spatio-temporal grid as defined in eq. (6.5) first. The double-sums in eqs. (4.10), (4.11), (4.17) and (4.18) can be efficiently realised through convolution of the input data with a discretised interpolation kernel β_{char}.

Two properties of the convolution [95] might be interesting for an efficient implementation of the weighted sums. Let $A \in \mathbb{R}^{N\times M}$ and $B \in \mathbb{R}^{N\times M}$ be two matrices; and let $b_1 \in \mathbb{R}^{N\times 1}$ and $b_2 \in \mathbb{R}^{1\times M}$ be two vectors. Firstly, as stated by the *convolution theorem*, the convolution in the spatio-temporal domain may be written as a multiplication in the frequency domain

$$\mathcal{F}\{A * B\} = \mathcal{F}^{-1}\big\{\mathcal{F}\{A\} \cdot \mathcal{F}\{B\}\big\}, \tag{6.10}$$

where $*$ denotes the convolution operator and $\mathcal{F}$ the Fourier transform. In this case, the computational costly convolution can be avoided by applying the *Fast Fourier Transform (FFT)* combined with a point-wise multiplication in frequency domain.

Secondly, the convolution is an *associative* operation

$$(A * b_1) * b_2 = A * (b_1 * b_2). \tag{6.11}$$

As a consequence of this associativity property a filtering kernel B may be decomposed into two simpler kernels with lower dimension in order for $B = b_1 * b_2$, then the convolution may be

computed with two simpler convolutions

$$A * B = (A * b_1) * b_2. \tag{6.12}$$

A decomposition of $B = b_1 * b_2$ requires that the rows of B are linearly dependent, in other words the rank of B has to be one. This is not the case for an anisotropic filter kernel and therefore not applicable for the ASM.

The two dimensional fast Fourier transform (FFT) [82] has a computational complexity of

$$\mathcal{F}(A) \in \mathcal{O}(NM \cdot \log(NM)). \tag{6.13}$$

As both, the preceding alignment of the input data on a spatio-temporal grid, and the multiplication have a lower computational complexity, the overall complexity is bounded by the complexity of the FFT which can be written as

$$ASM_{\mathrm{FFT}} \in \mathcal{O}\left(\frac{X^{\mathrm{rec}} \cdot T^{\mathrm{rec}}}{\Delta x^{\mathrm{rec}} \cdot \Delta t^{\mathrm{rec}}} \log\left(\frac{X^{\mathrm{rec}} \cdot T^{\mathrm{rec}}}{\Delta x^{\mathrm{rec}} \cdot \Delta t^{\mathrm{rec}}}\right)\right). \tag{6.14}$$

In addition, it should be noted that the advantage of the method in computational complexity also implies also a higher memory consumption.

In [109], Schreiter et al. proposed an efficient implementation of the ASM based on MATLAB's fast cross-correlation and convolution implementations. The authors propose to map the scattered input data $z_i(x_i, t_i)$ to the closest coordinate on the spatio-temporal grid from eq. (6.5). The additional error can be neglected, because the sampling distances Δx^{rec} and Δt^{rec} are comparatively low. Their experiments in MATLAB show, that for large datasets, and a fine-grained reconstruction resolution, an FFT based implementation outperforms a naive implementation by a speedup larger than 250, but the speedup declines significantly with increasing reconstruction resolution.

6.4 Summary

This chapter not only provided information about our actual implementation of data containers and traffic state estimation algorithms, but also mentioned considerations of feasible ways for an efficient implementation of the respective software components. The information on our implemented software was condensed, as it was realised with standard MATLAB data structures and programming techniques. An exception here were the implementation details about the validation system, where individual software components, the respective interfaces, and necessary data transformations have been described in detail.

The second main part of this chapter dealt with the computational complexity and different ways feasible for an efficient implementation of ASM-related traffic state estimation algorithms. These efficient implementations may be of special interest when they are used in large-scale applications, where both storage capacity and computing time turn into relevant algorithm properties.

On the basis of the implemented software, we were able to conduct a comprehensive series of experiments. Further information about those experiments and findings is provided in the subsequent chapter.

7 Experimental Evaluation and Findings

This chapter covers a description and the corresponding results regarding the experiments conducted within this work. All experiments are based on data from a microsimulation as well as from real world measurements – and attempt to clarify open issues previously raised in this work. This chapter is structured among three main sections: First, we examine the performance of finite-difference based velocity estimation in Sec. 7.1. Second, we describe the results of two different calibration approaches in Sec. 7.2. Finally, we give a comprehensive summary about our experiments regarding the ASM algorithm family in Sec. 7.3. The experiments in Sec. 7.1 and 7.3 are generally based on Monte-Carlo-simulations, whereas the experiment in Sec. 7.2 is based on a data-fitting procedure.

7.1 Sensitivity Analysis of GPS-based Velocity Estimation

This section covers a sensitivity analysis of velocity estimation methods based on finite differences as presented in Sec. 5.3.2. Simulated FC trajectories[24] within a range of 9 km constitute the basis for this experiment. In order to simulate trajectories as obtained in real world a bivariate Gaussian position error $\sim \mathcal{N}(0, I_2 \cdot \sigma_{x,y})$ was added[25]. Furthermore, the FC sampling time-interval ΔT and the spatial sampling interval Δx of the road model were varied as illustrated in Fig. 7.1. The simulated FC trajectory points also include the respective vehicle velocity, and

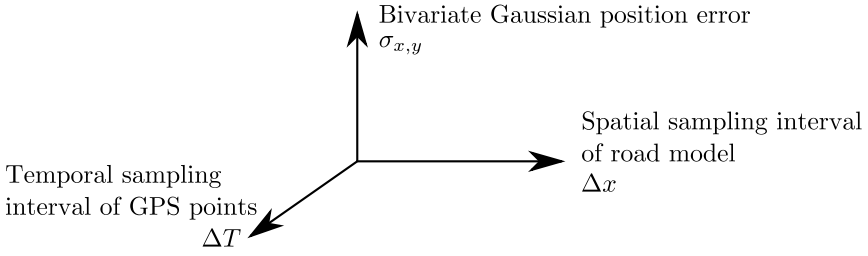


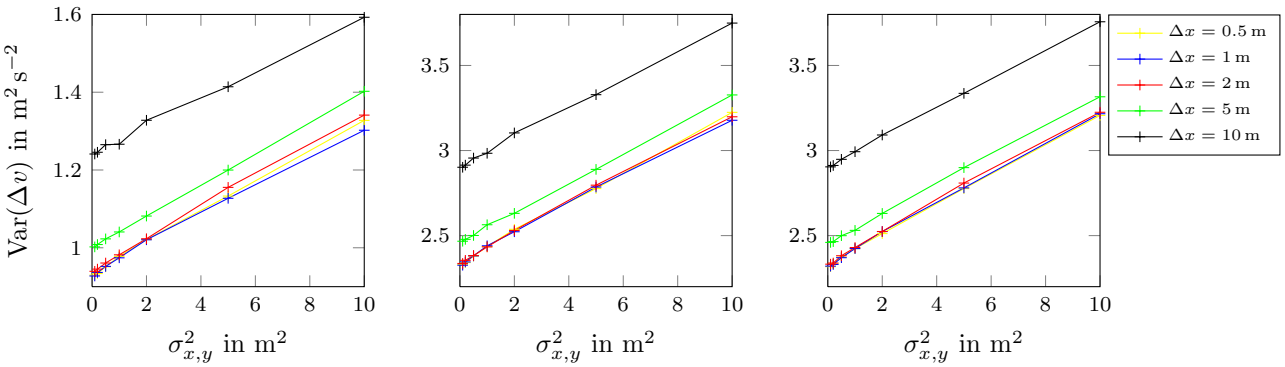
Figure 7.1: Parameter space of the sensitivity analysis of the velocity estimation.

[24] Around 4000 individual trips.
[25] I_2 is a 2×2 identity matrix.

thus it is possible to evaluate the absolute velocity error with

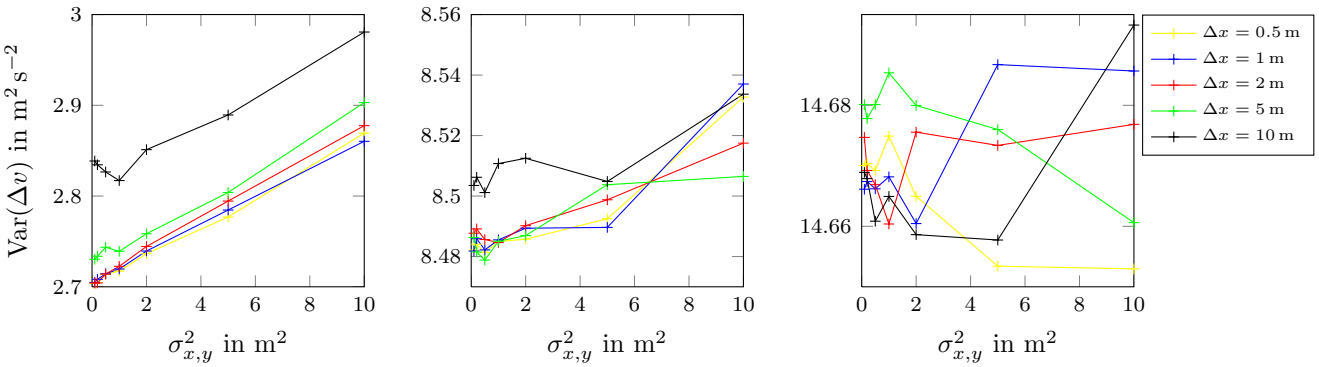
$$\Delta v_k = \hat{v}_k - v_{\mathrm{sim},k}, \tag{7.1}$$

where $\hat{v}_k$ represents the respective estimated velocity (i.e. v_Δ, v_∇, or v_{PCHIP}), and $v_{\mathrm{sim},k}$ the respective true velocity obtained from microsimulation. For the data-fusion algorithms, as presented in Sec. 4.5, the error variance is of great importance, and thus the error variances along the aforementioned parameter space were examined as illustrated in Fig. 7.2 to 7.4.



(a) v_{PCHIP} (b) v_Δ (c) v_∇

Figure 7.2: Comparison of velocity-error variances for a trajectory sampling-interval of $\Delta T = 5\,\mathrm{s}$. The variance depends linearly on the GPS position noise for all methods, although the error-variance of the PCHIP method in (a) is the half of the variances of the simple differences in (b)–(c).

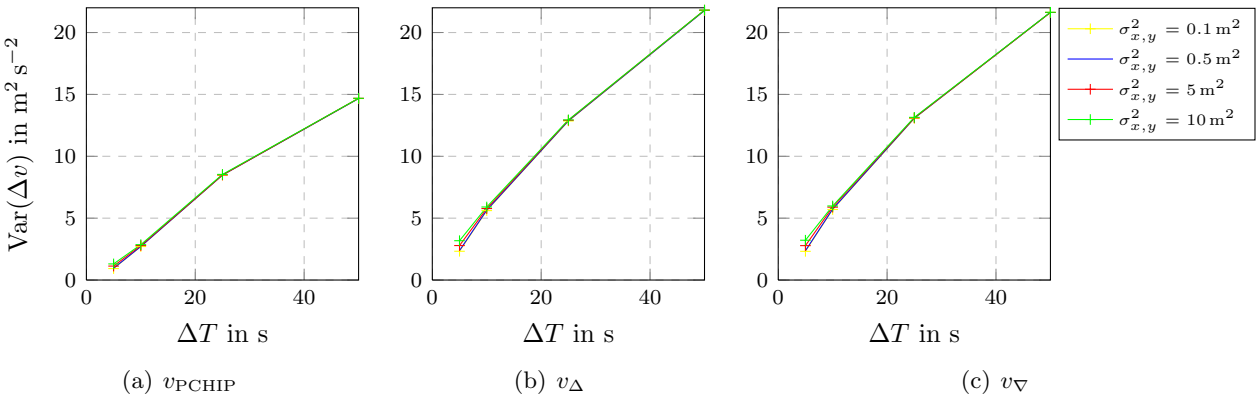


(a) $\Delta T = 10\,\mathrm{s}$ (b) $\Delta T = 25\,\mathrm{s}$ (c) $\Delta T = 50\,\mathrm{s}$

Figure 7.3: Comparison of velocity-error variances for different trajectory sampling intervals ΔT achieved with the PCHIP method, the behaviour of the two other methods is similar: With increasing time-intervals, the influence of both $\sigma^2_{x,y}$ and Δx becomes more and more negligible.

It is possible to draw the following conclusions from this experiment:

1. The temporal trajectory sampling interval ΔT has the highest impact on $\mathrm{Var}(\Delta v)$ when compared with all other parameters.
2. For a low ΔT, i.e. $\Delta T \leq 10\,\mathrm{s}$, the value of the parameters $\sigma^2_{x,y}$ and Δx should also receive attention.
3. With the PCHIP method it is possible to achieve a lower error variance when compared to the forward- and backward-difference based methods. This fact can be explained by the harmonic averaging performed by the PCHIP method in difference to the other methods.



(a) v_{PCHIP} (b) v_{Δ} (c) v_{∇}

Figure 7.4: Comparison of velocity error variances with respect to different GPS sampling time intervals for $\Delta x = 1\,\text{m}$ spatial resolution of the road model. When compared to the other two dimensions, ΔT has the highest impact on the error variance.

It has to be noted that all three methods underestimate the momentary vehicle velocity by a systematic error of $-0.5\,\text{m}\,\text{s}^{-1}$ and $-0.7\,\text{m}\,\text{s}^{-1}$ on average for the finite-difference methods and the PCHIP method, respectively. At the moment, the reasons for this bias have not yet been investigated. Consequently, not only these particular methods but also other methods for FC velocity estimation in general should be assessed further. However, this was not possible due to time reasons.

7.2 Calibration

Treiber et al. mention that a calibration of parameter regarding the ASM is not possible directly [53]. However, we try to make some suggestions considering that issue. The parameters can be divided into two groups: (1) c_{free}, c_{cong}, V_{thr}, and ΔV depend on the particular *road and traffic conditions*; (2) the parameters σ and τ depend on the *sampling density and noise-level* of the input data. In the following, we describe the calibration of the traffic specific parameters on the basis of stationary detector data in Sec. 7.2.1, whereas an approach for full calibration based on the concept of reconstruction followed by validation is given in Sec. 7.2.2.

7.2.1 Model-Based Calibration on Stationary Data Only

The traffic-specific parameters can be directly associated with specific properties of the fundamental diagram (introduced in Sec. 2.2): The transition between free-flow and congested flow is defined by the point of capacity; the characteristic velocities are represented by the slopes of the density-flow diagram. Accordingly, we propose a calibration of the mentioned parameters based on a steady-state car-following model. Several steady-state models for the fundamental-diagram have been proposed in the past, for an overview we refer to [6]. The model by Van Aerde and Rakha [113] is mentioned as one of the relatively simple models with the best representation of the traffic characteristics. Their model defines the density according to the velocity-depended gap between two successive cars, denoted by

$$k(v) = \frac{1}{h} = \frac{1}{c_1 + \frac{c_2}{v_0 - v} + c_3 \cdot v}, \tag{7.2}$$

where h denotes the average vehicle-gap, v_0 is the average free-flow velocity, and $[c_1, c_2, c_3]$ denote the three gap-coefficients. In order to determine the four unknown parameters we propose a least-squares fitting of eq. (7.2) on data from stationary traffic sensors. Due to the nature of the data it is not recommended to fit eq. (7.2) directly. In [114], Ponzelt gives practical recommendations: (1) Large vehicles, i.e. trucks and buses, require more space on the road. When the flow rates of those large vehicles are known (denoted by q_{truck}) they should be grossed up to car-equivalent rates by multiplication with a constant factor, i.e.

$$q = a_{\text{equiv}} \cdot q_{\text{truck}}. \tag{7.3}$$

For our calculations we use a grossing-up factor of $a_{\text{equiv}} = 2$. (2) The data is not equally distributed over all traffic regimes. Therefore, velocities should be first binned by their density values; Ponzelt recommends a bin-width of $0.001\,\text{Veh}\,\text{m}^{-1}$. A complete fundamental diagram, including raw data, binned data, the fitted van Aerde model and the point at capacity are illustrated in Fig. 7.5.

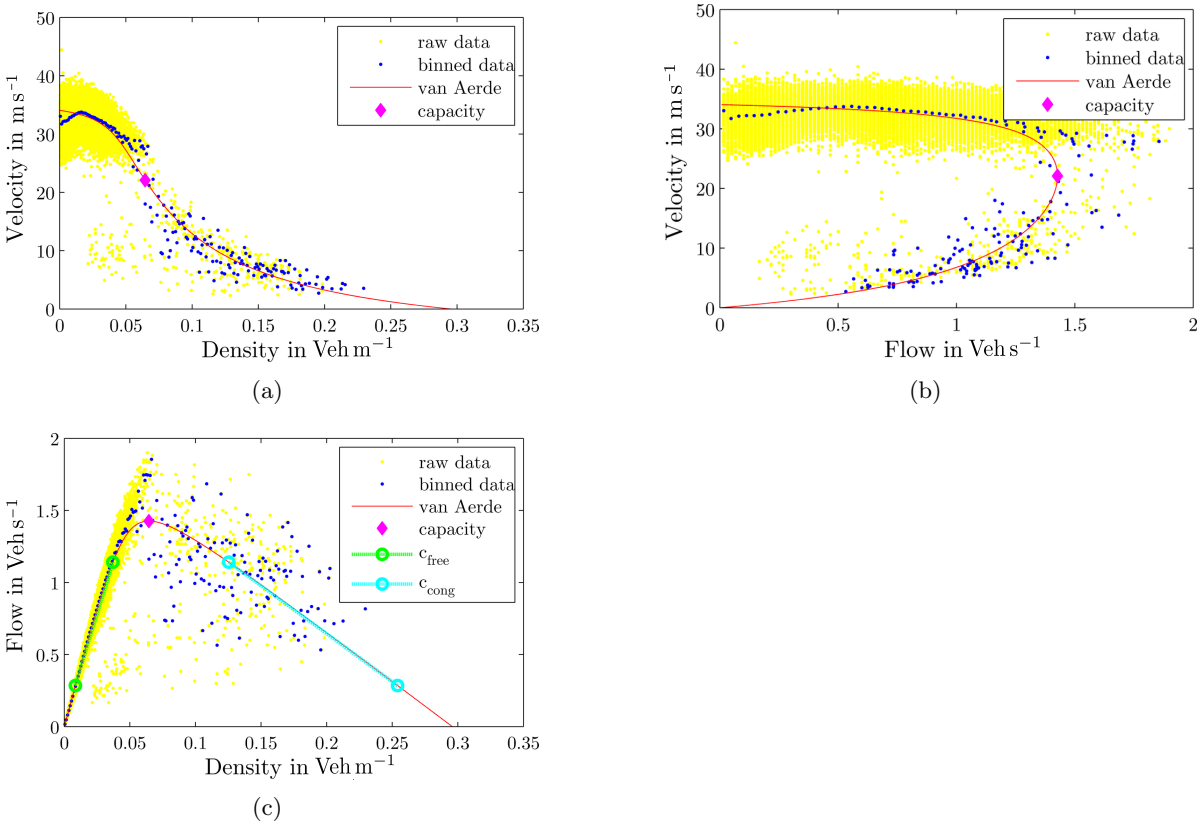


Figure 7.5: Fundamental diagrams created with stationary detector data from June 04, 2012 until June 27, 2012 motorway A4 at $x = 10\,\text{km}$ in driving direction east including raw data, binned data, and the calibrated steady-state traffic model after van Aerde. *In addition, the characteristic velocities c_{free} and c_{cong} are illustrated in (c).*

The transition between free-flow and congested traffic regimes happens at the inflection point in Fig. 7.5(a), which is defined by the maximum flow q_{max} and the optimal velocity v_{opt}. The maximum flow can be calculated by a first derivative test of eq. (7.2), which yields

$$q_{\text{max}} = \frac{-2 \cdot \sqrt{c_2}\sqrt{c_2 + c_1 \cdot v_0} + c_1 \cdot v_0 + 2 \cdot c_2 + c_3 \cdot v_0^2}{c_1^2 + 4 \cdot c_2 \cdot c_3 + 2 \cdot c_1 \cdot c_3 \cdot v_0 + c_3^2 \cdot v_0^2}. \tag{7.4}$$

As threshold between free-flow and congested traffic we propose the optimal velocity v_{opt} denoted by

$$V_{\text{thr}} \approx v_{\text{opt}} = \frac{1}{2}\left(v_0 + \frac{c_1 \cdot q_{\max}}{1 - c_3 \cdot q_{\max}}\right). \tag{7.5}$$

At the moment, the transition-width ΔV in eq. (4.20) is assumed to be constant. A full transition between free-flow and congested traffic is completed within a velocity difference of approximately $4 \cdot \Delta V$ symmetrically around V_{thr}.

For the characteristic velocities under free-flow and congested conditions we propose the use of eq. (3.11) in combination with the van Aerde model. The slopes of the density-flow diagram show qualitatively linear behaviour over a wide range, and thus an approximation for the characteristic velocities can be written as

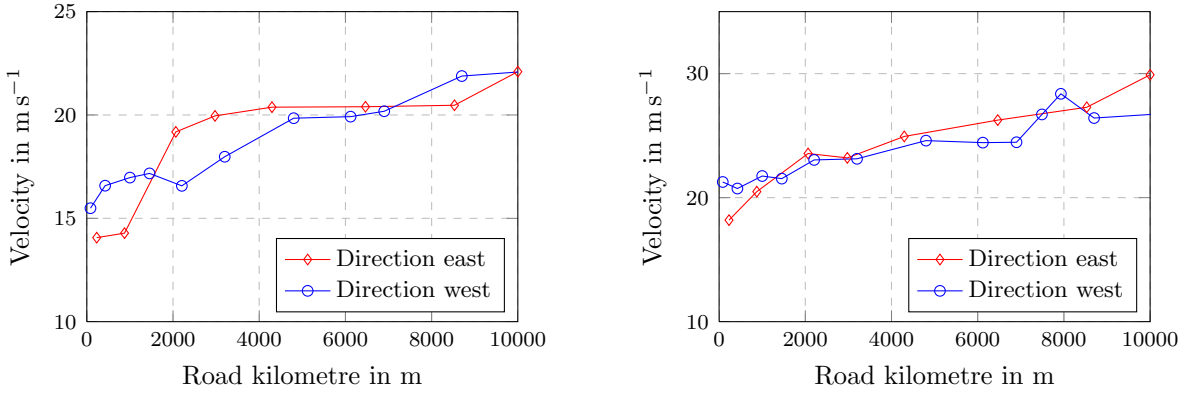
$$c_{\text{free}} = \frac{a_{q,\max} \cdot q_{\max} - a_{q,\min} \cdot q_{\max}}{k_{q,1}(a_{q,\max} \cdot q_{\max}) - k_{q,1}(a_{q,\min} \cdot q_{\max})} \tag{7.6}$$

and

$$c_{\text{cong}} = \frac{a_{q,\min} \cdot q_{\max} - a_{q,\max} \cdot q_{\max}}{k_{q,2}(a_{q,\min} \cdot q_{\max}) - k_{q,2}(a_{q,\max} \cdot q_{\max})}, \tag{7.7}$$

where $k_{q1,2}$ denotes the traffic density depending on the flow based on eqs. (2.7) and (7.2). The numerical indices refer to either one of the two possible solutions. Further, the qualitatively linear sections of the flow-density relation can be specified by the parameters $a_{q,\min}$ and $a_{q,\max}$. Good results were observed in the cases with values $a_{q,\min} = 0.2$ and $a_{q,\max} = 0.8$.

For the full calibration results of all stationary detectors based on data from June 04, 2012 until June 27, 2012 we refer to Appendix B. Although we used data from 24 days the amount of data obtained in congested traffic regimes is too small for most detectors and therefore, a plausible estimation of c_{cong} is impossible. However, it is possible to perform a plausible estimation of c_{free} and v_{opt} as illustrated in Fig. 7.6. From those calibration results it is possible to draw the



(a) v_{opt} (b) c_{free}

Figure 7.6: Calibration results for v_{opt} and c_{free}. Both quantities go in line with the road kilometre, but are assumed to be constant over the whole road stretch by the ASM algorithm.

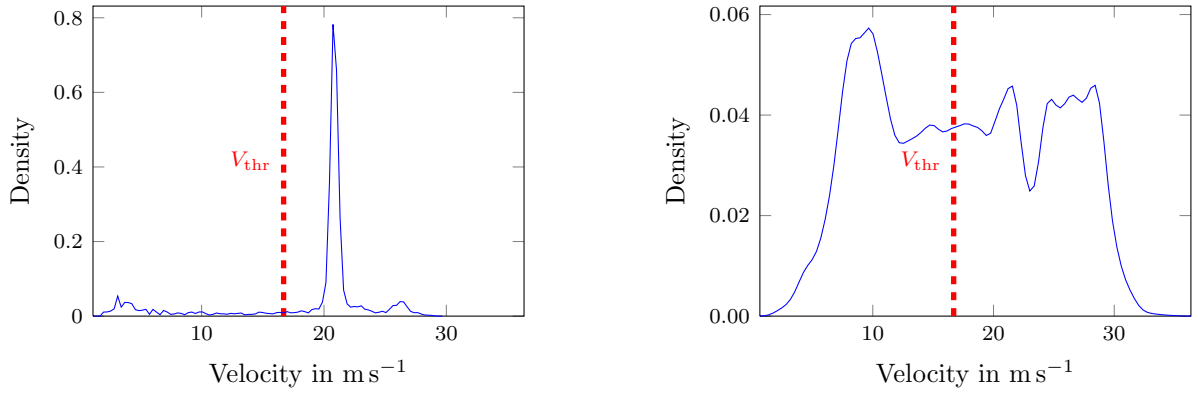
following conclusions considering the calibration procedure and the underlying traffic stream: v_{opt} and c_{free} are not constant over the entire road stretch. This fact corresponds to the obligated speed limits: they go in line with the road kilometre, i.e. $80\,\text{km}\,\text{h}^{-1}$ until $x \approx 4\,\text{km}$, $100\,\text{km}\,\text{h}^{-1}$ until $x \approx 8.3\,\text{km}$, and $130\,\text{km}\,\text{h}^{-1}$ afterwards.

Both quantities v_{opt} and c_{free} correspond to the traffic dependent algorithm parameters V_{thr}

and c_{free}, respectively. On the one hand it is not easy to incorporate a spatially varying c_{free} to the ASM method, but on the other hand it is easy to incorporate a spatially varying V_{thr}. As eq. (4.20) operates only point-wise on the spatio-temporal field it is possible to introduce a spatial varying $V_{\text{thr}}(x)$ as following:

$$w(x,t) = \frac{1}{2}\left[1 + \tanh\left(\frac{V_{\text{thr}}(x) - \min(V_{\text{free}}(x,t), V_{\text{cong}}(x,t))}{\Delta V}\right)\right]. \tag{7.8}$$

First experiments with the congestion pattern illustrated in Fig. 7.10(a) show that the use of $V_{\text{thr}}(x)$ yields lower and less sensitive RMS- and MAP-errors. (These error-measures are described in Appendix C.) However, the traffic pattern does not contain large areas with velocities near v_{opt} and therefore is ill-conditioned to investigate this particular topic (illustrated in Fig. 7.7).



(a) Simulated data

(b) Real-world data

Figure 7.7: Comparison of distribution of velocities estimated from simulated data in (a) and from real-world data in (b) (data from Fig. 7.10(a) and Fig. 7.9, respectively): for simulated data, the density of velocities around V_{thr} is less than 0.01, whereas in (b) the velocities are evenly distributed over a wide range. The respective densities were obtained with a kernel smoothing function estimate based on a normal kernel function.

7.2.2 Calibration by Reconstruction and Subsequent Validation

This section covers an approach for calibration of the parameters σ and τ, which depend on the density of the input data. Stationary detectors deliver data that is arranged lattice-like horizontally in the spatio-temporal plane; Thus, it is straight forward to specify an average sampling interval Δx and ΔT. On the other hand, floating car data is aligned like vertically askew pearl chains, where the alignment between the interpolation kernel and the input-data becomes suboptimal, and thus demand larger smoothing parameters.

According to the mentioned requirements, we performed a calibration on the basis of real-world data: Basically, a state-estimation based on floating car data (project REFEREE), followed by a validation against stationary detector data (ASFINAG). Exemplary input data is illustrated in Fig. 7.8. As no velocity information is included in the floating car data, the velocities are estimated with the aforementioned PCHIP method (see Sec. 5.3). Our final calibration approach can be summarised as follows:

1. Prune the available floating car data $\mathbf{d}_{\text{FC}} = [\{x_i, t_i, tr_i\}_{i=1}^{I}]$ and stationary detector data $\mathbf{d}_{\text{ST}} = [\{x_j, t_j, v_j\}_{j=1}^{J}]$ to a common spatio-temporal frame.

2. For all floating car trips, i.e. $tr_i = const.$, estimate the respective velocities v_i.

3. For every 4-tuple in the list of predefined filter parameters $\mathbf{p} = [\{\sigma_i, \tau_i, c_{\mathrm{f},i}, c_{\mathrm{c},i}\}_{i=1}^{P}]$:
   a) Filter the velocities $[\{v_i\}_{i=1}^{L}]$ from the stationary detector with the ASM as *ground-truth data*, i.e. "reconstruct" the velocities solely at spatio-temporal locations where stationary data is available.
   b) Reconstruct velocities from $\mathbf{d}_{\mathrm{FC}}$ at the spatio-temporal locations where stationary data is available.
   c) Calculate global error measures based on the data from the steps (3.a) and (3.b).

4. Select the 4-tuple with the lowest error combined over all global error measures.

Unfortunately, we were not able to achieve a plausible calibration result with this method. Possible reasons for the failed calibration may be too little available data covering all traffic states, i.e. free flow and congested flow. Another possibility may be a low agreement between floating car and stationary detector data, caused by the relatively high sampling intervals of the floating car data or the biased space-mean velocities obtained by the stationary sensors.

A different approach would be the calibration of the parameters based on the travel-time measurements. The travel times obtained by an ANPR system and the travel times obtained from a spatio-temporal reconstruction share a high degree of agreement. However, calibration based on travel times is not practical either, as the comparable quantities imply a high degree of aggregation. A velocity field reconstructed through the ASM by using floating car data from May 21, 2013 as well as the extracted travel times are illustrated in Fig. 7.9, whereby the parameters[26] have been selected with an "educated guess". The congestion pattern may be caused by capacity overload at Knoten Prater, i.e. at road kilometre $x \approx 0\,\mathrm{km}$, and agrees structurally with congestion patterns caused by permanent bottlenecks that were observed on German motorways [1, p. 100].

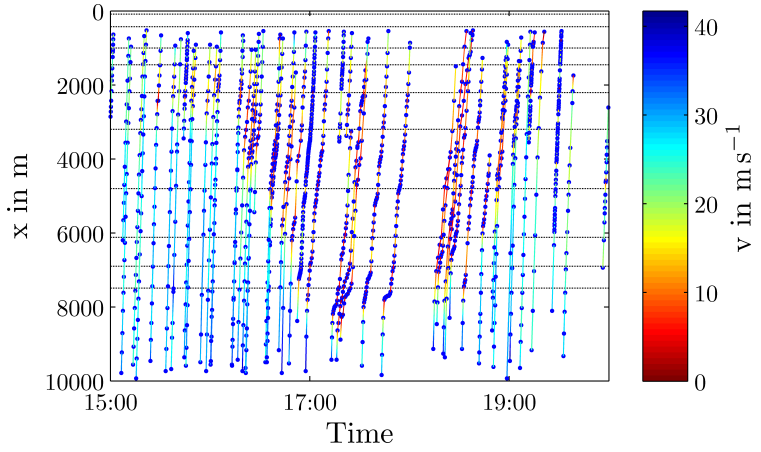
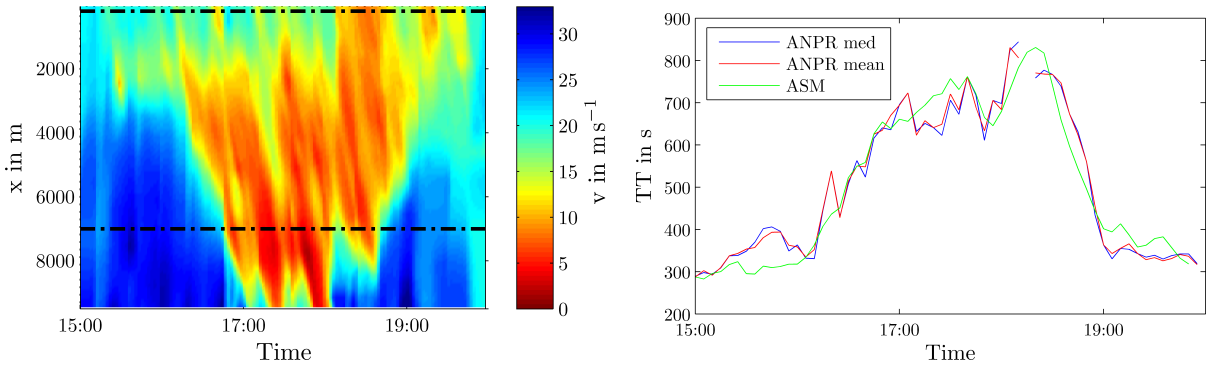


Figure 7.8: Illustration of available data from the project REFEREE, from May 21, 2013 in driving direction north-west: The blue dots represent the real data, the estimated velocities are represented by the coloured lines. Further, the dashed, horizontal lines indicate the positions of the stationary detectors.

[26] Parameters: $\sigma = 750\,\mathrm{m}$, $\tau = 150\,\mathrm{s}$, $c_{\mathrm{free}} = 70\,\mathrm{km\,h^{-1}}$, $c_{\mathrm{cong}} = -16\,\mathrm{km\,h^{-1}}$, $V_{\mathrm{thr}} = 60\,\mathrm{km\,h^{-1}}$, and $\Delta V = 20\,\mathrm{km\,h^{-1}}$.



(a) Reconstruction (b) Travel times

Figure 7.9: A reconstructed velocity field by the ASM is illustrated in (a) obtained on the basis of trajectory data illustrated in Fig. 7.8. The locations of the ANPR cameras are indicated by the dashed lines. The measured travel times in (b) agree with the travel times obtained from (a).

7.3 Behaviour of ASM-based Methods

In this section we describe several experiments focussing on the behaviour of ASM-based methods with respect to the "quality" of the input data. As the methods are applicable for both Eulerian and Lagrangian data, we evaluate the impact of (1) spatio-temporal input data density, (2) randomly missing input data, (3) biased measurements, and (4) the presence of measurement noise. The quality of reconstruction was evaluated with the well known error measures RMSE, MAPE, MPE, and SPE as summarised in Appendix C.

In order to investigate the impact of data-density without presence of measurement noise the experiments were performed on the basis of synthetically generated data obtained from the aforementioned microsimulation framework. The simulation model is calibrated to represent the traffic conditions on a normal weekday between 17:00 and 18:00. The simulated traffic congestion (illustrated in Fig. 7.10(a)) is triggered by a short speed restriction of $15\,\mathrm{km\,h^{-1}}$, taking effect over 1050 s. The main wave-front propagates with with less than $-8\,\mathrm{km\,h^{-1}}$ upstream, while the characteristic shock-wave velocity of $-16\,\mathrm{km\,h^{-1}}$ can be observed in the oscillations at the tail of the congestion pattern. Please note that typical congestion patterns on Austrian motorways may remain over several hours as illustrated in Fig. 7.9. Consequently, the used test-patterns have different shapes than observed patterns in real-world. In this context, the same traffic specific algorithm parameters[27] were used for each experiment described below.

Below, we describe experiments conducted *without presence of measurement noise*: on the basis of stationary detector data only in Sec. 7.3.1, on the basis of floating car data only in Sec. 7.3.2, and the combination of both data-sources in Sec. 7.3.3; followed by experiments with the presence of noise in Sec. 7.3.4.

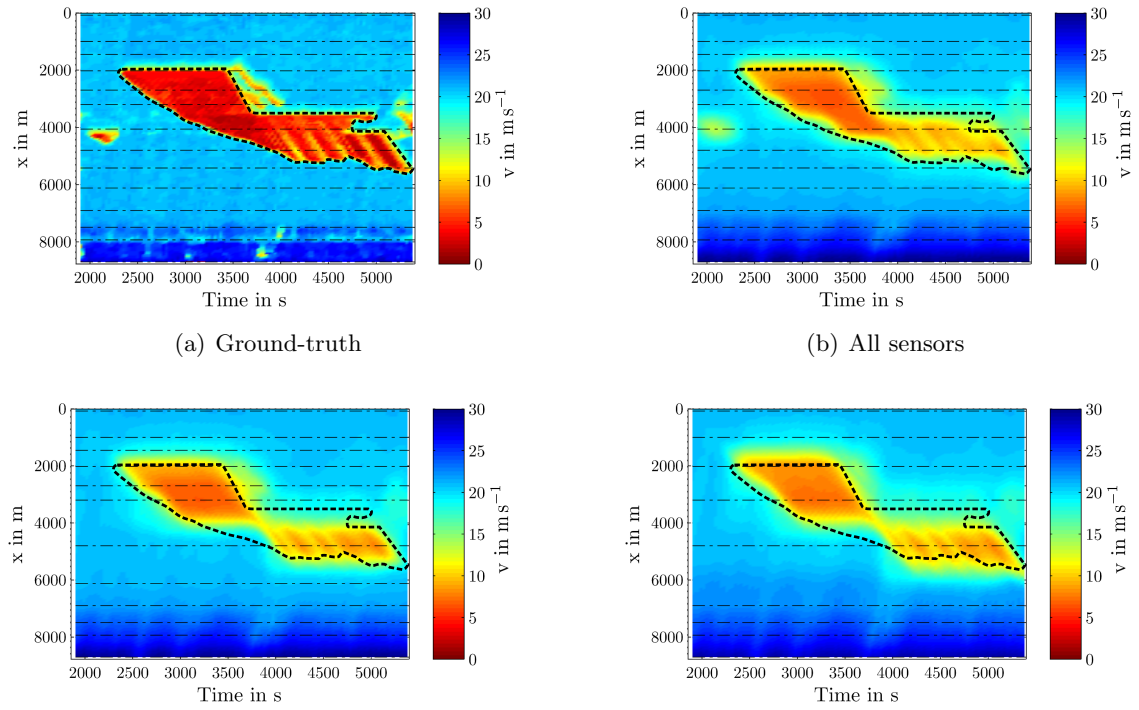
7.3.1 Stationary Input-Data Only

The subsequent section provides the description of experiments limited to stationary input-data only. In this context, the same data-dependent interpolation kernel parameters were used for each experiment with stationary detector data, namely $\sigma_{\mathrm{St}} = \overline{\Delta x/2}$ and $\tau_{\mathrm{St}} = 60\,\mathrm{s}$. The first two experiments can be seen as "case studies", whereas all other experiments rely on multiple runs of individual, random cases.

[27] Algorithm parameters: $c_{\mathrm{free}} = 70\,\mathrm{km\,h^{-1}}$, $c_{\mathrm{cong}} = -15\,\mathrm{km\,h^{-1}}$, $V_{\mathrm{thr}} = 60\,\mathrm{km\,h^{-1}}$, and $\Delta V = 20\,\mathrm{km\,h^{-1}}$; Reconstruction resolution: $\Delta x^{\mathrm{rec}} = 50\,\mathrm{m}$ and $\Delta t^{\mathrm{rec}} = 30\,\mathrm{s}$.

Availability of Stationary Detectors

This experiment evaluates the impact of the presence of particular stationary sensors. Starting with a detector density greater than at the real test-site, i.e. $\Delta x \approx 600\,\mathrm{m}$, stationary sensors are removed consecutively. The ground-truth and different reconstructed spatio-temporal velocity fields are illustrated in Fig. 7.10.



(a) Ground-truth

(b) All sensors

(c) Two less sensors

(d) Five less sensors

Figure 7.10: Spatio-temporal velocity fields, including the ground-truth in (a) and the respective reconstructions based on different number of stationary sensors in (b)–(d). The distinct area of low velocities in the ground-truth data is marked with the black dashed line.

Through this method we are able to reconstruct the congestion pattern, although the velocities in the congestion area are over-estimated. This phenomenon can also be observed for the extracted travel times as illustrated in Fig. 7.11, whereby the quality of extracted travel times is not satisfying regardless the number of present sensors. One reason for this may be the difference between time-mean and space-mean velocity as mentioned in Sec. 2.1.2. Furthermore, rising detector-distances imply increasing interpolation kernel parameters σ and τ and consequently softer transitions between the different traffic states.

In sum, the quality of reconstruction depends mainly on the density of the input data, as well as on the congestion pattern, and the relative alignment between sensors and the congestion pattern. As a consequence the error does not depend directly on the number of stationary sensors as illustrated in Fig. 7.12. For example, every error measure is higher with one removed sensor than with two removed sensors. These findings are also confirmed by the experiment described below.

Randomly Missing Data-Points

The main finding of the previous experiment is, that in some cases a higher amount of input data may produce inferior reconstruction results. Thus, the following experiment evaluates the impact

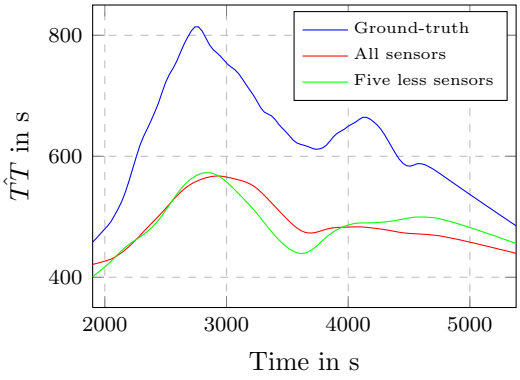


Figure 7.11: Extracted travel times from ground-truth data compared to travel times obtained from reconstructed velocity fields. In areas of congestion, the velocities are over-estimated and as a consequence the travel times are under-estimated.

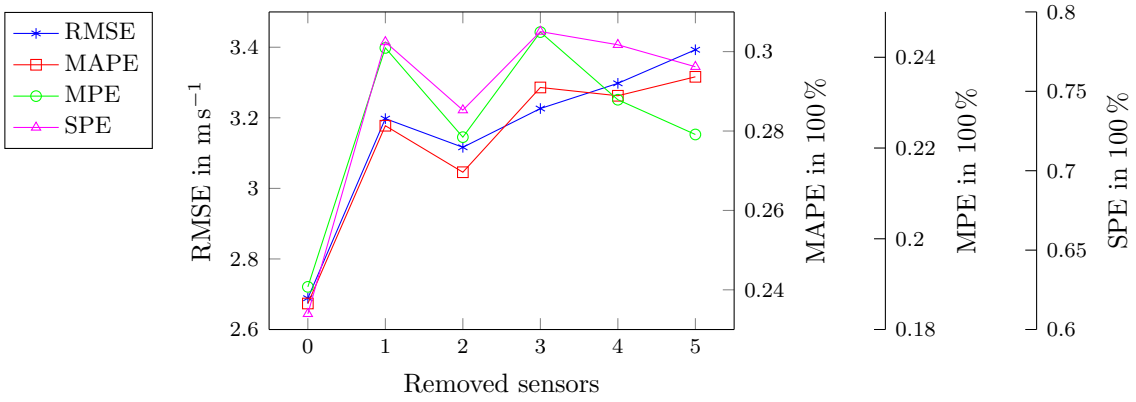


Figure 7.12: Global error measures in relation to the number of removed sensors. The overall error does not directly depend on the number of removed sensors, but rather on the relative alignment between the traffic sensors and the respective congestion pattern.

of randomly missing data. Starting with input-data from all stationary detectors as illustrated in Fig. 7.10(b), data-points are removed randomly at each stationary sensor where the location in time is equally probable on the overall time-range. Because the quality of reconstruction depends on the individual locations of the missing data points, each experiment was performed 200 times[28]. Box plots of global error measures subject to the relative amount of missing data are shown in Fig. 7.13.

It is possible to draw the following conclusions from this experiment: On average, the ASM is not sensitive against randomly missing data. Even with an increasing amount of randomly missing data, the performances of individual reconstructions are subject to great variation. Furthermore, the MPE and the extracted travel times from the previous two experiments indicate a positive bias in the estimated velocity field. This goes in line with the respective theory regarding time-aggregation of velocities. In the subsequent experiment, we examine a simple method for for bias correction.

Velocity Bias Correction.

In this experiment we investigate a simple method for correction of the velocity-bias as already introduced in Sec. 5.2. All mentioned estimation algorithms require as input data space-aggregated

[28] A higher number of runs would be better. However, this is not practicable due to computing time reasons.

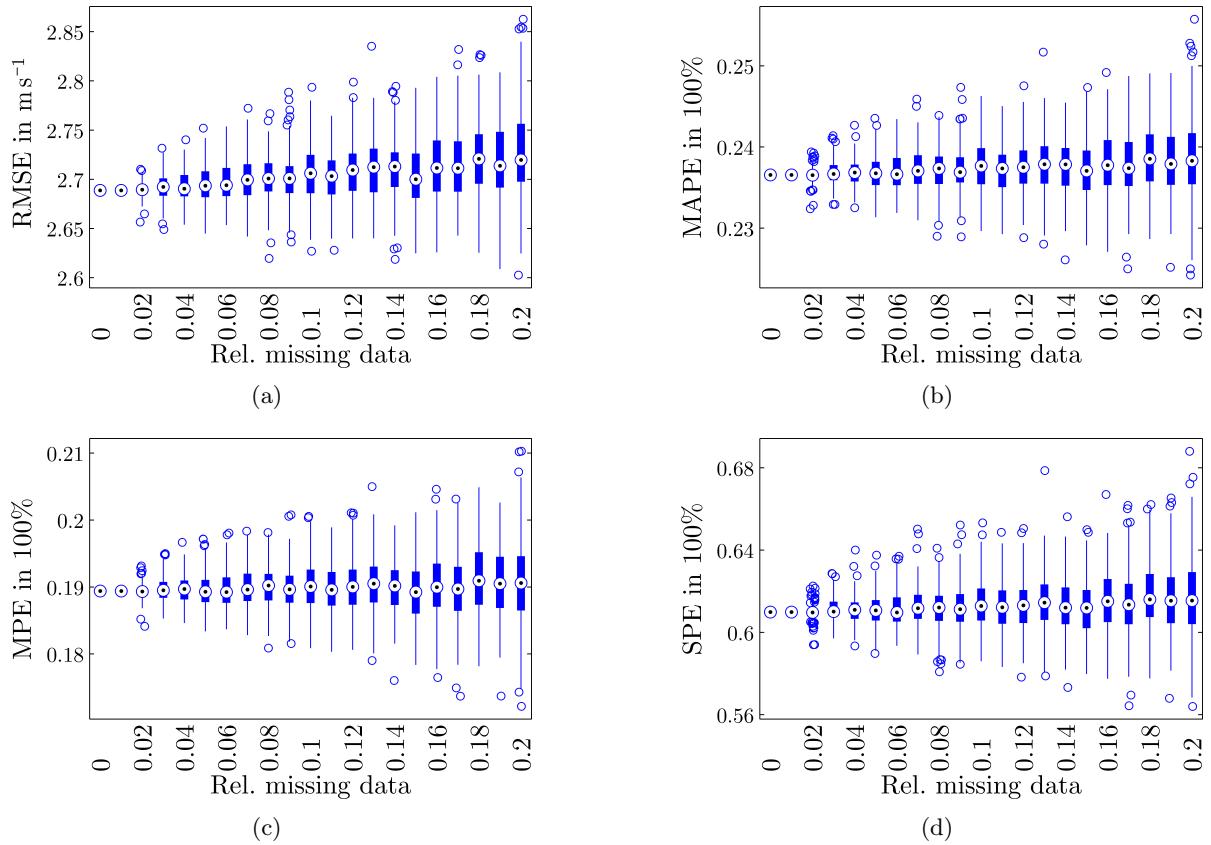


Figure 7.13: Box plots of estimator performance based on data from stationary detectors are subject to the relative amount of repeated randomly missing data (200 runs). The respective performance highly varies depending on the individual case. However, the average performance is not influenced by randomly missing data.

velocities V_s. Anyhow, stationary sensors aggregate individual vehicle velocities to time averages V_t. Let us recapitulate the relation stated in eq. (5.10)

$$V_\mathrm{s} = V_\mathrm{t} \cdot \left(\frac{1}{2} + \sqrt{\frac{1}{4} - CV_\mathrm{V}^2} \right),$$

when assuming a constant CV_V, it is possible to write

$$V_\mathrm{s} \approx V_\mathrm{t} \cdot \kappa_\mathrm{corr}, \tag{7.9}$$

where κ_corr is a constant bias-correction factor.

In this experiment, we performed a bias correction as stated in eq. (7.9) with input-data from stationary detectors as illustrated in Fig. 7.10(b). The correction-factor κ_corr is varied in the range [0.9, 1] which corresponds to a CV_V in the range of [0.3, 0]. The resulting global error measures and extracted travel times are shown in Fig. 7.14 and 7.15, respectively. As the global error measures do not share a common minimum, we propose the minimum RMSE for selection of a optimal correction factor, which yields $\kappa_\mathrm{corr} = 0.97$. However, with this method it is only possible to reduce the MAPE from approximately 19 % to 15 % which is a strong indicator for a remaining systematic error. This phenomenon can also be observed at the extracted travel times in Fig. 7.15: especially for low velocities (and high travel times) there still remains a large bias; only for free flow traffic conditions, the presented method can perform a sufficient

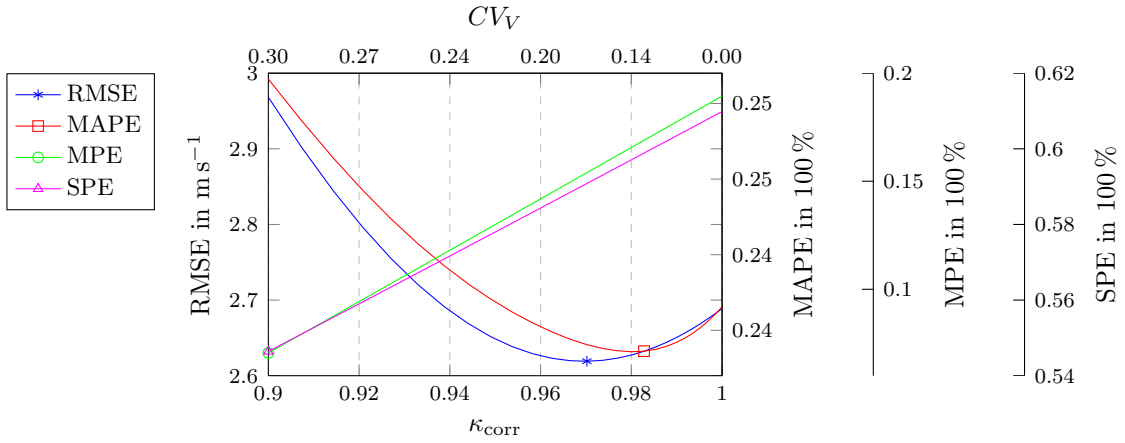


Figure 7.14: Global error measures in linear relation to a bias correction factor κ_{corr}, where the plot-mark indicates the respective minimum. The error measures do not have a common minimum and therefore we propose the use of $\kappa_{\text{corr}} = 0.97$ or the respective CV_{V} of 17 %.

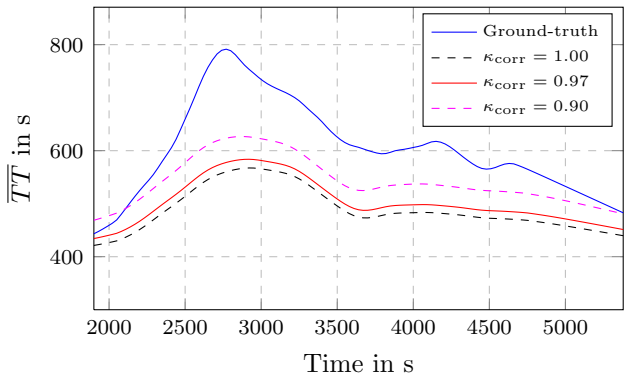


Figure 7.15: Illustration of the influence of different bias correction factors κ_{corr} on travel times obtained from a spatio-temporal velocity field: It is obvious, that the method can not compensate the bias for low velocities.

error correction. Hereinafter, we performed all remaining experiments with a bias correction factor of $\kappa_{\text{corr}} = 0.97$, although the use of a more sophisticated bias correction method would be advisable.

7.3.2 Floating Car Data

A description of experiments with the ASM, just limited to input-data from floating cars, is provided in the following. Those experiments try to evaluate the impact of the density of FC data on the reconstruction performance. The FC data was obtained from the same microsimulation as used in the previous experiments (illustrated in Fig. 7.10(a)). In general, the density of FC data is affected by two parameters, namely the GPS sampling time interval ΔT and the relative frequency of the FC. The FC frequency indicates the relative amount of GPS equipped vehicles present in the total flow. Moreover, all FC trajectories follow the momentary flow velocity in the spatio-temporal plane (cf. Fig. 6.5), and therefore it is difficult to formulate a rule for selection of the data-dependent interpolation kernel parameters σ and τ applicable for this category of data. In order to find appropriate parameters we performed a search based on ASM reconstruction

followed by the evaluation of

$$\{\sigma_{\mathrm{FC}}, \tau_{\mathrm{FC}}\} = \underset{\{\sigma_{\mathrm{FC}},\tau_{\mathrm{FC}}\}}{\arg\min}\ \mathrm{RMSE} \tag{7.10}$$

with respect to different FC frequencies in the range of 0.1-6 %. The resulting parameters are illustrated in Fig. 7.16. However, in some cases especially with relatively low FC frequencies, tra-

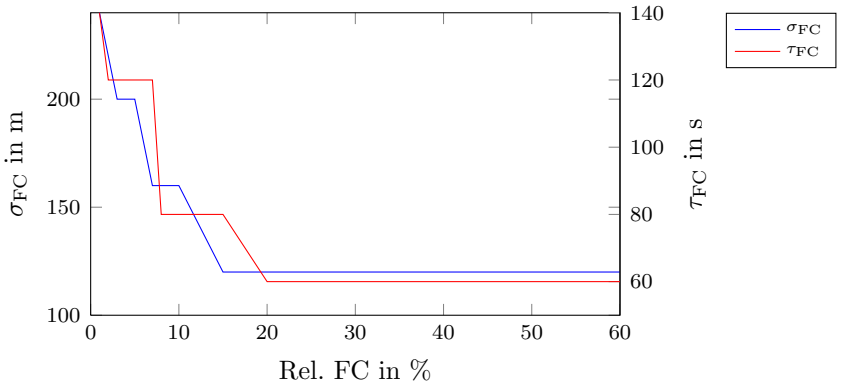


Figure 7.16: Illustration of the data-dependent parameters σ_{FC} and τ_{FC} with respect to the relative FC rate used in the subsequent experiments.

jectories may be distributed irregularly over the spatio-temporal plane and thus, the respective filter parameters may be too small.

The experiments below analyse not only the impact of the relative frequency of floating cars, but also the effect of different trajectory time sampling intervals ΔT. In order to draw meaningful inferences, every experiment is repeated 50 times[29] with a different sample of floating cars[30].

Relative Frequency of Floating Cars

This experiment evaluates the influence of the relative FC frequency. Thus, we performed an ASM reconstruction with trajectories having a time sampling rate of $\Delta T = 1\,\mathrm{s}$ where the frequency of the FC was varied in the range of 0.1-6 %. Both, individual outcomes, and the respective mean values of the global error measures are illustrated in Fig. 7.17.

The evolution of the algorithm performance is comparable for all error measures which can be summarised as follows: every error measure is nearly constant above FC frequencies larger than 1 %; below a frequency of 1 % the performance is deteriorating rapidly. At present, an FC frequency of 1 % is an unrealistic scenario on Austrian motorways. In reality however, a coarser scale of reconstruction over a larger spatio-temporal area may be of interest (see also: real-world congestion pattern in Fig. 7.9). The relative FC frequency has no direct impact on the ASM performance, rather the number of FC trajectories available in the reconstruction time period. The actual relation between the FC frequency and the average number of FC in the reconstruction time frame is shown in Fig. 7.18. In general, the gradient of the straight line depends linearly on the particular reconstruction time period, i.e. 3500 s in this case.

In contrast to ASM reconstructions obtained from stationary detector data, reconstructions solely based on FC data are not biased. This fact is indicated by the low MPE (illustrated in Fig. 7.17(c)) and the better agreement of the extracted travel times obtained from reconstructions and from ground truth data as illustrated in Fig. 7.19.

[29] A higher number of runs would be better. However, this is not practicable due to computing time reasons.
[30] Evenly sampled from a pool of FC trajectories reflecting a relative FC rate of 20 %.

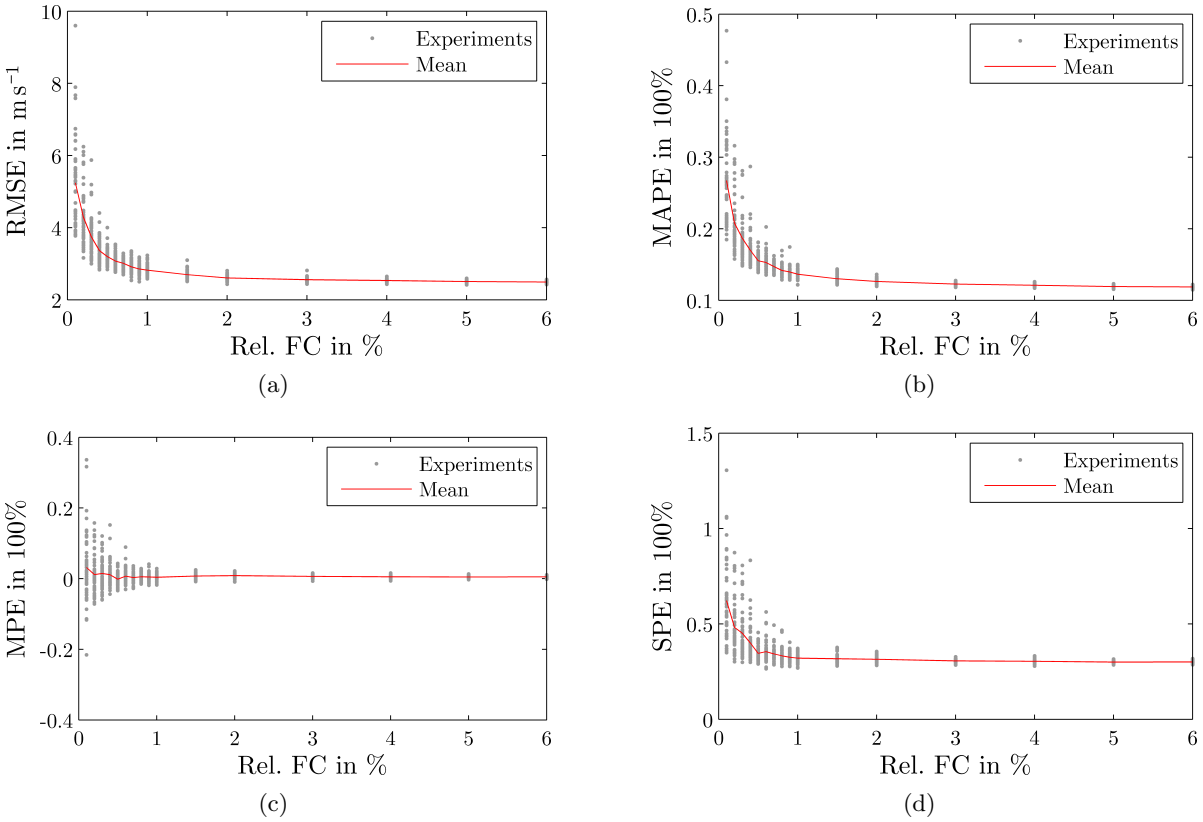


Figure 7.17: Illustration of estimator performance subject to the relative amount of floating car trajectories for repeated experiments, i.e. 50 times. The performance highly depends on the number of trajectories located within the reconstruction area.

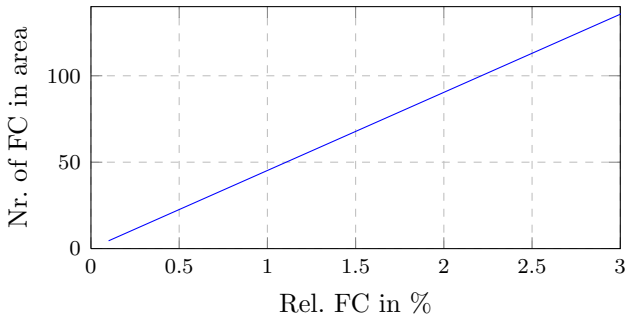
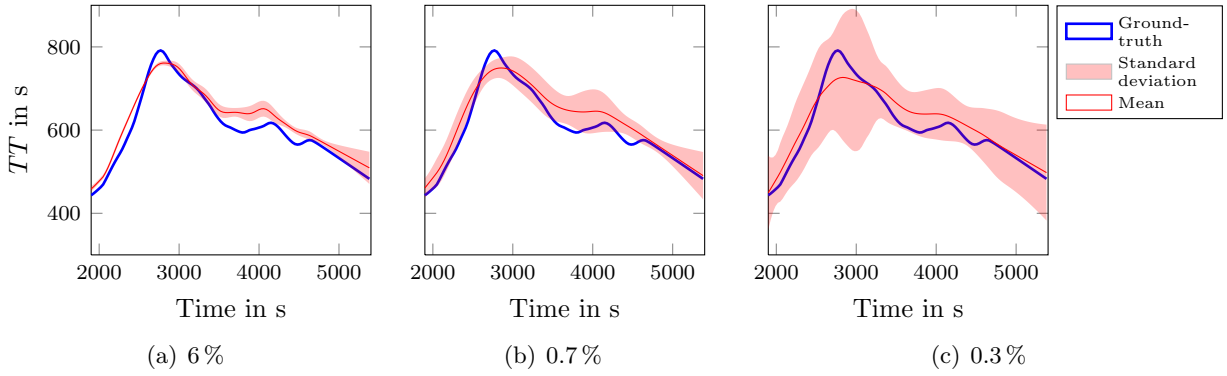


Figure 7.18: Relation between the relative rate of floating cars and the respective number of floating cars in the reconstruction region, i.e. in the valid simulation time-period from $t = 1900\,\mathrm{s}$ to $t = 5400\,\mathrm{s}$.

Time-domain GPS Sampling Rate.

This section describes an experiment regarding the influence of the FC trajectory sampling time-interval ΔT. The previous experiment was repeated with different ΔT, namely 5 s, 10 s and 20 s. We were not able to recognise a qualitative difference in mean and variance of the global error measures, and therefore no additional plots are included. A similar behaviour can be observed for extracted travel times, except that they show a slightly higher variation for higher ΔT. In sum, it is possible to state that ΔT has a low influence on the quality of a reconstruction by the ASM, except when FC velocities have to be estimated from coordinates as described in Sec. 7.1.



(a) 6 % (b) 0.7 % (c) 0.3 %

Figure 7.19: Calculated travel times obtained from a spatio-temporal reconstruction for different rates of floating cars (50 runs). On average the travel times correspond well with the travel times obtained from the ground-truth field.

7.3.3 Reconstruction with Combined Data-Sources – Data Fusion

This section describes our conducted experiments with focus on ASM variants suitable for heterogeneous input data, namely the GASM and the EGTF. As these experiments were performed without presence of measurement noise, all noise dependent algorithm parameters can be ignored for both algorithms.

Again, the input data was obtained from the same microsimulation as used in the previous experiments (illustrated in Fig. 7.10(a)) and can be characterised as follows: 14 stationary detectors provide one-minute aggregated velocity and flow data; floating car trajectories also have velocity information and a minimal sampling time interval of $\Delta T = 1\,\mathrm{s}$.

Fusion by Linear Combination – GASM

This section describes experiments regarding the GASM (described in Sec. 4.5.2). In short, the GASM principally performs data fusion by linear combination of spatio-temporal velocity fields, which can be written for two data sources as follows

$$V = \alpha \cdot V_{\mathrm{St}} + [1 - \alpha] \cdot V_{\mathrm{FC}}, \tag{7.11}$$

where V_{St} and V_{FC} are obtained by preceding reconstructions by the ASM from stationary and FC data, respectively. Treiber et al. propose a rule for selection of the weighting factor α in (eq. (4.27)) which mainly depends on the error-variance (i.e. $\mathrm{Var}(\Delta v)$) of the respective input data. In the current experimental setting no measurement noise is present, which yields equal weights of $\alpha = 0.5$.

However, that rule ignores the density of the input data. Therefore, the objective of this experiment is to evaluate the algorithm performance with respect to the actual value of α and different rates of floating cars; data from stationary sensors is constant. Once again, the experiments were performed repeatedly (50 runs) with a new sample of floating cars for every run. The resulting average error measures are illustrated in Fig. 7.20 and can be summarised as follows: Only for sufficiently high FC rates of approximately $\geq 5\,\%$, the minimum RMSE is reached with equal weighting of each data source, i.e. $\alpha = 0.5$. But with declining rates of available FC data, the minimum RMSE is moving continuously towards $\alpha = 1$ as illustrated in Fig. 7.20(a). All other error measures indicate an optimal performance with $\alpha = 0$ which is caused by the high residual bias in the stationary flow field V_{St} (see Sec. 7.3.1). This bias-phenomenon can also be observed at the mean travel times belonging to the experiment illustrated in Fig. 7.21: With an ascending weighting factor α, the bias in travel time increases steadily. One more time, the

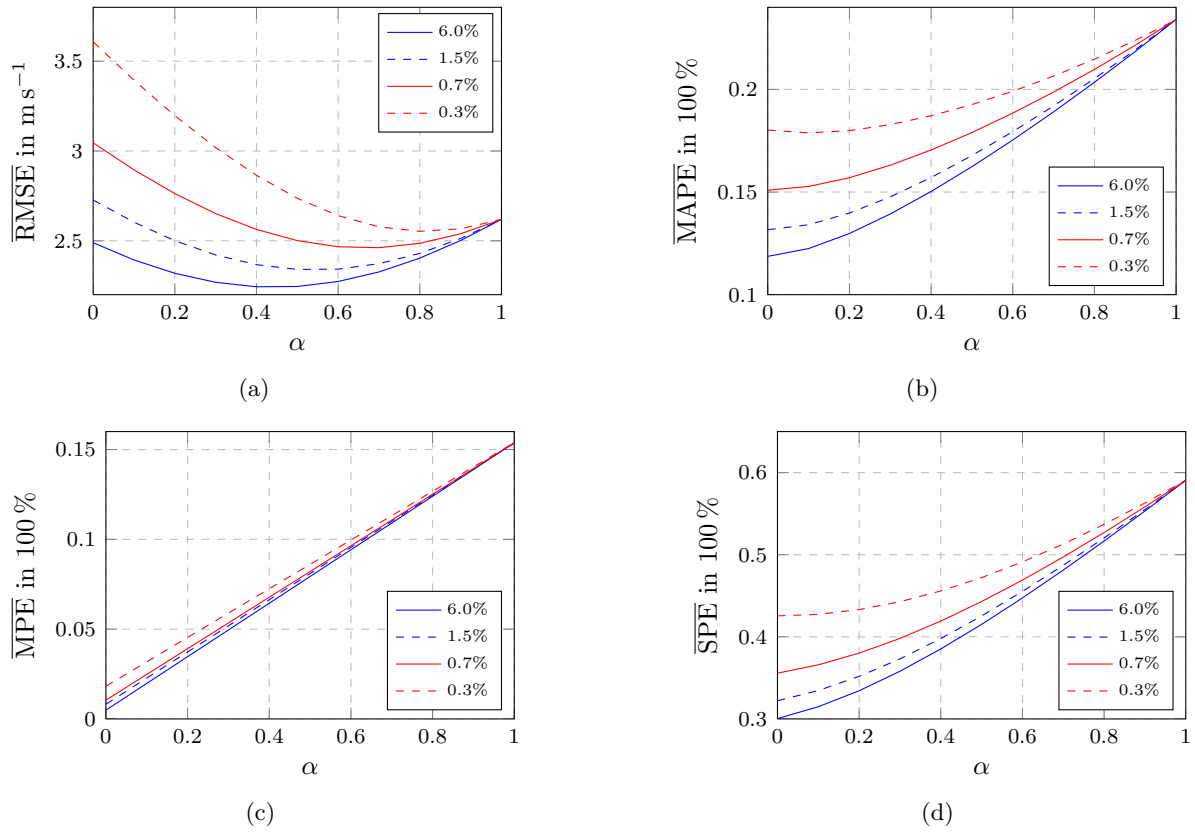


Figure 7.20: Average performance of the GASM for different weighting factors α without the presence of measurement noise (50 runs). The RMSE indicates a optimal performance with equal weights only for high FC rates (a), whereas the other error measures still indicate a large bias in the flow field obtained from stationary detector data (b)–(d).

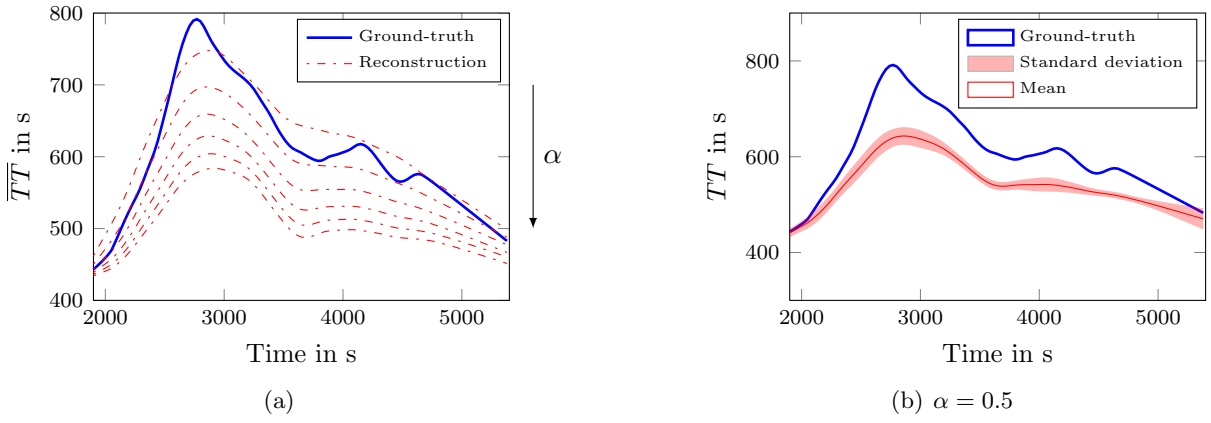


Figure 7.21: Illustration of extracted travel times from a GASM based reconstruction (FC frequency of 0.7 %, 50 runs): (a) mean travel times w.r.t. the weighting factor α (in the range of [0, 1]), and (b) mean travel times and standard deviation for $\alpha = 0.5$.

outcomes of the experiment are strongly affected by the bias of the velocity data obtained from stationary detectors.

Fusion with the EGTF

In this section we describe an experiment regarding the EGTF method (described in Sec. 4.5.3). The method is i.a. comprised of two noise dependent parameters for each data source j, namely $\Theta^{(j)}$ and $\mu^{(j)}$, which allow an adequate weighting with respect to the corresponding noise levels. As the data originating from microsimulation is not comprised of measurement errors one can set $\Theta^{(j)} = 1$ and $\mu^{(j)} = 0$. In contrast to the GASM, the EGTF is weighting each data source with respect to its actual data density. Thus, not only the relative FC rate but also the trajectory time sampling interval ΔT are of interest. A reconstruction based on the EGTF was performed using the same input data as in the previous section for a relative FC rate in the range of $0.1 - 2\,\%$ and time sampling intervals of ΔT of 1 s, 5 s, 10 s and 20 s. The respective mean global errors are illustrated in Fig. 7.22.

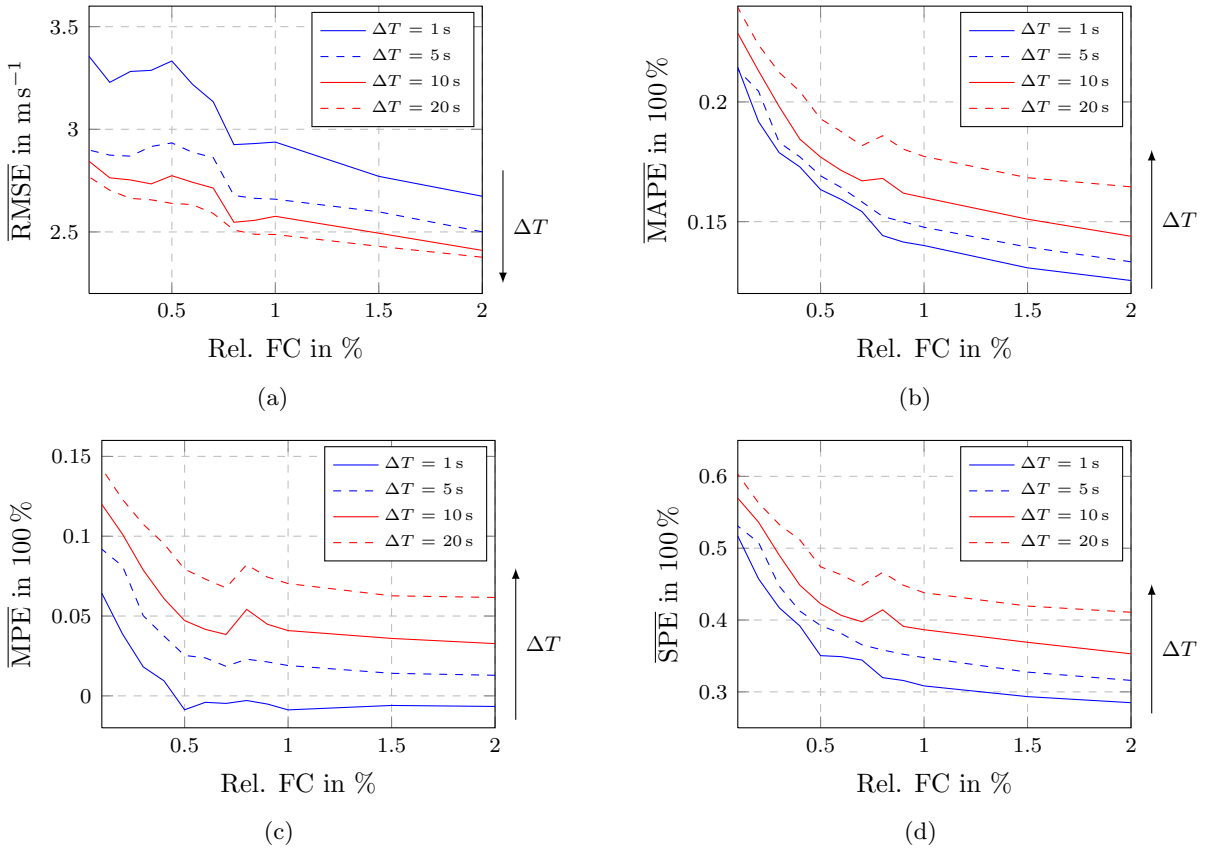
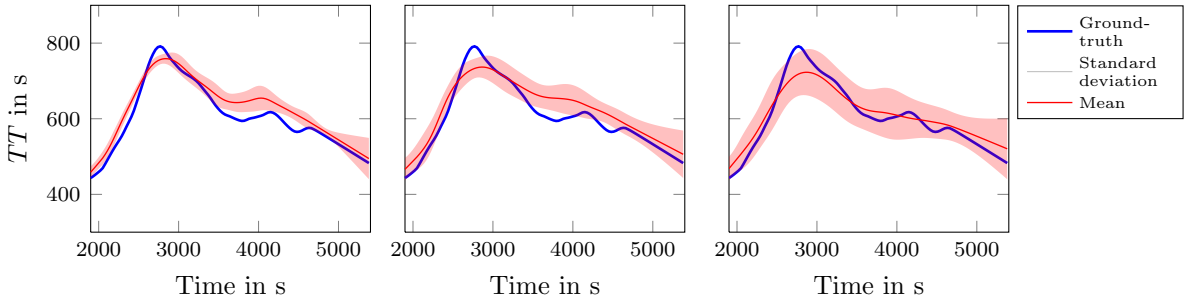


Figure 7.22: Average error measures of the EGTF method for different floating cars and trajectory sampling time intervals ΔT (50 runs). The RMSE surprisingly indicates the best performance with the largest ΔT (a), whereas all other measures indicate the best performance with the lowest ΔT.

As might be expected, almost every error measure reaches a minimum at the maximum relative flow of floating cars and the minimum ΔT. However, this is not true for the RMSE, where the minimum is achieved with $\Delta T = 20\,\text{s}$, which may be explained by the fact that on the basis of data from stationary detectors it is possible to achieve a lower RMSE when compared to FC data (compare Sec. 7.3.1 and Sec. 7.3.2). Furthermore, the trajectory sampling interval ΔT has an influence on the number of data points in the MBR $A(x,t)$, and thus the respective weighting gets influenced implicitly. Extracted travel times for a constant $\Delta T = 1\,\text{s}$ and different flow rates are illustrated in Fig. 7.23. Even for very low FC rates of 0.3 % travel times extracted from a reconstruction show a relatively low structural bias indicated by an MPE of approximately 3 %.



(a) 2.0 % (b) 0.7 % (c) 0.3 %

Figure 7.23: Illustration of mean travel times obtained from an EGTF based reconstruction for different rates of floating cars and a trajectory sampling time interval of $\Delta T = 1\,\mathrm{s}$ (50 runs). Even with a low FC rate the method is able to estimate the travel time adequately.

.

In summary, with both data-fusion methods it is possible to achieve notable enhancements in the reconstruction quality when compared to similar methods performed with data from stationary detectors only. For the GASM on the other hand, such improvements are only possible under the provision that the floating car data is available with an adequately high spatio-temporal resolution. In general, the EGTF outperforms the GASM, especially with respect to the density of the required FC data and its ability to correct the velocity bias.

7.3.4 ASM Performance under the Presence of Noise

In this section we provide information about the performance of the ASM method with respect to the presence of noise. As already mentioned in Chapter 5, it is difficult to gather information about the real measurement errors from sensor manufacturers. In order to be able to perform a simulation of measurement noise nonetheless, it is necessary to make certain assumptions: (1) The probability density (pdf) of the noise sources is assumed to be known, i.e. Gaussian distributed in the most cases; (2) The indicated uncertainty-ranges are interpreted as the $\pm 3\sigma$ region of the Gaussian curve.

Due to time reasons, we solely examined the performance of the ASM, i.e. the reconstruction based on a single data source only. A description of the experiments with stationary- and floating car data is provided in the following two sections.

Erroneous Stationary Detector Data

This section describes our experiments regarding the ASM with respect to erroneous measurements[31] performed by stationary velocity sensors. At a stationary detector, individual vehicle velocities v_α are aggregated to one-minute time averages v_i. When assuming a Gaussian measurement noise model, denoted by

$$v_i \approx \overline{v}_{\alpha,i} + \overline{\eta}_{\mathrm{v},i} \qquad \text{with} \qquad \overline{\eta}_{\mathrm{v},i} \sim \mathcal{N}(0, \sigma_{\eta,i}), \tag{7.12}$$

$\overline{v}_{\alpha,i}$ represents the time-aggregated velocity absence of noise. Further, $\sigma_{\eta,i}$ depends on the single-vehicle error variance σ_η and the momentary flow rate in the respective aggregation interval as

[31] Only random errors, to be precise.

follows:

$$\sigma_{\eta,i} = \frac{\sigma_\eta}{\sqrt{q_i}}. \tag{7.13}$$

The experiments were performed repeatedly (i.e. 200 runs) on the basis of 14 stationary detectors as illustrated in Fig. 7.10(a) whereby the error variance σ_η was varied in the range of $0-0.46\,\mathrm{m\,s^{-1}}$ which corresponds to specified uncertainties in the range of $0-\pm10\,\mathrm{km\,h^{-1}}$ (cf. Tab. 5.2). The respective mean and standard deviation of the global error measures are shown in Fig. 7.24. The

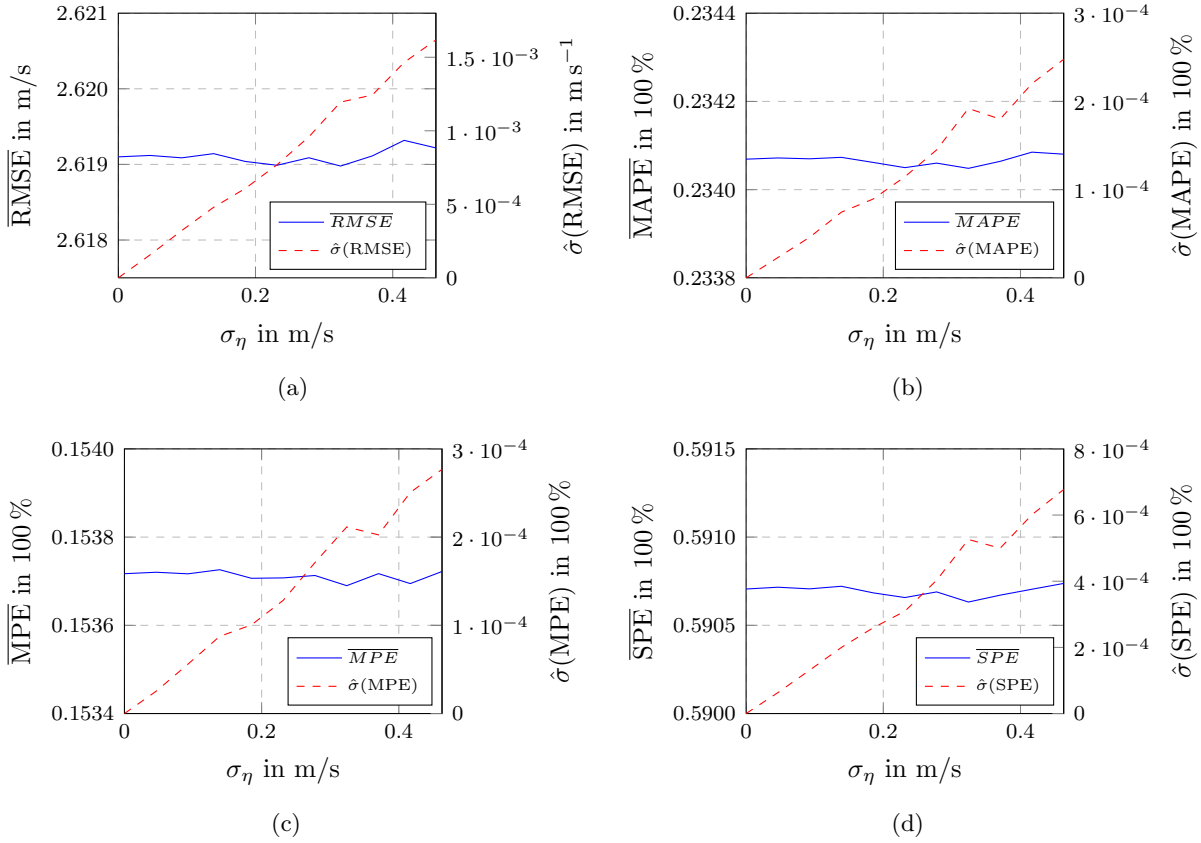


Figure 7.24: Mean and standard deviation s of the global error measures obtained by ASM based reconstruction from noisy input data with respect to the levels of measurement noise.

.

outcomes are comparable for each error measure and can be summarised as follows: The average performance may be assumed as independent, while the standard deviation $\hat{\sigma}$ almost linearly depends on σ_η. Nevertheless, the influence of the noise is comparably low in contrast to the influence of the relative alignment between the respective congestion pattern and the available data (cf. Fig. 7.13).

Erroneous Floating Car Data

This section describes our experiments regarding the ASM with respect to erroneous measurements performed by GPS receivers under the assumption that both, position and velocity information are available. The CEP is a common measure for the position error, which is specified at around 3 m for modern GPS receivers. Furthermore, when GPS position errors are assumed

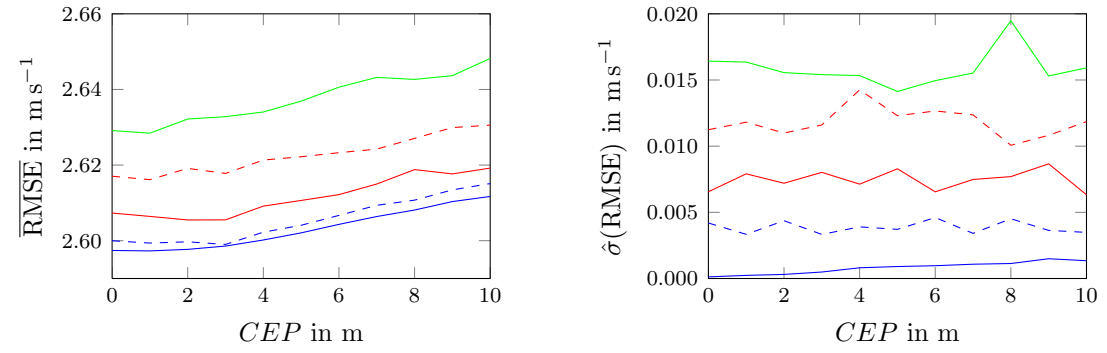
to be bivariate Gaussian, the standard deviation $\sigma_{x,y}$ is directly related to the CEP by

$$\sigma_{x,y} = 0.85 \cdot CEP. \tag{7.14}$$

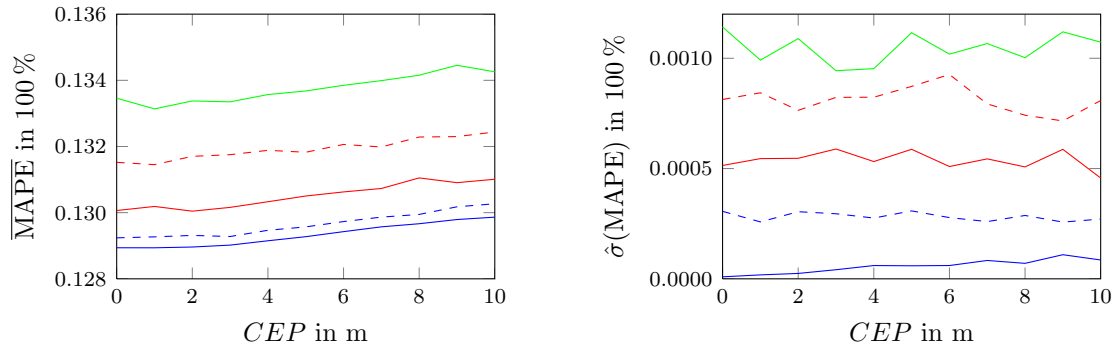
For velocity measurements based on the Doppler effect, one may expect a measurement uncertainty of $\pm 0.1\,\mathrm{m\,s^{-1}}$, which corresponds to a $\sigma_{\Delta v}$ of $1/30\,\mathrm{m\,s^{-1}}$ under the assumption of Gaussian distributed measurement errors. On the other hand, with finite difference based velocity estimation methods one has to expect much larger standard deviations, e.g. $\sigma_{\Delta v} \approx 4\,\mathrm{m\,s^{-1}}$ for high ΔT.

On the basis of floating car data with a constant flow rate of 2 % and a time sampling interval $\Delta T = 1\,\mathrm{s}$, we performed a repeated reconstruction by the ASM, where the parameters CEP and $\sigma_{\Delta v}$ were varied in the ranges of $0 - 10\,\mathrm{m}$ and $1/30 - 4\,\mathrm{m\,s^{-1}}$, respectively (50 runs for every tuple of CEP and $\sigma_{\Delta v}$). Mean and standard deviation of the obtained global error measures are illustrated in Fig. 7.25.

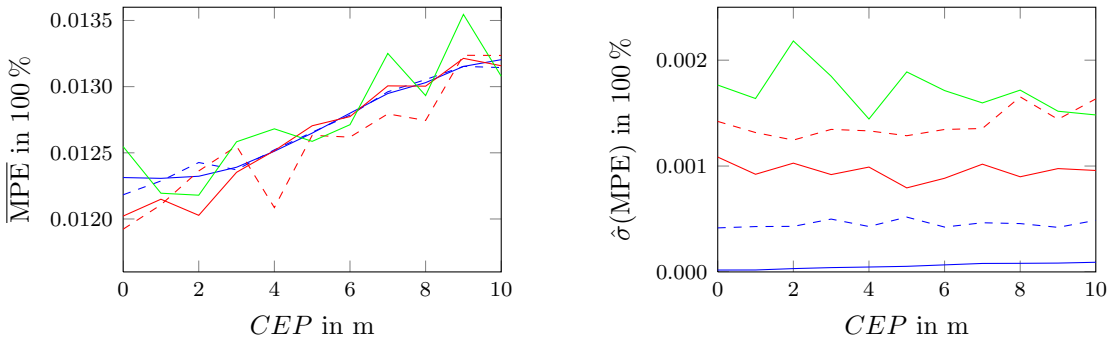
The findings of the experiment can be summarised as follows: Both, the average RMSE and



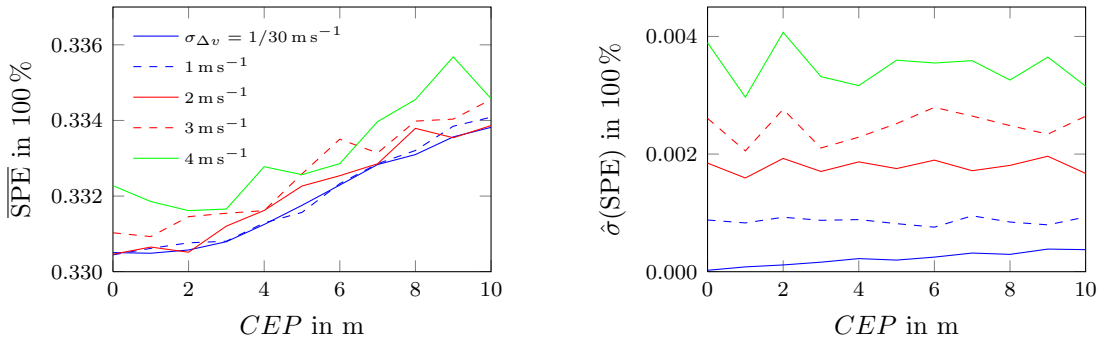
(a)



(b)



(c)



(d)

Figure 7.25: Comparison of mean and variance of global error measures obtained from repeated ASM reconstructions (50 runs) with respect to the CEP and σ_v plotted against the horizontal axis and as particular curves, respectively.

MAPE increase with the CEP and $\sigma_{\Delta v}$, whereas the MPE as an indicator of the structural bias remains relatively constant. Overall, these experimental findings are comparable with those of the preceding experiment based on stationary detector data. It is possible to notice an impact of measurement noise on the reconstruction quality. However, this impact is comparably low in contrast to the density of available input data which has the largest impact on the final reconstruction quality.

7.4 Summary

This chapter gave an overview of most of the relevant experiments conducted within this thesis. First, we evaluated the performance of relatively simple floating car velocity estimation methods based on finite differences. In general, such methodologies can be used where floating car velocity information is not available directly, as it was the case for floating car trajectories collected within the project REFEREE. The experiments show that the expected error variances may be remarkably higher when compared to velocities obtained by a GPS device immediately. Hence, we absolutely recommend the instantaneous acquisition of floating car velocities, either by GPS or by an odometer.

Second, we described the results of two different approaches for calibration: (1) The steady state traffic flow model after Van Aerde and Rakha (velocity-density relation) was calibrated separately for each cross-sectional sensor. (2) An attempt has been made to find optimal parameters for an ASM-based velocity field reconstruction with floating car data from the project REFEREE as input-data and cross-sectional data as validation-data. Unfortunately, the used real-world data was not very suitable for the proposed calibration approaches. On the one hand, the amount of data from stationary detectors representing congested traffic conditions was too low. On the other hand, the floating car trajectories were available without velocity information and only with a relatively low trajectory density. Furthermore, the used floating car data and cross-sectional data were insufficiently consistent with each other. Nevertheless, with the first calibration approach, we were able to indicate a point of improvement of the ASM, namely the use of a spatially varying $V_{\text{thr}}(x)$.

Finally, we provided comprehensive information about experiments with the ASM algorithm family with special focus on their sensitivity concerning the quality of the input data. We were able to identify (1) the bias of time aggregated velocities, (2) the density of the available data, and (3) the relative alignment between the data and the particular congestion patterns as those

properties of the input data with the highest impact on the estimation quality. With this in mind, it is obvious, that with a reconstruction obtained by the EGTF it is possible to achieve better results when compared to estimation from single data sources.

8 Conclusion

This thesis not only covered a comprehensive literature review, but also provided information about the efficient implementation, as well as an evaluation of traffic state estimation methods suitable for the Austrian motorway system. As required, we discussed state estimation methods suitable for homogeneous as well as heterogeneous data sources, such as data from stationary detectors and GPS equipped floating cars.

Our literature review focussed on the following aspects: an introduction to the basic terms of traffic engineering, followed by an overview of state-of-the-art methodologies for dynamic traffic flow modelling, as well as an overview of applicable data-fusion concepts. Under consideration of our research objective we introduced the categories of dynamic traffic flow models and covered the whole spectrum of first-order macroscopic traffic flow models, as they are most likely applicable for fusion of heterogeneous traffic data. Building on those models, we not only provided different concepts of data-fusion well-known in traffic engineering (e.g. Kalman filtering methods), but also outlined concepts popular in other fields of science (e.g. Newtonian relaxation methods and 4D-Var methods). As the thesis was developed in cooperation with two university institutes of different specialised backgrounds[32], it was taken into account that people with either one of the backgrounds are able to comprehend this work.

Furthermore, we provided an overview of our test-site, the Austrian A4 motorway including a detailed discussion of all available data sources. As those sources were important for a final data-fusion, we put special focus on required data preprocessing steps and the common measurement uncertainties. Beyond this, we implemented the heuristic traffic state estimation algorithms ASM, GASM, and EGTF, with some of them being suitable for incorporating spatio-temporal velocity information from heterogeneous sources. In order to evaluate those methods we implemented a complete evaluation system on the basis of PTV-VISSIM and MATLAB, which allows the generation of synthetic traffic data for both traffic state estimation and subsequent validation purposes. Building on this evaluation system, we conducted a comprehensive experimental evaluation of (1) required preprocessing methods and (2) the implemented state estimation methods. Lastly, we outlined two approaches for calibration of the ASM on the basis of real world data.

Although open questions regarding the methods under investigation still remain, further research within this thesis is not possible due to time issues. The same also applies to the proposed implementation of a second traffic state estimation method.

[32] Namely civil engineering and electrical engineering.

Before finishing this thesis with a final outlook we provide a number of practical considerations about working with heterogeneous data and a summary of our main findings and results.

Lessons Learned – Practical Considerations

1. The general idea of data-fusion is to combine information from different data sources in order to improve the quality of knowledge about a dynamic system. By its nature, heterogenous data is likely to be obtained by different external organisations, e.g. stationary data by the ASFINAG, and thus it is not possible to gather data on demand. Furthermore, a data inquiry may go along with long delivery times or might not even cannot be fulfilled at all.

2. Even if sufficient quantities of data from different sources are available in the same spatio-temporal frame, it is not guaranteed that the data is well posed for conducting experiments. For example, when the data is obtained solely under free-flow conditions the spatio-temporal field may contain too little information in order to carry out certain analyses.

3. In the beginning, the required effort for data-preprocessing (e.g. trajectory matching, trajectory velocity estimation, and velocity bias correction) had been underestimated. Even if these tasks are not connected directly to our main research objective, they have a significant effect on the performance of the discussed data-fusion methods.

4. Because of the aforementioned reasons we propose microsimulation software as the tool of choice for a comprehensive evaluation of traffic state estimation methods, with the only drawback that data from microsimulation may not agree in detail with real-world measurements.

Main Contributions and Results

1. The detailed record of contemporary first-order macroscopic traffic modelling techniques and respective estimation algorithms was created on the basis of a variety of individual scientific publications. At the moment no similar survey about this topic is known. Therefore, this thesis should enable the readers to become a more holistic view of that topic.

2. It was possible to show, that a velocity field reconstruction on Austrian motorways based on the ASM can produce satisfactory results when using a single data-source. Even better results can be achieved by combining more data-sources by using the EGTF (stationary detectors and floating cars). However, both categories of data-sources imply advantages and disadvantages. Considering data from stationary detectors it is easy to specify appropriate interpolation kernel parameters σ and τ on the basis of average data spacings. However, a velocity bias correction is mandatory. In the case of floating car data, which is irregularly distributed the mentioned parameters cannot be tuned by a rule of thumb. But the positive aspect here is that floating car data is not comprised of any bias.

3. For the ASM, Treiber and Helbing propose a spatio-temporally constant threshold velocity V_{thr} for the differentiation between free-flow and congested traffic conditions (cf. 4.20). But our experiments show that for our test-site the optimal velocity v_{opt} is increasing continuously with rising distance to the city centre. (This behaviour may be generally expected for urban motorways.) By introducing a spatially varying $V_{\text{thr}}(x)$ we were able to slightly improve the results under the presence of homogeneous congested traffic patterns.

Even greater improvements may be expected under the presence of triggered stop or go patterns or oscillating congested traffic.

4. The evaluation of reconstructed velocity fields obtained with data from the real traffic system may be difficult, as the acquisition of reference measurements along the road section is an expensive task. The comparison of extracted travel times from a spatio-temporal velocity field with travel time data obtained with an ANPR system may be a feasible and more economic way to evaluate the quality of a velocity field reconstruction with regard to traffic dynamics and bias (cf. Fig. 7.9 and 7.19).

Outlook

As mentioned throughout the thesis, several open issues regarding our research still remain: (1) The velocity bias correction scheme by multiplication is not able to reduce the bias adequately. Thus, we propose the implementation of a more sophisticated correction method, e.g. the method proposed by Yuan [85]. (2) The congestion pattern extensively used for evaluation has a different structure than the congestion patterns observed in the real world. In this regard, the conducted experiments may not be representative of real-world traffic conditions. Therefore, the range of validity of the VISSIM model used has to be extended by several hours. In order to generate realistic congestion patterns, it may be necessary to control VISSIM via the COM programming interface. (3) Floating car velocity informations can be measured by the Doppler effect with relatively low measurement errors, but such velocity information is not available sometimes. On the other hand, simple finite difference based methods, i.e. forward/backward difference or PCHIP, come with comparatively high error variances. A systematic investigation of different state-of-the-art vehicle velocity estimation methods would provide further assistance to that research objective. (4) Finally, we were able to identify several promising data-fusion methods within our literature review. However, it is not possible to evaluate all of them within a single master's thesis. Consequently, we propose the scheduling of follow-up themes with focus on evaluation of modern data fusion concepts, e.g. based on model the formulation in Lagrangian coordinates or on variational formulation.

Bibliography

[1] M. Treiber and A. Kesting, *Verkehrsdynamik und -simulation - Daten, Modelle und Anwendungen der Verkehrsflussdynamik*. Springer, 2010. 18, 26, 27, 28, 32, 40, 42, 44, 55, 83

[2] H. Lieu, "Traffic Flow Theory: A state of the art report - Revised Monograph on Traffic Flow Theory," 2002, transportation Research Board, Washington. [Online]. Available: http://www.fhwa.dot.gov/publications/research/operations/tft/ 18, 26

[3] W. Brilon, M. Großmann, and H. Blanke, "Verfahren für die Berechnung der Leistungsfähigkeit und Qualität des Verkehrsablaufes auf Straßen," *Schriftenreihe Forschung Straßenbau und Straßenverkehrstechnik*, vol. 669, 1994. 18, 54, 55

[4] V. Knoop, S. Hoogendoorn, and H. van Zuylen, "Empirical Differences between Time Mean Speed and Space Mean Speed," in *Proceedings of Traffic and Granular Flow 07*, 2007. 18

[5] L. Chu, J.-S. On, and W. Recker, "Adaptive Kalman Filter Based Freeway Travel time Estimation," in *84th TRB Annual Meeting*. Transportation Research Board, Washington D.C., Jan. 2005. 19

[6] M. Regler, "Verkehrsablauf und Kapazität auf Autobahnen," Ph.D. dissertation, Ruhr Universität Bochum, 2004. 20, 79

[7] L. Immers and S. Logghe, "Course Notes for the Lecture Verkeerskunde Basis: Traffic Flow Theory," May 2002. [Online]. Available: https://www.mech.kuleuven.be/cib/verkeer/dwn/H111part3.pdf 20

[8] B. Kerner, "The physics of traffic," *Physics World Magazine*, vol. 12, pp. 25–30, Aug. 1999. 20

[9] W. Brilon and H. Zurlinden, "Überlastwahrscheinlichkeit und Verkehrsleistung als Bemessungskriterium für straßenverkehrsanlagen," in *Forschung Straßenbau und Straßenverkehrstechnik*. Bundesministerium für Verkehr, Bau- und Wohnungswesen, Abteilung Straßenbau, Straßenverkehr, Bonn, August 2010, vol. 9, no. 870. 20

[10] D. Helbing, M. Treiber, A. Kesting, and M. Schönhof, "Theoretical vs. empirical classification and prediction of congested traffic states," *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 69, no. 4, pp. 583–598, 2009. [Online]. Available: http://EconPapers.repec.org/RePEc:spr:eurphb:v:69:y:2009:i:4:p:583-598 21

[11] A. Kesting and M. Treiber, "Datengestützte Analyse der Stauentstehung und -ausbreitung auf Autobahnen," *Straßenverkehrstechnik*, vol. 1, pp. 5–11, 2010. 21, 22, 39, 40, 52

[12] M. J. Lighthill and G. B. Whitham, "On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads," *Royal Society of London Proceedings Series A*, vol. 229, pp. 317–345, May 1955. 23, 27

[13] S. P. Hoogendoorn and P. H. L. Bovy, "State-of-the-art of Vehicular Traffic Flow Modelling," in *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 215, no. 4, Jun. 2001, pp. 283–303. 23, 26

[14] D. Ni, "A spectrum of traffic flow modeling at multiple scales," in *Simulation Conference (WSC), Proceedings of the 2010 Winter*, Dec. 2010, pp. 554–566. 24

[15] L. Leclercq, "Hybrid approaches to the solutions of the "Lighthill–Whitham–Richards" model," *Transportation Research Part B: Methodological*, vol. 41, no. 7, pp. 701 – 709, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191261506001433 24

[16] W. Huang, "HyTran: A New Approach for the Combination of Macroscopic and Microscopic Traffic Flow Models," Ph.D. dissertation, Graz University of Technology, Austria, 2013. 24

[17] C. F. Daganzo, "Requiem for second-order fluid approximations of traffic flow," *Transportation Research Part B: Methodological*, vol. 29, no. 4, pp. 277 – 286, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/019126159500007Z 26

[18] M. Papageorgiou, "Some remarks on macroscopic traffic flow modelling," *Transportation Research Part A: Policy and Practice*, vol. 32, no. 5, pp. 323 – 329, 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0965856497000487 26

[19] N. H. Gartner, C. J. Messer, and A. Rathi, *Revised Monograph on Traffic Flow Theory*. Transportation Research Board, 2001. [Online]. Available: http://www.fhwa.dot.gov/publications/research/operations/tft/ 26

[20] G. Newell, "A simplified theory of kinematic waves in highway traffic, part I: General theory," *Transportation Research Part B: Methodological*, vol. 27, no. 4, pp. 281 – 287, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/019126159390038C 28, 29

[21] ——, "A simplified theory of kinematic waves in highway traffic, part II: Queueing at freeway bottlenecks," *Transportation Research Part B: Methodological*, vol. 27, no. 4, pp. 289 – 303, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/019126159390039D 28, 29

[22] ——, "A simplified theory of kinematic waves in highway traffic, part III: Multi-destination flows," *Transportation Research Part B: Methodological*, vol. 27, no. 4, pp. 305 – 313, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/019126159390040H 28, 29

[23] R. Ansorge, "What does the entropy condition mean in traffic flow theory?" *Transportation Research Part B: Methodological*, vol. 24, no. 2, pp. 133 – 143, 1990. [Online]. Available: http://www.sciencedirect.com/science/article/pii/019126159090024S 29

[24] R. Herbin and L. Leclercq, ""A Note on the Entropy Solutions of the Hydrodynamic Model of Traffic Flow" Revisited," *Transportation Science*, vol. 45, no. 1, pp. 138–142, Feb. 2011. [Online]. Available: http://dx.doi.org/10.1287/trsc.1100.0342 29

[25] I. S. Strub and A. M. Bayen, "Weak formulation of boundary conditions for scalar conservation laws: an application to highway traffic modelling," *International Journal of*

Robust and Nonlinear Control, vol. 16, no. 16, pp. 733–748, 2006. [Online]. Available: http://dx.doi.org/10.1002/rnc.1099 29

[26] C. F. Daganzo, "A variational formulation of kinematic waves: basic theory and complex boundary conditions," *Transportation Research Part B: Methodological*, vol. 39, no. 2, pp. 187 – 196, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191261504000487 29, 47

[27] ——, "A variational formulation of kinematic waves: Solution methods," *Transportation Research Part B: Methodological*, vol. 39, no. 10, pp. 934 – 950, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191261505000068 29, 30, 47

[28] ——, "On the Variational Theory of Traffic Flow: Well-Posedness, Duality and Applications," *Networks and Heterogeneous Media*, vol. 1, pp. 601–619, 2006. 30

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004. 30

[30] C. G. Claudel and A. M. Bayen, "Lax-Hopf Based Incorporation of Internal Boundary Conditions Into Hamilton-Jacobi Equation. Part I: Theory," *IEEE Trans. Autom. Control*, vol. 55, no. 5, pp. 1142–1157, 2010. 30

[31] ——, "Lax-Hopf Based Incorporation of Internal Boundary Conditions Into Hamilton-Jacobi Equation. Part II: Computational Methods," *IEEE Trans. Autom. Control*, vol. 55, no. 5, pp. 1158–1174, 2010. 30, 47

[32] P.-E. Mazaré, A. H. Dehwah, C. G. Claudel, and A. M. Bayen, "Analytical and grid-free solutions to the Lighthill–Whitham–Richards traffic flow model," *Transportation Research Part B: Methodological*, vol. 45, no. 10, pp. 1727 – 1748, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191261511001044 30, 47

[33] "Mobile Millennium," Website, University of California, Berkely, [Visited on 02.08.2013]. [Online]. Available: http://traffic.berkeley.edu/ 30

[34] C. F. Daganzo, "The Cell Transmission Model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269–287, 1994. 31, 46

[35] ——, "The Cell Transmission Model, part II: Network Traffic," *Transportation Research Part B: Methodological*, vol. 29, no. 2, pp. 79–93, 1995. 31

[36] S. K. Godunov, "A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations," *Math. Sbornik*, vol. 47, pp. 271–306, 1959, translated US Joint Publ. Res. Service, JPRS 7225 Nov. 29, 1960. 31

[37] L. Leclercq, J. Laval, and E. Chevallier, "The Lagrangian coordinate system and what it means for first order traffic flow models," in *Proceedings of the 17th International Symposium on Transportation and Traffic Theory, London*, 2007. 32

[38] F. van Wageningen-Kessels, J. van Lint, S. Hoogendoorn, and C. Vuik, "Lagrangian Formulation of a Multi-class Kinematic Wave Model," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2188, pp. 29–36, 2010. 32, 46

[39] Y. Yuan, J. van Lint, S. Hoogendoorn, J. Vrancken, and T. Schreiter, "Freeway traffic state estimation using extended Kalman filter for first-order traffic model in Lagrangian coordinates," in *Networking, Sensing and Control (ICNSC), 2011 IEEE International Conference on*, Apr. 2011, pp. 121 –126. 32, 46

[40] K. Han, T. Yao, and T. L. Friesz, "Lagrangian-based Hydrodynamic Model: Freeway Traffic Estimation," 2012," arXiv:1211.4619. [Online]. Available: http://arxiv.org/abs/1211.4619 32, 47

[41] D. Work, O.-P. Tossavainen, Q. Jacobson, and A. Bayen, "Lagrangian Sensing: Traffic Estimation with Mobile Devices," in *American Control Conference, 2009. ACC '09.*, Jun. 2009, pp. 1536–1543. 35, 46

[42] J. W. C. van Lint and S. P. Hoogendoorn, "A Robust and Efficient Method for Fusing Heterogeneous Data from Traffic Sensors on Freeways," *Comp.-Aided Civil and Infrastruct. Engineering*, vol. 25, no. 8, pp. 596–612, 2010. 35, 40, 45, 47, 123

[43] D. Work, O.-P. Tossavainen, S. Blandin, A. Bayen, T. Iwuchukwu, and K. Tracton, "An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, Dec. 2008, pp. 5062 –5068. 36, 46, 123

[44] J. C. Herrera and A. Bayen, "Traffic flow reconstruction using mobile sensors and loop detector data," in *87th TRB Annual Meeting*, 2008. 36, 46, 123

[45] G. Welch and G. Bishop, "An Introduction to the Kalman Filter," University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, Tech. Rep., 1995. 37

[46] Z. Chen, "Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond," McMaster University, Tech. Rep., 2003. 38

[47] F. Bouttier and P. Courtier, "Data Assimilation Concepts and Methods," European Centre for Medium-Range Weather Forecasts (ECMWF), Tech. Rep., Mar. 1999, lecture Notes. 38

[48] Z. Li and I. M. Navon, "Optimality of variational data assimilation and its relationship with the Kalman filter and smoother," *Quarterly Journal of the Royal Meteorological Society*, vol. 127, no. 572, pp. 661–683, 2001. [Online]. Available: http://dx.doi.org/10.1002/qj.49712757220 38

[49] D. Volpi, "Estimation of parameters in traffic flow models using data assimilation," Master's thesis, University of Reading, United Kingdom, 2009. 38, 46

[50] D. Jacquet, C. C. de Wit, and D. Koenig, "Traffic Control and Monitoring with a Macroscopic Model in the Presence of Strong Congestion Waves," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, Dec. 2005, pp. 2164 – 2169. 38, 47

[51] D. Jacquet, M. Krstic, and C. C. de Wit, "Optimal control of scalar one-dimensional conservation laws," in *Proceedings of the 2006 American Control Conference, IEEE ACC'06*, Jun. 2006, p. 6 pp. 38, 47

[52] G. Cressman, "An operational objective analysis system," *Mon. Wea. Rev.*, vol. 87, pp. 367–374, 1959. 39

[53] M. Treiber and D. Helbing, "Reconstructing the spatio-temporal traffic dynamics from stationary detector data," *Cooper@tive Tr@nsport@tion Dyn@mics*, vol. 1, pp. 3.1–3.24, 2002. 39, 40, 42, 43, 44, 47, 79

[54] M. Treiber, A. Kesting, and R. E. Wilson, "Reconstructing the traffic state by fusion of heterogeneous data," *Comp.-Aided Civil and Infrastruct. Engineering*, vol. 26, no. 6, pp. 408–419, 2011. 40, 43, 44, 47, 123

[55] E. W. Weisstein, "Fourier Transform–Exponential Function. From MathWorld—A Wolfram Web Resource," visited on October 15, 2013. [Online]. Available: http://mathworld.wolfram.com/FourierTransformExponentialFunction.html 43

[56] T. Schreiter, "Vehicle-class Specific Control of Freeway Traffic," Ph.D. dissertation, Delft University of Technology, 2013. 44

[57] J. Sau, N.-E. E. Faouzi, A. B. Aissa, and O. de Mouzon, "Particle filter-based real-time estimation and prediction of traffic conditions," *Applied Stochastic Models and Data Analysis*, vol. 12, 2007. 46

[58] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez, "Traffic Density Estimation with the Cell Transmission Model," in *Proceedings of the American Control Conference, 2003.*, vol. 5, Jun. 2003, pp. 3750–3755. 46

[59] X. Sun, L. Munoz, and R. Horowitz, "Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, HI, 2003, pp. 9–12. 46

[60] ——, "Mixture Kalman filter based highway congestion mode and vehicle density estimator and its application," in *American Control Conference*, Boston, MA, 2004, pp. 2098–2103. 46

[61] M. Papageorgiou, J.-M. Blosseville, and H. Hadj-Salem, "Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris," *Transportation Research Part B: Methodological*, vol. 23, no. 1, pp. 29–47, 1989. 46

[62] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended Kalman filter: a general approach," *Transportation Research Part B*, vol. 39, no. 2, pp. 141–167, Feb. 2005. 46

[63] A. Hegyi, D. Girimonte, R. Babuska, and B. De Schutter, "A comparison of filter configurations for freeway traffic state estimation," in *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, Sep. 2006, pp. 1029 –1034. 46

[64] A. Hegyi, L. Mihaylova, R. Boel, and Z. Lendek, "Parallelized Particle Filtering for Freeway Traffic State Tracking," in *Proceedings of the European Control Conference 2007*, Kos, Greece, Jul. 2007, pp. 2442–2449. 46

[65] J. Abdi, B. Moshiri, E. Jafari, and A. Sedigh, "Traffic state variables estimating and predicting with extended Kalman filtering," in *Power, Control and Embedded Systems (ICPCES), 2010 International Conference on*, 29 2010-dec. 1 2010, pp. 1 –4. 46

[66] ——, "Traffic state variables estimating and predicting with Neural Network via Extended Kalman filter algorithm with estimated parameters as offline," in *Intelligent Systems and Informatics (SISY), 2010 8th International Symposium on*, Sep. 2010, pp. 383 –388. 46

[67] L. Mihaylova, A. Hegyi, A. Gning, and R. Boel, "Parallelized Particle and Gaussian Sum Particle Filters for Large-Scale Freeway Traffic Systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 36–48, Mar. 2012. 46

[68] R. Boel and L. Mihaylova, "Modelling Freeway Networks by Hybrid Stochastic Models," in *Intelligent Vehicles Symposium, 2004 IEEE*, Jun. 2004, pp. 182–187. 46

[69] ——, "A compositional stochastic model for real time freeway traffic simulation," in *Transportation Research*, vol. 40, no. 4, 2006, pp. 319–334. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191261505000573 46

[70] L. Mihaylova, R. Boel, and A. Hegyi, "An Unscented Kalman Filter for Freeway Traffic Estimation," in *Proc. of the 11th IFAC Symposium on Control in Transportation Systems*, The Netherlands, 2006, pp. 31–36. 46

[71] ——, "Brief paper: Freeway traffic estimation within particle filtering framework," *Automatica*, vol. 43, no. 2, pp. 290–300, Feb. 2007. 46

[72] H. Payne, "Models of freeway traffic and control," in *Mathematical Models of public systems, simulations councils proc.*, ser. 1. Beckey, G.A. (Ed.), 1971, pp. 51–60. 46

[73] D. B. Work, S. Blandin, O.-P. Tossavainen, B. Piccoli, and A. M. Bayen, "A Traffic Model for Velocity Data Assimilation," *Applied Mathematics Research eXpress*, vol. 2010, no. 1, pp. 1–35, 2010, http://amrx.oxfordjournals.org/content/2010/1/1.full.pdf+html. [Online]. Available: http://amrx.oxfordjournals.org/content/2010/1/1.abstract 46

[74] F. van Wageningen-Kessels, Y. Yuan, H. van Lint, S. Hoogendoorn, and C. Vuik, "Sinks and sources in Lagrangian coordinates: derivation, interpretation and simulation results," in *TRB Annual Meeting Online - TRB 90th Annual Meeting, January 23-27, 2011.* Washington DC: Transportation Research Board, 2011, pp. 1–16. 46

[75] Y. Yuan, J. van Lint, R. Wilson, F. van Wageningen-Kessels, and S. Hoogendoorn, "Real-Time Lagrangian Traffic State Estimator for Freeways," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 59–70, Mar. 2012. 46

[76] "Arterial traffic management products," Website, Smart Microwave Sensors GmbH, [Visited on 23.04.2013]. [Online]. Available: http://www.smartmicro.de/index.php/traffic-radar/arterial-management 52

[77] "TDC4 data sheet," Swarco AG. [Online]. Available: http://www.swarco.com/en/Products-Services/Traffic-Management/Detection-and-Enforcement/Traffic-Counting/TDC4-Traffic-detectors 52

[78] "TDC1 data sheet," Swarco AG. [Online]. Available: http://www.swarco.com/en/Products-Services/Traffic-Management/Detection-and-Enforcement/Traffic-Counting/TDC1-Traffic-detectors 52

[79] "TDC3 data sheet," Swarco AG. [Online]. Available: http://www.swarco.com/en/Products-Services/Traffic-Management/Detection-and-Enforcement/Traffic-Counting/TDC3-Traffic-detectors 52

[80] "UMRR Traffic Sensor Type 29 data sheet," Smart Microwave Sensors GmbH. [Online]. Available: http://www.swarco.com/en/Products-Services/Traffic-Management/Detection-and-Enforcement/Traffic-Counting/TDC1-Traffic-detectors 52

[81] "MultaStat CRM NG data sheet," Multanova AG. [Online]. Available: http://www.multanova.ch/de/produkte/einsatzgebiet.37/ 52

[82] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-time signal processing (2nd ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1999. 54, 76

[83] S. M. Kay, *Modern Spectral Estimation: Theory and Application*, ser. Prentice-Hall signal processing series. Englewood Cliffs N.J: Prentice Hall, 1988. 54

[84] J. W. C. V. Lint, "Reliable travel time prediction for freeways," Ph.D. dissertation, Netherlands TRAIL Research School, 2004. 54

[85] Y. Yuan, J. W. C. Van Lint, T. Schreiter, S. Hoogendoorn, and J. L. M. Vrancken, "Automatic speed-bias correction with flow-density relationships," in *Networking, Sensing and Control (ICNSC), 2010 International Conference on*, 2010, pp. 1–7. 54, 55, 101

[86] M. Linauer, "Generierung streckenbezogener Verkehrsdaten als Basis für den Einsatz in Verkehrstelematiksystemen," Ph.D. dissertation, Universität für Bodenkultur, Wien, Österreich, 2005. 55

[87] S. Turner and of Highway Information Management, United States. Federal Highway Administration. Office and Institute, Texas Transportation, *Travel time data collection handbook*. Office of Highway Information Management, Federal Highway Administration, US Dept. of Transportation, 1998. 55

[88] W. Mansfeld, *Satellitenortung und Navigation: Grundlagen, Wirkungsweise und Anwendung globaler Satellitennavigationssysteme*. Vieweg+Teubner Verlag, 2009. [Online]. Available: http://books.google.at/books?id=U_CrAa09i7UC 56, 57, 59

[89] United States. Assistant Secretary of Defense for Command, Control, Communications, and Intelligence, *Global Positioning System Standard Positioning Service Performance Standard*, 4th ed. Assistant Secretary of Defense for Command, Control, Communications, and Intelligence, 2008. 56

[90] R. B. Langley, "Dilution of precision," *GPS World*, pp. 52–59, 1999. 56

[91] F. van Diggelen, "GNSS Accuracy – Lies, Damn Lies and Statistics," *GPS World*, vol. 18, no. 1, pp. 26–32, Jan. 2007. 57

[92] "BT-Q1000X Travel Recorder Users Manual," Qstarz International Co., Ltd. [Online]. Available: http://www.qstarz.com/download.php?t=4&m=BT-Q1000X 57, 59

[93] "HOLUX Wireless GPS Logger M-241 user's manual," HOLUX Technology Inc. [Online]. Available: http://www.holux.com/JCore/en/products/products_content.jsp?pno=341 57, 59

[94] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang, "Map-matching for low-sampling-rate GPS trajectories," in *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ser. GIS '09. New York, NY, USA: ACM, 2009, pp. 352–361. [Online]. Available: http://doi.acm.org/10.1145/1653771.1653820 57

[95] H.-J. Bartsch, *Taschenbuch mathematischer Formeln*, Fachbuchverlag Leipzig, Ed. Carl Hanser Verlag, München, 2004. 58, 75, 113

[96] W. Niemeier, *Ausgleichsrechnung: Statistische Auswertemethoden*. De Gruyter, Berlin, 2008. 60, 70

[97] E. Tzoreff and B.-Z. Bobrovsky, "A Novel Approach for Modeling Land Vehicle Kinematics to Improve GPS Performance Under Urban Environment Conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 344–353, 2012. 60, 61

[98] Y. Cui and S. S. Ge, "Autonomous vehicle positioning with GPS in urban canyon environments," *Robotics and Automation, IEEE Transactions on*, vol. 19, no. 1, pp. 15–25, 2003. 61

[99] C. Barrios and Y. Motai, "Improving Estimation of Vehicle's Trajectory Using the Latest Global Positioning System With Kalman Filtering," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 12, pp. 3747–3755, 2011. 61

[100] C. B. Moler, *Numerical computing with MATLAB.* SIAM, 2004. 61

[101] D. Ni and H. Wang, "Trajectory reconstruction for travel time estimation." *J. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 113–125, 2008. 63

[102] J. W. C. van Lint and N. J. van der Zijpp, "Improving a travel-time estimation algorithm by using dual loop detectors," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1855, pp. 41–48, 2003. 63

[103] J. Barceló, Ed., *Fundamentals of Traffic Simulation*, ser. International Series in Operations Research & Management Science. Springer, 2011, vol. 145. 65

[104] "EPSG-Übersicht: Geodätisches Datum und Projektionen in Österreich," Bundesamt für Eich- und Vermessungswesen. [Online]. Available: http://www.bev.gv.at 66, 70, 71

[105] *Technische Lieferbedingungen für Streckenstationen: TLS 2012.* Deutschland. Bundesministerium für Verkehr, Bau- und Stadtentwicklung, 2012. [Online]. Available: http://www.bast.de/nn_42742/DE/Aufgaben/abteilung-v/referat-v5/v5-tls/tls-streckenstationen.html 67

[106] "SQLite Home Page," Website, The SQLite Team, [Visited on 05.08.2013]. [Online]. Available: https://www.sqlite.org/ 69

[107] M. Kortmann, "mksqlite Website," Website, [Visited on 05.08.2013]. [Online]. Available: http://mksqlite.berlios.de/ 69

[108] "PROJ.4 - Cartographic Projections Library," OSGeo - Open Source Geospatial Foundation, [Visited on 12. May 2013]. [Online]. Available: http://proj.osgeo.org 71

[109] T. Schreiter, H. van Lint, M. Treiber, and S. Hoogendoorn, "Two fast implementations of the Adaptive Smoothing Method used in highway traffic state estimation," in *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, 2010, pp. 1202–1208. 74, 76

[110] A. Guttman, "R-trees: a dynamic index structure for spatial searching," *SIGMOD Rec.*, vol. 14, no. 2, pp. 47–57, Jun. 1984. [Online]. Available: http://doi.acm.org/10.1145/971697.602266 75

[111] "GML LidarK Library," Website, Graphics & Media Lab, Lomonosov Moscow State University, Russia, [Visited on 06.08.2013]. [Online]. Available: http://graphics.cs.msu.ru/en/science/research/3dpoint/lidark 75

[112] "The SQLite R*Tree Module," Website, The SQLite Team, [Visited on 05.08.2013]. [Online]. Available: https://www.sqlite.org/rtree.html 75

[113] M. Van Aerde and H. Rakha, "Multivariate calibration of single regime speed-flow-density relationships [road traffic management]," in *Vehicle Navigation and Information Systems Conference, 1995. Proceedings. In conjunction with the Pacific Rim TransTech Conference. 6th International VNIS. 'A Ride into the Future'*, 1995, pp. 334–341. 79

[114] M. Ponzlet, "Auswirkungen von systematsystem und umfeldbedingten Schwankungen des Geschwindigkeitsverhaltens und deren Beschreibung in Verkehrsflußmodellen," *Schriftenreihe des Lehrstuhl für Verkehrswesen der Ruhr-Universität*, vol. Buch 16, 1996. 80

[115] Wikipedia, "Geographic coordinate system — Wikipedia, The Free Encyclopedia," 2013, [Online; accessed 25-July-2013]. [Online]. Available: http://en.wikipedia.org/w/index.php?title=Geographic_coordinate_system&oldid=564360867 113

[116] D. Pilone and N. Pitman, *UML 2.0 in a Nutshell*. O'Reilly Media, Inc., 2005. 125

A WGS84 Datum Resolution

The WGS84 datum is denoted in polar coordinates, namely longitude and latitude, and thus represents angles in the range of 0-180° and 0-90°, respectively. This implies that the numerical resolution of the geospatial datum does not reflect the spatial resolution on the spherical surface directly, i.e. the spacing of lines of longitude varies with the latitude as illustrated in Fig. A.1.

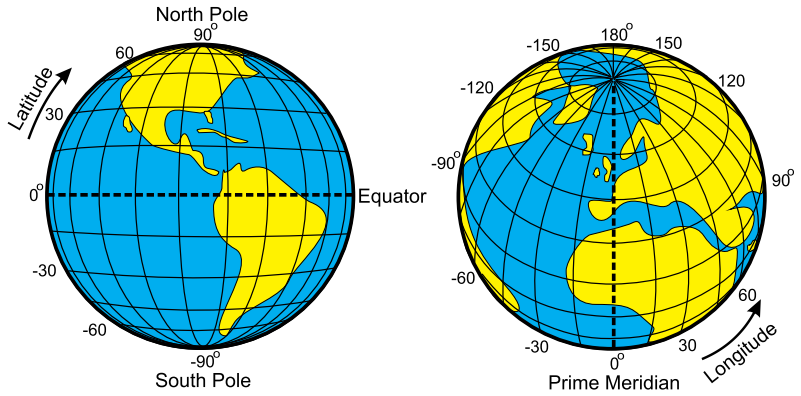


Figure A.1: The WGS84 coordinate system with polar coordinates (Source: [115])

In the following, we want specify the spatial resolution of a geodetic datum depending on the number of decimal places. Hereafter, the shape of the earth is simplified with an exact sphere. By the definition of the sea-mile [95], the spatial distance between two meridians differing in one arc-minute at a great circle, e.g. the equator, is $d_{\text{minute,lon}} \approx 1\,\text{seamile} = 1852\,\text{m}$. Consequently, the distance between two lines of latitude differing in one degree is

$$d_{\text{deg,lat}} = 60 \cdot d_{\text{minute,lat}} = 111\,120\,\text{m} = const. \tag{A.1}$$

The circumference of small circles for a given latitude is given by the great-circle circumference multiplied with the cosine of the latitude. Thus, the distance between two meridians also depends on the cosine of the particular latitude lat as

$$d_{\text{deg,lon}} = d_{\text{deg,lat}} \cdot \cos(lat). \tag{A.2}$$

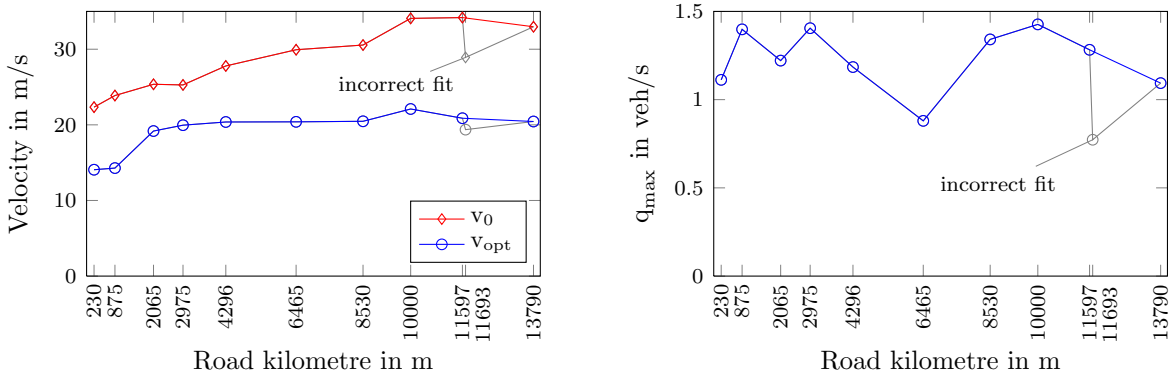
The spatial resolutions for equatorial regions and the Vienna region, i.e. with a latitude of approximately 48°, are summarised in Tab. A.1.

		Resolution in m		
			Longitude	
Decimal places	**Step-width**	Latitude	0°	48°
0	1°	111120	111120	74353
1	0.1°	11112	11112	7435
2	0.01°	1111	1111	743
3	0.001°	111	111	74.3
4	0.0001°	11.1	11.1	7.43
5	0.000 01°	1.11	1.11	0.743
6	0.000 001°	0.11	0.11	0.074

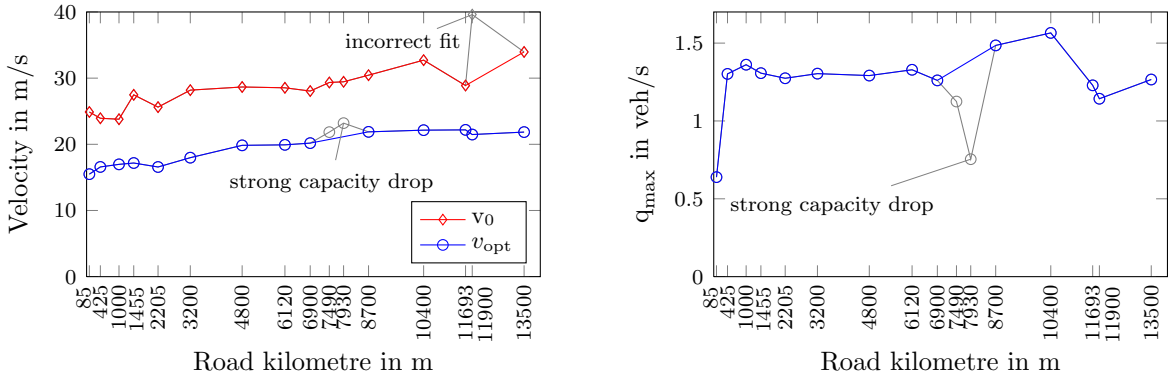
Table A.1: Spatial resolution of WGS84 dates in decimal format depending on the number of decimal places for equatorial and Vienna region (latitude of approx. 48°).

B Calibration

In this section we provide the complete results of the calibration methodology described in Sec. 7.2. The diagrams in Fig. B.1 and B.2 illustrate the calibration results, while Tab. B.1 provides the information in tabular form. Additionally, the fitted curves for every particular stationary sensor are illustrated in Fig. B.3 and B.4.

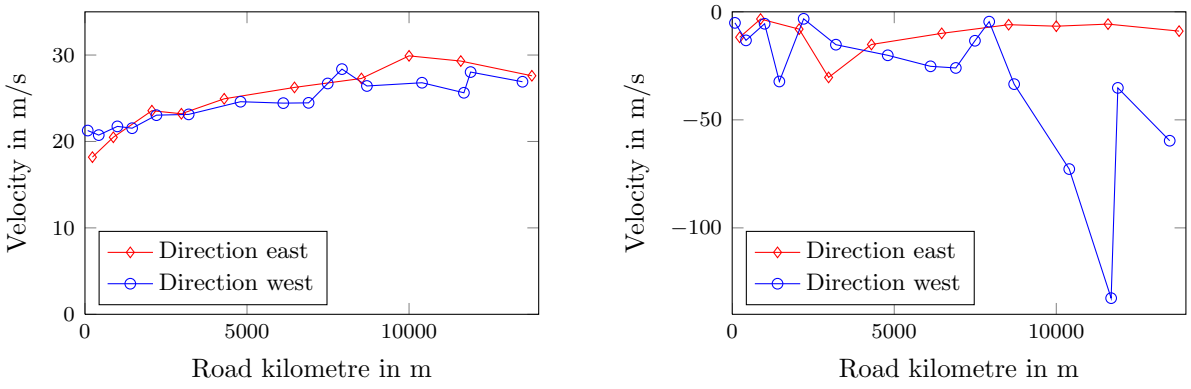


(a) Direction east



(b) Direction west

Figure B.1: Graphs representing the relevant results v_0, v_{opt}, and q_{max} of the mentioned calibration procedure, where invalid results are colored in grey.



(a) c_{free} (b) c_{cong}

Figure B.2: Illustration of the characteristic velocities c_{free} and c_{cong} derived from the calibrated van Aerde model: The methodology is able to provide realistic calibration results for c_{free} (a). However, there was not enough data to achieve plausible calibration results for c_{cong} (b).

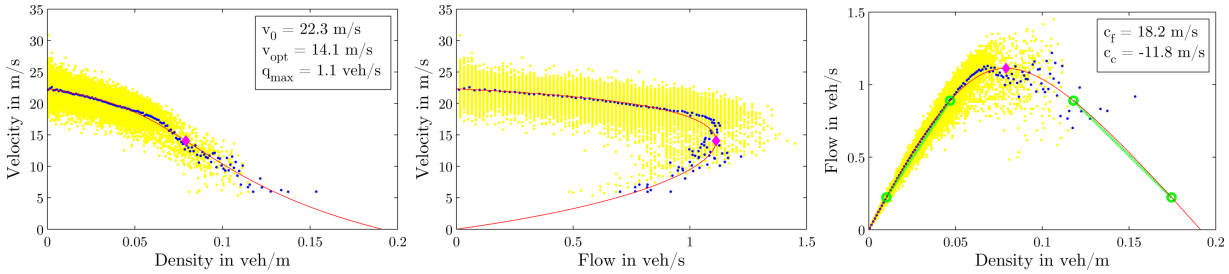
(a) Direction east

x in m	c_1 in m	c_2 in $\text{m}^2\,\text{s}^{-1}$	c_3 in s^{-1}	v_0 in $\text{m}\,\text{s}^{-1}$	v_{opt} in $\text{m}\,\text{s}^{-1}$	q_{max} in veh/s	c_f in $\text{m}\,\text{s}^{-1}$	c_c in $\text{m}\,\text{s}^{-1}$
230	3.42	40.45	0.31	22.34	14.07	1.11	18.19	-11.80
875	1.00	19.69	0.50	23.87	14.28	1.40	20.48	-3.33
2065	3.93	11.57	0.52	25.36	19.18	1.22	23.56	-8.02
2975	7.34	14.21	0.21	25.28	19.96	1.40	23.21	-30.36
4296	5.90	24.98	0.39	27.79	20.38	1.18	24.94	-15.12
6465	5.54	46.36	0.63	29.93	20.40	0.88	26.25	-9.97
8530	2.47	24.14	0.51	30.55	20.47	1.34	27.29	-5.95
10000	2.39	33.87	0.47	34.08	22.10	1.43	29.91	-6.65
11597	1.95	45.60	0.52	34.18	20.87	1.28	29.30	-5.68
11693	1.37	12.77	1.15	28.92	19.37	0.77	27.48	-1.55
13790	3.29	65.22	0.50	32.97	20.44	1.09	27.60	-8.93

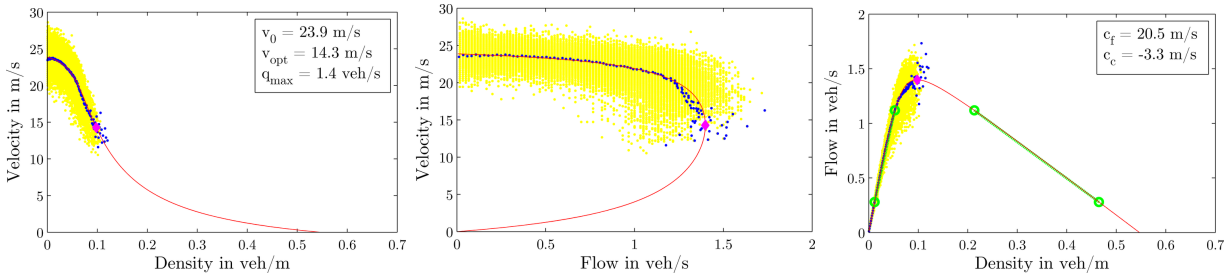
(b) Direction west

x in m	c_1 in m	c_2 in $\text{m}^2\,\text{s}^{-1}$	c_3 in s^{-1}	v_0 in $\text{m}\,\text{s}^{-1}$	v_{opt} in $\text{m}\,\text{s}^{-1}$	q_{max} in veh/s	c_f in $\text{m}\,\text{s}^{-1}$	c_c in $\text{m}\,\text{s}^{-1}$
85	3.56	51.55	0.98	24.88	15.49	0.64	21.26	-5.11
425	4.16	24.44	0.32	23.94	16.58	1.30	20.74	-13.22
1000	2.44	11.17	0.49	23.79	16.97	1.36	21.74	-5.53
1455	4.84	74.94	0.06	27.48	17.17	1.31	21.53	-32.25
2205	1.43	15.55	0.59	25.62	16.57	1.27	23.04	-3.24
3200	3.89	52.09	0.27	28.19	17.98	1.30	23.13	-15.21
4800	5.66	39.91	0.26	28.66	19.84	1.29	24.60	-20.12
6120	6.17	40.59	0.21	28.54	19.92	1.33	24.44	-25.23
6900	6.95	34.99	0.23	28.05	20.18	1.26	24.47	-25.98
7490	6.09	23.91	0.46	29.34	21.84	1.12	26.71	-13.40
7930	4.54	10.46	1.06	29.45	23.20	0.75	28.37	-4.55
8700	6.84	37.63	0.16	30.45	21.88	1.49	26.42	-33.47
10400	7.38	71.51	0.00	32.72	22.14	1.57	26.79	-72.78
11693	12.60	36.53	0.00	28.88	22.18	1.23	25.63	-132.52
11900	2.93	287.47	0.00	39.60	21.48	1.14	28.03	-35.20
13500	7.69	115.80	0.00	33.96	21.85	1.27	26.91	-59.67

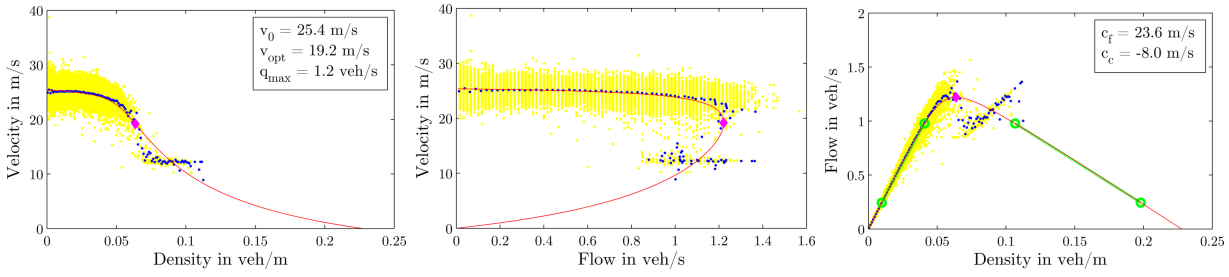
Table B.1: Numerical results of the performed calibration methodology based on the model after van Aerde and Rakha.



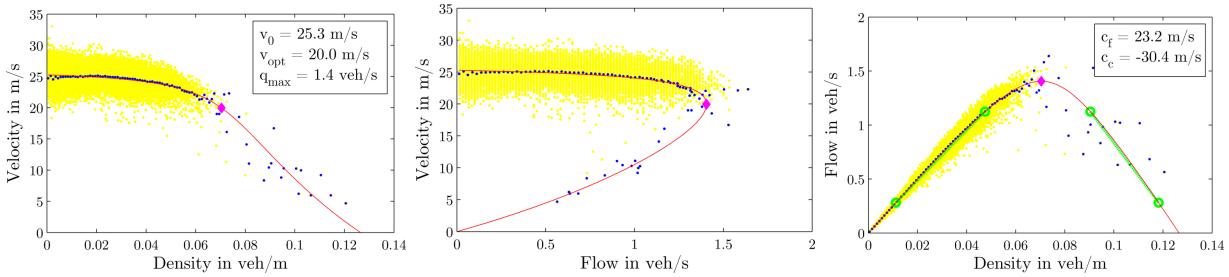
(a) 230



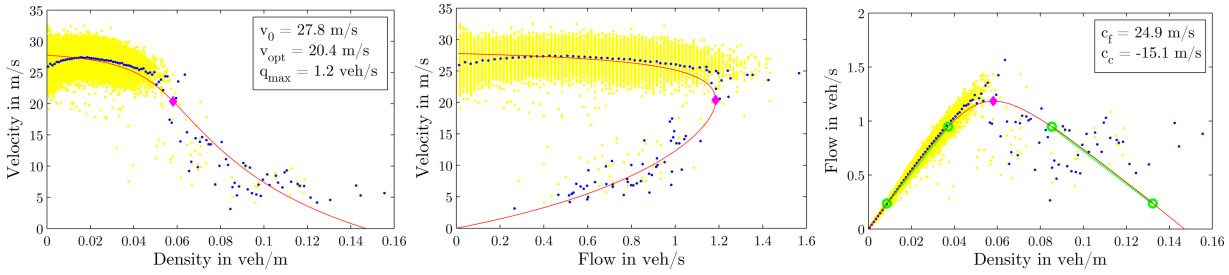
(b) 875



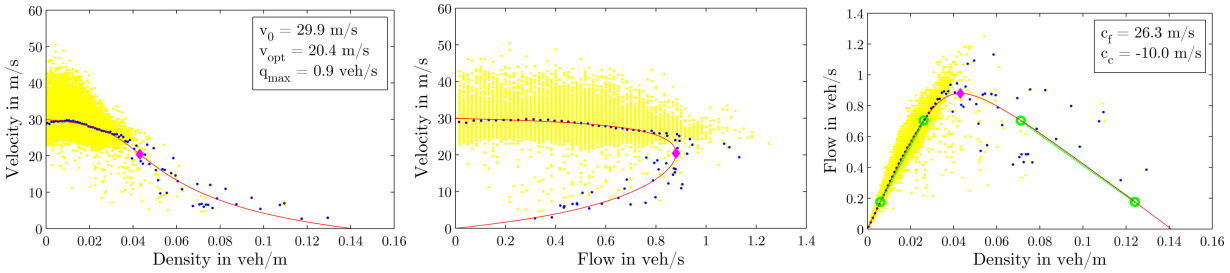
(c) 2065



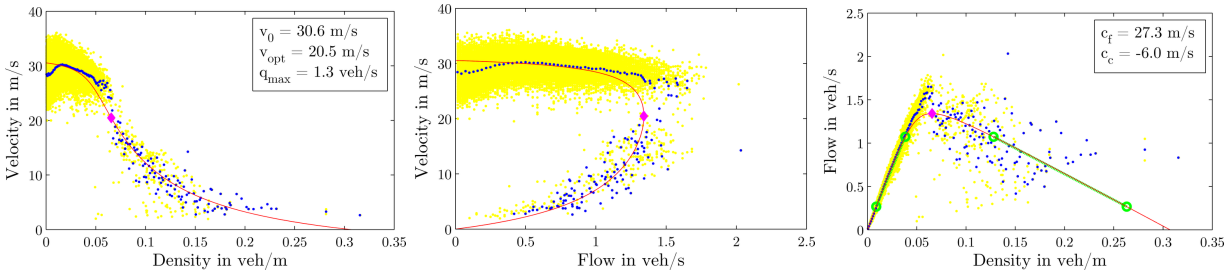
(d) 2975



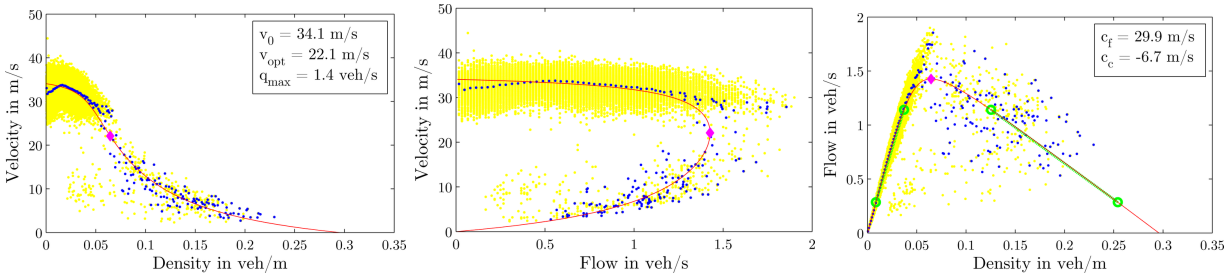
(e) 4296



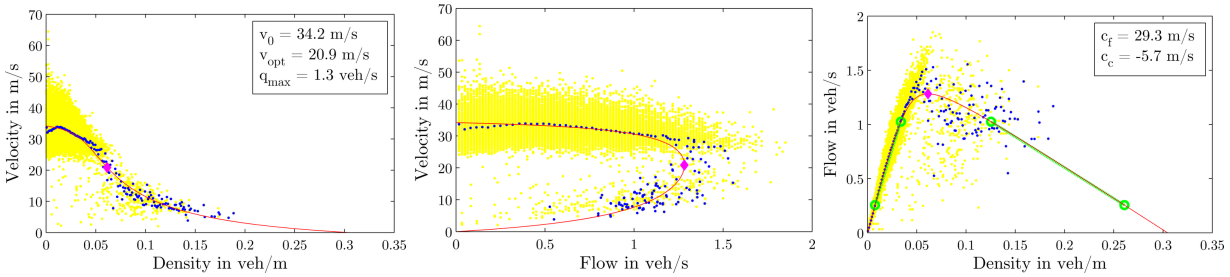
(f) 6465



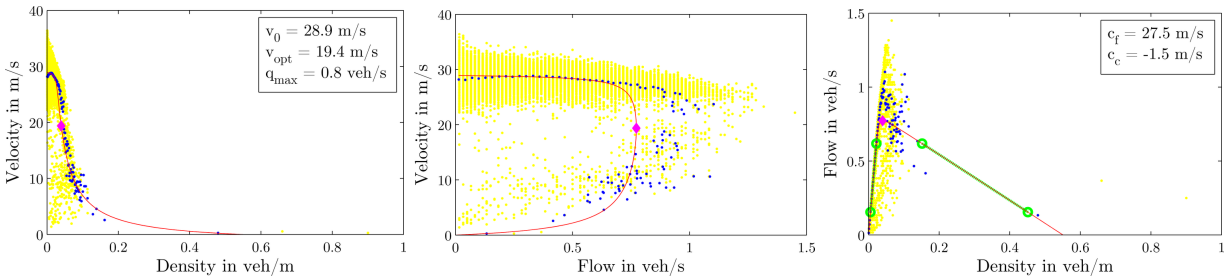
(g) 8530



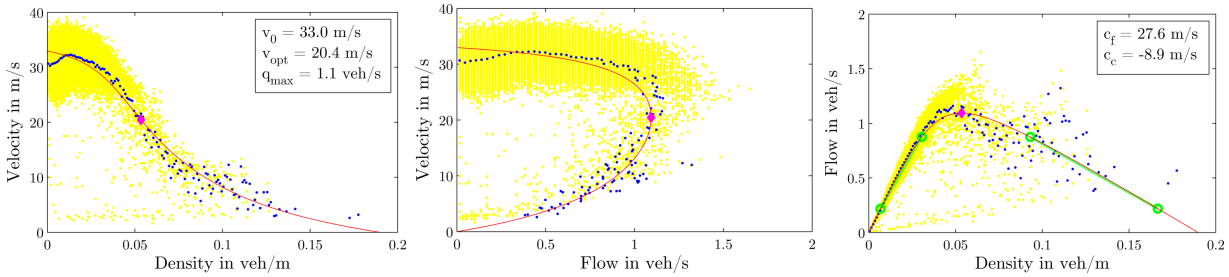
(h) 10000



(i) 11597

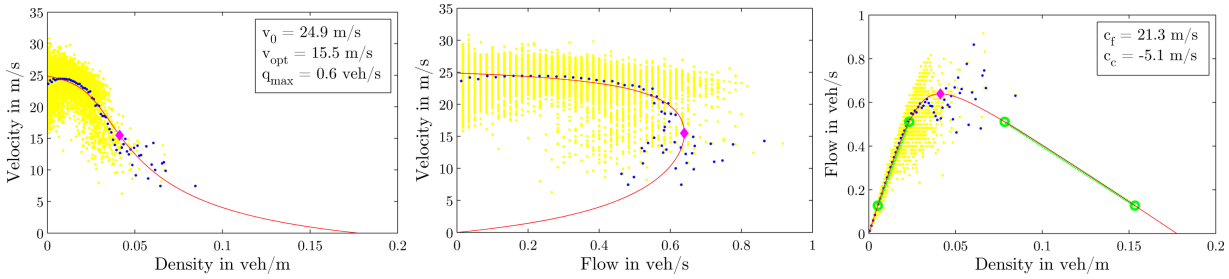


(j) 11693: incorrect fitting

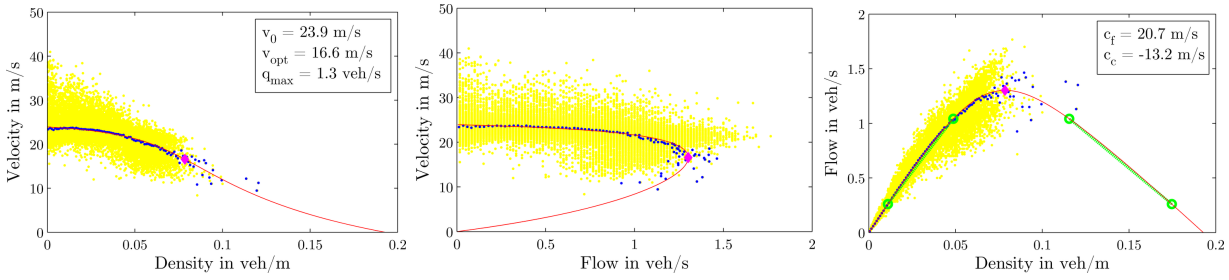


(k) 13790

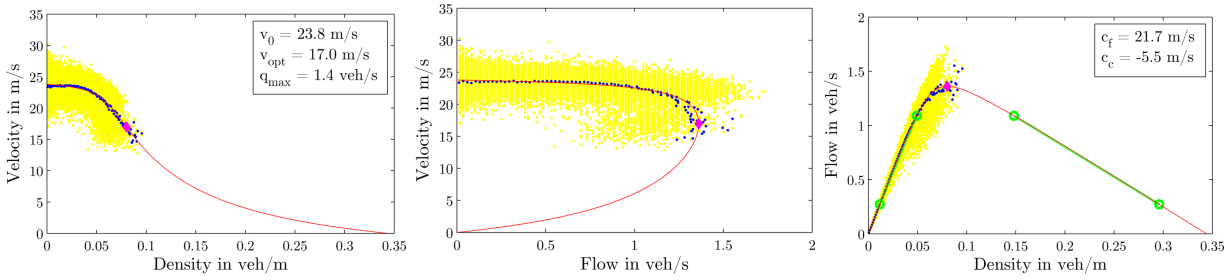
Figure B.3: Detailed overview of fitted models for every single stationary detector in driving direction east.



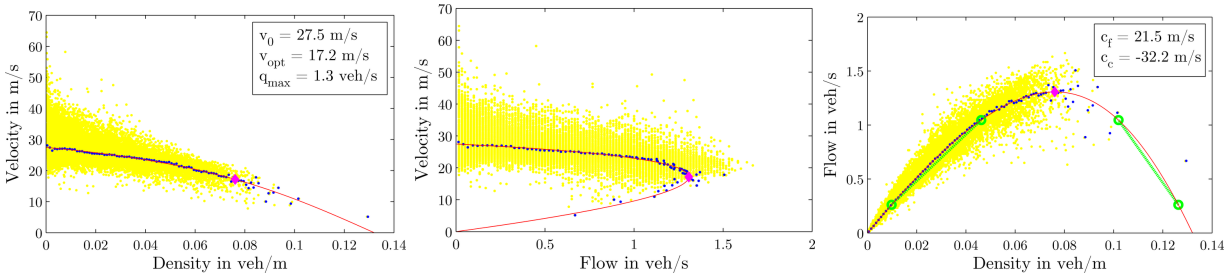
(a) 85



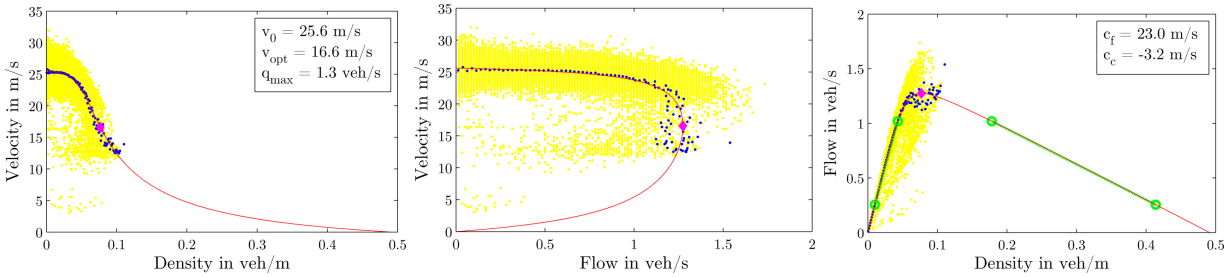
(b) 425



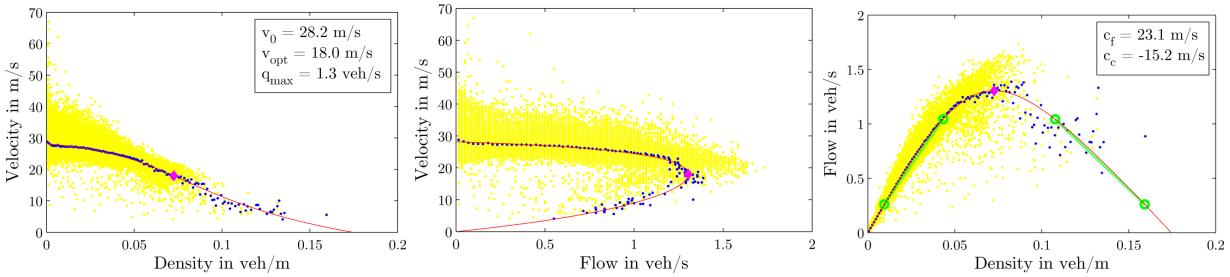
(c) 1000



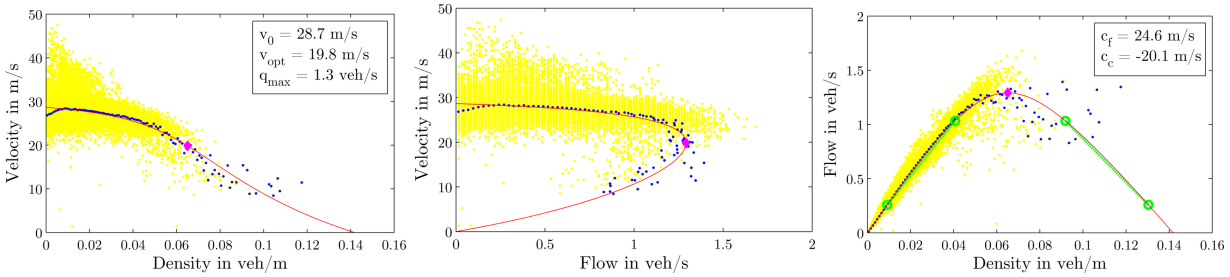
(d) 1455



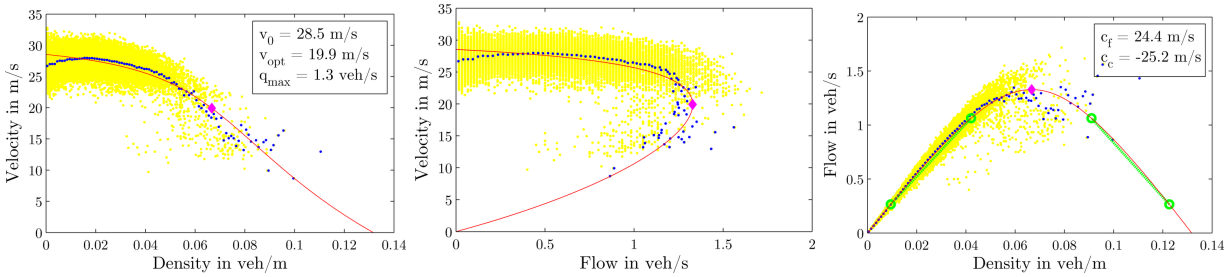
(e) 2205



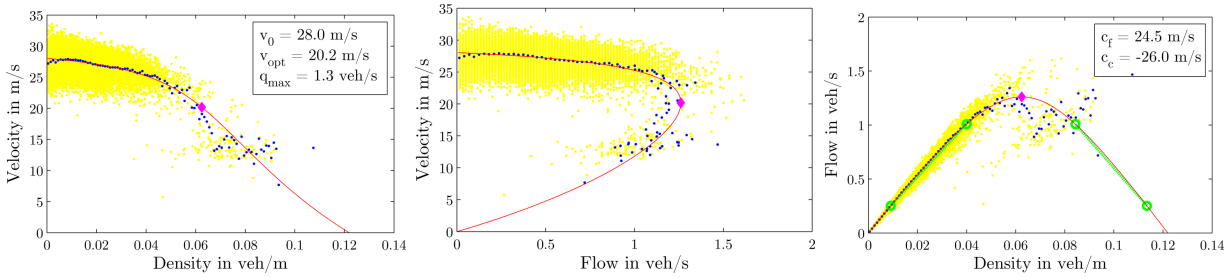
(f) 3200



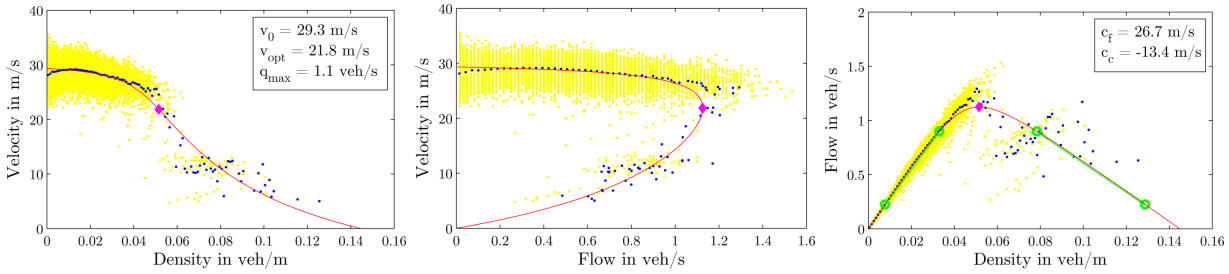
(g) 4800



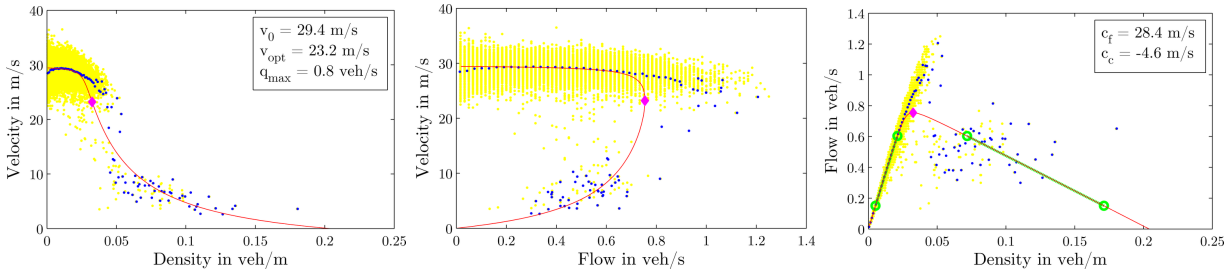
(h) 6120



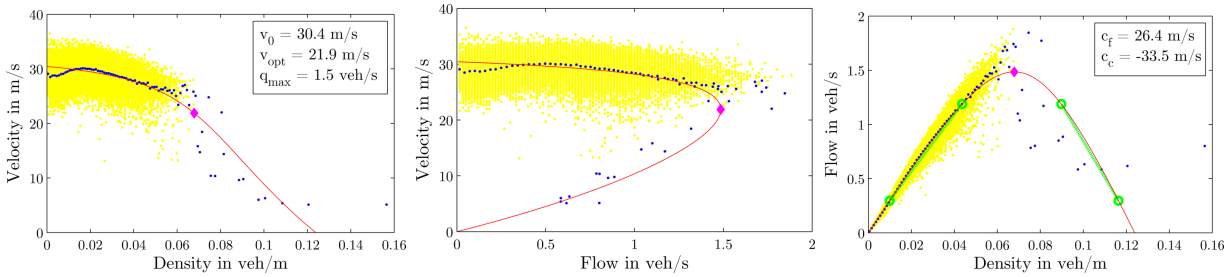
(i) 6900



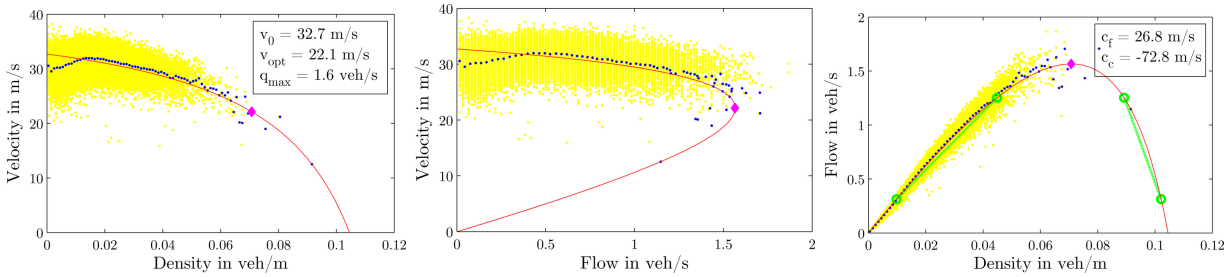
(j) 7490



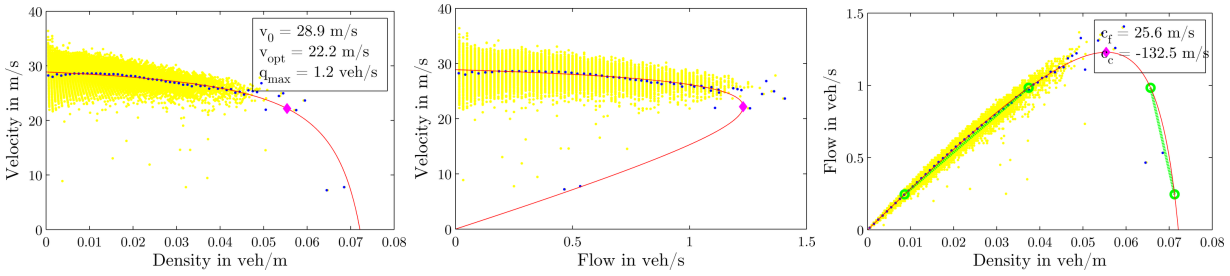
(k) 7930: strong capacity-drop



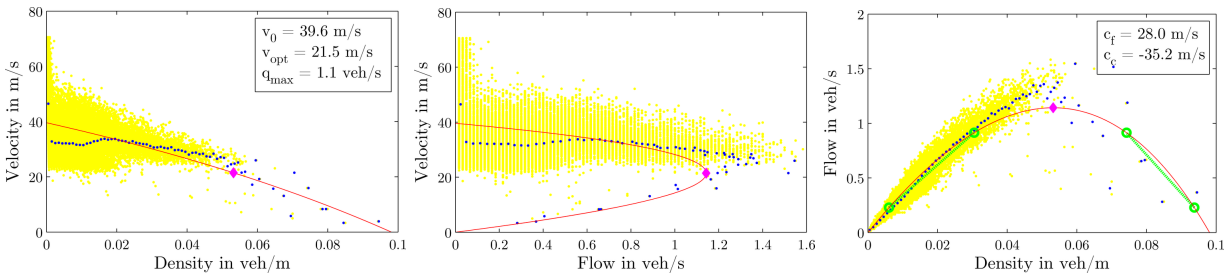
(l) 8700



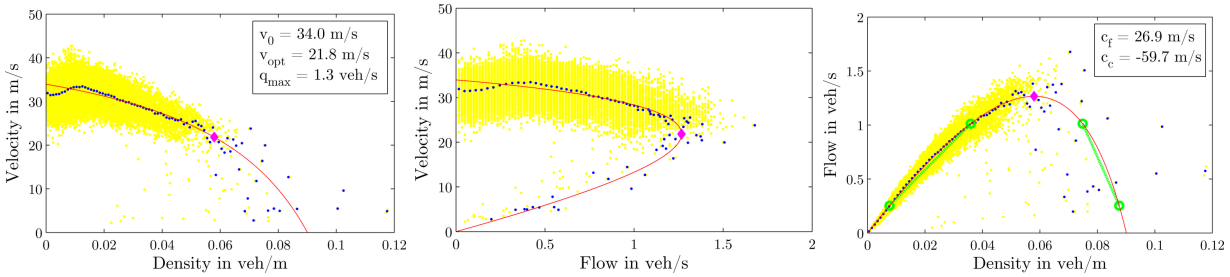
(m) 10400



(n) 11693



(o) 11900: incorrect fit of v_0



(p) 13500

Figure B.4: Detailed overview of fitted models for every single stationary detector in direction west.

C Global Performance Measures

Global performance measures evaluate the error of an *estimated* quantity $\hat{z}_i$ compared to a (estimated) *true* quantity z_i. We propose the following error measures widely used in traffic engineering [42–44, 54] for comparison of estimated spatio-temporal fields with ground-truth data:

Mean Square Error (MSE) and Root Mean Square Error (RMSE) Both errors are strongly related to the standard deviation and the variance of the error between $\hat{z}_i$ and z_i, respectively and can be written as

$$\mathrm{RMSE} = \sqrt{\mathrm{M}SE} \qquad \text{with} \qquad \mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{z}_i - z_i\right)^2. \tag{C.1}$$

The MSE is the second moment of the error, and thus involves both bias and variance of the error. Because of the square, the (R)MSE puts heavy weight on high errors, and is not robust against outliers. As the observed quantities and the RMSE share the same units, it is often interpreted as the "expected" error.

Mean Absolute Percentage (MAPE) The mean of the absolute percentage error (MAPE) can be written as

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{z}_i - z_i}{z_i}\right| \times 100\%. \tag{C.2}$$

Similar to the (R)MSE above, the MAPE indicates both, variance and bias between estimate and true value.

Due to the division by z_i, the MAPE is only defined for quantities $z_i \neq 0$. Further, for z_i close to zero the MAPE gets unreasonably high. This fact also applies for the MPE and the SPE below.

Mean Percentage Error (MPE) In contrast to the MAPE, the MPE is calculated without taking the absolute value with

$$\mathrm{MPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{z}_i - z_i}{z_i} \times 100\%, \tag{C.3}$$

where positive and negative errors compensate each other. Consequently, the MPE can be used as a measure for structural bias.

Standard Deviation of the Percentage Error (SPE) The SPE is the standard deviation of the MPE, and therefore a measure for the variability around the MPE.

$$\mathrm{SPE} = \sqrt{\frac{1}{N}\left(\frac{\hat{z}_i - z_i}{z_i} - \mathrm{MPE}\right)^2} \times 100\% \tag{C.4}$$

All relative error measures are normalised against the true z_i. Especially a small z_i (e.g. low velocities caused by a traffic jam) gives a higher weight on the error compared to higher velocities under free-flow conditions. Every particular error measure comes with drawbacks but is able to indicate special aspects of the performance. Thus, we recommend a combined use of the presented error measures.

D
Software Documentation

In this chapter information about the implemented MATLAB classes is provided. As mentioned before, those classes provide the core functionality of the implemented algorithms. Our implemented functions are well documented and provide basic help wich is accessible via the commands `help` or `doc`.

A graphical representation of the implemented classes in the *Unified Modelling Language (UML)* [116] is provided in Fig. D.1 to D.4.

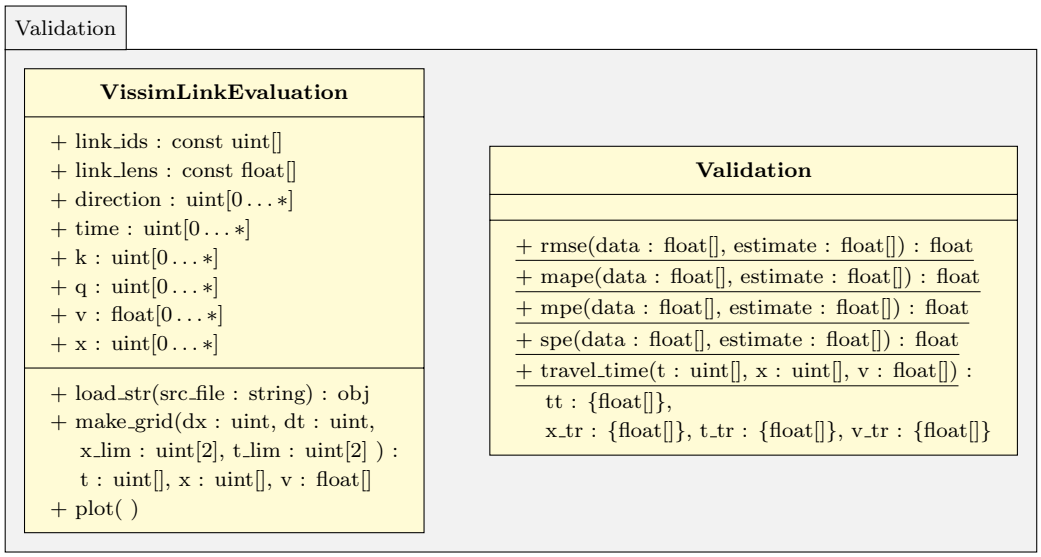


Figure D.1: UML class diagrams of implemented MATLAB classes required to perform evaluation and validation tasks.

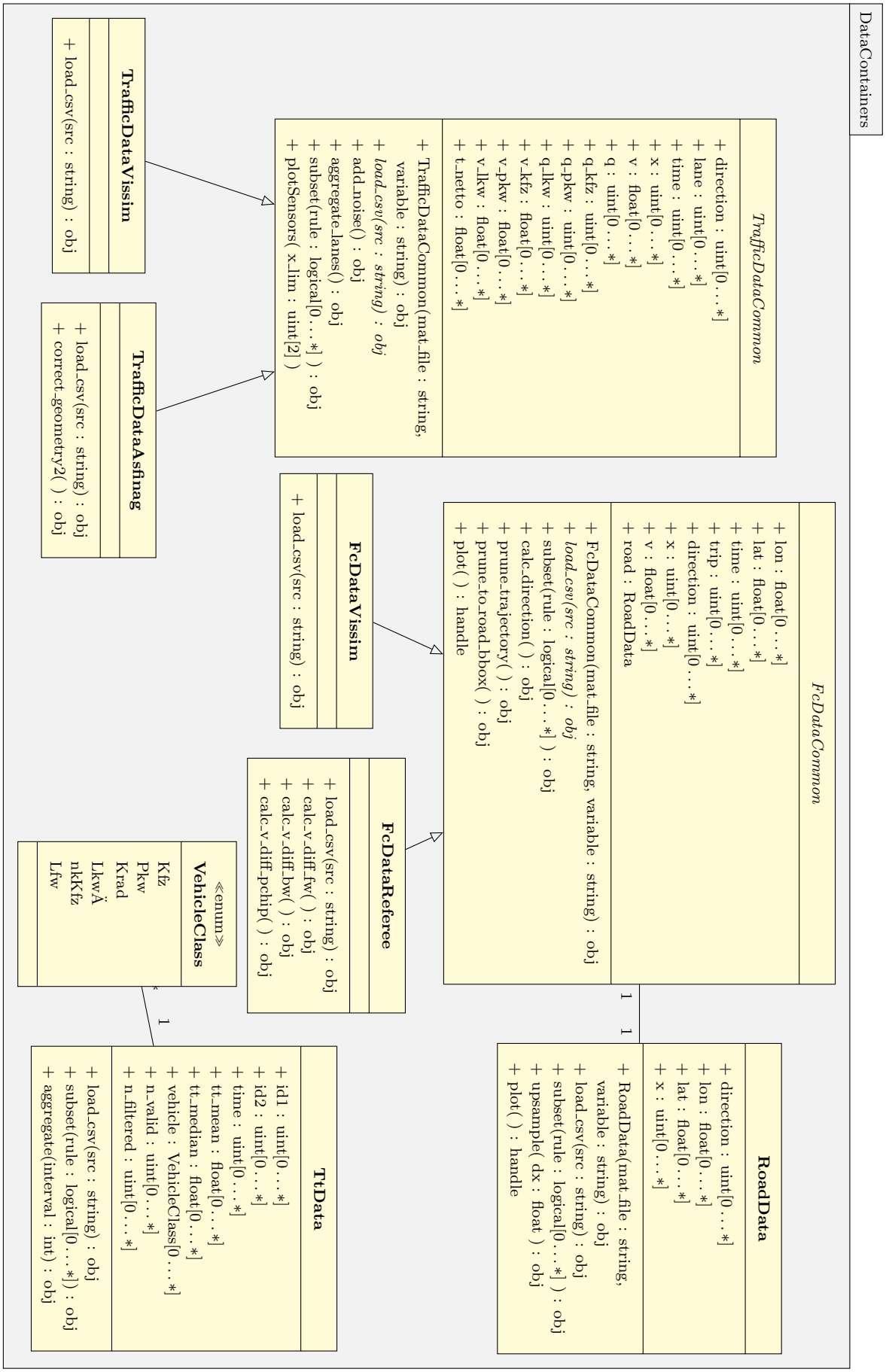


Figure D.2: UML class diagrams of implemented MATLAB classes for the purpose of storing an preprocessing data records.

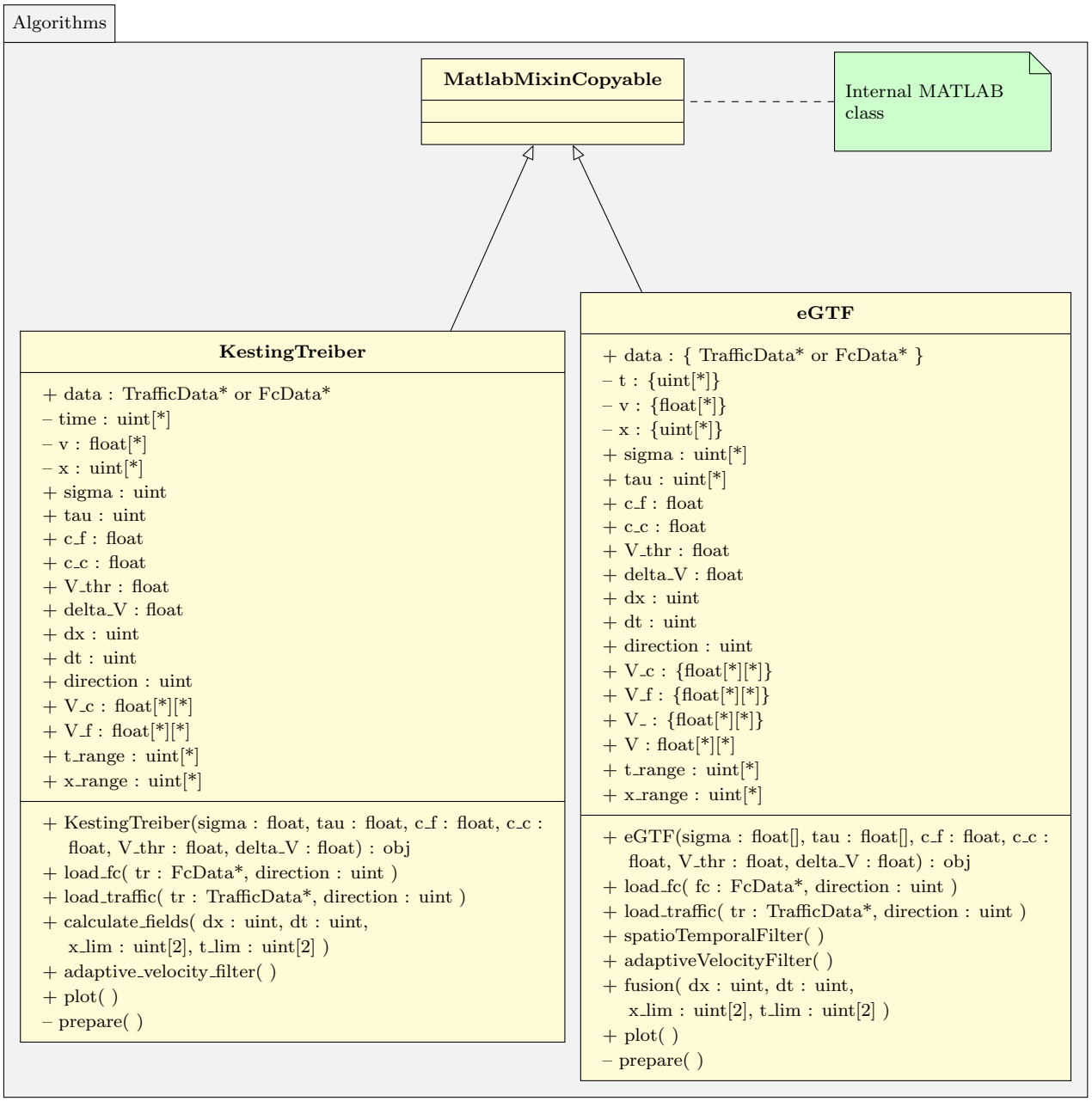


Figure D.3: UML class diagrams of implemented MATLAB classes for performing traffic state estimation and data fusion.

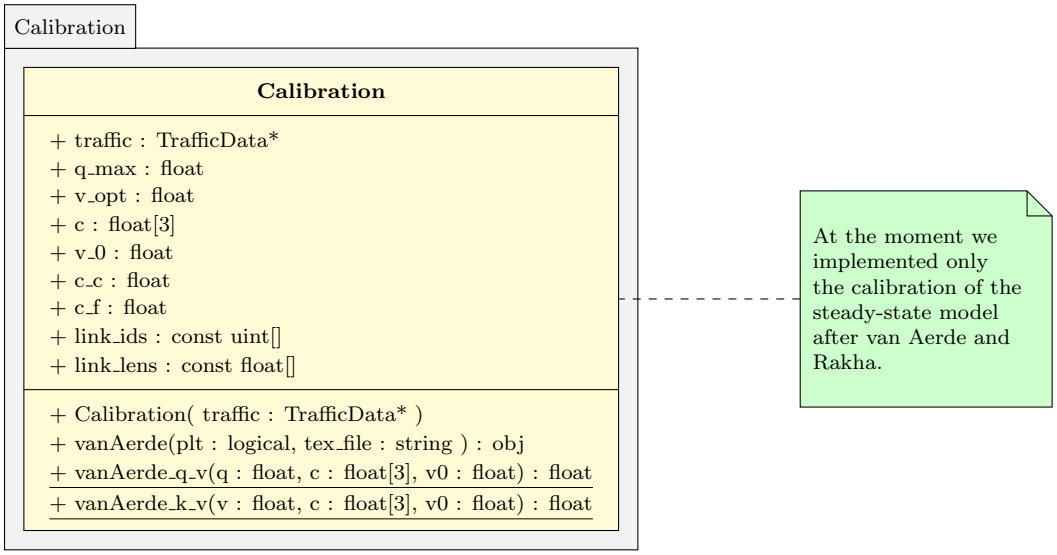


Figure D.4: UML class diagram of a implemented MATLAB class in order to perform a calibration of steady state traffic flow model on the basis of stationary detector data.