Technische Universität München
TUM School of Engineering and Design

Strategies for Detection of Congestion Patterns Using Multiple Sensor Technologies

Lisa Kessler

Vollständiger Abdruck der von der TUM School of Engineering and Design der Technischen Universität München zur Erlangung des akademischen Grades einer *Doktorin der Ingenieurwissenschaften (Dr.-Ing.)* genehmigten Dissertation

Vorsitz: Prof. Dr.-Ing. Rolf Moeckel

Prüfer der Dissertation:

1. Prof. Dr.-Ing. Klaus Bogenberger
2. Prof. Dr.-Ing. Axel Leonhardt
3. Prof. Dr. Constantinos Antoniou

Die Dissertation wurde am 08.06.2021 bei der Technischen Universität München eingereicht und durch die *TUM School of Engineering and Design* am 17.11.2021 angenommen.

Executive Summary

This dissertation introduces a novel approach to extending congestion definition and detection strategies in the context of traffic flow optimization on freeways. The aim is to represent congestion adequately and to provide a tool box to aid traffic planners and controllers in achieving optimized traffic flow. The thesis contributes data fusion approaches, congestion detection methodologies, and performance measures while comparing multiple traffic sensor technologies.

Traffic data are represented in a space-time speed field along a road stretch over a time period. In this speed distribution, congestion clusters are identified, where *congestion* is defined as occurring when the speed falls below a certain threshold. A *cluster* is defined as a spatially and temporally coherent congested area both in time and space that is isolated by free-flow conditions. Each cluster has a certain shape of convex hull and a certain area. A 4+1 congestion classification is developed, with four of these congestion types defined based on the mainline speed distribution: *Jam Wave*, *Stop and Go*, *Wide Jam*, and *Mega Jam*, ranging from a short speed breakdown to a widely congested area. Congestion type *Lane Jam* is identified based on the lane-by-lane speed distribution. It indicates congestion which does not affect all lanes equally.

Virtual vehicles that traverse the space-time domain are simulated, resulting in *virtual trajectories*. The speed profile of each trajectory is analyzed and categorized according to certain parameters, and the shape of this profile determines the congestion type of the trajectory. Several trajectories traversing a particular cluster with an equidistant offset determine an overall congestion type. The result is a set of congestion events in the space-time domain with each cluster having a shape, a size, and a congestion type.

All developed methodologies are examined on data collected using four different sensor technologies: local speed measurements (1) from induction loops and (2) from infrastructure radars, (3) travel time measurements from repeated Bluetooth device re-identification, and (4) floating-car trajectories from probe vehicles. All of them are gathered on a three-lane freeway in Germany with a stretch length of approximately 160 km.

To obtain a speed distribution from sparse space-time travel times such as those derived from Bluetooth device re-identification, a section-based low-resolution travel time smoothing method is developed. It aims at fusing noisy raw travel times into a consistent speed distribution. Moreover, a fusion approach of all four data sources is given. Taking the strengths of each sensor technology into account while overcoming their respective weaknesses, this multi-sensor data fusion yields an improved speed distribution over time and space with regard to a symmetric square inverse mean percentage error.

The four data sets are compared against three mainline congestion metrics. The first metric measures which sensor technology is able to detect emerging congestion earliest and how much delay is present in the others. The results show that data from floating cars show an

average delay of 2-3 min compared to induction loop data in areas where the sensor spacing is small. The second metric assesses the suitability of each detection technology for recognizing each congestion type. Among the findings is that *Stop and Go* traffic is best detected by floating car data, while *Jam Wave* events cannot be reliably recognized by Bluetooth sensors. In addition, data from Bluetooth detections overestimate the congestion type *Stop and Go* compared to induction loops. The third metric investigates locations and times of spatio-temporal congestion pattern hot spots for each of the data sources. The results include that congestion frequently emerges upstream of interchanges, specifically upstream of off-ramps.

The lane-based analyses assess the performance in identifying *Lane Jam* clusters. The methods consider spatio-temporal hot spots of said cluster occurrences and average travel times per lane. Moreover, the share of each lane, the first occurrence of congestion in space and time per lane, and the temporal and spatial offsets of all lanes are investigated using the introduced congestion clusters. The results demonstrate a high share of congestion on the right lane, restricted to particular localities, and indicate that congestion often starts on the right lane, spreading to the middle and left lanes with a delay of few minutes. The middle lane shows the longest average travel times over the entire duration of an incident.

The methodology of congestion type definition and congestion quantification has proven to effectively operate on multiple data sources and to provide quantitative, machine-interpretable insights and statistics on freeway traffic incidents.

Contents

Chapter 1

Motivation

Congestion both on freeways and in urban areas is rapidly increasing. The commercial traffic service providers INRIX and TomTom state in their yearly published reports (INRIX Global Scorecard [INRIX, 2021] and TomTom Traffic Index [TomTom, 2021]) the heavy growth of amount of congestion. Metropolitan areas are especially affected by congestion, which means increased travel times, higher greenhouse gas emissions, and more fuel costs, among other negative effects. In Germany, the costs related to congestion amounted to EUR 2.8 billion in 2019 [INRIX, 2021].

1.1 Research Context

Incidents and congestion are caused by various triggers. The main reasons are construction sites, accidents, and high density, which yield a bottleneck activation, negatively influencing traffic flow and speed. So far, both in literature and in practice, several approaches have been investigated to detect and to describe congestion events. Here, it is of great importance to apply congestion detection approaches which act automatically and computationally efficiently, as it is crucial to detect congestion as early and as precisely as possible. However, the sensor equipment along a road stretch also influences the detection quality.

Historically, traffic information was broadcast via radio and mainly consisted of the existence of an incident together with its approximate locality. Information messages such as Traffic Message Channel (TMC) feeds were likely to include freeways only. Some years ago, the road network covered by traffic information messages extended to highways but the transmitted information still focused on the existence, location, and sometimes an additional travel time of each congestion, instead of their exact properties.

Applying powerful detection strategies enables the determination of exact travel times by simulating vehicles moving through the considered space-time domain. Also, the opportunity to propose alternative routes helps in traffic control and prediction. Furthermore, special requirements such as the estimated energy consumption of Electric Vehicles (EVs) can be derived if all congestion properties are recognized reliably. An accurate traffic state reconstruction is used to calibrate traffic flow simulation models, to predict traffic states, and to improve the back-end of single probe vehicle fusion approaches. The identification of congestion hot spots supports the redesign and planning of roads.

To identify congested regimes, an exact definition of *congestion* is required. In literature, several approaches exist to separate congested from free-flow traffic, including studies with a fixed or variable velocity, or with a travel time threshold. Other possibilities include a certain

minimum spatial or temporal extent threshold for a congested region. This dissertation involves a 4+1 congestion pattern definition, ranging from a single jam wave to an extended broad congested area.

Every congestion event involves a varying number of vehicles and yields specific characteristics. Several patterns of congestion can be identified, each of which needs a different treatment to optimize the effects on individuals and the entire traffic system, for example minimizing the increase in travel time. To this end, it is necessary to know the particular type of congestion precisely and to recognize, describe, and predict occurring congestion as exactly as possible. This dissertation deals with various properties of congestion events on freeways and with strategies to detect congestion patterns.

One major research topic is the identification of congestion type hot spots. The concentrated occurrence of each congestion pattern is investigated spatio-temporally. For example, considering stop and go waves, the waves are separated by free-flow sections. In these, the speed is higher than in a widely congested area. When investigating congestion hot spots without a particular differentiation into patterns, the average speeds are overestimated and the average travel times underestimated. This provides motivation for the consideration of congestion *type* hot spots instead.

A precise and complete estimate of traffic states on major roads is essential for various applications in traffic engineering such as traffic control or road planning. Congestion, and traffic in general, is detected by sensor measurements. Data can be collected using different detection technologies [Bickel et al., 2007; Listl, 2003]. On the one hand, stationary sensors exist, such as induction loops that detect local speeds and flows. These conventional traffic data measurements from stationary devices have played a dominant role in traffic observations over the last decades. This data source has the advantage of registering every passing vehicle, however installation and maintenance costs are relatively high and require construction sites in the street which results in temporary road closures. Another infrastructure-based approach for traffic monitoring is radar sensors, often located outside of the road next to the shoulder lane. These are not as cost-intensive as loop detectors, however their data quality is outperformed by conventional sensors (see section 3.1).

On the other hand, there are section sensors. For instance, using Bluetooth (BT) technology, the unique hardware address of a device that passes a sensor located next to the road is registered. Once the device is re-identified by a second sensor located downstream, the travel time and therefore also the average speed can be derived. This is a quite effective approach to collecting data since most vehicles are equipped with or transport devices which are equipped with WIFI or BT antennas. Installation and maintenance costs are moderate, however the sensors capture only a subset of the vehicle fleet. Also, effects such as a non-representative vehicle sample need to be treated. Additionally, the sensor spacing influences the data quality, leading to a less exact average speed if the sensor distance is large.

Another travel time data collection method is speed measurements gathered from probe vehicles or smartphones – so-called *Floating Car Data (FCD)*. In this approach, probe vehicles determine their position via Global Navigation Satellite System (GNSS) and continuously report these positions to a central server. Again, this is a rather convenient method of data collection, but only a sub-sample of all vehicles is observed. The representativeness of FCD is influenced by the data provider. For example with a car manufacturer, trucks would not be

included, which would artificially raise the average speed in uncongested conditions.

In either data collection method, the result is a traffic state comprising speed, flow, and density. This reconstruction is the basis for traffic planning. In terms of traffic control, the online traffic state would be taken into account, basically derived from the same sensor technologies but using different interpolation and especially extrapolation approaches. This dissertation considers retrospective offline data, however in general, all algorithms proposed are also applicable to online data. This has however not yet been tested, and is left for future research as described in section 6.2.

Both individual data sets and a data fusion of all available data sources support an understanding of where and when congestion occurs frequently and which congestion type is most likely to be assigned to an identified congestion cluster. Thus, it provides traffic controllers with a valuable tool for traffic planning, e.g., for optimizing sensor placement of stationary detectors. Also, they might estimate and categorize the space-time extensions of occurring congestion more appropriately, resulting in reliable incident warnings, alternative route recommendations, and online traffic control. Furthermore, these results enable a fast and real-time capable prediction of emerging congestion based merely on a small online data set because the set of historical information is large and can be processed once with machine learning approaches.

Various types of road require different congestion definitions. In Germany – where a general speed limit does not hold on freeways – it might be different from situations where speeds between individual lanes do not differ much in general. In this dissertation, congestion is treated as speed below a fixed velocity threshold. Specifically, congestion is identified if and only if a vehicle's speed profile reaches a velocity below 40 km/h on a one-minute average. As a matter of course, these values can be parameterized, but have been identified to be realistic in practice.

1.2 Research Gap

In literature, comprehensive smoothing algorithms to interpolate collected data from Stationary Detector Data (SDD) and from FCD are described. The former can be interpolated using the Adaptive Smoothing Method (ASM) [Treiber and Helbing, 2002], the latter using the Phase-Based Smoothing Method (PSM) [Rempe, Franeck, et al., 2017] (section 2.1.3). To the best of the author's knowledge, an algorithm to fuse low-resolution travel time measurements such as those gathered by BT sensors into a consistent speed distribution has not yet been published.

Multi-sensor data fusion approaches have been widely developed. However, a computationally efficient algorithm regarding the fusion of 4+ vehicle detection methods along a road stretch of 150+ km validated on real traffic data has not been proposed so far.

Automatic Incident Detection (AID) methodologies mostly involve significant speed drops in data from stationary measurements or travel time increases in section-based measurements (section 2.2.4). The concept of simulating vehicles by virtual trajectories corresponding to probe vehicles is a rather novel approach, which has been tested so far mainly on FCD. Travel times are typically given as time losses or additional amounts of time needed to pass a congested region. However, a speed profile, or more precisely an acceleration profile, is often

omitted. Providing specific traffic information such as an acceleration profile might be helpful for non-conventional vehicles such as EVs [Kessler and Bogenberger, 2015; Kessler and Bogenberger, 2016; Kessler and Bogenberger, 2019]. Virtual vehicles provide such speed or acceleration profiles, however real-world data are too sparse to serve as input data for their simulation, therefore any model requires an extensive calibration. The derivation of congestion clusters as common unit of a congested space-time area by virtual trajectories has not been investigated.

Various classifications of *congestion* have already been described. Kerner et al. [Kerner and Rehborn, 1996; Kerner, 1999] proposed a three-phase traffic theory whereof two phases indicate congestion. Treiber and Helbing suggested six states of congestion [Treiber, Kesting, and Helbing, 2010]. In a project report [Transver GmbH, 2010], four types are introduced. Depending on causes, exact locations, and further impacts, different patterns of congestion emerge and appear on freeways (see section 2.2.1). Definitions are mainly described using the spatial or the temporal extensions (length or duration) but combinations of both towards a spatio-temporal extent introducing a size of a congested area have not been studied so far.

The common standard is to investigate mainline congestion. Lane-specific jam analyses are just described at particular locations or in special situations. The phenomenon of *Lane Jam* independently from off-ramp spillback is omitted in most works considering congestion patterns so far. *Mainline* in this work refers to the entirety of all lanes of the cross-section, independently from ramps, shoulders, and other auxiliary lanes, that allow for traveling on a roadway to the same target direction. In contrast to *mainline* congestion detection, *lane-by-lane* patterns are investigated.

1.3 Research Objective and Methodology

This dissertation aims at providing six novel methodologies: two fusion methods, two congestion detection contributions, and two comparison metrics.

The two fusion approaches include one for low-resolution section-based travel time measurements, and one as a multi-sensor data fusion approach to derive a consistent speed distribution out of several individual data sets.

Further, it investigates congestion detection and classification approaches. To this end, virtual trajectories are introduced that start at an arbitrary location and time. These aim at identifying and isolating congested areas with certain properties. These trajectories can be interpreted as virtual vehicles following the speed profiles recorded from real data sources. The approach is model-free, and hence, calibration is not necessary. A special focus is set to the shape of the resulting area. Grouping similar shapes leads to a cluster classification of 4+1 congestion types, consisting of four mainline and one lane-by-lane congestion pattern.

The metrics compare the performances of both congestion and congestion type detection across multiple sensor technologies. Here, loop speeds, radar speeds, BT travel times, and FCD are investigated. One metric measures which sensor technology is able to detect emerging congestion earliest and how much delay is present in the others. The other metric deals with congestion pattern detection and is divided into two research objectives: (i) the assessment of the suitability of each detection technology for recognizing each congestion type, and (ii) the

investigation of locations and times of spatio-temporal congestion pattern hot spots for each of the data sources.

The evaluation of all proposed methodologies is executed based on data from the German freeway A9 between Munich and Nuremberg. Figure 1.1 shows a map of Germany with the considered test stretch of 157 km in magenta.

Figure 1.1: Location of autobahn A9 stretch inside Germany

Briefly, this study separates single congestion events and quantifies them precisely by providing the following characteristics:

- Shape of a congested area as convex hull, including its start and end points in space and time
- Size of a congested area for quantitative comparisons
- One of 4+1 congestion types

Furthermore, common features of several congestion events are characterized, such as:

- Temporal and/or spatial hot spots
- Proneness for congestion at certain road stretches
- Suitability of congestion detection by various sensor technologies

Research Objectives of this Dissertation

- Development of an interpolation method for low-resolution travel time measurements yielding a speed distribution
- Concept and prototype approaches to fuse multiple traffic detection data sources
- Implementation of an automatic congestion detection methodology using virtual trajectories
- Comprehensive congestion definition of congestion types including lane-by-lane traffic analysis
- Analysis of the suitability of various detection technologies and combinations to recognizing congestion types
- Analysis of speed propagation on individual lanes and general recommendation of lane selection when approaching congestion

Figure 1.2 outlines the methodological approach of the thesis. Starting from raw data of one of four detection technologies, the speed or the travel time data are interpolated using an appropriate smoothing method. A multi-sensor data fusion approach creates a fifth speed distribution.

Virtual trajectories then pass through the speed contours, identifying coherent congestion clusters, each with a certain shape and size. Again, virtual trajectories are utilized to classify the cluster into one of the congestion types.

Based on these recognized congestion clusters, the data sets are compared. Spatial and temporal hot spots are investigated in a congestion proneness analysis. Lane-specific traffic analysis provides lane-by-lane speed information and answers the question of which lane is best for ensuring minimum time loss in congestion. The congestion type recognition method

computes which detection technologies are able to detect which congestion types, and vice versa. Congestion recognition gives a hint as to which data source is able to detect congestion earliest and computes an average delay of the other sources in various situations.

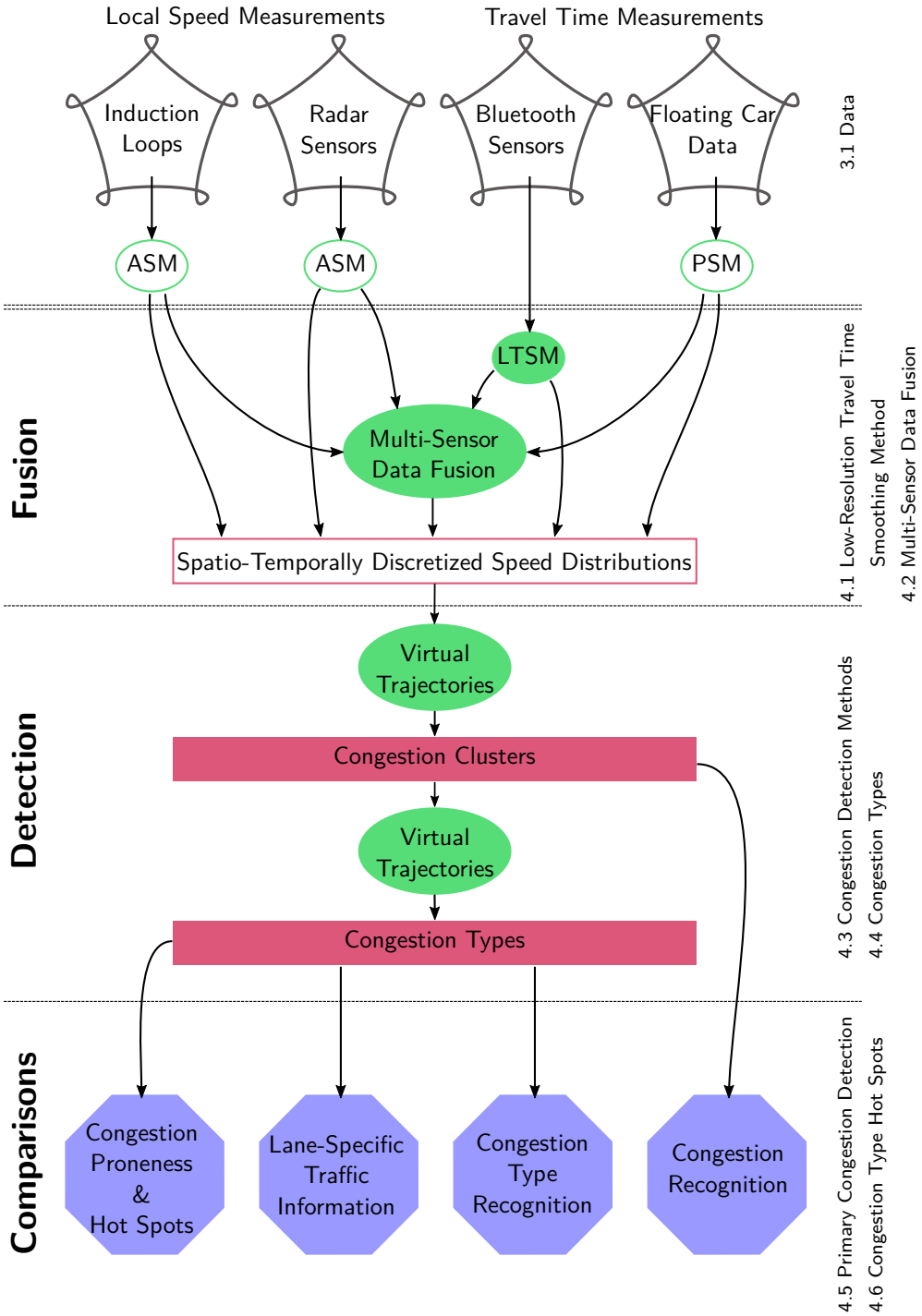


Figure 1.2: Structure of dissertation

1.4 Outline of the Dissertation

This dissertation is structured as follows. Chapter 2 discusses the state of the art of research on traffic state reconstruction based on several sensor technologies, and congestion detection strategies.

In chapter 3, all utilized data are presented. It starts with a description of the freeway stretch and the relevant vehicle detection sensors. It then compares the data quality resulting from each sensor technology. The last section gives insights into the comprehensive software implementation of all data processing.

Chapter 4 describes the theory of all proposed methodologies. It introduces the fusion method for low-resolution travel times and multi-sensor data fusion approaches. Then, methodologies are investigated to recognize congestion clusters and to identify congestion types. At the end, several data sets are compared with regard to (i) which data source is able to detect emerging congestion earliest and (ii) where congestion type hot spots can be found.

In chapter 5, the presented methodologies are applied to real-world data from four vehicle detection approaches. Many examples show the results of the fusion methods, the congestion detections, and the detection rate of congestion types per data set qualitatively. The comparison metrics are analyzed quantitatively and the results are discussed, such as the general suitability of the detection technologies to recognize different congestion types.

In the last chapter, chapter 6, the conducted research is summarized and a conclusion is drawn. Also, an outlook is given on further open research questions that were raised by this thesis. Furthermore, a vision of the future with a high penetration rate of automatically driven vehicles is discussed.

At the end of each chapter, the key findings and takeaways are indicated in a colored box.

Chapter 2

State of the Art

This chapter describes the state of the art of traffic state estimation and congestion detection. Section 2.1 reviews traffic state reconstruction approaches. It is divided into four parts. First, a theoretical description of traffic using the fundamental diagram and its representation as contour plots is explained. In section 2.1.2, in-use traffic detection systems such as local speed measurements and section-based travel time measurements are described. Section 2.1.3 presents current interpolation methods such as the Adaptive Smoothing Method (ASM) and Phase-Based Smoothing Method (PSM). In section 2.1.4, comparison and fusion approaches of multiple data sources are reviewed.

Section 2.2 investigates the state of the art of congestion detection. It is also divided into four parts. Section 2.2.1 summarizes several congestion definitions. In section 2.2.2, approaches to classification into different traffic states are treated, such as Kerner's Three-Phase Theory. Section 2.2.3 discusses lane-by-lane traffic and congestion detection. The chapter concludes with section 2.2.4, where congestion detection strategies are reviewed.

2.1 Traffic State Estimation and Reconstruction

The traffic state is reconstructed based on sparse observations. In a space-time domain, the three parameters flow, speed, and density are detected. To this end, traffic detection systems (section 2.1.2) measure traffic data, and smoothing methods (section 2.1.3) interpolate these observations to a consistent traffic state. A fusion of data sources may improve the space-time reconstruction (section 2.1.4).

2.1.1 Traffic Flow Theory

Flow, speed, and density distributions describe the macroscopic traffic state entirely. The traffic flow q measures the number of vehicles per time unit (in $[veh/h]$), speeds v are measured per vehicle (in $[km/h]$), and the traffic density k measures the number of vehicles per road section (in $[veh/km]$). These three parameters are connected via the fundamental equation $q = v \cdot k$, the basis of the fundamental diagram.

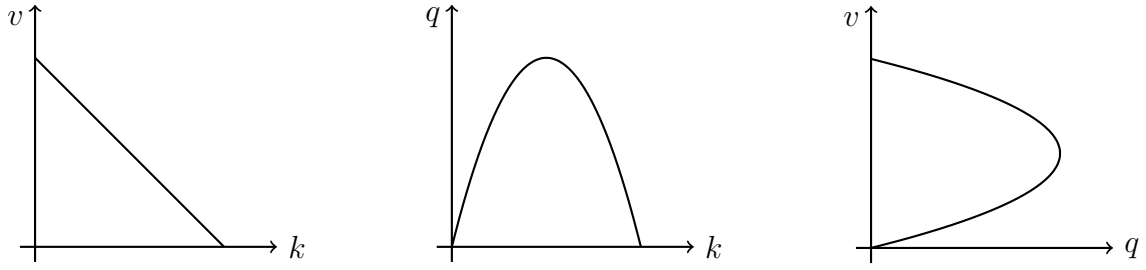
Fundamental Diagram

A macroscopic traffic model considering these three parameters is based on the fundamental diagram, originally introduced by [Greenshields et al., 1935]. It models reliable traffic states

and transitions, and distinguishes between two states: stable and unstable. The transition from stable to unstable marks a traffic breakdown which happens at a critical traffic density and its corresponding critical speed. The fundamental diagram (Figure 2.1) is used to determine the capacity of the road or to control inflow/outflow of sections along a road stretch.

The relationship between speed and density is linear with a negative slope. As the density increases, the speed of a freeway decreases. The relationship between q and k is parabolic, where small q values correspond to either small or large k values. Large q values correspond to medium k values. The flow-density diagram yields a time-space diagram from which travel times, delays, and queue lengths of a road section can be estimated.

The $q - v$ curve is not a function. This allows q to exist at two different speeds v_1 and v_2. This occurs when either v is higher and k is lower or when v is lower and k is higher. Both cases yield the same q. The speed-flow diagram consists of two branches. The free-flow branch corresponds to high v values while increasing q until the optimum flow is reached. After the optimum flow, the diagram switches to the congested branch. With decreasing q, v also decreases. The Level of Service (LOS) for a freeway is based on the speed–flow fundamental diagram [J. Li and H. M. Zhang, 2011].



(a) $k - v$ diagram (b) $k - q$ diagram (c) $q - v$ diagram

Figure 2.1: Fundamental diagrams

Antoniou et al. [2013] developed a dynamic data-driven framework which is able to estimate and to predict the traffic state. The framework clusters speed observations according to the properties of the fundamental diagram and applies a Markov process training. The authors validate their model using freeway data from California and Israel and show that their approach outperforms other existing models.

Representation Using Speed Contour Plots

A spatio-temporal representation of speeds or flows is the so-called *speed contour plot*. It depicts a space-time domain where the horizontal axis represents the time line (increasing time from left to right) and the vertical axis represents the space (direction of travel from bottom to top). Figure 2.2 shows an example.

Speeds are colored as indicated on the color bar on the right. Here, a cell-based speed distribution is plotted with segments of uniform lengths. Each cell has a spatial extent of 500 m and a temporal extent of 1 min. The speed in a cell is constant. The segments are clearly visible and appear as pixels.

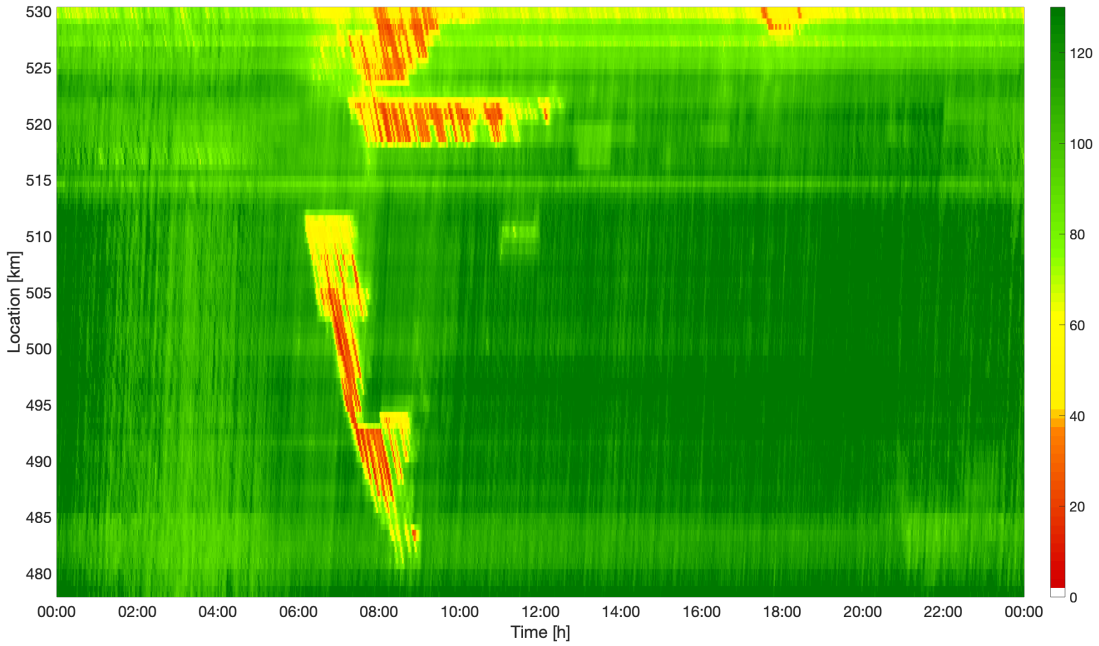


Figure 2.2: Example of a speed contour plot

This diagram indicates free-flow areas in green and yellow, and congestion in red. Stationary bottlenecks appear as red areas with a horizontal border at the top (downstream congestion front), whereas moving congestion appears as red stripe moving upstream in time (towards the lower right). A speed contour plot is a simple way of graphically visualizing the speed distribution, which is one of the main focuses of this thesis.

Observations measured on the road, specifically on a freeway, such as speed fill the space-time diagram sparsely. The methodology to derive a speed contour plot varies depending on the detection system. In the following, multiple traffic detection systems are introduced and the derivation of their respective space-time diagram is explained.

2.1.2 Traffic Detection Systems

Several sensor technologies are in place that gather data on how vehicles travel on freeways. Each comes with advantages and disadvantages. The following sections introduce relevant technologies utilized in this dissertation.

Local Measurements (Speeds)

A well-established option is to measure local velocities at a fixed point of the road. Induction loops that are buried in the road surface provide very exact and reliable speed and flow information. A recent extension is the combination of stationary ultrasonic, radar, and infrared camera detection. These combined detectors are usually installed on gantries and are therefore referred to as *gantry detection* in this dissertation. Like induction loops, they detect very exact and reliable speed and flow information. The main disadvantages are the high installation and

maintenance costs, which limits the use of this and other infrastructure-based measurement technologies to just a few corridors.

Another local speed detection mechanism is radar sensors mounted on poles beside the freeway facing a certain segment of the road. Like gantries and loops, they gather speed and flow information, but are less cost-intensive. Depending on their exact installation position, not all lanes can be detected equally and some vehicles are hidden when passing the sensor. An overview of lane-by-lane radar usage as a basis for a simulation model can be found in [Schöpplein, 2013].

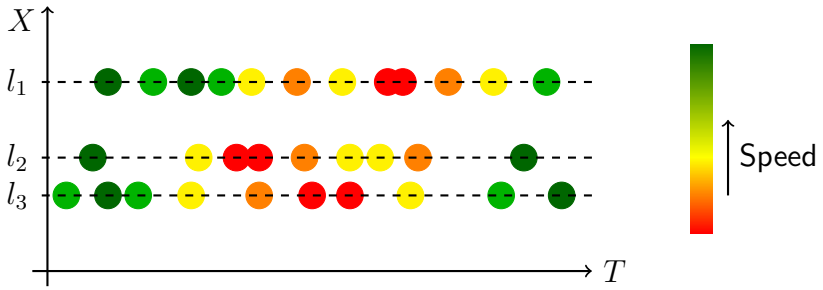


Figure 2.3: Local speed measurements

Figure 2.3 illustrates the method of data collection by local detectors. Let $X \times T$ be a space-time domain. Every detector (at locations l_1, l_2, l_3, respectively) observes passing vehicles continuously. Their speed values are indicated using the color bar on the right. Spatially between sensors and temporally between single measurements, no speed information is available.

Section Measurements (Travel Times)

Using WiFi or Bluetooth (BT) sensor technology, the hardware address, specifically a media access control (MAC) address, as a unique identifier of a device that passes the sensor is registered. If re-identified at a further sensor downstream, the matching of both hardware addresses allows the derivation of a travel time and therefore the average speed of the device [Haghani et al., 2010; Barcelo et al., 2010; Martchouk et al., 2011; Margreiter, Spangler, et al., 2015; Margreiter, 2016b; Lesani et al., 2016]. Figure 2.4 illustrates the space-time domain of resulting speeds. Each line corresponds to one vehicle or device re-identified at two or more locations l_1, l_2, l_3. The slope and the color of the line corresponds to the average segment speed of the vehicle.

BT installation is cost-efficient, as no exposed sensor position is needed and both driving directions can be covered by one sensor solely. Moreover, origin-destination relations can be derived while tracking a device over a long distance. However, the receivers collect information from a sub-sample of the vehicle collective only, and additionally, since they are conceivably placed several kilometers apart from each other, the average speed can be less granular. BT sensors detect in a certain range in their surrounding, which involves a fuzzy location of the detected device. Furthermore, no lane-by-lane information is available, and traffic flow values cannot be detected. Also, the method is dependent on the penetration rate and only measures a travel time at the end of a section. This results in delayed detection of increasing travel

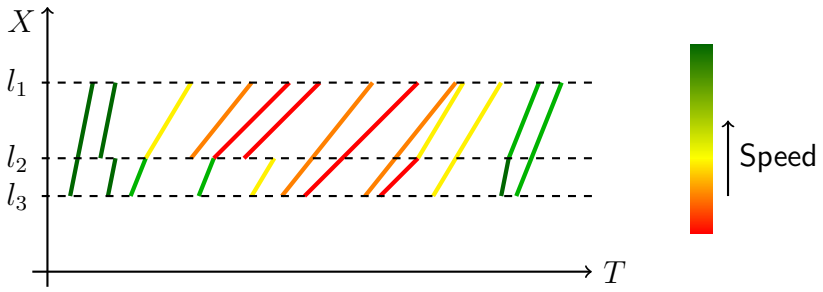


Figure 2.4: Travel time measurements

times during congestion, which is a major drawback for accurate real-time information [Jong, 2012].

Margreiter et al. [Margreiter, Spangler, et al., 2015; Margreiter, 2016b; Margreiter, Busch, et al., 2019] show that the penetration rate of BT devices on the autobahn considered in section 3.1 amounts to 20-40 % compared to the total traffic flow determined by induction loop detectors. Also, the authors investigate the share of heavy-duty vehicles which is approximately five times higher than that of cars. Furthermore, they identify a higher equipment rate for commuters, leading to both time-of-day- and day-of-week-dependent usage. Moreover, the authors introduce data filtering approaches. For example, one device can be detected multiple times at one detector due to the sender's range, or multiple devices on board of the same vehicle could be detected. Both cases involve data to be ignored, and hence, to be filtered. Additionally, stopping or detours of vehicles distort the data.

The work published by [Dion and Rakha, 2006] presents a low-pass adaptive filtering algorithm for predicting mean travel times using automatic vehicle identification data. The authors of [Martchouk et al., 2011] describe travel time variability derived from BT data. In [Barcelo et al., 2010], the reliability of travel time predictions and dynamic origin-destination estimations on freeways are analyzed based on BT traffic monitoring. Yildirimoglu [2019] use BT data to simultaneously estimate vehicle paths and travel times but deal in an urban context. Cotten et al. [2020] investigate novel BT devices and assess the data quality in comparison to video, radar, and Floating Car Data (FCD).

Floating-Car Based Speed Values

FCD, also called probe data, are gathered from vehicles or smartphones that determine their current position via the Global Navigation Satellite System (GNSS) and report it on a regular basis to a central server. Time and space differences allow for reconstructing the probe's speed profile on its route. Basically, FCD are also travel time measurements but due to the small segment spacing, average velocities of measured travel times can be assumed to be local speeds. Figure 2.5 illustrates the space-time domain resulting from FCD. Each trace corresponds to an FCD trajectory. Its color is smooth due to the frequent location updates every few seconds, depending on the fleet.

FCD are available wherever traffic is flowing, which allows for new applications in network-wide traffic monitoring and management. Several approaches focus on traffic state estimation and prediction with (mainly) FCD [Bekiaris-Liberis et al., 2016; Work et al., 2010;

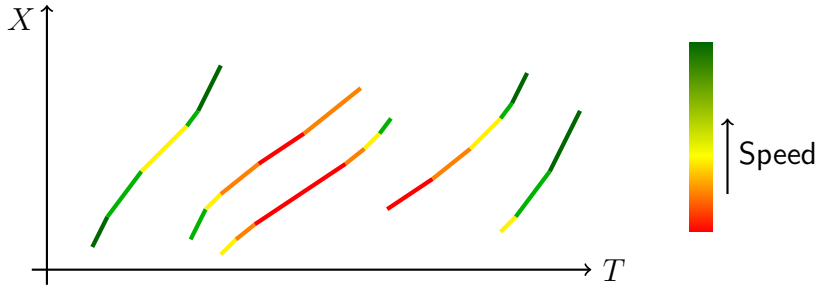


Figure 2.5: FCD measurements

Palmer et al., 2011]. Another advantage is that floating-car based technology does not require any measurement infrastructure except the telecommunication network.

However, the amount of FCD observations depends on the number of passing vehicles, since FCD represent only a sub-sample of the whole fleet. This limitation becomes less and less relevant with sample rates of some percent of the total traffic volume, in particular on highways during peak hours.

Another aspect is the time delay. FCD-based traffic state estimation is quite convoluted with many influences such as update frequencies from vehicles to the back-end server, the fleet size of floating cars, the current traffic flow, and the provider treatment. A drawback of FCD is that GNSS position and map matching errors cause a great impact [Y. Li and Q. Li, 2010]. FCD are not gathered lane-by-lane.

In [Day and Niloy, 2021], the reliability of FCD-based speeds on highways as a ground truth is assessed. The authors conclude that estimated distributions have substantial differences compared to the ground truth distribution, with congested conditions exhibiting less deviations than in uncongested traffic. The authors of [Sunderrajan et al., 2016] estimate the traffic state – including density and flow – based on FCD.

Further Traffic State Detection Technologies

Besides the already discussed detection technologies, on which the evaluations in this thesis are based, there are a few more worth to mention. An overview of in-use sensors can be found in [Listl et al., 2019]. Specifically, cameras such as the Automatic Number Plate Recognition (ANPR) and Light detection and ranging (LiDAR) should be mentioned. Both measure local speeds, and ANPR cameras additionally measure travel times. LiDAR sensors generally produce high quality data. The reconstruction of the traffic state works reliably if well calibrated [Nedoma et al., 2018; Tsogas et al., 2011].

Most public data sets found in the literature discuss either fewer data sources or shorter road stretches. This thesis chose to investigate the autobahn A9 test field with measurement equipment as previously mentioned. Considering a short freeway stretch of several kilometers does not allow the identification of congestion patterns, hot spots, and propagation reliably. Comparing fewer than three data sources does not make sense with regard to both assessments of congestion detection and fusing strategies. Therefore, the main focus of this work is set to the previously described technologies along the mentioned experimental site.

2.1.3 Methods of Reconstruction

Traffic data derived from one of the described detection technologies have to be interpolated to derive an entire space-time speed reconstruction. Smoothing methods interpolate between known mesh points of the speed distribution. This dissertation mainly utilizes two interpolation methods, one smooths speed and flow measurements from local sensors, the other one speeds from probe vehicles. Hence, a short synopsis of both follows.

ASM

The ASM approach is an anisotropic interpolation method introduced by Treiber and Helbing in 2002 [Treiber and Helbing, 2002; Treiber and Helbing, 2003; Treiber, Kesting, and Wilson, 2011]. The basic idea is that raw data of a sparse input source are smoothed in two traffic-characteristic directions: v_{cong} denominating the wave speed in congested traffic conditions, and v_{free} denominating the wave speed in free-flow conditions. In a discrete time-space domain with t and x representing time and space, respectively, the resulting complete speed matrices $V_{cong}(t,x)$ and $V_{free}(t,x)$ are combined cell-wise:

$$V_{ASM}(t,x) = w(t,x)V_{cong}(t,x) + (1 - w(t,x))V_{free}(t,x) \tag{2.1}$$

The weight $w(t,x)$ is adaptive and favors low speeds:

$$w(t,x) = \frac{1}{2}\Big(1 + \tanh\big(\frac{V_{thr} - \min(V_{cong}(t,x), V_{free}(t,x))}{\Delta V}\big)\Big) \tag{2.2}$$

with V_{thr} a threshold where weight $w(t,x)$ equals to 0.5 and ΔV is a parameter to control the steepness of the weight function.

The spatio-temporal smoothing approach interpolates the time-series data spatially between detector positions but also in time at a fixed sensor location. Additionally, it allows for considering speeds and flows at *virtual* detector positions. Several studies assess the accuracy of the method with positive results, e.g. [Lint, 2010; Huber and Bogenberger, 2013; Rempe, Franeck, et al., 2016]. The approach was proven to be accurate, efficient, robust, and flexible with respect to the input data, e.g. [Treiber and Helbing, 2003; Lint and Hoogendoorn, 2010; Schreiter et al., 2010; Treiber, Kesting, and Wilson, 2011; Kessler, Huber, et al., 2018a].

PSM

Speed data from probe vehicle trajectories, so-called FCD, can be interpolated using the PSM introduced by [Rempe, Franeck, et al., 2017]. The PSM is an approach that is based on concepts of the ASM. It was developed to reconstruct space-time traffic speeds with higher accuracy given only FCD. It utilizes findings summarized by the *Three-Phase* traffic theory [Kerner, 1999; Kerner, 2008] in order to distinguish between localized and moving congestion. The method outperformed the ASM in recent studies [Rempe, Franeck, et al., 2017; Kessler, Rempe, et al., 2021].

The basic idea is that in the first step of the PSM, raw data are smoothed in the direction of typical speed propagation of each traffic phase. v_{cong} is assumed to be the propagation speed

of moving congestion with low vehicle speeds (also called *Wide Moving Jam* in the *Three-Phase* traffic theory). Congestion that is caused by a bottleneck, e.g., a construction site or an on-ramp, is often localized and its downstream front is attached to the bottleneck location. In order to account for the locality, data are smoothed in only the temporal direction for the so-called *synchronized* traffic flow phase. Based on the speeds and the amount of available data, each cell (t, x) is classified into one of the three phases: Free flow, synchronized flow, or wide moving jam using methods from probability theory.

In the second step, phase-specific speed estimates are computed. Raw speed data that are assigned to a specific phase are smoothed using either a free-flow kernel parameterized with v_{free} or a congested kernel parameterized with v_{cong}. The phase-specific speed estimates are aggregated into a final speed estimate using a weighted average. This method was shown to be efficient and accurate [Rempe, Kessler, et al., 2017; Kessler, Rempe, et al., 2021].

Further Traffic State Reconstruction Methods

Besides the mentioned interpolation methods, several further approaches are known in literature. The so-called *ASDA/FOTO* approach identifies three traffic phases in space and time and subsequently forecasts congestion fronts using the shock-wave equation of traffic [Kerner, Rehborn, et al., 2004]. The traffic speed reconstruction method *TravRes* uses GNSS-based travel times and was introduced by Ou et al. [2010b]. A recent traffic state reconstruction based on moving probe vehicles is presented in [Barreau et al., 2020]. Z. He et al. [2017] propose a space-time grid of speed distributions composed not of rectangular cells, but rather parallelograms inclined in the direction of a backward-moving wave, which yields promising results, especially in congestion.

In summary, Table 2.1 gives an overview of recent interpolation methods. As can be seen, a gap exists for low-resolution travel time measurements. An appropriate interpolation method deriving a consistent speed distribution out of section-based travel times is developed in section 4.1.

Detection Strategy	Suitable Reconstruction Methods
Local measurements	ASM, ASDA/FOTO
Section measurements	—
FCD	PSM, TravRes

Table 2.1: Overview of interpolation methods

2.1.4 Comparisons and Fusion Approaches of Multiple Data Sources

This section presents an overview of research conducted if two or more data sets are available. It first describes comparison approaches where different data sources are compared with regard to varying aspects, such as the quality of traffic information. Later on, current fusion approaches are presented where data sets from multiple sources are superimposed.

Comparison of Data Sources and Quality of Traffic Information

Comparisons of different traffic detection technologies have been widely studied. The authors of [Bachmann, Roorda, et al., 2013] compared BT measurements and loop detector data in the Greater Toronto Area on a stretch of several kilometers. The authors of [Yuan et al., 2014] describe a network-wide traffic state estimation while they combine loop detector and FCD. The authors of [Leonhardt and Steiner, 2012] describe an instance-based learning method for the estimation and prediction of traffic state variables using spatio-temporal traffic patterns while they conjointly use FCD and local sensor data. In [Cohen and Christoforou, 2015], the authors statistically analyze the differences between loop detectors and FCD in the area of Lille, France. In [Wang et al., 2011], a comparison of BT with vehicle license plate recognition is conducted and errors are estimated statistically.

In literature, data sources are mainly compared to assess the quality of traffic information. Regarding the time delay of speed information, Rainer and Müllner [2011] state that for traffic messages broadcast in Austria in 2011 via the Traffic Message Channel (TMC), the average time of delay was about 10 min. In [Kim and Coifman, 2014], traffic information provided by INRIX is assessed. The results indicate that the reported INRIX speeds tended to lag stationary detector data by almost 6 min on average. The authors of [Rakha et al., 2013] also assess INRIX data and compare them to stationary sensor data while focusing on travel time and travel time reliability.

Other studies such as [Bogenberger, 2003; Bogenberger and Hauschild, 2009; Lux, 2011; Rehborn et al., 2011] discuss quality of traffic information services via indices (see [Bogenberger and Weikl, 2012; Huber, Bogenberger, and Bertini, 2014] for an overview of these studies). The fundamental idea of all these approaches is to interpret *quality* as the level of similarity between the broadcast traffic information and a ground truth representing the real traffic situation adequately. This ground truth data are either generated by tracking test vehicles ([Bogenberger and Hauschild, 2009; Lux, 2011]) or by applying traffic state reconstruction methods based on stationary detector data ([Bogenberger, 2003; Rehborn et al., 2011; Huber, Bogenberger, and Bertini, 2014]).

Fusion of Data Sources

Differences and characteristics of each technology challenges the fusion of the sources. The aim of fusion approaches is to develop methods that make use of all information hidden in the data and compute a combined result that outperforms estimates based on a single source.

The authors of [Faouzi and Klein, 2016] give a detailed overview of the state-of-the-art in fusion of data in Intelligent Transportation Systems (ITS) and its advantages. A review of multi-sensor data fusion for traffic is described in [Zhao and D. Zhang, 2018]. In [Klein, 2020], a comprehensive summary of available sensors and fusion techniques is given. Bachmann [2011] and [Bachmann, Abdulhai, et al., 2013] compared seven fusion methods for traffic speeds and travel time estimations. One key finding is that a simple convex combination of loop detectors and BT measurements is one of the best fusion strategies.

Data are usually noisy, sensors are sparse in space and time, and prone to outages. All of this limits the applicability of many published approaches in practice. The article by Klein [2019] presents three widely applied data fusion techniques and describes their relevance to ITS:

Bayesian inference, Dempster–Shafer evidential reasoning, and Kalman filtering. In [Zeng et al., 2008], an evidence-theory-based data fusion approach for traffic incident detection is described. Data from loop detectors, camera observation and FCD are fused on a rather short stretch of few hundred meters on an urban highway.

Corsi and Capitanelli [2011] applied data fusion techniques for traffic planning and control in a setting with satellite images, acoustic, and GNSS data. In [Ou et al., 2010a], a method fusing loop data consisting of both speed and flow, and in-car travel times is presented. In [Zhou and Mirchandani, 2015], a framework for the fusion of loop detector and GNSS data is described. This framework is able to distinguish between lane-based traffic states. The authors of [Faouzi, Klein, and Mouzon, 2009] study the fusion of loop data and toll collection data using a Dempster-Shafer approach to get an improved travel time estimate.

Leonhardt [2008b] describes a method to estimate and predict travel times based on probe vehicles and local detector data using regression techniques. Yuan et al. [2014] present an approach to network-wide traffic state estimation combining loop detector and FCD. The article [T.-H. Chang et al., 2016] describes a data fusion from travel time measurements from toll collection stations and stationary vehicle detectors in Taiwan. S. He et al. [2016] discuss a data fusion approach for mobile phone probes and microwave sensors, and give a sensitivity analysis on impact factors. The authors of [Rempe, Kessler, et al., 2017] developed a model to fuse FCD and loop detector data to forecast upstream congestion fronts on a freeway. A comparison of two model-based approaches on filtering methods is conducted in [Trinh et al., 2019]. The results are approved using synthetic data from a simulation. Liu et al. [2018] describe a Kalman filter method for freeway traffic state estimation fusing two data sources: wireless communication records combined with microwave sensor detection. Another Kalman filter based approach is given in [Fulari et al., 2015].

Rostami-Shahrbabaki et al. [2018] propose a fusion of loop data and FCD at urban intersections to estimate queue lengths and outflows. In [Ambühl and Menendez, 2016] and [Dakic and Menendez, 2018], a fusion of loop data and FCD is described with the goal to approximate the macroscopic fundamental diagram of urban networks. However, data stem from micro-simulations which tend to idealize real data. The authors of [Hegyi et al., 2013] fuse loops, FCD, and camera data using the ASM to improve the speed of jam detection and respective control measures. An evaluation is performed again using simulated data.

A recent study by [Kessler, Rempe, et al., 2021] proposes a conjoint optimization of accurate traffic speed and travel time reconstruction using multi-sensor data fusion. The authors compare different fusion approaches such as a section-average, ASM, PSM, and a novel approach, the so-called *PSM-W*, a weighted integration of BT measurements into a harmonic mean fusion of loop detectors and FCD. The weighting factor corresponds to the travel time of a BT device. If it is short, the speed is likely to be high. Analogously, if the travel time is long, the speed is supposed to be low. In both cases, the speed value is reliable and the weighting factor in the data fusion is high. In case that the travel time is something in between, several realizations of the segment speed profile are possible. This leads to an unreliable average speed estimation and therefore, the weighting factor in the fusion is low. The authors conclude that a data fusion with all sources will not necessarily improve the overall traffic reconstruction. They recommend a speed retrieving based on FCD only or on FCD plus local speed measurements which yields best results.

2.2 Congestion Definition and Detection

This section describes existing approaches of congestion definitions and classification possibilities. Both are used as a basis to categorize several congestion patterns. Further, the section discusses strategies for the detection of congestion including both mainline and lane-by-lane traffic information.

2.2.1 Congestion Definition

In literature, several approaches exist to define emerging congestion patterns and congested areas. Most studies utilize either a fixed or a variable velocity threshold or a certain amount of travel time increase.

Traffic jams in Switzerland are defined below a threshold of 10 km/h [ASTRA, 2021]. INRIX [2021] define congestion if the speed in one of their segment-based intervals falls below a certain percentage of the average free-flow velocity at normal traffic. The online information service of Bavaria named BayernInfo [2021] proposes traffic jams below 32 % of the free-flow velocity which is at most 130 km/h. TomTom [2021] define congestion as certain amount of additional travel time during peak hours compared to one hour of driving in free-flow conditions.

The German Federal Highway Research Institute BASt recommends to treat congestion if the traffic state fulfills one of the following three conditions: (1) measured speeds are below 30 km/h and measured density is above 70 veh/km, (2) lane occupation is larger than 50 % and measured speeds are below 45 km/h, (3) the moving average of measured speeds reaches a threshold of 35 km/h at most, measured flow is at least 1800 veh/h, both car and truck lanes have a flow unequal to zero, and the speed difference of both lane groups amounts to 25 km/h maximum [BASt, 2018].

The online control algorithm INCA was developed to control variable speed limits and message signs. It identifies the time when a congestion warning should be shown on gantries. It uses three parameters to define congestion: A threshold of the traffic flow q depending on the number of lanes, the speed differences between the speed shown as a warning and the driven speed at the gantry, and a threshold of the standard deviation of speeds [Vukanovic, 2006; Denaes et al., 2009a; Denaes et al., 2009b].

In this dissertation, congestion prevails if the speed of a cell in a discretized space-time domain is set to a value less than a critical velocity $v_{\text{crit}} = 40$ km/h, combined with some additional requirements to the neighbored cells. The threshold of 40 km/h is chosen to cover both halting/stopped vehicles and slow moving traffic. A sensitivity analysis was conducted and showed that if the speed value is set too low, fewer congestion events are recognized appearing as single artifacts, specifically just a few are detected by section-based travel times. If the speed value is set too high, many widely congested areas are detected while the occurring congestion patterns do not allow for a fine classification of several types. Stop and Go waves usually oscillate around low speeds. The number of 40 km/h has proven to be appropriate in practice [Kessler and Bogenberger, 2021]. This value ensures to detect a sufficiently large number of congestion events to test the algorithms proposed in this thesis.

2.2.2 Congestion Classification

A basic and well-known traffic state classification is the three-phase traffic theory developed by Kerner et al. [Kerner, 2004; Kerner, 2009]. This theory distinguishes three phases of traffic: free-flow, synchronized flow and wide moving jam. Hence, two of these phases indicate congestion. The evaluation is based on the analysis of extensive data sets of congestion patterns on German and international freeways [Kerner and Rehborn, 1996; Kerner, 1999; Kerner, 2002; Kerner, 2004; Kerner, 2008]. The wide moving jam phase is a congested stretch that propagates upstream with a constant mean velocity as one coherent structure through all other traffic states or bottlenecks. In synchronized flow, the downstream front is usually fixed to a certain location as for example bottlenecks.

Helbing et al. extended the prevailing traffic conditions in congestion to five phases. In [Helbing et al., 2009; Schönhof and Helbing, 2007; Treiber and Kesting, 2011], the authors derive conditions for congested traffic states, starting from the instability diagram of a traffic flow model. They analyze the occurrence, appearance, spreading in space and time, and the related increase in travel time. The terminology of traffic phases is discussed and an empirical evidence of the existence of a phase diagram of traffic states is given. In contrast to other presented phase diagrams, it is shown that *widening synchronized patterns* are possible if the maximum flow is located inside of a metastable density regime. Apart from different discussions and investigations, it is pointed out that combinations of on- and off-ramps create different patterns than a single isolated on-ramp.

Leonhardt [2008a] investigates traffic prediction using spatio-temporal traffic patterns. The author compares recently measured traffic states with entries in a historical database and matches the most similar pattern to predict the traffic situation.

Four of the congestion types used in this thesis were first developed in [Transver GmbH, 2010]: *Jam Wave*, *Stop and Go*, *Wide Jam*, and *Mega Jam*. The single congestion wave is a thin stripe within the space-time diagram implying a temporarily low velocity. *Stop and Go* describes several narrow stripes representing congestion waves separated by free-flow sections. *Wide Jam* is a broad area with predominant congestion velocity. An extensive area with the domination of congestion speed values is *Mega Jam*. It represents a wide-spread traffic breakdown and mainly happens during severe vacation traffic or major incidents. In this report, traffic jams are analyzed based on GNSS traces of individual vehicles and stationary detectors. Different characteristics have been investigated and typical congestion formations have been identified. The driven velocity and the duration of each vehicle driving this velocity are two significant parameters for their classification. In [Bursa et al., 2018], a similar congestion classification on alpine freeways is described.

2.2.3 Lane-Specific Traffic Information

Bursa et al. [2019] analyze congestion patterns on alpine motorways. The authors compare lane-specific and aggregated lane traffic parameters. They conclude that congestion patterns vary across the lanes and that in regions with severe uphill slopes such as the alps congestion should be detected lane-by-lane. The study by Coifman and Ponnu [2020] analyzes the influence of a lane-by-lane congestion on adjacent lanes in the fundamental diagram.

The authors of [Sasahara and Elefteriadou, 2018] investigate lane-based speed differences upstream of an off-ramp causing spillback. They conclude that congestion from an off-ramp queue does not affect all lanes similarly, and breakdown does not necessarily occur on all freeway lanes. Their work is continued in [Sasahara, Carvalho, et al., 2020] where they predict lane-by-lane flows and speeds for freeway segments and estimate capacities per lane.

Xie et al. [2021] analyze lane capacities with a focus on freeway diverge sections in a semi-congestion environment. Traffic control strategies specifically for a lane-based online control are developed in [G.-L. Chang et al., 2020], especially ramp metering at off-ramps. A dynamic queue length estimation for long freeway off-ramps is simulated in [Heshami and Kattan, 2021]. The authors discuss control strategies for spillback on the off-ramp queue.

2.2.4 Congestion Detection Strategies

Incident detection algorithms in online traffic control have been researched since decades. An early system is the COMPASS system [Masters et al., 1991], installed in the Metropolitan Toronto area. It combines two algorithms, the All Purpose Incident Detection and the Double Exponential Smoothing algorithms which are able to make use of short-term forecasting techniques to detect irregularities of traffic variables such as volume, occupancy, or speed.

Karantanos and Chow [2015] discuss an advanced incident detection algorithm with filters and extracts non-recurrent traffic features from recurrent congestion. In [C. Chen et al., 2004], freeway bottlenecks are identified. The authors use a speed difference threshold as an indicator of bottleneck identification. Chiabaut and Faitout [2021] predict traffic congestion and travel times based on a clustering of historical congestion maps and identify consensual days based on 10 months of data on a French freeway. Y. Chen et al. [2020] examine spatio-temporal congestion patterns identification on urban freeways based on FCD.

Blumthaler et al. [2020] compared FCD and Stationary Detector Data (SDD) on congestion recognition using the congestion types from [Transver GmbH, 2010]. They state that on an alpine freeway with a high share of heavy-duty vehicles and low sampling rates of FCD trajectories, FCD do not allow for identifying short-term congestion patterns. The paper by [Sheikh et al., 2020] investigates the estimation of traffic incidents from a hybrid observer method. Traffic incidents are detected by using an improved Automatic Incident Detection (AID) technique based on lane-changing speed.

In [Margreiter, 2016a], the author suggests an AID algorithm based on a comparison of BT travel times. It first filters out too low and too high travel speed values and, second, executes a time-dependent comparison to neighbor values on the same road stretch. This algorithm determines the start and end time of congestion or incident and hence provides the times when a warning should be shown to passing drivers on variable message signs. Their algorithm works fast and reliable in terms of incident detection, still the authors do not conduct a general traffic state reconstruction. Mercader and Haddad [2020] describe a novel AID method for freeways, based on the use of data provided by BT sensors and an unsupervised anomaly detection approach. The authors validate their proposal using real data from a highway in Tel Aviv, Israel. Their method is based on characterizing anomalous traffic conditions by exploiting the fact that anomalies tend to be isolated.

Takeaways from Chapter 2: State of the Art

- Speed contour plots are an appropriate representation of speed distribution
- Congestion speed threshold of 40 km/h per cell in a space-time domain
- Usage of four mainline congestion patterns as a basis for classification
- Application of ASM and PSM as suitable interpolation methods for SDD and FCD

Identified research gaps:

- Fusion method for section-based travel times to consistent speed distribution
- Fusion method for traffic sensor data with multiple sources on a long road stretch
- Consideration of lane-by-lane congestion as part of a comprehensive congestion pattern classification
- Coupling of congestion patterns and data sources to determine detection suitability

Chapter 3

Traffic Data Collection

This chapter presents all the data sources used in this thesis with a discussion on their respective measurement principles along with their strengths and weaknesses. The experimental site is a German freeway stretch of 157 km, situated in central Bavaria. As time range, two periods of data are analyzed, from 2015 and from 2019. Data are gathered using four detection methodologies: gantries, radars, Bluetooth (BT) sensors, and Floating Car Data (FCD). Gantries and stationary radar sensors measure local speeds. BT sensors re-identify a passing vehicle at two or more locations and therefore yield travel times in between. FCD contain trajectories of single vehicles aggregated to space-time speeds.

The chapter is organized as follows. In section 3.1, the freeway stretch and each of the data sets are introduced. Section 3.2 discusses the strengths and weaknesses of each data set to point out the base line for all comparisons. In the last section 3.3, the software implementation and data processing are explained. In the end, every set of traffic data is represented as spatio-temporally discretized speed matrix with a constant velocity v per cell.

3.1 Data

This section gives an overview of all provided data and the considered road stretch along the German freeway A9 between Munich and Nuremberg. Data are available from the four detection technologies (1) gantry, (2) side radar, (3) BT, and (4) FCD.

In the following, each detection technology is presented and the available data are introduced. Where necessary, steps of preprocessing are discussed, such as outlier detection, filtering, and smoothing. Additionally, for all hardware installation (that is gantry, radar, and BT sensors), the exact positions along the road stretch are given. FCD are available through – moving – probe vehicles without fixed locations.

3.1.1 Experimental Site: Autobahn A9 in Germany

As a test site, the German autobahn A9 between Munich and Nuremberg was chosen. It is a 157 km stretch of freeway and one of Germany's most used roads. In 2019, more than $150,000$ vehicles per day were counted on average at interchange *Munich-Nord* in both directions [Autobahndirektion Südbayern, 2021]. This high traffic volume ensures the emergence of congestion, the basis of all conducted analyses. Munich is located at km 529, Nuremberg at km 372. A larger city in between is Ingolstadt situated around km 460. Data are available for

both driving directions, in Northbound direction (NB) (Munich towards Nuremberg, decreasing kilometers) and in Southbound direction (SB) (Nuremberg towards Munich, increasing kilometers). The direction of travel in the contour plots is always illustrated 'from bottom to top', therefore NB direction plots are depicted with a reversed vertical axis.

Figure 3.1 shows the localization of the considered road stretch in magenta. The Bavarian road network is depicted with bold orange lines (highest functional road class). The stretch extends between the interchange freeway A3 in the north and the Munich city highway B2R in the south. In between, the road network shows five interchanges with other Bavarian freeways. The freeways A99, A92, and A6 affect both Westbound and Eastbound direction. The interchange A93 leads only to the East, whereas the interchange A73 leads only to the West.

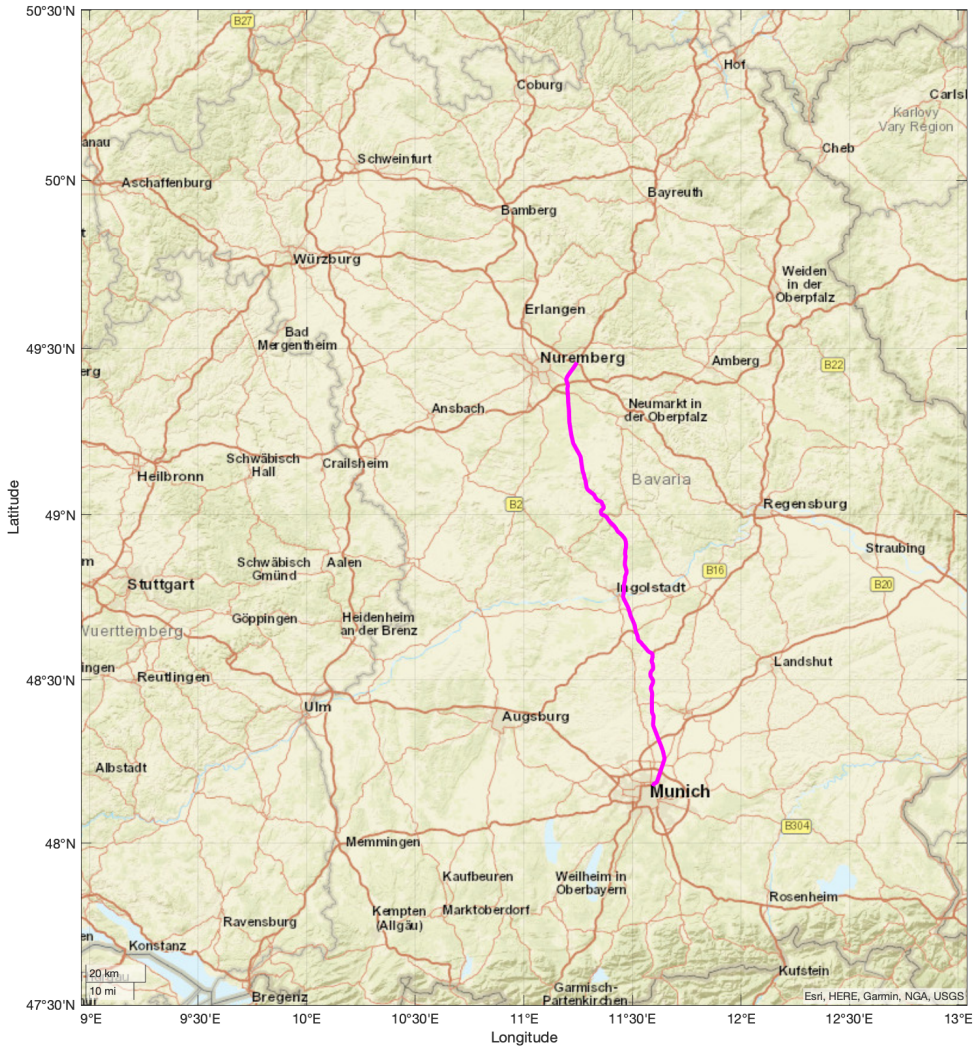


Figure 3.1: Bavarian road map with considered freeway stretch

Several areas of this freeway are equipped with variable message signs and variable speed limits, and the shoulder lane is in part-time use when the traffic flow is high. Figure 3.2 sketches the 157 km long road stretch with its varying lane configurations. On the bottom

and top, Munich and Nuremberg are indicated, respectively. In cyan, the interchanges are illustrated. All entrance and exit ramps are depicted in magenta in the middle of the road, always affecting both driving directions. Except from the very end of the freeway towards Munich, the street consists of at least three continuous lanes. The part-time use of the hard shoulder lane is hatched. It temporarily increases the number of available lanes from 3 to 4.

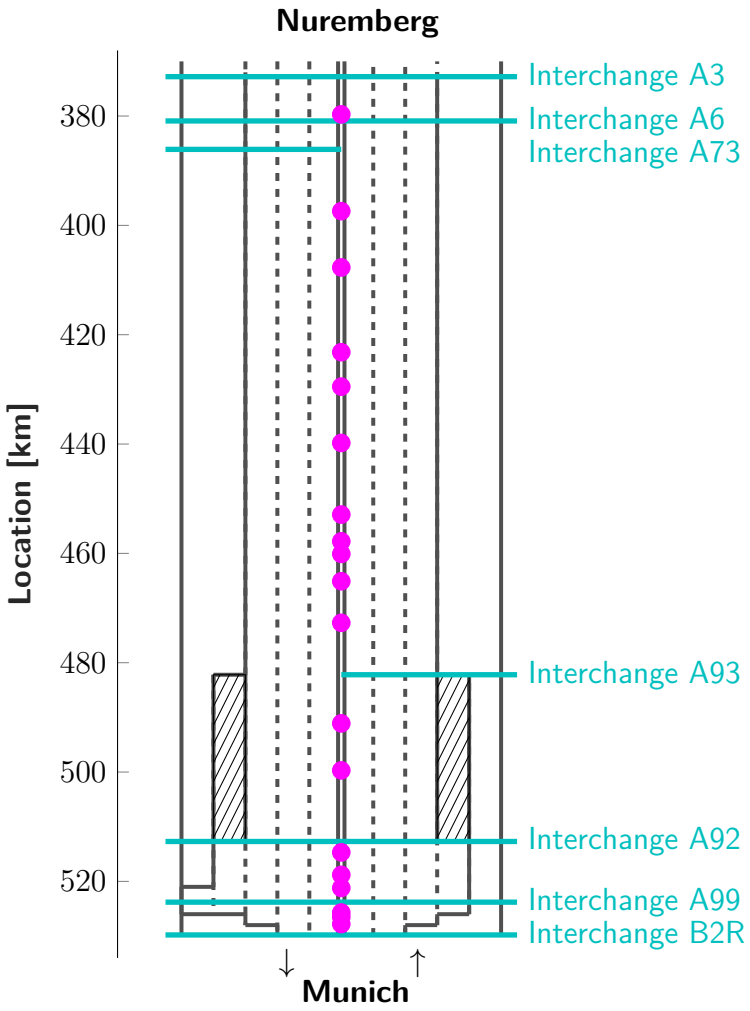


Figure 3.2: Sketch of autobahn A9 stretch with separate lanes: interchanges (cyan), ramps (magenta), part-time hard shoulder (hatched)

3.1.2 Traffic Data from Gantries

One traffic detection technology is sensors installed on gantries over-head of the road (see Figure 2.3). Their detection works locally and consists of three parts: a radar, an ultrasonic, and an infrared part. The number of vehicles passing the detector (flow) and their corresponding speeds are measured lane-specifically. Usually, flow and speed values are aggregated

to intervals of 1 min. Flows are summed up during the interval over all vehicles per lane, while speeds are averaged (arithmetic mean) during the considered minute over the number of vehicles ($v = \sum_i v_i / n_{veh}$). Minute-by-minute speed values are available both per lane and as an averaged speed value over all lanes.

Gantry measurement data are provided for two separate sections of the road, one coming from the road authority of *Northern Bavaria* in the range of km 372 to km 400, the other from *Southern Bavaria* in the range of km 480 to km 529 (green in Figure 3.3). In NB direction, 15 locations (in Northern Bavaria) and 38 locations (in Southern Bavaria) measure flows and speeds with average spacings of 1.9 km and 1.3 km, respectively. In SB direction, 20 locations (in Northern Bavaria) and 44 locations (in Southern Bavaria) measure data with average spacings of 0.8 km and 1.2 km, respectively. Speed and flow data were recorded during two periods, in spring 2015 and in spring and summer 2019.

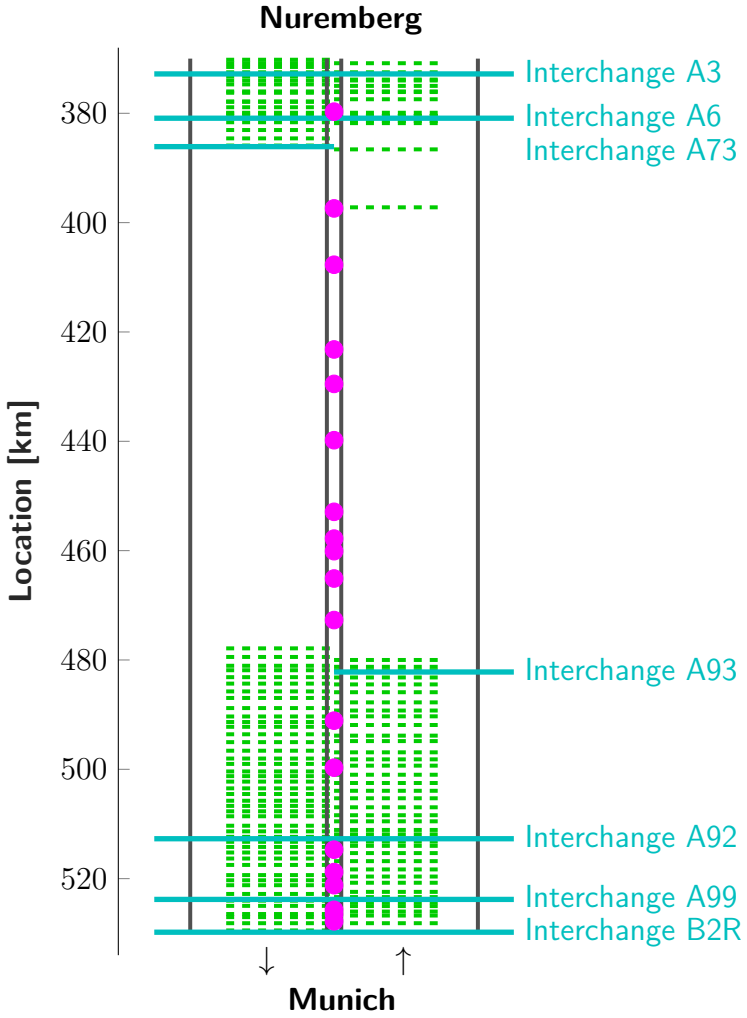


Figure 3.3: Sketch of autobahn A9 stretch with gantry locations: interchanges (cyan), ramps (magenta), gantries (dashed green)

3.1.3 Traffic Data from Radars

Radar sensors on this freeway are devices located nearside the road, installed on top of an approximately 4 m tall mast to monitor all lanes of one driving direction. Like gantry measurements, radars provide flow and speed values lane-specifically, also aggregated to one minute intervals (see also Figure 2.3). Six sensors in NB direction at km 411 and in the area between km 453 and km 463 as well as six locations in SB direction between km 452 and km 469 detect three lanes each. The locations of the radars are sketched in Figure 3.4 (orange). Minute-by-minute data from spring and summer 2019 are available.

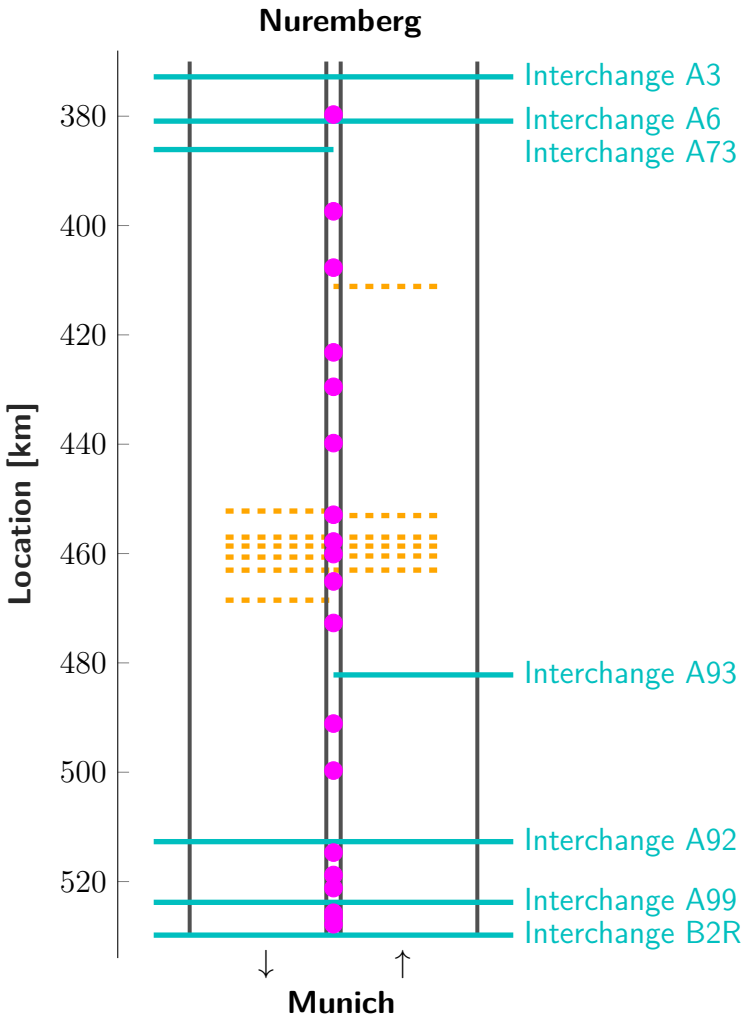


Figure 3.4: Sketch of autobahn A9 stretch with radar locations: interchanges (cyan), ramps (magenta), radars (dashed orange)

3.1.4 Traffic Data from Bluetooth

A section-based detection technology is BT sensors (see Figure 2.4). At various locations, BT sensors are placed nearside the shoulder lane and register the unique hardware address of passing BT devices. If a device is re-identified at a downstream sensor, the in-between travel time is calculated. Data are captured second-by-second at these locations. On the considered road stretch of autobahn A9, 35 BT sensors are installed with an average spacing of 4.6 km, each recognizing vehicles on both driving directions. Figure 3.5 shows the BT sensor locations in blue.

Data are available from spring 2019 along the entire road stretch. Note that the driving direction cannot be determined based on a single BT measurement solely. By matching a corresponding entry in the database for an individual hardware address hash value, the direction of travel is known. Traffic statistics derived from BT along this autobahn stretch can be found in [Kessler, Karl, et al., 2019].

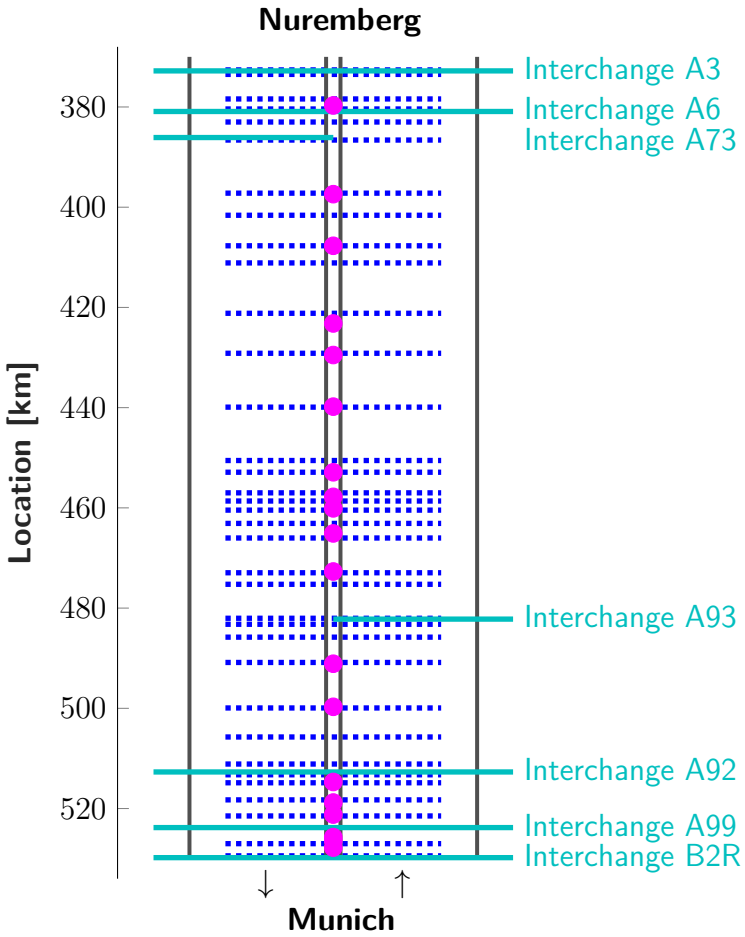


Figure 3.5: Sketch of autobahn A9 stretch with BT locations: interchanges (cyan), ramps (magenta), BT (dotted blue)

3.1.5 Traffic Data from Floating Cars

Floating Car Data (FCD) are derived by observing probe vehicles (see Figure 2.5). The probe speed measurements are gathered from map-matched Global Navigation Satellite System (GNSS) positions. Vehicles traveling along a road stretch send their GNSS positions second-by-second to a server and an algorithm fuses these data to consistent travel times or speeds per link. Usually, this service is not free of charge and data providers sell their fused data as dynamic traffic information. In this study, two FCD sets are available. One comes from the traffic information provider *TomTom* and comprises three months of data in 2015. The other one comes from measurements by the car manufacturer *BMW* fleet and comprises two months of data in 2019, both explained in detail in the following.

TomTom Data Set

TomTom offers a live data feed reporting current speeds for link segments in a road network with a time update of 1 min. The spatial resolution of the FCD are segments with an average link length of about 100 m and a maximum length of 198 m. The reported speed is calculated in a *fusion engine*, which determines a consistent speed for each link in each update cycle from various input sources. Consistency has to be ensured both spatially and temporally and cannot be assessed by the end user. The speed estimation is driven by a weighted average over the last observations from individual probes plus some black box refinements under the control of the proprietary provider. The general workflow is depicted in Figure 3.6.

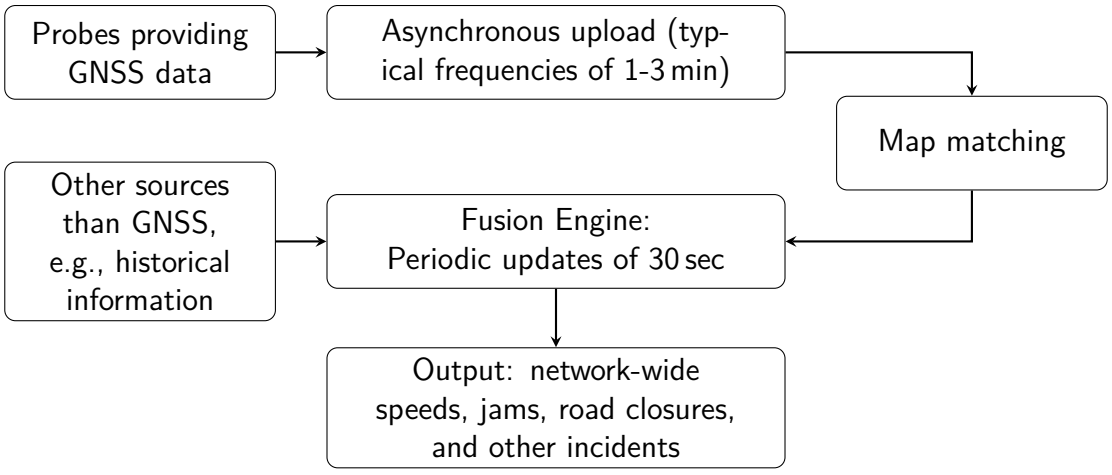


Figure 3.6: Processing chain of TomTom FCD

The provided data set comprises speed data from spring 2015 for the road stretch between Munich and interchange A93 in both driving directions. Data are available as minute-by-minute speed values. Streamed FCD are not smoothed spatially or temporarily but are calculated by the fusion engine per segment per time update interval. The timestamp corresponds to the output time of the fusion engine.

BMW Data Set

The second FCD set is provided by BMW. Trajectories from individual BMW cars are collected as a list of (x, t) tuples for each vehicle that reports data along the road based on their sent GNSS positions. First, travel times between subsequent tuples are computed. Second, both the road network and the time range are segmented. Third, for this given segmentation and cell discretization, average velocities are assigned to all grid cells which the vehicle passes through. If several vehicles cross the same spatio-temporal cell, the harmonic mean velocity is considered [Rempe, Kessler, et al., 2017].

Data provided for this study are discretized speed cells with a size of 1 min $\times$ 100 m. They were recorded on several days between April and December 2019 on the entire road stretch between Munich and Nuremberg, in total 41 and 45 days in NB and SB direction, respectively.

3.2 Quality of all Data Sets

This section compares all available data sets with regard to the data quality that was experienced during the analysis. Also, it gives an overview of the strengths and weaknesses to define the base line, refer to Table 3.1.

3.2.1 Gantry Data Set

The strength of gantry measurements is to detect exact and reliable local speeds. In general, gantry data require no filtering to exclude certain locations or times from the analyses. Just few outages occurred during the data recordings due to data server breakdowns. One location was on a construction site, which led to either no traffic or congestion (if – slow – construction site vehicles passed the sensors). Therefore, either a subset of or all lanes were omitted for some periods. These data have a high and easy availability for any period of the analyses.

3.2.2 Radar Data Set

If well calibrated, radar speed measurements generally provide accurate data. Yet, in the investigated data set, there are some locations where speed detections are less plausible. These locations are not as easy to detect as gantry construction sites, for example because their speeds do not decrease much such that congestion is measured. During the analyses, data were filtered and cleaned continuously. Most work had to be done manually, for example identifying a side radar rotated to an exit ramp instead of to the mainline.

A weakness of this radar data set is the small number of available sensors. Six sensors per driving direction is not sufficient to detect congestion reliably on the majority of locations along the road stretch. Smoothing methods such as the Adaptive Smoothing Method (ASM) require a certain number of mesh points for the interpolation to obtain realistic resulting speeds. Additionally, some sensors are so far apart from others such that interpolation over all locations is not feasibly possible.

Like gantry data, the radar data set has a high and easy availability for any period of analyses.

3.2.3 Bluetooth Data Set

Travel time measurements using BT are mainly plausible compared to the other data sources. The strength of this detection technology is its cheapness meaning that the number of sensors may increase continuously where hardware equipment requires concentration. Most of the sensors operate energy self-sufficiently by solar photovoltaics. This enables that a sensor may be moved to an optimized place, e.g., if it recognizes devices on a road nearby.

A disadvantage is a relatively large time delay when reconstructing the traffic state. Each BT recognition has to await its pairing in the data base at the end of each section. If congestion emerges within a section, it is supposed to be detected later than by other technologies because a sufficiently high number of vehicles has to pass the congested area first.

Also, a weakness is that high travel times, indicating a slow segment speed, occur from time to time, that originally come from trip interruptions or detours. These trips are challenging to filter since they do not necessarily imply congestion. During the day, most of these trips get averaged by higher speeds and are therefore not relevant to be filtered in advance. During night hours, when flow is little, slow speeds do not get averaged. Hence, it may happen that congestion is detected because two 'slow' vehicles followed after another and no 'fast' vehicle (indicating free-flow) drives in between. Then, any data smoothing algorithm will assume congestion between these two vehicles which challenges the implementation.

The minimum number of vehicles to be detected to compute a reliable speed reconstruction and to determine congestion from this is not addressed here. Before traffic is controlled online on a BT basis only, a reliable algorithm with a low false-positive rate of congestion detection has to be proven (section 6.2), which is out of the scope of this work.

Like gantry and radar data, BT data are easily accessible.

3.2.4 Floating Car Data Set

Both data sets, the TomTom and the BMW data set, consist of exact speeds. The TomTom data set is prepared for customers. It therefore provides reliable speed values, also when flow is little (see section 5.3). In such a case, the data set is enriched with historical speed information. The BMW data set is not thought for a broad set of customers, hence, it is a raw data set without preprocessing. This leads to many empty and undefined speed values in the sparse velocity matrix. Depending on the purpose of the analyses, these gaps are assumed to be either Not-a-Number or free-flow velocity (indicating uncongested areas).

Both data sets are not easily accessible. Since these data are neither public nor maintained by a public authority, the period of data collection and provision is barely extendable.

3.2.5 Summary

Table 3.1 shows an overview of the type of measurements and the provided information per data source. The first three detection technologies are collected by the public road authority and available as raw data while FCD are provided by private companies which generally do not publish any single vehicle information.

Measurements by gantry and radar sensors result in a similar type of data structure. Both technologies detect locally (in contrast to section-based traffic data) and measure aggregated

	Gantries	Radars	Bluetooth	FCD
Speed measurement	✓	✓	✗	✓
Flow measurement	✓	✓	✗	✗
Travel time measurement	✗	✗	✓	✓
Location-based information	✓	✓	✗	✗
Segment-based information	✗	✗	✓	✓
Lane-based information	✓	✓	✗	✗
Single vehicle information	✗	✗	✓	✗

Table 3.1: Overview of all available data sets used in this thesis

flows and averaged speeds minute-by-minute per lane. Spatially, they do not intersect in their detection areas. Gantry data are available in the Greater Munich Area and the Greater Nuremberg Area, radar data in the region of Ingolstadt. This motivates to combine both data sets. In the following, the combination is referred to as *Stationary Detector Data (SDD)*. Figure 3.7 shows the consolidation: gantry locations in green and radar locations in orange. SDD are available from February, April, May, June, and July 2019.

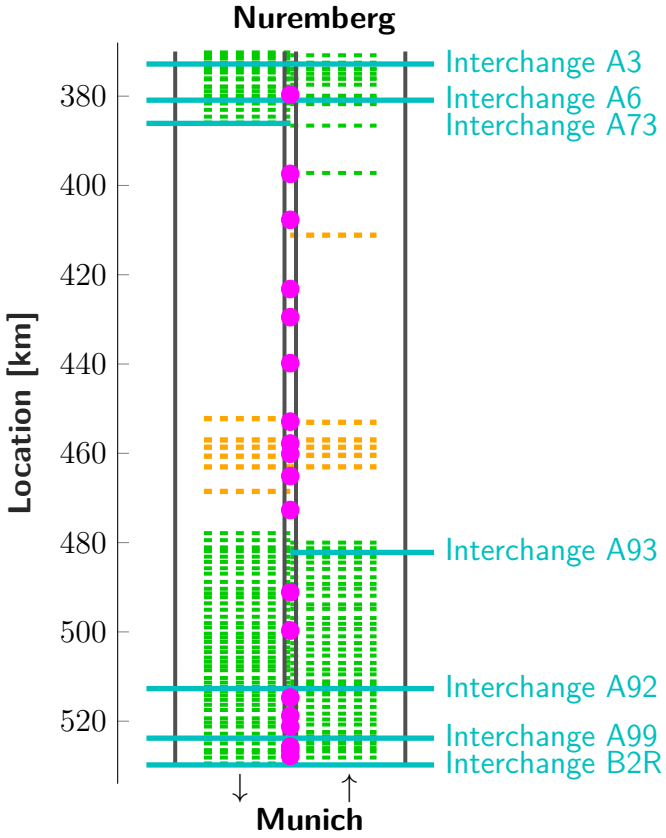


Figure 3.7: Sketch of autobahn A9 stretch with stationary detector locations: interchanges (cyan), ramps (magenta), gantries (dashed green), radars (dashed orange)

Summarizing, Figure 3.8 shows the locations of all detection technologies along the experimental site, SDD (green and orange) and BT (blue). FCD are available throughout the entire freeway section.

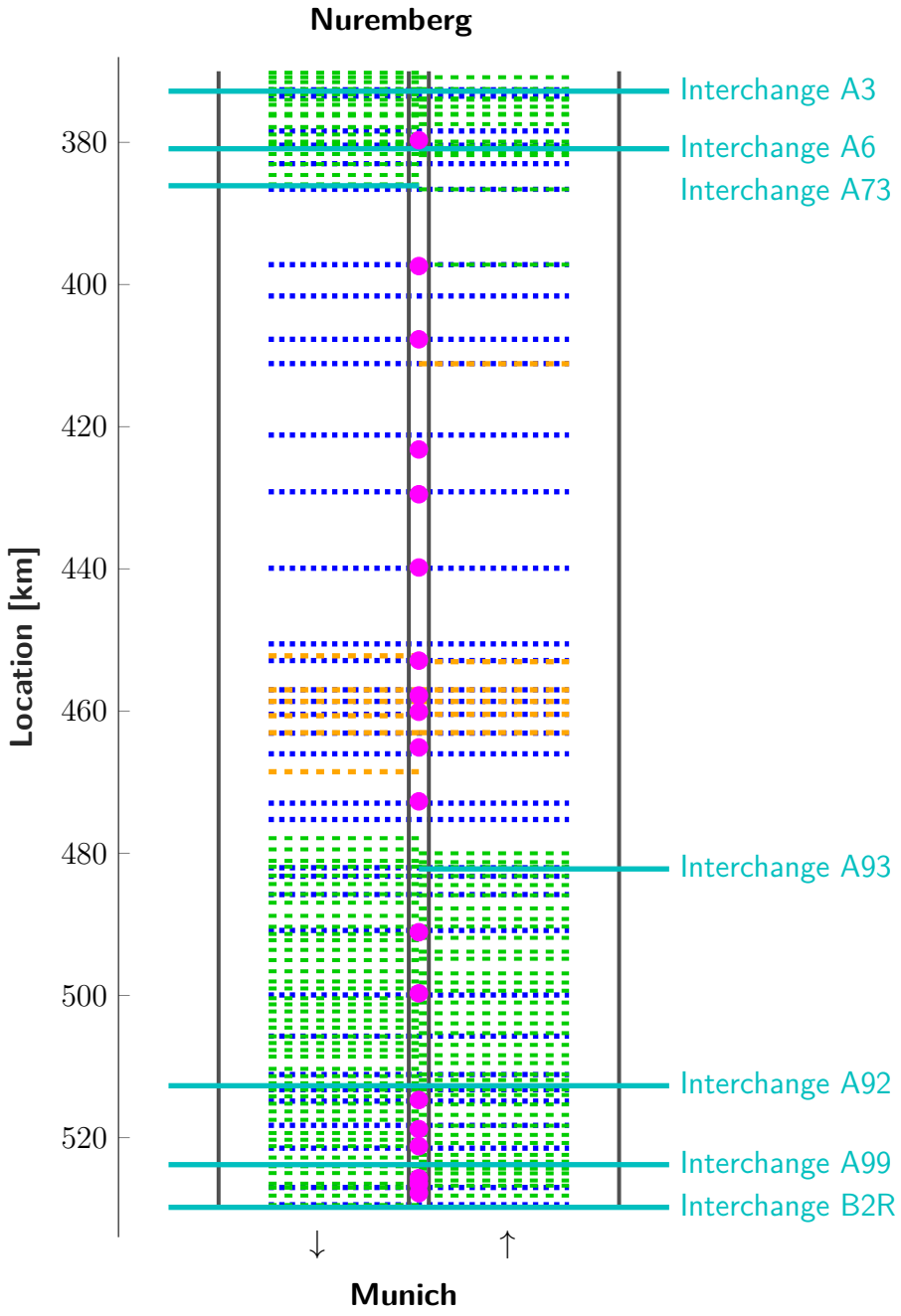


Figure 3.8: Sketch of autobahn A9 stretch with all detection locations: interchanges (cyan), ramps (magenta), gantries (dashed green), radars (dashed orange), BT (dotted blue); FCD available throughout the entire stretch

3.3 Software Implementation

This section describes the software implementation. It starts with the representation of traffic data as a space-time speed matrix, followed by the data processing steps of each data set. Furthermore, common data collection time periods are given. The general workflow of data processing is implemented as shown in the following diagram (Figure 3.9).

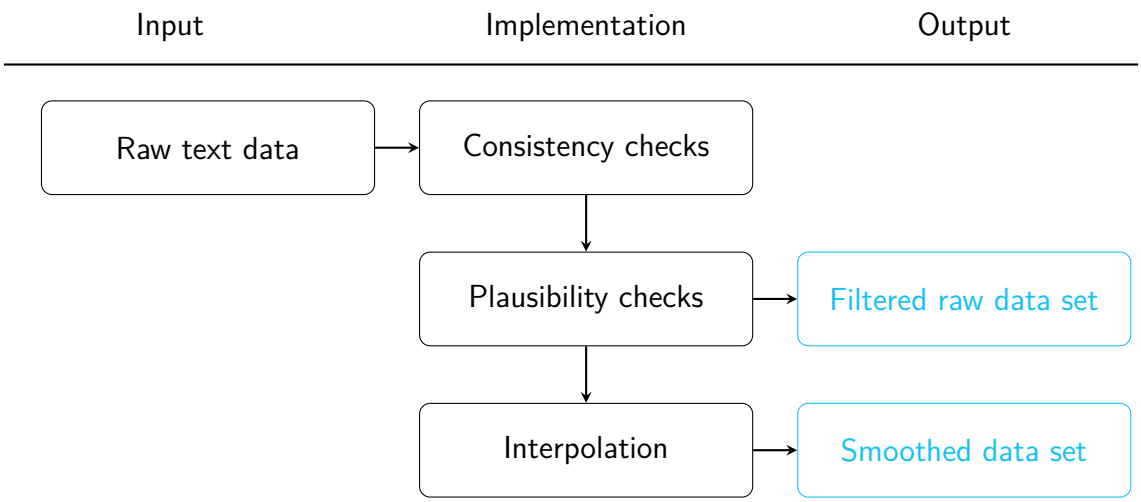


Figure 3.9: Workflow of data set processing

3.3.1 Traffic Data Representation

All analyses in this thesis are centered around a spatio-temporally discretized speed matrix. Sensors (either by direct local speed measurements or by deriving speeds from travel time measurements) generate a speed distribution over time for certain locations or sections on the road. Let speed measurements be available for a road stretch X and a time period T. A spatio-temporal speed function $v = v(x,t)$ returns the speed value for any location $x \in X$ and any time $t \in T$. Furthermore, let X be separated into a set of sections $x_1, \ldots, x_n$, the spatial resolution of the speed data. Similarly, let T be separated into a set of time intervals $t_1, \ldots, t_m$. The sets $\{x_i\}_{i=1,\ldots,n}$ and $\{t_j\}_{j=1,\ldots,m}$ separate the spatio-temporal area $X \times T$ into a grid, and the function v is constant inside each cell $x_i \times t_j$ of this grid.

Cells of interest for this study are only those that contain a speed value below a predefined threshold since they denote congestion. There exist other definitions of *congestion* apart from a velocity threshold (confer section 2.2). In this thesis, congestion prevails if any spatio-temporally discretized cell is assigned a speed value below the critical velocity v_{crit}. Table 3.2 gives an overview of the introduced parameters.

3.3.2 Gantry Data Set

Speed and flow measurements by gantry sensors come from two data hubs, analogously to the responsibilities in northern and southern Bavaria.

Name	Description
X	Road corridor
$x \in X$	Location in X
$\{x_i\} = \{x_1, \ldots, x_n\}$	Set of sections (spatial resolution)
T	Time period
$t \in T$	Time in T
$\{t_j\} = \{t_1, \ldots, t_m\}$	Set of time intervals (temporal resolution)
$x_i \times t_j$	Spatio-temporal speed cell
v_{ij}	Speed in cell $x_i \times t_j$
v_{crit}	Speed threshold to distinguish free-flow vs. congested traffic

Table 3.2: Parameters and variables

Input In the north, the traffic management center *Fischbach* collects data and provides them as text files, one .csv-file per day per driving direction containing all sensors, hence, two files per day. In the southern center *Freimann*, speeds and flows are stored in .xml-files, one file per lane per location per day per driving direction. In sum, 372 files per day need to be processed. As a first step, all text files are imported into a software environment. There, the relevant information such as speeds of each lane, locations, and timestamps are extracted.

Consistency Checks The gantry data consistency check eliminates negative speed values, for example coming from construction site vehicles driving in the opposite direction. Further, it excludes data, that do not obey the conditions of the test field, for example values from a ramp lane sensor. Missing time stamps (e.g., due to a system breakdown or a detector outage) are added to the time vector and corresponding speeds are set to Not-a-Number values.

Plausibility Checks The plausibility check detects and eliminates data lacks such as construction sites. In such a case, speed data measured by gantries are not desired though feasible, e.g. observed construction site vehicles driving very slow. Considering a construction site as a congested region would yield a hot spot although it exists only temporarily.

The considered road stretch comprises sections with a varying number of lanes (between 2 and 5 per driving direction, Figure 3.2). For comparisons with other data sets, the data processor sets up all cross-sections as three-lane traffic. Two parallel lanes are equipped only at a negligibly small number of locations, therefore, a separate treatment is not necessary. Wherever more than three lanes are equipped with sensors, the lanes are aggregated to exactly three. Without this aggregation, comparisons with data from other detection technologies would be less tractable. Furthermore, no stringent consideration of one lane would be possible (e.g., shoulder lane).

The plausibility checker captures additional lanes by combining two existing lanes to one lane, making use of local knowledge. Generally, wherever part-time shoulder usage is allowed (which is in regions with 3 regular lanes), the shoulder lane and the right lane are combined; the middle and left lanes are kept individually. It is verified in the data set that the speeds between the shoulder lane and the right lane do not differ much. In sections with 4 regular lanes without shoulder usage, the two left lanes are combined while the two right lanes are kept

separately. Again, the speeds of the two left lanes are seen to be relatively similar. Locations with 5 lanes are treated individually since they appear rarely, and lanes are grouped according to their target directions. Especially regarding the consideration of the novel congestion type *Lane Jam*, a careful proceeding is essential.

The combination is executed as follows. Let l_1 and l_2 be two lanes at one sensor location for which an aggregation should be executed. Let $n(l_1)$ and $n(l_2)$ the corresponding number of vehicles per lane within a time interval δt of one minute of detection. Furthermore, let $v(l_1)$ and $v(l_2)$ be their average speeds measured on each lane per minute, respectively. The flow $q(l) = n(l)/\delta t$ of the combined lane l is given as

$$q(l) = q(l_1) + q(l_2) = \frac{n(l_1) + n(l_2)}{\delta t} \tag{3.1}$$

which is the sum of all vehicles passing either lane l_1 or lane l_2. The speed $v(l)$ amounts to

$$v(l) = \frac{v(l_1) \cdot q(l_1) + v(l_2) \cdot q(l_2)}{q(l)} \tag{3.2}$$

which is the flow-weighted average speed of both lanes.

All information is represented as a speed matrix together with space and time vectors. The cell-based speed information is ordered according to location and timestamps.

Interpolation Additionally to these 3+1 measured data, 3+1 interpolated speed matrices are computed (3 lane-based, 1 mainline) using the ASM approach.

Output As a last step, all information from Fischbach and Freimann is merged and stored in one common file consisting of eight speed matrices per direction and day.

3.3.3 Radar Data Set

Though the input speeds are represented differently, the radar data set is processed similarly.

Input Radar data are available as .kuz-files where all information such as timestamp, lane, speed, and flow is stored in one file per sensor per month. First, each file is imported into the software environment and broken into smaller parts comprising one day.

Consistency Checks Consistency checks identify negative speed values and set them to plausible numbers. They may appear depending on the range of byte storage, either from 0 to $255\,\mathrm{km/h}$ or from -128 to $+127\,\mathrm{km/h}$.

Plausibility Checks The lane numbering is validated and only data coming from one of the three main lanes per direction are included. Since radar measurements are gathered in a 3-lanes region only, a lane reduction is not necessary. The result is a set of speed vectors per sensor per day per driving direction.

A location sorter fills these speed vectors into a space-time matrix per day per driving direction, where it orders all sensors by location according to the direction of travel. Moreover, the software creates sorted location and time vectors.

Interpolation Again, in addition to the filtered raw data, also an ASM-smoothed data set is created per day per driving direction.

Output Analogously to the gantry data set, eight speed matrices are stored in one output file per direction and day: 3 lane-based and one mainline, each real and interpolated.

3.3.4 Bluetooth Data Set

BT data are stored in a MySQL data base containing relevant information of each BT recognition (timestamp, location, hash value of hardware address, predecessor and successor recognitions, among others).

Input Applying a MySQL dump from the productive operating system writes the BT detections within the test field into one .dsv-file per month (16+ million lines), not separated into driving directions. A local MySQL file processor splits the monthly .dsv-file into daily .csv-files. Each .csv-file is then imported in the software environment.

Consistency Checks As a first step, average speeds out of the measured travel times are computed between the previous and the current detection, taking into account the difference of both detection locations. If no predecessor is available, the speed is set to Not-a-Number. The consistency checker investigates if all resulting speeds are positive (note that $v = 0$ cannot happen for travel time measurements). Additionally, it checks if the previous and succeeding registrations of the BT devices are filled consistently in the data base, that is, if the given timestamp and location of a supposedly preceding/following recognition is equal to timestamp and location of the data base entry indicating the previous/next recognition at the sensors.

Plausibility Checks The plausibility checker filters all BT detections to exclude measurements of the following types ([Margreiter, 2016a]):

- Several devices carried in the same vehicle (temporal offset between detections ≤ 1 sec at each location)
- Multiple detections of the same device in the range of a single BT sensor
- Implausible detections of the same hardware address in the network due to wrong hash value matchings
- Travel times leading to speed values > 250 km/h for the segment (high speed)
- Travel times leading to speed values < 5 km/h for the segment (low speed)
- Hopping devices at several locations indicating different travel directions due to missing detections or detours

The data processor runs through all *measurements*, which is a single device detection. It creates *trips* from the detections which means that it identifies two measurements having the same hash value of a hardware address and belonging to the same travel. If more than two measurements describe the same trip, the travel is split into a sequence of trips, each with a dedicated start and end point both in time and space. Hence, the direction of travel can be determined. If a measurement is a single detection without any predecessor or successor, it is excluded from the analyses because the driving direction of this measurement is not known. Further, a list of all timestamps and all locations is created where each row corresponds to one trip.

Interpolation The fusion of these individual travel times and their determined speeds to an overall space-time speed reconstruction is described in section 4.1.

Output In total, one file per direction and day is created, each consisting of a list as (location1, location2, time1, time2, speed).

3.3.5 Floating Car Data Set

The TomTom and the BMW data set are set up similarly. Therefore, also their processing workflow is nearly equal except from the interpolation.

Input FCD based speed values are stored as .csv-files, one file per day and driving direction. Data are organized in (x, t) cells with a cell size of 100 m and 1 min. For each cell containing a speed value, the x and t cell indices and the corresponding speed are indicated. First, a data importer converts the raw data into the software environment.

Consistency Checks The consistency checker filters for nonnegative speeds. Note, that in contrast to BT speeds, $v = 0$ may occur in these high-resolution travel time measurements because each device sends its updated GNSS position every few seconds and is not re-identified at certain locations only.

Plausibility Checks The plausibility checker filters realistic speed values which are not too low or too high. Moreover, the space and time grid vectors are created.

Interpolation The TomTom data set contains speed values for every minute and segment. An interpolation is not necessary. In contrast, the BMW data set contains several undefined values. In addition to the raw data set, an interpolated speed matrix per day and per direction is created for BMW data using the Phase-Based Smoothing Method (PSM) approach.

Output For each direction and day, a file is stored containing one and two speed distributions, a raw and an interpolated one, in case of TomTom and BMW, respectively.

3.3.6 General Processing

In general, each data set is stored in the scheme of one file per day and per driving direction. If any speed data are implausible or need to be excluded for some reason, this value is omitted from the raw data set, that is, it is set to either Not-a-Number or zero or free-flow, depending on the application specification. Only afterwards, raw data are interpolated to smoothed speed data such that data errors are not propagated.

Every data set contains all kinds of days such as normal working days, weekends, vacation days, public holidays, etc. in order to enable congestion emergence analyses on a broad data basis.

These data sets consist of different days per data source. Gantry sensors have been set up a long while ago whereas radar and BT sensors were recently installed on this road stretch. The calibration of both led to several sensor outages and maintenance phases. The time periods of data collection are chosen as depicted in Table 3.3. Unfortunately, the recorded time periods coincide only little. The set of FCD could not be enlarged due to the proprietary ownership. When it comes to data fusion, several days have to be excluded from the analyses due to different reasons, such as data lacks of any data source. In total, only 10 days exist where fine speed data from all sources are available. Hence, a fusion approach cannot be realized on a large-scale data base.

Detection Technology	**Available Time Periods**
Gantry data set	April, May, June, 2015; February, April, May, June, July, 2019
Radar data set	February, April, May, June, July, 2019
BT	April and May, 2019
TomTom FCD	April, May, June, 2015
BMW FCD	Single days in April and May, 2019; November and December, 2019

Table 3.3: Available time ranges

Takeaways from Chapter 3: Traffic Data Collection

- Large data sets from gantry, radar, BT, and FCD measurements available for evaluations
- 157 km long freeway stretch on the German autobahn A9 equipped with various detection technologies, see Figure 3.8
- Speed distributions from 2 to 5 consecutive months available, varying time ranges per data source
- Specific properties of each data source yield different strengths and weaknesses, see Table 3.1
- Enormous processing of raw data necessary to generate comparable, comprehensive, and plausible data sets for scientific evaluations
- Raw and interpolated speed matrices created by an integrated toolchain

Chapter 4

Extended Congestion Definition and Detection Strategies

This chapter contains all methodologies of the thesis. It discusses automatic congestion identification, differentiating various traffic detection technologies. Each detection technology has advantages and disadvantages making them suitable for different use cases. The chapter is structured in three major parts: Fusion, congestion detection, and comparison approaches of data sources. Each part consists of two sections.

The fusion part consists of two algorithms. First, a novel interpolation method, the *Low-Resolution Travel Time Smoothing Method (LTSM)*, fuses section-based travel times such as Bluetooth (BT) data to a consistent space-time speed field and is described in section 4.1. Second, multi-sensor data fusion approaches fuse speed data from all available detection technologies to a combined speed distribution and are described in section 4.2.

The congestion detection part explains congestion identification and classification. It first introduces virtual trajectories that simulate vehicles in the space-time domain, and a clustering approach to find coherent congested areas (section 4.3). Then, a 4+1 congestion classification is explained in section 4.4. It consists of four mainline congestion patterns and one lane-specific type. The result is a set of congested areas, so-called *congestion clusters*, where each cluster has three properties: a shape (convex hull), a size (area of the convex hull), and a congestion type.

The last part is comparison metrics of the data sources. Section 4.5 presents an analysis which data source is able to recognize an incident earliest ex-post and which delay the other data sources show. In section 4.6, congestion hot spots are investigated using automated pattern recognition. For every combination of detection technology and congestion type, spatio-temporal hot spots are identified.

4.1 Low-Resolution Travel Time Smoothing Method

All data derived from local speed measurements, i.e., gantry sensors and radars, can be interpolated applying the *Adaptive Smoothing Method (ASM)* (section 2.1). The input is speeds at certain locations which the methodology extends to hold at intervals. A spatio-temporal speed matrix out of high-resolution travel time measurements such as Floating Car Data (FCD) can be obtained using the *Phase-Based Smoothing Method (PSM)* (section 2.1). Hereby, speeds are reported for segments on the road stretch.

Low-resolution travel time measurements derived BT or WIFI sensors (section 2.1.2), where the detectors are placed several kilometers apart from each other, can be transferred into speed data deploying the methodology explained in the following. If an equipped device was re-identified at a first location l_1 and later at a second location l_2, the average speed v between l_1 and l_2 can be computed as

$$v = \frac{l_2 - l_1}{t(l_2) - t(l_1)}. \tag{4.1}$$

Equation 4.1 is only a rough estimation in case that l_1 and l_2 are far apart. In the following, an approach is described to derive a spatio-temporally discretized speed matrix from segment-wise travel time measurements.

4.1.1 Preliminaries

Let a *measurement* be defined as a single device (or vehicle) detection at one sensor. This device does not need to be re-identified at a downstream detector. A single measurement does not inform about the direction of travel. A *trip* is the travel between exactly two sensors. From a trip, the average vehicle speed can be calculated as the difference of both detector locations divided by the corresponding travel time (equation 4.1). This refers to the linear speed interpolation of both placements. If a vehicle passes by several sensors, it is said to be a *sequence of trips* since the speed may vary along the detections. The velocity of a trip is constant.

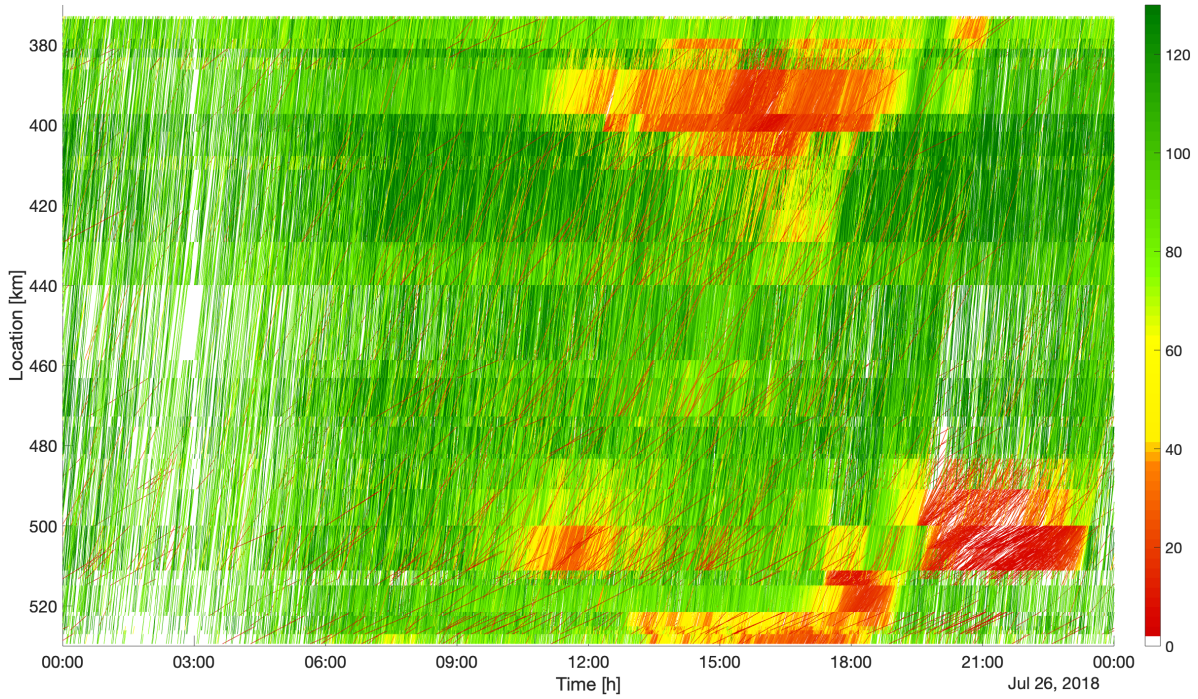


Figure 4.1: Travel times (lines); color corresponds to averaged speed (in [km/h])

Figure 4.1 illustrates a speed contour plot with the horizontal axis being the time line (increasing time from left to right) and the vertical axis representing the space (direction of

travel from bottom to top). Here, raw data of BT travel times from one day in July 2018 in Northbound direction (NB) are depicted. The vehicles travel starting on the lower left corner going towards the upper right (decreasing kilometers while increasing time). Each linear line corresponds to a single trip. The color depicts the speed of the trip.

The speed distribution can be approximately recognized showing phases with both free-flow and congested regimes. Red lines in an area of mainly green lines correspond to trips with a very low speed compared to neighbored vehicles (e.g., travel interruptions). These raw data are not yet filtered according to the plausibility checks described in section 3.3.4.

The algorithm follows a three-step approach. First, measured travel times are discretized into a spatio-temporal grid. Each trip corresponds to a linear line with a dedicated start and end point both in space and time. Second, for each trip, the algorithm identifies all affected grid cells and matches trips and cells, respectively. Third, all trip segments inside a grid cell are utilized to compute the average speed of this cell.

4.1.2 Spatio-Temporal Discretization

Let travel time measurements be available for a road corridor X and a time period T. Furthermore, let X be separated into a set of location intervals $x_1, \ldots, x_n$ and let T be separated into time intervals $t_1, \ldots, t_m$. Figure 4.2 illustrates an equidistant discretization of $X \times T$ into a grid with cells $x_i \times t_j$.

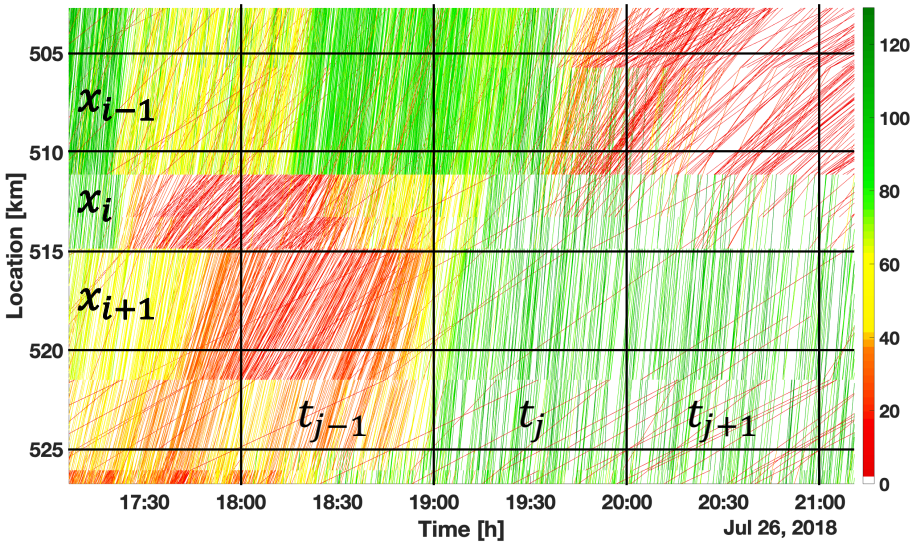


Figure 4.2: Travel times in a spatio-temporally discretized grid

4.1.3 Matching Between Trips and Cells

Each cell $c = c(x_i, t_j)$, abbreviated by c_{ij}, is crossed by several trips and each trip p traverses a certain number of (neighbored) cells. In order to obtain an average speed value v_c per cell c, the algorithm considers the share of each trip which lies inside c (trip segment p_c). It distinguishes five cases:

(a) Trip starts before and ends behind the cell

(b) Trip starts before and ends inside the cell

(c) Trip starts inside and ends behind the cell

(d) Trip starts inside and ends inside the cell

(e) No intersection between trip and cell

The distinctions are necessary to determine the exact start and end points of the spatio-temporal trip segment. Figure 4.3 shows how c_{ij} is traversed by trips from cases (a) – (d).

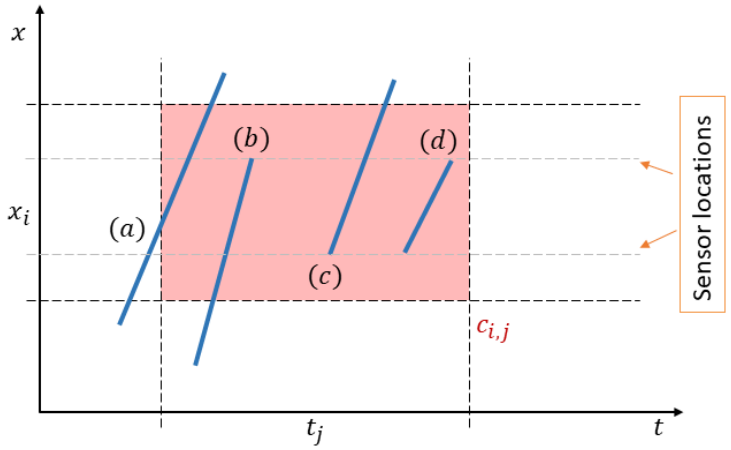


Figure 4.3: Cell with intersecting trips for cases (a) – (d)

4.1.4 Average Speed in Cells

To compute the average speed value v_c for c, the speeds of all k affected trips $p_{c,k} = p_{c,k}(x_i, t_j)$ are weighted combined with a weighting factor α_k.

$$v_c = \frac{\sum_k \alpha_k \cdot v(p_{c,k})}{\sum_k \alpha_k} \tag{4.2}$$

The proposed method distinguishes three variants of setting α_k (see Figure 4.4):

(1) distance-weighted average, $\alpha_k = dist_{p,c}$ (distance of the trip segment)

(2) duration-weighted average, $\alpha_k = dur_{p,c}$ (duration of the trip segment)

(3) distance-duration-weighted average, $\alpha_k = dist_{p,c} \cdot dur_{p,c}$

The spatio-temporally discretized cell grid of $X \times T$ contains the average speed value v_c for each cell c.

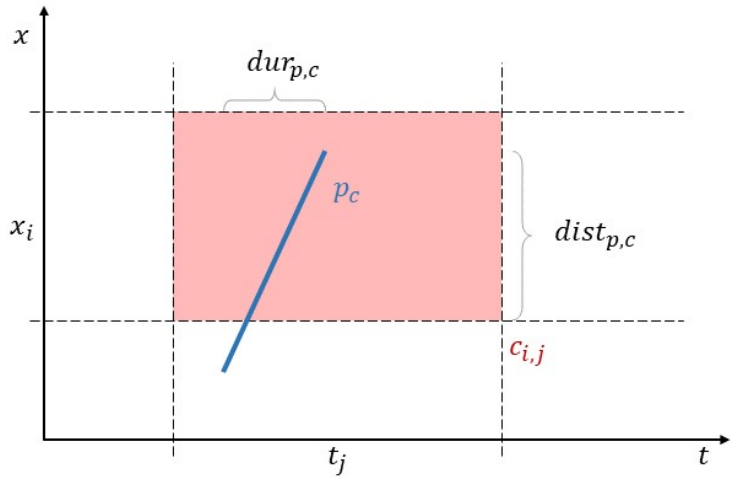


Figure 4.4: Possible settings for α: distance and/or duration weights per trip segment p_c

4.2 Multi-Sensor Data Fusion

In this section, a multi-sensor data fusion strategy is developed to mitigate the shortcomings of individual detection technologies. Let at least one data set from each detection technology explained in section 3.1 be available. Some data sets provide speed, others travel time information. An information fusion of available data sets enables the obtainment of a more detailed insight into the prevailing traffic state.

This section describes several approaches to fuse data from various sensors such as gantry detectors, radars, BT, and FCD. The assessment of the fused data set is executed using the novel comparison measure *Symmetric Square Inverse Mean Percentage Error (SSIMPE)*.

4.2.1 Preliminaries

Again, let $X \times T$ span the space-time domain with intervals $x_i, i = 1, \ldots, n$ and $t_j, j = 1, \ldots, m$. Data of all detection technologies are represented as smoothed speed values in a uniform grid with step size $|x_i| = 500\,\text{m}$ and $|t_j| = 1\,\text{min}$. Each matrix consists of cells c_{ij} having a constant velocity v_{ij}. Given a number of S sensor technologies on the considered road stretch, $v^s, s = 1, \ldots, S$ denote the speed matrices, respectively.

Since all data are sparse in time and space, there are cells in each velocity matrix v^s_{ij}, that do not contain a velocity value. These entries are treated as *undefined* values. Let $\exists v^s_{ij}$ denote that the speed value at (x_i, t_j) of data set s is defined, and $\neg\exists v^s_{ij}$ denote the value being undefined.

4.2.2 Fusion Methods

This section presents three fusion approaches. First, a simple weighted average is described, followed by a hierarchical ordering of all data sources. The third method applies a standard fusion strategy involving neighborhood information to the fused matrices derived from the first two methods.

Method 1: Weighted Average Speed per Cell

The first approach is the simple convex combination. Essentially, this is the average speed per cell per data source weighted with a reliability factor α_s:

$$v_{ij}^{\text{fused}} = \sum_{s=1}^{S} \alpha_s v_{ij}^s \tag{4.3}$$

with $\sum_s \alpha_s = 1$. If α_s is set to $1/S \; \forall s$, all data sets are weighted equally and the fused speed corresponds to the arithmetic mean of all speed matrices.

Method 2: Fill-up of Gaps Ordered by Reliability

The second method processes data according to their estimated reliability. Without loss of generality, let sensor technology 1 be the most reliable and technology S the least reliable data source. The most reliable data in a cell c_{ij} dominates over any other data source and defines the speed v_{ij}^{fused}. If not available (i.e. $\neg\exists v_{ij}^1$), the second most reliable data source is dominant and so forth. The method is given as:

$$v_{ij}^{\text{fused}} = \begin{cases} v_{ij}^1 & \text{if } \exists v_{ij}^1 \\ v_{ij}^2 & \text{if } \neg\exists v_{ij}^1 \text{ and } \exists v_{ij}^2 \\ v_{ij}^3 & \text{if } \neg\exists v_{ij}^1, \neg\exists v_{ij}^2 \text{ and } \exists v_{ij}^3 \\ \dots & \\ v_{ij}^S & \text{if } \neg\exists v_{ij}^1, \dots, \neg\exists v_{ij}^{S-1} \text{ and } \exists v_{ij}^S \\ \text{undefined} & \text{if } \neg\exists v_{ij}^1, \dots, \neg\exists v_{ij}^S \end{cases} \tag{4.4}$$

The reliability of all detection technologies may vary in time and/or space. For example, BT data are less sound during night, or direct gantry measurements are more trustworthy than interpolated speeds between gantries. Therefore, the reliability can be assigned more complex, depending on the purpose. With the data sets treated in this dissertation, the simple approach of a constant reliability is sufficient, for example, because of a small gantry spacing or no other data than BT available nightly.

Method 3: Fused Speeds with Neighbor Speed Information

The third methodology is an ASM-based approach to find average speeds per cell. First, all speed data are calculated according to one of the other two fusion approaches (equations 4.3 or 4.4). Thereafter, all speed values inside the cells are optimized using the ASM approach.

To this end, additional speed information from the neighborhood of each cell is taken into account, analogously to both smoothing kernels of the ASM [Treiber and Helbing, 2003]. This enables an interpolation in both spatial and temporal directions and favors to also fill undefined cells from the previous fusion methods.

4.2.3 Assessment of Approaches

The average deviation between any two spatio-temporally discretized data sets, both individual and fused data, is compared using the *SSIMPE*. For two equally sized matrices y and z, each with cells $c_{ij} \in X \times T$, SSIMPE is defined as:

$$\text{SSIMPE}(y, z) = \frac{1}{n \cdot m} \cdot \sum_{x_i \in X} \sum_{t_j \in T} \left(\frac{y_{ij}^{-1} - z_{ij}^{-1}}{0.5(y_{ij}^{-1} + z_{ij}^{-1})} \right)^2 \tag{4.5}$$

The symmetry does not prioritize one of the matrices over the other. The square ensures to only sum nonnegative values while higher deviations receive a larger weight than smaller ones. The inverse mean indicates that a deviation in the range of smaller speeds is more harmful than any deviation in the range of higher speeds. The optimum would be SSIMPE $= 0$. Only cells with defined speed can be compared, otherwise $(\neg\exists y_{ij} \vee \neg\exists z_{ij})$ the addend representing the deviation of that cell (x_i, t_j) is set to 0 and the denominator $n \cdot m$ representing the amount of summed cells is reduced.

In order to assess the reconstruction accuracy, two aspects need to be considered. One is the Ground Truth (GT) that is used to assess the accuracy of a fusion result. Second is the methodology to determine an accuracy, given the wish to use all sensor data without sparing one as the GT. In order to select a GT, the following considerations are done:

- As visible in speed contour plots, the measurements of each data source are sparse in time and space.
- Gantry and radar sensors provide accurate speed measurements but are limited to certain locations.
- FCD have a low number of samples at nightly hours, but provide relatively accurate speed estimates for varying times and spaces.
- BT-based travel times measurements are abundant, though, especially when averaged over congested segments, the cell-based velocities are inaccurate due to large distances between neighboring sensors.

Albeit, BT data are a valuable source to assess travel times over sections, they are less suited for accurate time-space velocity estimates. Although FCD do not resemble macroscopic but microscopic speeds, the vehicle's velocities are sufficiently good representatives, specifically in congested traffic where vehicle speeds are synchronized [Kerner, 1999]. In this study, a combination of gantry, radar and FCD speeds is used as approximation of the GT, using fusion method 1 (equation 4.3) to combine the sources into a sparse matrix v^{GT}.

A commonly used approach in model training and evaluation is to divide available data into a *training* and a *test* data set. Figure 4.5 depicts the methodology applied: First, all data in sparse matrices of gantry detectors, radars, and FCD are randomly separated into one of both sets with probability of 50 % each. Next, the mentioned fusion techniques are applied to the training matrices (green) of different combinations of sensors. The test matrix (blue) consists of the combined gantry, radar and FCD data. The quality assessment applies the SSIMPE error measure (equation 4.5). For robustness of the results, this random procedure is repeated for 50 times and average errors are considered.

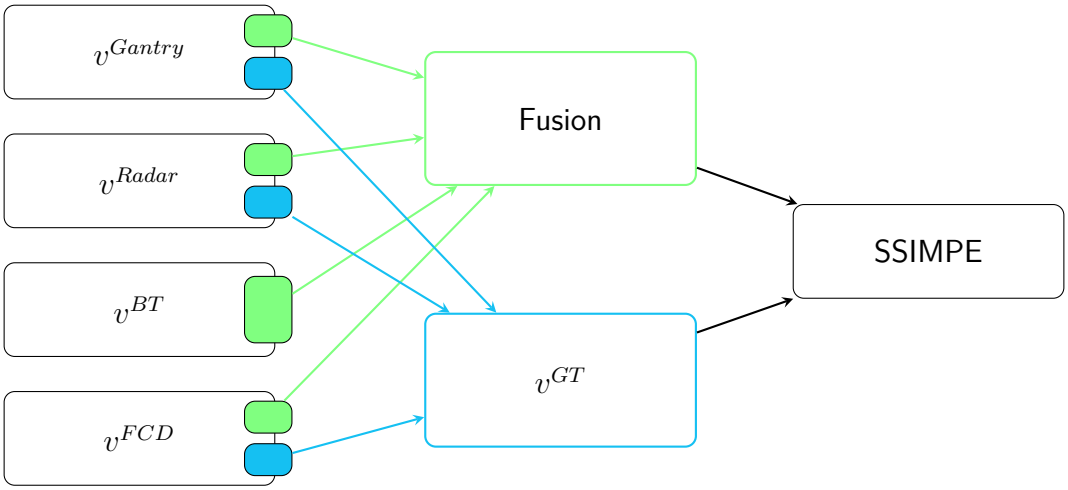


Figure 4.5: Information flow of sensor data velocity matrices to fusion and quality assessment

4.3 Congestion Detection Methods

This section describes a strategy to detect occurring congestion on a given spatio-temporal speed distribution. First, the concept of virtual trajectories is introduced. Second, the identification of congestion clusters is explained.

4.3.1 Virtual Trajectories

Virtual trajectories aim at simulating vehicles crossing the speed matrix. The foundation is again a spatio-temporally discretized speed matrix v along the considered road stretch X and the time frame T with cells $x_i \times t_j$.

A virtual trajectory corresponds to a trace throughout the speed matrix which a vehicle would take if starting at an arbitrary point in the matrix.

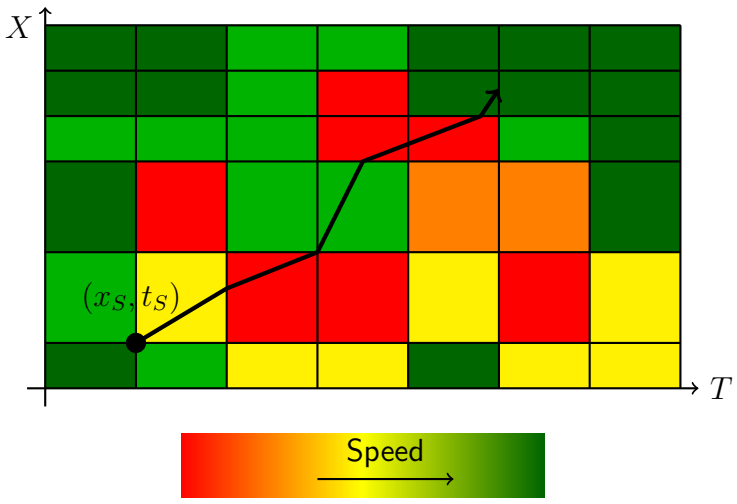


Figure 4.6: Sketch of the functionality of virtual trajectories

This can be done for some starting point $(x_S, t_S) \in X \times T$ by solving the ordinary differential equation

$$\frac{dx}{dt} = v(x(t), t) \tag{4.6}$$

with initial condition $x(t_S) = x_S$ (x as function returning a location in dependency of time). Whenever the trajectory reaches the bounds of a cell, it continues to the neighbored cell with the prevailing speed there. Figure 4.6 sketches the methodology for one virtual trajectory.

This way, from each trajectory, a velocity profile, the traveled time, and the traveled distance can easily be extracted given any starting point (x_S, t_S).

4.3.2 Congestion Clusters

This section introduces a novel congestion detection strategy. Each congestion event is determined as a *cluster* on a spatio-temporally discretized speed matrix. A cluster is defined as the convex hull of congested cells. In the following, a methodology is described to automatically detect congestion patterns. Figure 4.7 illustrates an example of an arbitrary speed function.

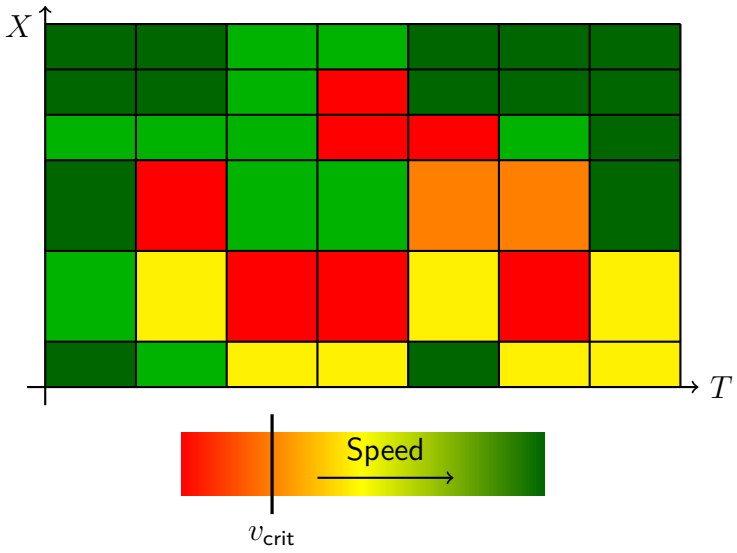


Figure 4.7: Computation of congestion clusters, part 1: Speed function v

Since only cells defined with a speed value below v_{crit} are relevant, the spatio-temporal speed is divided into cells with speed $v \geq v_{\text{crit}}$ and cells fulfilling the condition $v < v_{\text{crit}}$ (marked in red in Figure 4.8). Two such cells $x_{i_1} \times t_{j_1}$ and $x_{i_2} \times t_{j_2}$ are denoted as *connected* if the following two conditions hold simultaneously:

$$\begin{aligned} i_1 &\in \{i_2 - 1, i_2, i_2 + 1\} \\ j_1 &\in \{j_2 - 1, j_2, j_2 + 1\} \end{aligned} \tag{4.7}$$

Let this one-hop neighborhood be suitably defined at the boundaries. By applying the speed threshold v_{crit} to the spatio-temporal speed area, non-congested regimes are dropped from the

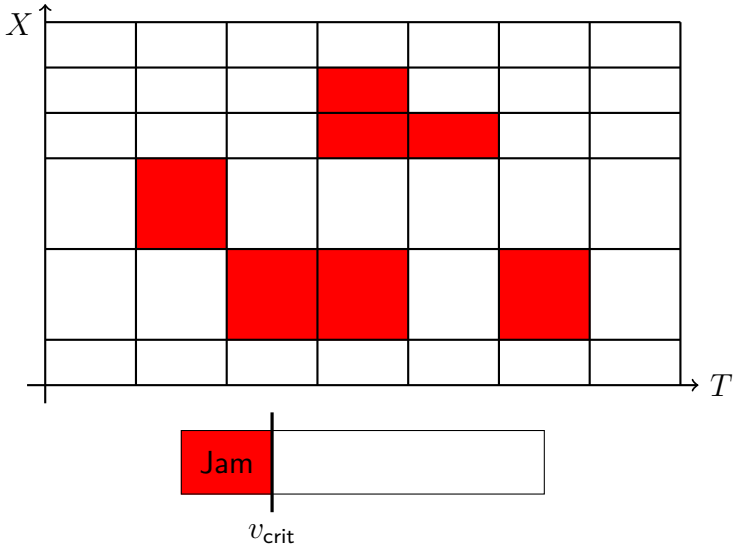


Figure 4.8: Computation of congestion clusters, part 2: Identification of congested cells

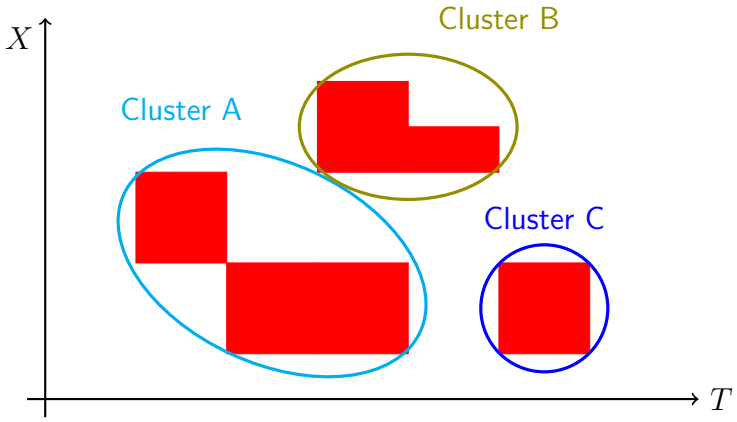


Figure 4.9: Computation of congestion clusters, part 3: Detection of congested areas

grid. In the example of Figure 4.7, seven cells remain as congested (Figure 4.8). Applying the connection relation of equations 4.7, three separate clusters remain, see Figure 4.9.

This definition of connected congestion clusters leads to many very small clusters. Furthermore, it can be observed that these clusters are frequently located close to each other, for example, if during a congested phase single cells exceed v_{crit}. In order to achieve a smoothing effect, clusters which are located close to each other are merged into one common cluster. *Close to* means that the travel time of a virtual trajectory is not larger than $t_{\text{merge}} \in \mathbb{R}_{\geq 0}$ between one cluster and the other.

To this end, all of the clusters located in $X \times T$ are iterated over, and a set of simulated trajectories is computed for each, with each trajectory starting from a different corner of a cell belonging to the current cluster using equation 4.6. This strategy is visualized in Figure 4.10 for cluster A.

The calculation of virtual trajectories is terminated as soon as time $t_S + t_{\text{merge}}$ is reached for

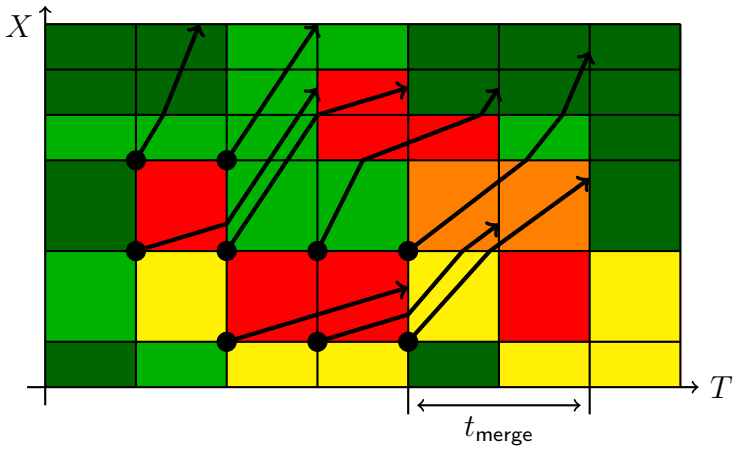


Figure 4.10: Computation of congestion clusters, part 4: Generation of trajectories (illustrated for cluster A)

any starting point t_S. All clusters touched by one of the generated trajectories are assigned to the original cluster. This means that all congested cells belonging to one of these clusters are considered to be a member of the same cluster. Figure 4.11 sketches the resulting two clusters A and C where A and B merged to A solely.

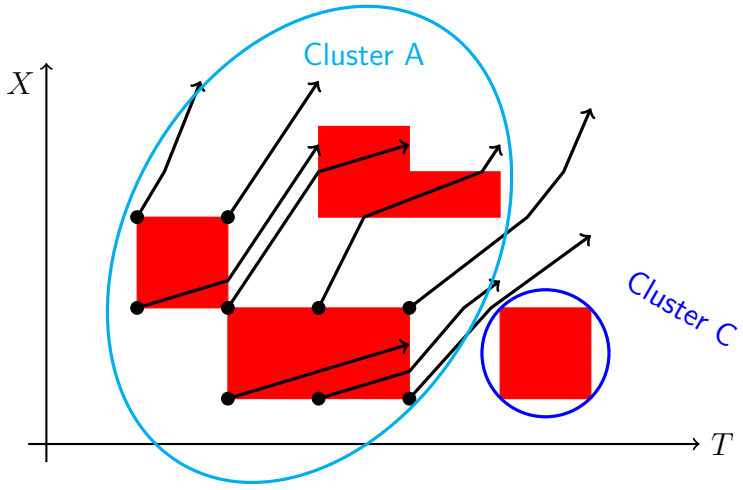


Figure 4.11: Computation of congestion clusters, part 5: Merging of clusters

The shape of the final cluster k is defined as the convex hull of all congested cells c_{ij} belonging to the identified congestion cluster (Figure 4.12). A convex hull makes the cluster's shape and moreover the cluster identification unique.

Each so computed cluster has a certain area A_k in a space-time unit. Very small clusters, that do not satisfy a minimum size of A_{min}, are eliminated. If these single cells were not merged into a more extended cluster, they are likely to be data inaccuracies, outliers, or correspond to single slow vehicles not implying congestion.

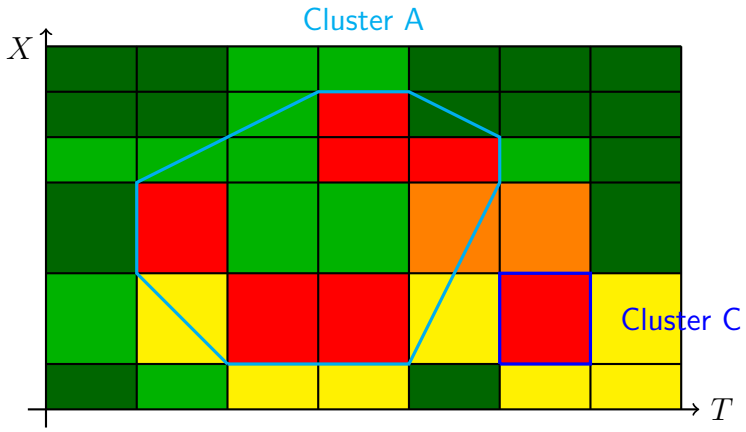


Figure 4.12: Computation of congestion clusters, part 6: Convex hull of identified cluster

4.4 Congestion Types

Considering resulting patterns of congestion clusters, different shapes, forms, and characteristics occur. This motivates to group similarly shaped clusters and to define congestion types. To optimize traffic planning and control, each type of congestion needs a different treatment to receive a more accurate GT and to predict traffic states precisely. In the following, a 4+1 congestion classification is introduced. Four of them relate to mainline, one to lane-specific traffic. Their automated assignment to a given congestion cluster is presented.

4.4.1 Mainline Congestion Types

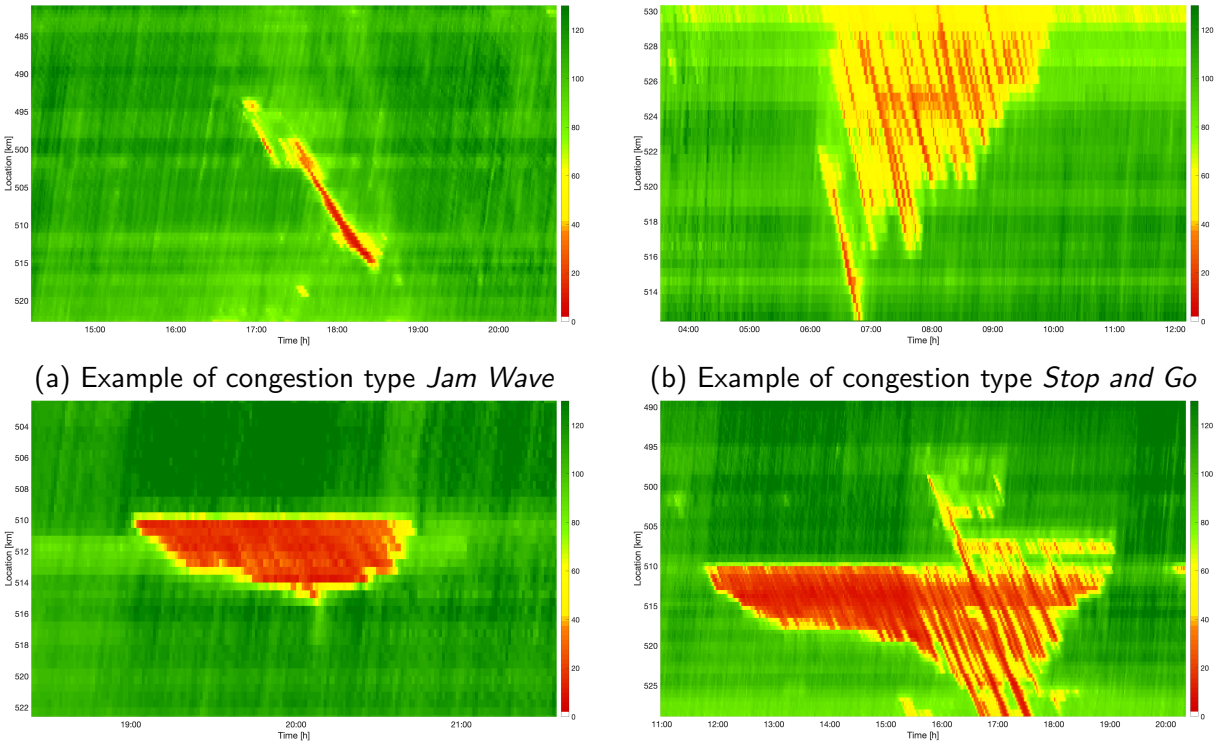
In the following, these four mainline congestion types are defined formally. Then, two methodologies are explained, first how a particular congestion type is assigned to a virtual trajectory, thereafter, how an appropriate congestion type is assigned to a cluster. Both are discussed by an example.

Definition of Four Congestion Types

A project report [Transver GmbH, 2010] introduced four congestion types which were also critically reviewed in [Bursa et al., 2018; Bursa et al., 2019]: *Jam Wave*, *Stop and Go*, *Wide Jam*, and *Mega Jam*. This definition is also used in this thesis. Table 4.1 briefly explains how the jam types are distinguished and Figure 4.13 shows examples as speed contour plots.

	Congestion Type	Description
(a)	Jam Wave	Single speed breakdown, thin stripe
(b)	Stop and Go	Several narrow stripes separated by free-flow sections
(c)	Wide Jam	Broad area with predominant congestion velocity
(d)	Mega Jam	Widely spread traffic breakdown, extensive congested area

Table 4.1: Definition of four congestion types



(a) Example of congestion type *Jam Wave*

(b) Example of congestion type *Stop and Go*

(c) Example of congestion type *Wide Jam*

(d) Example of congestion type *Mega Jam*

Figure 4.13: Examples of four congestion types

Each cluster found by the algorithm in section 4.3.2 is assigned one congestion type according to several conditions. To this end, virtual trajectories traverse the entire space-time area starting from the spatial beginning of the considered road stretch and continue with the prevailing speed in each discretized cell (equation 4.6). In contrast to the previous section, the simulated vehicles do not stop after a travel time of t_{merge} but continue until the end of the space-time domain to traverse the entire area. The resulting speed profile of each individual trajectory is analyzed according to certain parameters explained in the following.

Assignment of a Congestion Type to a Virtual Trajectory

Let $t_{\text{merge}} \in \mathbb{R}_{\geq 0}$ – analogously to section 4.3.2 – be defined as the minimum time spent in free-flow conditions by a trajectory (meaning that its speed is at least v_{crit}). This parameter distinguishes between several individual congestion clusters and separates them. Let t_0 be the first point in time at which a trajectory turns its speed v below v_{crit} before which v was at least v_{crit} for longer than t_{merge} (t_0 as beginning of first speed breakdown). Let similarly t_1 be the last point in time at which the speed of this trajectory is below v_{crit} and directly afterwards turns to $v \geq v_{\text{crit}}$ for a duration of more than t_{merge} (t_1 as end of last speed breakdown). Summarizing, this trajectory experiences congestion between t_0 and t_1.

$$t_0 = min\{t \mid v(t) < v_{\text{crit}} \;\&\; v([t - t_{\text{merge}}, t[) \geq v_{\text{crit}}\} \tag{4.8}$$

$$t_1 = max\{t \mid v(t) < v_{\text{crit}} \;\&\; v(]t, t + t_{\text{merge}}]) \geq v_{\text{crit}}\} \tag{4.9}$$

Additionally, the following parameters need to be introduced:

Parameter	Description
v_{crit}	Velocity threshold indicating congestion
t_{merge}	Minimum time break indicating new congestion if speed in between is at least v_{crit}
$t_{JamWave}$	Maximum duration of a *Jam Wave*
$t_{MegaJam}$	Minimum duration of a *Mega Jam*
$n_{StopandGo}$	Minimum number of speed drop downs

Table 4.2: Parameter values to define congestion types of a trajectory

If the trajectory experiences a short incident ($t_1 - t_0 \leq t_{JamWave}$), the congestion type is set to *Jam Wave*. If $t_1 - t_0$ is greater than $t_{MegaJam}$, it is a *Mega Jam*. Otherwise, the number of speed drops below v_{crit} between t_0 and t_1 is calculated. If there are fewer than $n_{StopandGo}$ speed breakdowns, the congestion type is a *Wide Jam* otherwise *Stop and Go* traffic. Figure 4.14 illustrates the differentiation for one virtual trajectory experiencing congestion.

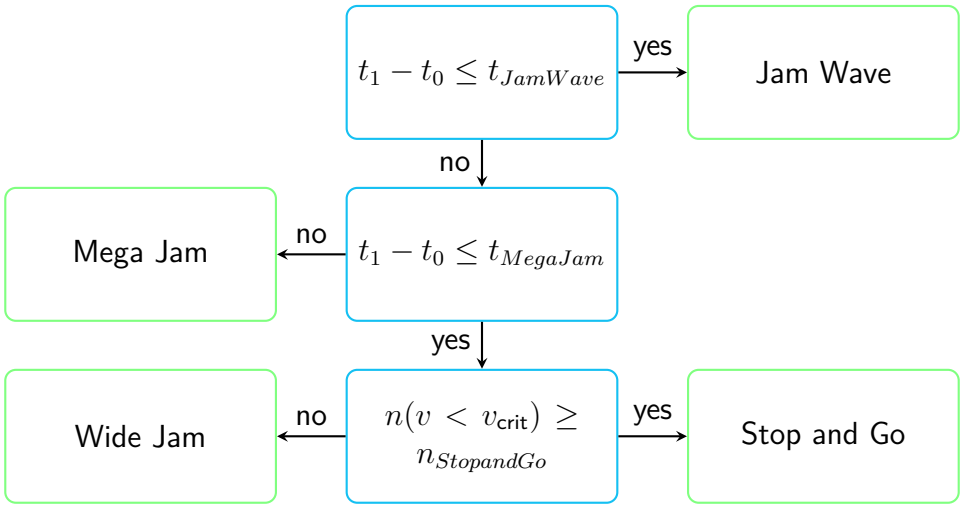


Figure 4.14: Distinction of congestion types for individual trajectories

Assignment of a Congestion Type to a Cluster

A congestion type is not only determined per trajectory but also per cluster. To this end, the clusters defined in section 4.3.2 are isolated and embedded in an environment of free-flow conditions (Figure 4.15).

Several trajectories cross the congested region one after another. The first trajectory starts at time t_S where t_S is the temporal start point of the considered space-time domain. The other trajectories start uniformly distributed with a time offset t_r. The f-th trajectory ($f \geq 0$) starts at time $t_S + f \cdot t_r$. Each trajectory undergoing the critical speed v_{crit} is analyzed, all others are ignored. The speed profile of each analyzed trajectory determines its congestion

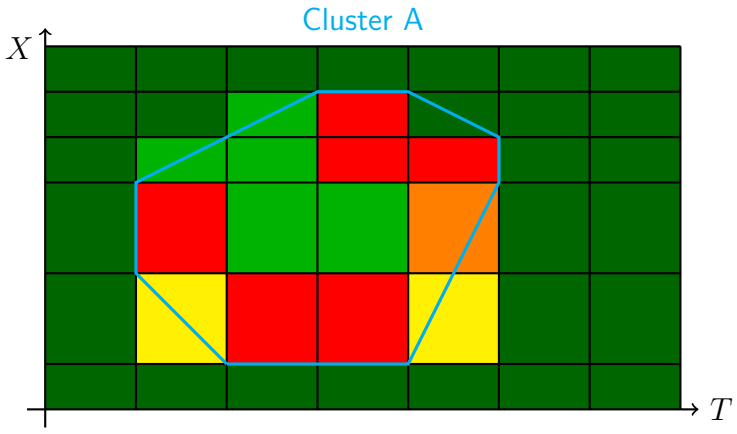


Figure 4.15: Computation of congestion types, part 1: Isolation of congestion cluster

type according to the flow chart in Figure 4.14. Figure 4.16 shows the virtual trajectories inside the cluster: those which undergo v_{crit} (i.e., intersect the congestion cluster) in solid lines, all others dotted. Each trajectory experiences a different speed profile. The profile of the bold trajectory ($f = 2$) is sketched in Figure 4.17.

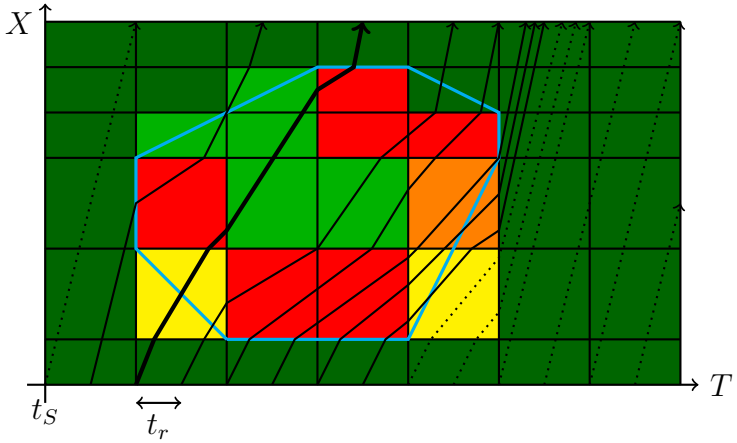


Figure 4.16: Computation of congestion types, part 2: Creation of virtual trajectories inside the cluster

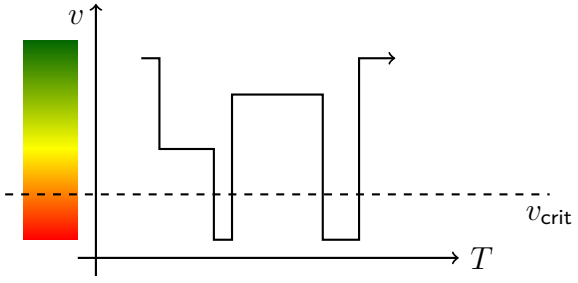


Figure 4.17: Computation of congestion types, part 3: Analysis of the trajectory's speed profile

The congestion type of a cluster is determined basically driven by a frequency-based assignment. Let $B = \{b \mid v(b) < v_{\text{crit}}\}$ be the set of all trajectories b for cluster k undergoing v_{crit} and therefore to be considered. One congestion type ct is assigned to each $b \in B$, denoted as $ct(b)$. Let further $n_{\text{2types}} > 0.5$ and $n_{\text{3types}} > 0.33$ denote the minimum percentage of assigned congestion types if two or three different types are classified among all trajectories in B, respectively (see Table 4.3).

Parameter	Description
t_r	Temporal offset of starting trajectories
n_{2types}	Minimum share of congestion types if two different types occur in k
n_{3types}	Minimum share of congestion types if three different types occur in k

Table 4.3: Parameter values to define the congestion type of a cluster k

Let a function $unique(x)$ for any array x return the same values without repetitions. Let $|x|$ denote the cardinality of an array x. The methodology of the congestion type assignment works with the differentiation as shown in Figure 4.18.

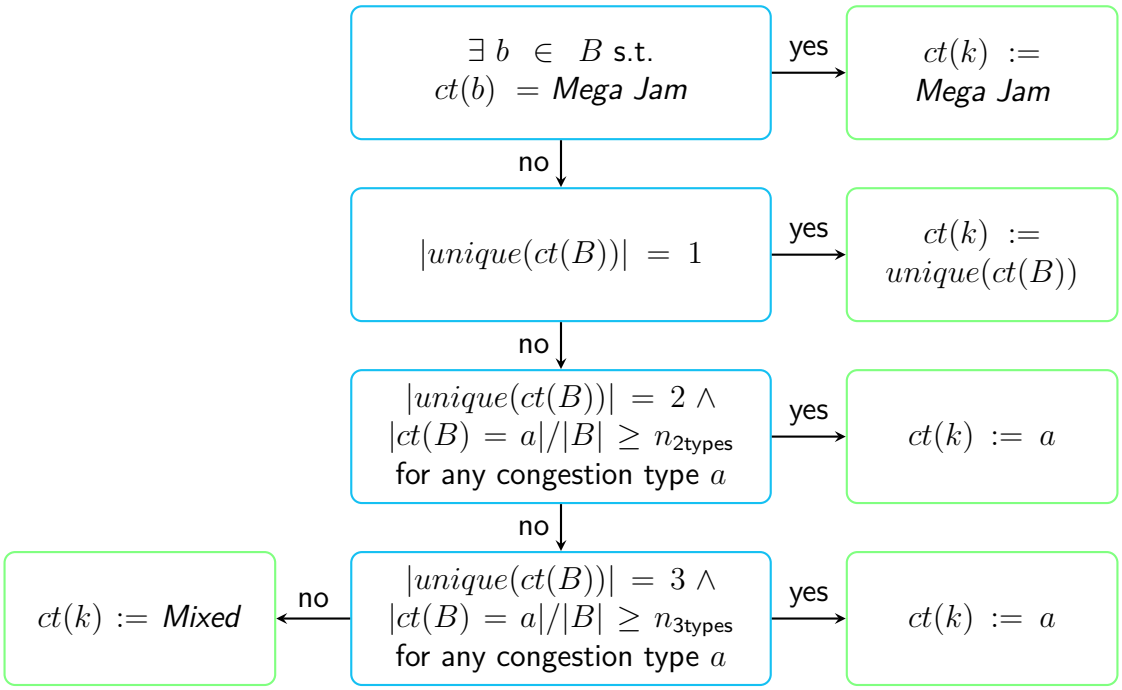


Figure 4.18: Distinction of congestion types for clusters

Example

The following figures discuss an example. The explained procedure is applied to data derived from gantry sensors on the Southbound direction (SB) from April 13, 2019 with an appropriate

parameter set (confer chapter 5 for application and evaluations). The speed values are interpolated using the ASM. Figure 4.19 shows the spatio-temporal area with congested regions and the identified separated clusters in blue. In each cluster, the top most cells appear as to be not part of the cluster. This is due to the graphical data representation structure. Indeed, the computation considers all congested cells.

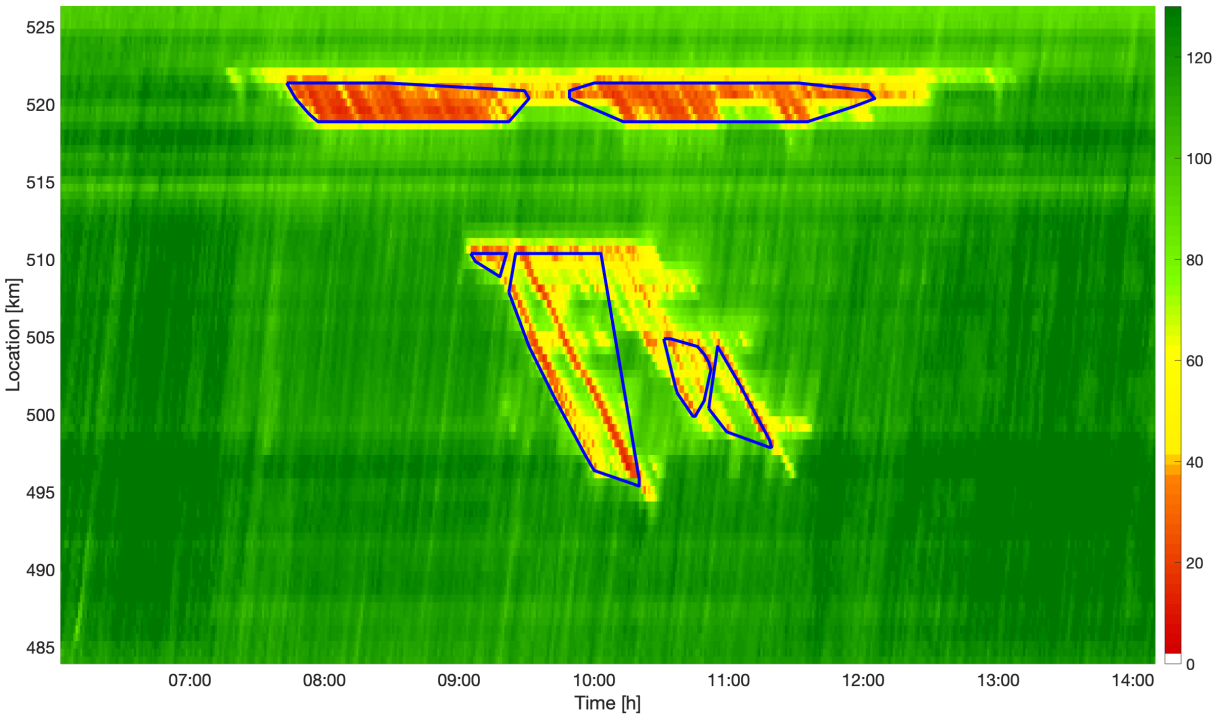
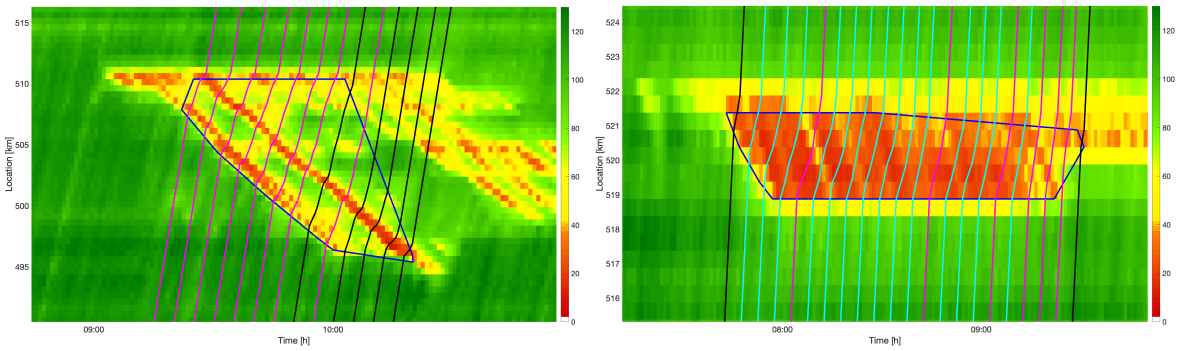


Figure 4.19: Identification of clusters; speeds in [km/h]



(a) One cluster with 2 types: *Stop and Go* (magenta), *Jam Wave* (black)

(b) One cluster with 3 types: *Wide Jam* (cyan), *Stop and Go* (magenta), *Jam Wave* (black)

Figure 4.20: Virtual trajectories crossing identified clusters; speeds in [km/h]

Virtual trajectories cross the congestion clusters. In total, 13 trajectories pass the cluster in Figure 4.20a whereof 8 are assigned to *Stop and Go* (magenta), 5 assigned to *Jam Wave* (black). The share of *Stop and Go* trajectories amounts to 62 %, therefore, the overall

cluster type is set to *Stop and Go*. In Figure 4.20b, another congestion cluster is crossed by several *Wide Jam* trajectories (cyan). Just at the boundaries, two trajectories are classified as *Jam Wave* (black). They touch the congested regime barely. Six trajectories (magenta) are assigned to the type *Stop and Go* since their speed trends drop below v_{crit} several times. In total, this cluster's congestion type is set to *Wide Jam*.

The overall congestion type classification of all clusters is depicted in Figure 4.21. It results in three *Jam Wave*, two *Stop and Go*, and one *Wide Jam* clusters.

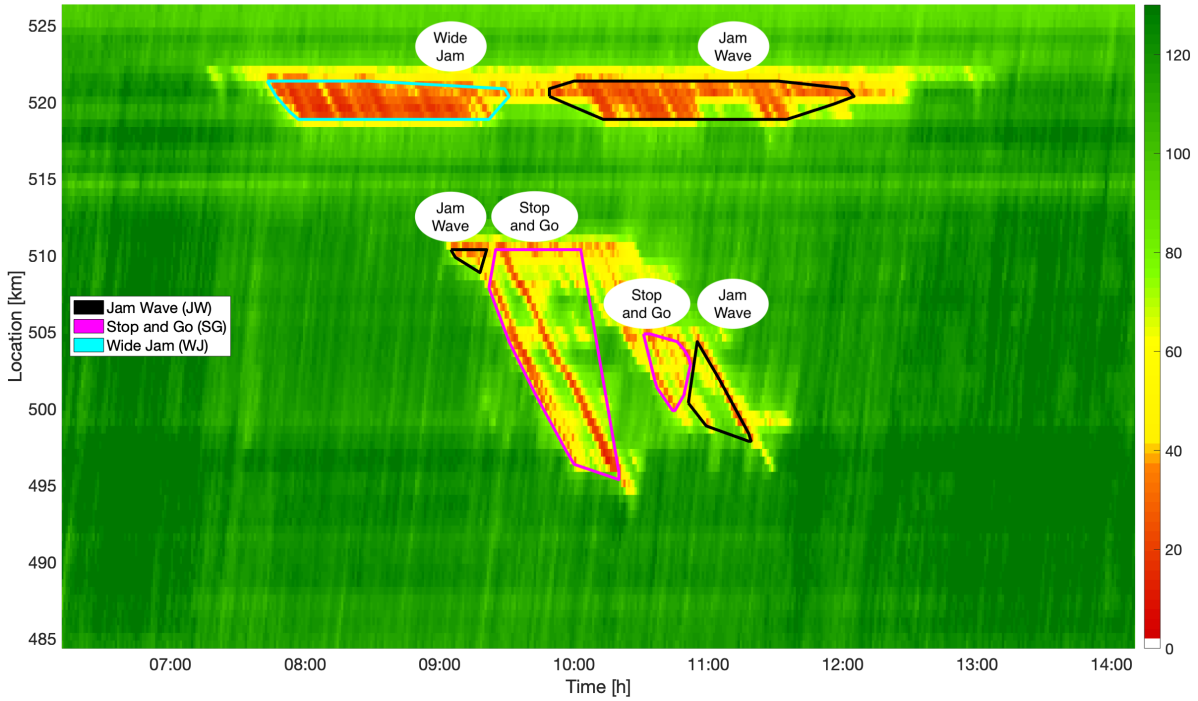


Figure 4.21: Assignment of congestion types

4.4.2 Lane-Specific Congestion

Besides the identified four congestion types, jams in which speed propagation behaves differently on each lane can be experienced. This lane-dependent congestion propagation can affect only a subset of all lanes. For example, upstream of a congested interchange or exit ramp, vehicles could be changing their lanes to exit respectively to continue on the freeway. If the spillback from congestion on the crossing freeway reaches approaching vehicles on the main freeway, the speeds of the lanes can oscillate individually. Besides spillback, another cause of this congestion type is heavy road grades. Significantly different speeds prevail on truck and car lanes while driving uphill with a high share of heavy-duty vehicles [Bursa et al., 2018; Bursa et al., 2019].

Naturally, this lane-specific congestion *Lane Jam* can only be detected if the sensor technology on a road stretch is capable of measuring speeds on each lane individually (e.g. induction loops, radars, cameras). A systematic derivation of other causes than those previously mentioned and the determination of *Lane Jam* by non-lane-based detection are left for future research.

Definition of *Lane Jam*

On a spatio-temporal field $X \times T$ let speed matrices of lanes $l_1, \ldots, l_u, u \geq 2$ exist. For each lane, the congestion cluster algorithm (section 4.3.2) identifies congested regimes. Let K_1 be the set of found clusters on lane 1, K_2 the set of clusters on lane 2, and so forth. $|K_i|$ denotes the number of clusters on lane $i \in \{1, \ldots, u\}$. To find matching pairs of congestion over lanes, the proposed methodology simultaneously iterates over all clusters $k_1 \in K_1$, all clusters $k_2 \in K_2$, and so on. For each of these combinations $k_1, k_2, \ldots, k_u$, the algorithm differentiates between the following cases:

(a) Clusters of all lanes have a nonempty intersection, i.e. $k_1 \cap k_2 \cap \cdots \cap k_u \neq \emptyset$, indicating occurrence of congestion on all lanes simultaneously

(b) Clusters of a subset of all lanes have a nonempty intersection while the remaining lanes do not have congested cells, i.e. $\exists z \in \{2, \ldots, u-1\} : k_1 \cap \cdots \cap k_z \neq \emptyset \wedge \bigcup_{i=1}^{z} k_i \cap \bigcup_{i=z+1}^{u} k_i = \emptyset$

(c) The intersection between each two sources is empty, i.e. $k_i \cap k_j = \emptyset \ \forall i, j \in \{1, \ldots, u\}, i \neq j$

Let a *region* R be defined as the union of all intersecting clusters over the lanes. In case (a), $R = k_1 \cup \cdots \cup k_u$; in case (b), $R = k_1 \cup \cdots \cup k_z$. It may happen that more than one cluster from a lane is included in R if a cluster $k_1 \in K_1$ intersects both k_{2a} and $k_{2b} \in K_2$. Then, $R = k_1 \cup k_{2a} \cup k_{2b} \cup \ldots$.

In case (c), if none of the identified clusters have any intersection, the algorithm checks if any of $K_2 \backslash \{k_2\}$ or ... or any of $K_u \backslash \{k_u\}$ has a nonempty intersection with k_1 (such that k_1 matches any other cluster to be in cases (a) or (b)). If not, k_1 solely is a *region* itself affecting only one lane. Analogously, the algorithm proceeds with all other lanes $2, \ldots, u$ and their clusters.

The result is a set of regions, each affecting a (sub)set of lanes $1, \ldots, u$. Each region R has a particular shape (not necessarily a convex hull since it is the union of several convex hulls) and a size $A(R)$ of its area. The percentage of lane i on region R is denoted by $p_i(R) \in [0, 1]$.

$$p_i(R) = \frac{A\Big(\bigcup_{j \in K_i, k_j \cap R \neq \emptyset} k_j\Big)}{A(R)} \tag{4.10}$$

Regions of which the size is smaller than A_{min} are omitted from the analyses. These clusters correspond to very short breakdowns on a subset of the lanes and therefore are not relevant.

To detect congestion on a subset of all lanes only, each found region needs to fulfill the following requirement:

$$\exists i \in \{1, \ldots, u\} : p_i(R) \leq p_{max} \tag{4.11}$$

where i runs through the subset of all lanes (not only affected ones). If congestion was not detected on lane i, then $p_i(R) = 0$ and condition 4.11 is fulfilled. p_{max} corresponds to a maximum share of the region per lane. If the intersection and the union of all affected clusters

in R are relatively equal (i.e. $p_i(R) > p_{max}\ \forall i$), this region describes congestion but no *Lane Jam*. If the intersection and the union across the lanes differ, which means that at least one lane i exists that has a share $p_i(R)$ of at most p_{max}, then it is said to be a *Lane Jam*.

The analyses over all days, all lanes, and all clusters include an average share of congestion per lane indicating which lane is frequently concerned by *Lane Jam* and by congestion at all.

Example

Figure 4.22 shows an example of *Lane Jam* from June 27, 2019 in SB direction. In Figure 4.22a, the identified convex hulls per lane are illustrated: On the right lane (magenta), one congestion cluster is found. On the middle lane (cyan), three separated clusters are found. On the left lane (black), two clusters are identified. The union of all these congestion clusters is a region. Figure 4.22b shows cells with speeds below v_{crit}. *Stop and Go* waves are clearly recognizable on the right lane over a period of more than four hours. Most of the time, the right lane only is affected. In the second half of the congestion, phases exist where only the middle and the left lane (green) are affected.

Least Travel Time

Average travel times in the regions per lanes are evaluated. Using the concept of virtual trajectories (section 4.3.1), each region R is embedded in a free-flow environment, similarly to the identification of mainline congestion types in section 4.4.1. Virtual trajectories starting at the same time and place traverse the speed matrix per lane and then, their travel times are compared. Free-flow is assumed to be driven in two areas: (i) everywhere outside of R, and (ii) inside R but outside of any cluster belonging to lane i. Inside R intersected with any cluster of lane i, the trajectory experiences real speeds of lane i.

The analyses over all days, all lanes, and all clusters include which lane is most efficient when traveling through congestion in general.

Primary Lane Congestion Detection

Each region R is assigned a first temporal and a first spatial occurrence. This corresponds to the time and the location of the earliest and the farthest downstream point, respectively. These points must always be vertices of one lane-specific cluster $k_1, \ldots, k_z$ with $z \in \{1, \ldots, u\}$ due to the principle of optimality. Hence, without loss of generality, a region can be restricted to all vertices of affected clusters to determine lane, time, and location where congestion is detected earliest.

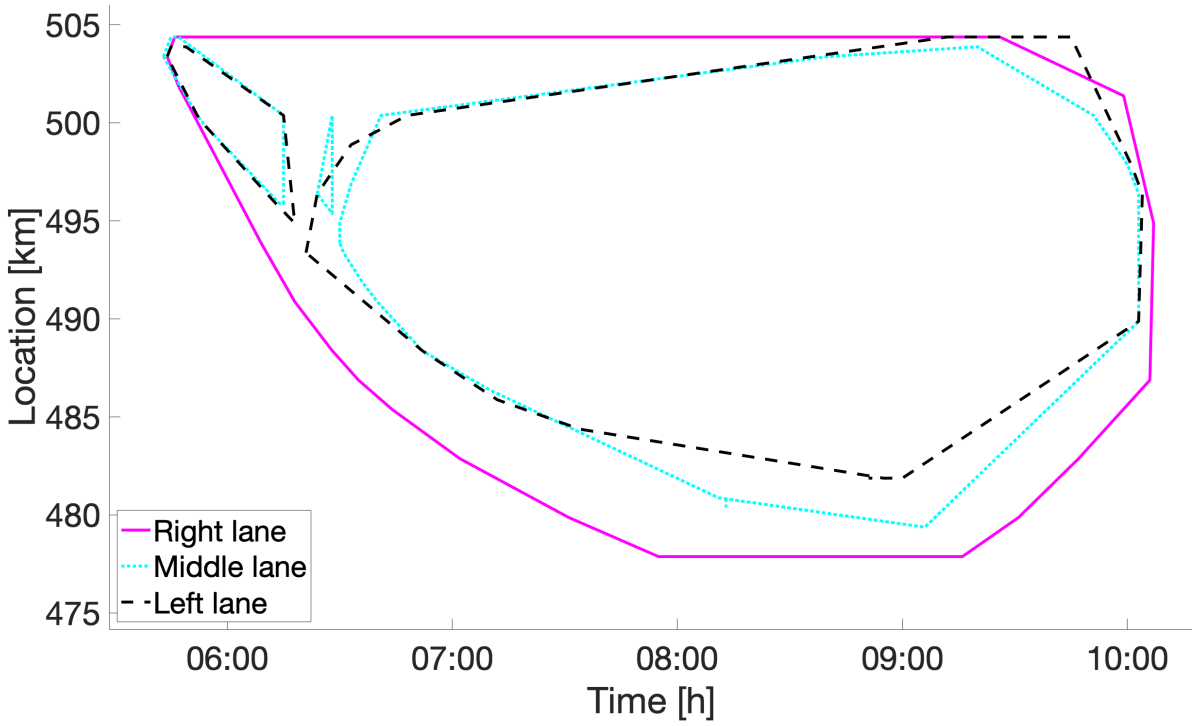
When lane j detected congestion first, the algorithm determines the temporal delay of the other lanes by simply computing the time difference

$$t_{\text{diff}} = t_{min}(k_i) - t_{min}(k_j) \tag{4.12}$$

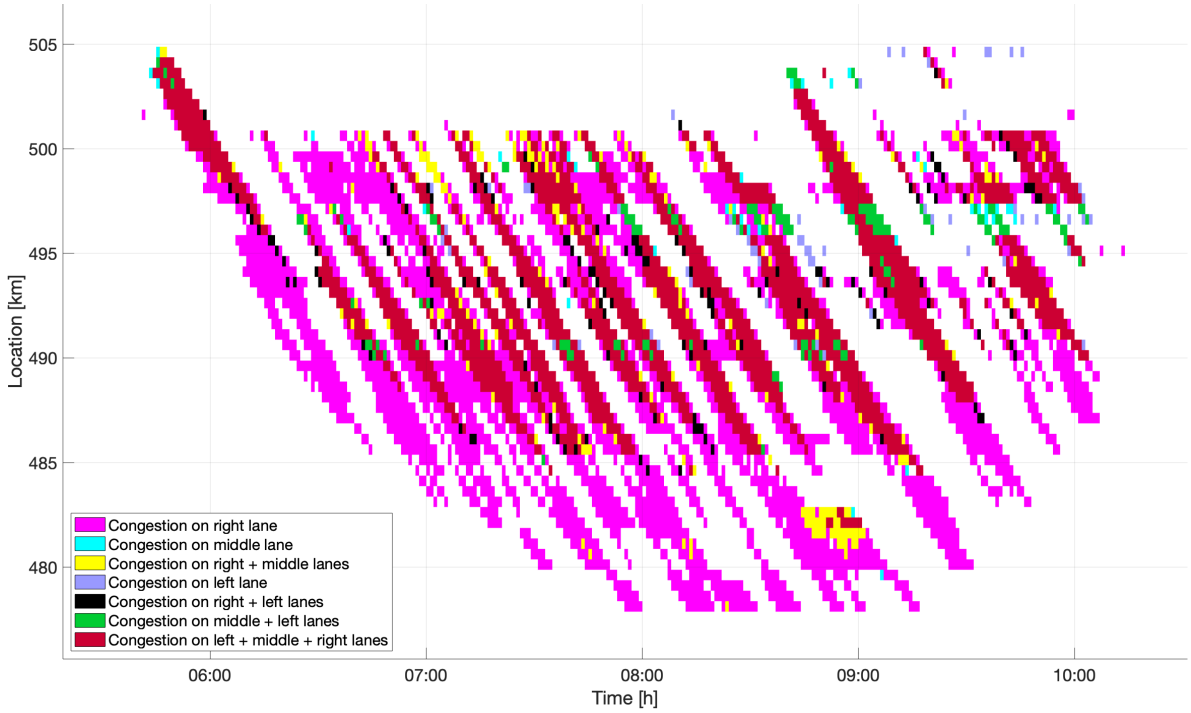
for all lanes i affected by the identical congestion, hence, in the same region.

Analogously, the spatial offset between the recognizing lane j and all affected lanes i can be computed as difference

$$x_{\text{diff}} = x_{max}(k_i) - x_{max}(k_j) \tag{4.13}$$



(a) Convex hulls in a region



(b) Congested cells in a region

Figure 4.22: Example of *Lane Jam*, detected on June 27, 2019 in SB direction

where x_{max} is the farthest location affected by congestion of region R.

The analyses over all days, all lanes, and all clusters include statistics which lane is able to detect *Lane Jam* first and which delay and which offset the other lanes experience on average.

Hot Spots

Lastly, *Lane Jam* hot spots are investigated. Each *Lane Jam* event is restricted to its first spatial and temporal occurrence, independently of its overall extensions, such that it can be interpreted as single dot in the space-time domain. By analyzing several days, all occurrences of *Lane Jam* types can be compared and lead to lane-specific congestion hot spots. All regions identified as congestion type *Lane Jam* are considered, independently if a *Lane Jam* varies its state and transforms to a mainline congestion type. Each region with its dedicated start and end points both in time and space, together with its size of the area are computed. The methodology then shows the spatio-temporal hot spots of *Lane Jam* occurrences, together with the size of the regions. Additionally, it gives a distribution of hot spots over the days of week.

4.5 Primary Congestion Detection

This section describes an approach to identify which detection technology is able to recognize congestion earliest. Speed data derived from local measurements are available at sparse positions. However, they usually have a high temporal accuracy since no data processing is necessary as it is for travel time measurements.

FCD have a higher spatial resolution with very short segment lengths but need to be aggregated online to obtain the full speed knowledge. BT measurements have to await their respective pairing before a travel time value can be inverted into a speed value. From the spatially low-resolution sections, any congestion cannot be localized uniquely (both in space and time).

This motivates the question which advantages are more dominant, temporally or spatially high accuracy. The ability to detect emerging traffic congestion promptly is essential for the provision of advanced congestion warnings to drivers. In the following, an approach is presented which identifies clusters per data source, matches them and determines the first occurrence of congestion. Additionally, it computes the delay of the other sources after congestion was detected by the first data set.

The algorithm works in three steps (Figure 4.23). First, single congestion clusters are identified, separately for each data source to be compared, based on the methodology presented in section 4.3.2. In a second step, all computed congestion clusters across different data sources are associated with each other. The third step comprises the comparison between the temporal starts of the resulting pairs of congestion clusters. Hence, it can be decided based on which data source a congestion warning could have been provided earlier in the respective situation.

As in previous sections, let speed data from different sources be available for a road corridor X and a time period T. Again, let $X \times T$ be separated into cells $x_i \times t_j$. $v(c_{ij}) = v_{ij}$ denotes the speed in cell $c_{ij} = x_i \times t_j$. Cells of interest are only those that contain a speed value below the speed threshold v_{crit}.

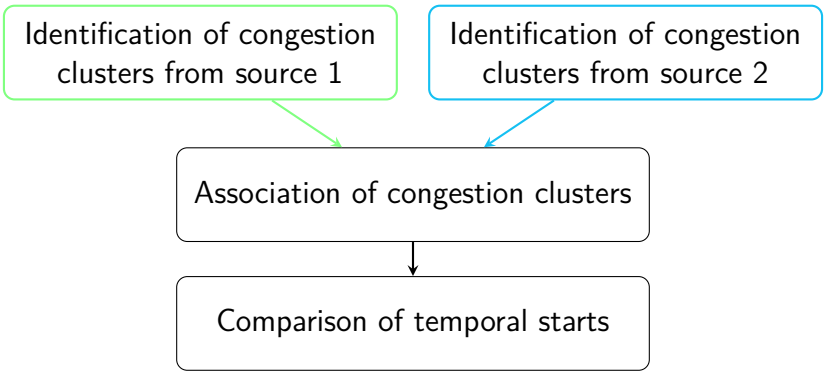


Figure 4.23: Methodological concept of primary congestion detection

One aspect in this context is that local speed measurements should not be restricted to detector positions. If one of the data sets contains data derived from fixed local sensors, instead, interpolated speed distributions (section 2.1) enables finding coherent congestion clusters. For the analysis of the first temporal occurrence of congestion (step 3 of the algorithm), the original speed measurements from the detection positions solely are taken.

If the grids $\{x_i\}$ or $\{t_j\}$ differ between the data sources, a preprocessing algorithm sets the respective grid to the union of the grid of the separate data sets. For example, if the spatial grid of source 1 comprises segments of lengths 113 m and 98 m and the spatial grid of source 2 comprises equidistant segments with lengths 100 m, then the union contains segment lengths of 100 m, 13 m, 87 m, and 11 m. The first speed value of source 1 holds in x_1 and x_2, the second one in x_3 and x_4. The first speed value of source 2 is valid in x_1 only and the second speed value keeps constant in segments x_2 and x_3 before it changes for x_4. The common grids are denoted as $\{x_i\}$ and $\{t_j\}$ after the preprocessing union.

4.5.1 Identification of Congestion Clusters

As a first step, the algorithm presented in section 4.3.2 – identifying congestion clusters – is applied to each data set separately. The input is spatio-temporally discretized speed matrices, the output is sets of congestion clusters per data source, respectively. Each congestion cluster k has defined spatial and temporal start and end points. In the following, the approach is formulated to compare two data sources for the ease of reading. Similarly, more data sets can be compared.

4.5.2 Association of Congestion Clusters

In order to match congestion clusters from source 1 and source 2, the algorithm iterates over the set of all congestion clusters derived from source 1. All source 2 congestion clusters, which show any intersection with a certain source 1 congestion cluster, are assigned to the source 1 cluster. The source 2 cluster with the earliest beginning is chosen for the match. However, to reduce the risk of incorrectly matched congestion clusters, the identified pairs of clusters need to fulfill additional conditions. Let $C_1 \subseteq X \times T$ denote the spatio-temporal area of a source 1

cluster, $C_2 \subseteq X \times T$ of a source 2 cluster. Furthermore, let $|A|$ denote the size of an area $A \subseteq X \times T$.

Two clusters C_1 and C_2 can only be assigned to each other if they fulfill the following three conditions:

$$\frac{|\, C_1 \cap C_2 \,|}{|\, C_1 \,|} \geq P_{min} \tag{4.14}$$

$$\frac{|\, C_1 \cap C_2 \,|}{|\, C_2 \,|} \geq P_{min} \tag{4.15}$$

$$|\, C_1 \cap C_2 \,| \geq A_{min} \tag{4.16}$$

Hereby, $P_{min} \in [0, 1]$ denotes the minimum percentage of the area of both C_1 and C_2 that has to be covered by both clusters. $A_{min} \in \mathbb{R}_{\geq 0}$ denotes the minimum area that needs to be covered by the intersection of both clusters. For appropriately chosen parameters, conditions (4.14) and (4.15) ensure that significant parts of the clusters are congested according to both data sources, condition (4.16) ensures that small clusters are ignored. The result is a set of isolated and coherent associated clusters throughout the entire speed matrix.

Different data sources yield jam areas with a different level of detail. For example, interpolated speed data result in more coherent clusters due to the smoothing (fewer clusters in total), FCD in more separate small clusters due to direct measurements (more clusters in total). Therefore, it can happen that one congestion cluster of source 1 is matched to several separate (not connected or merged) congestion clusters of source 2. The algorithm preserves the different number of jam clusters in an $1 : m$ matching, since otherwise the conditions (4.14) – (4.16) would not hold, specifically conditions (4.14) and (4.15) would yield wrong first timestamps.

Each association $a_{1,2} = \{C_1, C_2\}$ is defined as set of two matched congestion clusters C_1 and C_2 from data sources 1 and 2, respectively. In case of the $1 : m$ matching, C_2 itself consists of a set of separate clusters. If data sets from more than two detection technologies are compared, the association of congestion clusters is repeated for another pair of data sources.

4.5.3 Comparison of Temporal Starts

As third step of the algorithm, the temporal starts of the resulting pairs of congestion clusters are compared. In case that one data set was interpolated out of local speed measurements, the congested times from the original detector positions $x_1, \ldots, x_D$ are taken to compare the first congested timestamps instead of the smoothed speed values, again with the threshold v_{crit}.

To this end, all touched detector positions $x_{i_1}, \ldots, x_{i_2}$ in the spatial area of one found interpolated congestion cluster are considered and from those, the first timestamp is taken which fulfills $v(x) < v_{\text{crit}}$ for $x = x_{i_1}, \ldots, x_{i_2}$ and lies in the spatio-temporal area of cluster C. The detector position $x_k \in \{x_{i_1}, \ldots, x_{i_2}\}$ with the least first congested time is taken as location for the primary congestion analysis with the corresponding first temporal occurrence $t_{min}(x_k)$. This represents the spatio-temporal start of the found congestion cluster in case of Stationary Detector Data (SDD).

Let t_{min} be the earliest time at which congestion occurred and a warning could have been sent.

$$t_{min}(C_1) := \min\{t \in T \mid t \in C_1\} \tag{4.17}$$
$$t_{min}(C_2) := \min\{t \in T \mid t \in C_2\} \tag{4.18}$$

Let furthermore be

$$t_{\text{diff}} := t_{min}(C_1) - t_{min}(C_2) \tag{4.19}$$

the difference measure between the first congestion occurrences of both detected clusters. In case $t_{\text{diff}} = 0$, both sources detected the associated congested regions at the same time, $t_{\text{diff}} < 0$ means that source 1 detected congestion earlier. Analogously, $t_{\text{diff}} > 0$ implies that source 2 registered congestion first where source 1 shows a delay of t_{diff}.

The described approach presents one strategy to determine the data source based on which a congestion warning could have been triggered earliest. The result also depends on the choice of parameters. However, the findings of the computational analysis indicate that different parameter sets only have little influence on the conducted comparison, see section 5.3.

4.6 Congestion Type Hot Spots

This section describes the determination of congestion type hot spots. Locations and times with an accumulated occurrence of individual congestion types are identified and analyzed.

The existence of different types of congestion prompts a need for a methodology for their differentiation, as traffic information and control measures should be adjusted based on the type of congestion. This research focuses on the differentiation of several congestion types and proposes a methodology to automatically determine their respective accumulated occurrences. An algorithm is presented which determines – again for identified coherent congestion clusters with their assigned congestion types – the temporal and spatial occurrences of the respective congestion start and end points. While considering spatio-temporal areas of several similar time frames (e.g. days), congestion hot spots can be investigated.

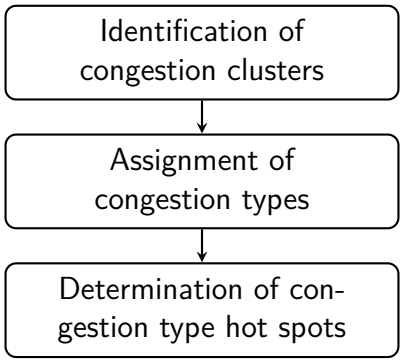


Figure 4.24: Methodological concept of hot spot detection

The algorithm works in three steps (Figure 4.24). First, coherent congested clusters are isolated in a space-time speed distribution. Thereafter, one congestion type is assigned to

each cluster. As a third step, the location and time of the emerged congestion with its assigned congestion type and its size are determined. Hence, it can be decided for example at which locations detectors should be condensed or at which times traffic control should be intensified.

Again, the basis is a spatio-temporally discretized speed matrix v with cells $c_{ij} \in X \times T$. Similarly to the previous sections, only cells with speeds below v_{crit} are interesting. Unlike the algorithm presented in section 4.5, different data sources are not compared directly. Therefore, it is not necessary to union the grids $\{x_i\}$ and $\{t_j\}$ to consistent grids. Each data set is considered separately, and the congestion type hot spots are identified individually.

4.6.1 Identification of Congestion Clusters and Assignment of Congestion Types

The first step is again the application of the congestion cluster identification and isolation algorithm presented in section 4.3.2.

As a second step, to each of the found clusters, a congestion type is assigned (section 4.4.1). It is also possible to consider locations with non-continuously available data, for example, if several segments on a road are not equipped with traffic sensors. Clusters can be identified reliably if locations without any speed information are filled up with a free-flow speed value because also the convex hull of the identified cluster is embedded in a neighborhood of free-flow conditions. The described algorithm can then find clusters in separate location sets.

4.6.2 Determination of Congestion Type Hot Spots

In the third step, the algorithm calculates the congestion type ct_k and the size A_k as well as the locations and the timestamps of the start (x_{k_S}, t_{k_S}) and the end (x_{k_E}, t_{k_E}) of each congestion cluster k. An analysis of several similar spatio-temporal areas leads to the identification of hot spots. For example, the evaluation of several days of a month can be conducted and results in spatial and temporal congestion type accumulations in a certain time period. The result is a distribution of the congestion types over the considered road stretch throughout the given times.

Takeaways from Chapter 4: Extended Congestion Definition and Detection Strategies

Developed Fusion Approaches

Low-Resolution Travel Time Smoothing Method (LTSM)	Fusion method to consistently reconstruct section-based travel time measurements to speed distribution
Multi-sensor data fusion	Superimposition of multiple data sources, each discretized in a space-time domain, to reconstruct a common speed distribution

Basic Strategies for Congestion Detection and Classification

Virtual trajectories	Simulation of vehicles through a spatio-temporally discretized speed matrix to obtain speed profiles and travel times
Congestion clusters	Identification of coherent regions of congested cells as convex hulls with certain areas
Mainline congestion types	Classification of congestion patterns: *Jam Wave*, *Stop and Go*, *Wide Jam*, *Mega Jam*
Lane-by-lane congestion	Assessment of lane-specific congestion proneness and travel times

Algorithms and Metrics for Comparisons

Earliest congestion detection	(1) Identification of clusters per data source (2) Association of clusters across data sets (3) Determination of first occurrences
Congestion pattern hot spots	(1) Identification of congestion clusters (2) Computation of congestion types (3) Determination of accumulated occurrences per congestion type per data source

Chapter 5

Application to Various Traffic Detection Technologies

This chapter describes the evaluation of the proposed methodologies. All developed algorithms from chapter 4 are applied to data sets described in chapter 3. Each section is organized in two parts, evaluation and conclusion.

First, the fusion algorithms are evaluated. Section 5.1 describes the qualitative results of the Low-Resolution Travel Time Smoothing Method (LTSM), a fusion method for section-based travel time measurements. It compares different parameter sets when the method is applied to data from one day and assesses the results. Multi-sensor data fusion approaches are evaluated in section 5.2. Based on a qualitative analysis involving the error measure Symmetric Square Inverse Mean Percentage Error (SSIMPE), a recommendation of a parameter set is given.

Then, in section 5.3, the results of the primary congestion detection are explained. Data sources are compared with regard to their earliest incident detection. Section 5.4 contains a discussion of the recognition ability of congestion types per detection technology. Data from the same day are evaluated per source. Section 5.5 presents the congestion type hot spot results. Per data source, focus areas with accumulated occurrences of each congestion type exist both spatially and temporally. After the consideration of mainline congestion types, lane-by-lane traffic is analyzed. Section 5.6 describes the evaluation of lane-specific congestion proneness, travel times, earliest congestion detection, and spatio-temporal hot spots. The chapter is concluded by a discussion in section 5.7. It summarizes the suitability of each detection technology to detect each congestion type.

5.1 Low-Resolution Travel Time Smoothing Method

Section 4.1 introduced an interpolation method called LTSM, which computes consistent fused speeds out of low-resolution travel time measurements. In this thesis, the methodology is approved using data gathered by Bluetooth (BT) sensors on the autobahn A9 between Munich and Nuremberg.

5.1.1 Evaluation

The proposed method is applied to data from July 26, 2018 in Northbound direction (NB). Raw travel time data are shown in Figure 4.1. In total, approximately 145,200 trips were

recorded on that day and direction after data filtering. The application involves two degrees of freedom. One is the cell size (grid discretization) of the space-time domain, the other the weighting factor α which can be chosen as distance, duration, or distance-duration weight for each trip.

For sensitivity, the evaluation is executed with three configurations of each, the cell size and weighting factor. As spatio-temporal grid discretization, two uniform (equidistant) parameter sets are chosen, a larger one and a smaller one, both in time and space. As a third variant, the cell distances are set to the non-equidistant BT sensor locations together with a uniformly distributed temporal grid. For each of the mentioned three cases, three possibilities of setting α are tested and presented in the following.

Equidistant Speed Reconstruction with Cell Size 2 km × 5 min

First, a traffic speed reconstruction with parameters (space(c_{ij}), time(c_{ij})) = (2 km, 5 min) is conducted. This yields 80 spatial intervals x_i and 289 temporal intervals t_j, in total 23,120 cells c_{ij}. The distance-weighted, the duration-weighted, and the distance-duration-weighted speed interpolations are as illustrated in Figures 5.1a – c. All three figures show a similar picture: while the duration-weighted speeds do not exceed the distance-weighted velocities, the distance-duration-weighted speed distribution averages both values and therefore gives a more exact reconstruction compared to the raw data.

Figure 5.1d plots the difference between the distance-weighted and the duration-weighted speed interpolation per cell with the speed color bar limited to 40 km/h. The mean deviation for all cells amounts to 12.0 km/h with a median of 9.4 km/h. The deviation is only positive, which means that the distance-weighted interpolated speeds are always greater than or equal to the duration-weighted speeds (instantaneous velocity in contrast to local speed). In congested regimes, the resulting speeds of both weighting factors are very similar, the difference amounts to nearly zero. In free-flow speed situations, mainly slow velocities are illustrated, there are fewer speed deviations for higher speeds. This is due to the fact that if a slow trip traverses a cell, the distance is equivalent to other trips crossing this cell but its duration is significantly increased.

Equidistant Speed Reconstruction with Cell Size 500 m × 1 min

Second, the speed distribution is reconstructed with a smaller cell size of (space(c_{ij}), time(c_{ij})) = (500 m, 1 min). This results in 315 spatial intervals x_i and 1441 temporal intervals t_j. In total, 453,915 cells c_{ij} are processed. Due to the smaller cell size, the speed contour reconstruction is more precise. Slow vehicles that were not properly filtered in advance, have only little influence. Again, the distance-weighted interpolation computes larger speed values than the duration-weighted interpolation. The distance-duration-weighted smoothing averages both results.

Compared to the larger cell size parameter set, some differences exist. For instance, in this setting, the freeway capacity decreases downstream of the bottleneck at 20:00 at km 480-490. Therefore, no probe vehicles pass this section at this point in time resulting in a white patch where no data are available. This white stretch is not recognizable in the first setting. The interpolated speeds applying the three variants of weighting factors as well as the difference

between distance-weighted and duration-weighted speeds are as illustrated in Figure 5.2. The average deviation amounts to 8.8 km/h per cell with a median of 3.6 km/h.

Non-Equidistant Speed Reconstruction with Cell Sizes Equal to Bluetooth Sensor Locations $\times$ 3 min

As a third variant, the speed reconstruction is executed with cells whose spatial cell discretization matches the locations of the BT sensors. This corresponds to a non-equidistant interval definition for X (referred to as *BT spacing*). Due to the average sensor distance of 4.9 km, the temporal cell discretization is set to 3 min which is in turn chosen to be uniformly distributed since the BT measurements detect in steps of one second. This results in an average speed of 100-120 km/h to traverse a cell c_{ij}. This discretization yields 33 spatial intervals x_i and 481 temporal intervals t_j, in sum 15,873 cells c_{ij}.

The resulting speed reconstruction is as illustrated in Figure 5.3. Single trips are not recognizable as they are in variants 1 and 2. The entire cell always appears colored constantly without any gradient of lines because the cell size is relatively rough. However, *Stop and Go* waves coming across between 13:00 and 15:00 around km 520-530 are of contrasting color, even though they move to upstream sections. The BT sensor locations are compact in these segments.

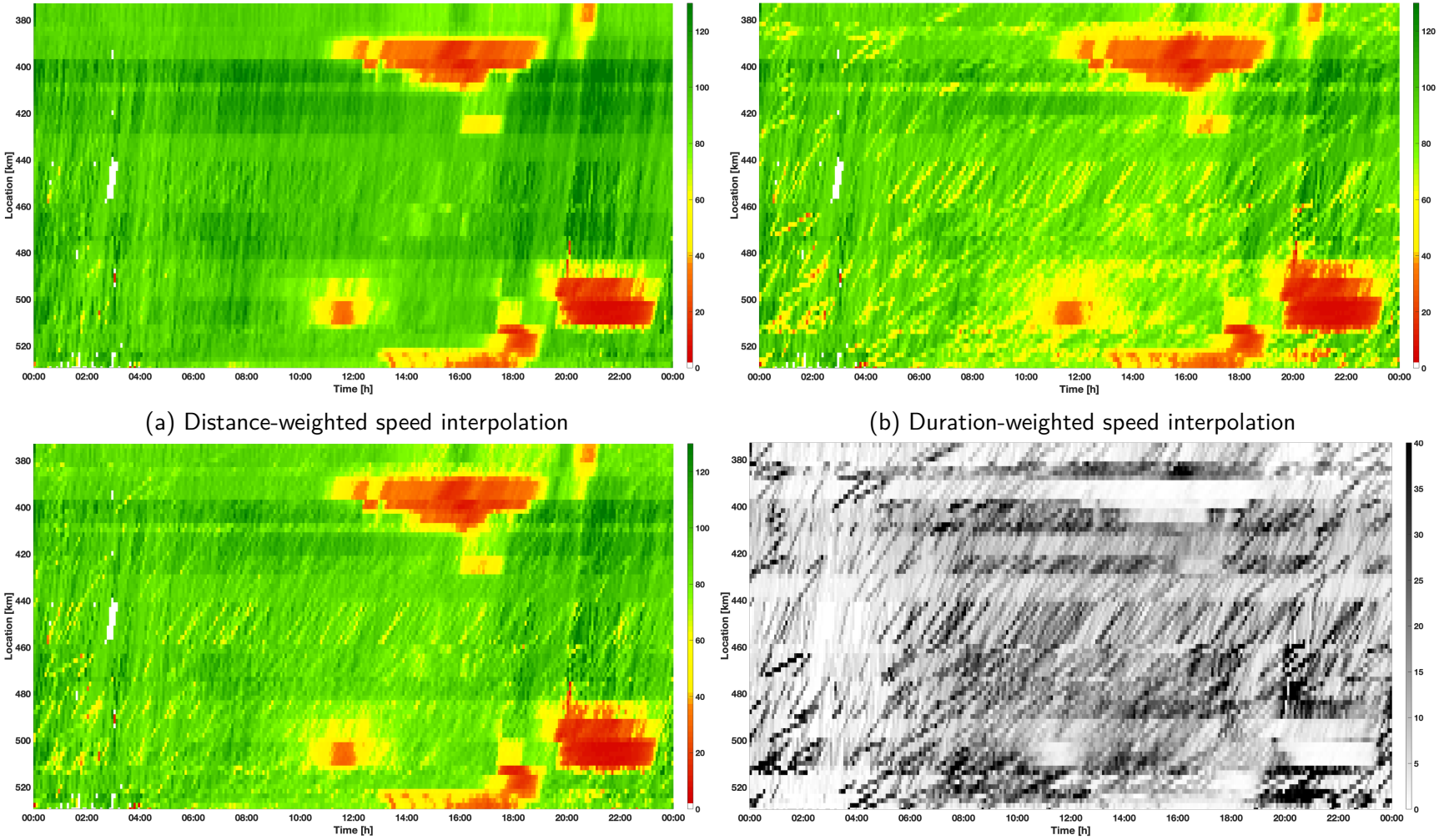
The mean deviation in Figure 5.3d amounts to 13.2 km/h per cell with a median of 10.8 km/h which makes this simple approach of a smoothing parameterization rather unlikely to be applied in practice.

5.1.2 Conclusion

The duration-weighted speed interpolation underestimates the measured raw speed data while a distance-weighted interpolation overestimates them. Therefore, it is recommended to set the interpolation weights α (equation 4.2) to the combined distance-duration measure. This corresponds to the basic understanding that the average speed of a cell in the grid can be computed by the sum of all trip (or trip share) distances divided by the sum of all trip (or trip share) times in traffic theory. In the following, specifically for the fusion of data sets in section 5.2, the distance-duration-weighted speed interpolation is used to smooth BT data.

The cell size for interpolation is set to an equidistant measure since the other data sources, especially all Floating Car Data (FCD) and gantry measurement interpolations are also discretized uniformly. Furthermore, the application of variant 3 (*BT spacing*) as parameter setting results in a more inexact speed reconstruction since the cell sizes are quite large. The deviations of the difference plot shows the highest variance for the third parameter set. In the following, the second approach, an equidistant cell discretization with cells of size 500 m $\times$ 1 min is chosen (resulting in Figure 5.2c).

Still, some red stripes occur in the smoothed speed matrix. The filter algorithm is optimized for considerations in the scope of this dissertation. For an operational use in real-time traffic control, the filtering should be adapted towards its purpose in freeway operation. This is left for future research.



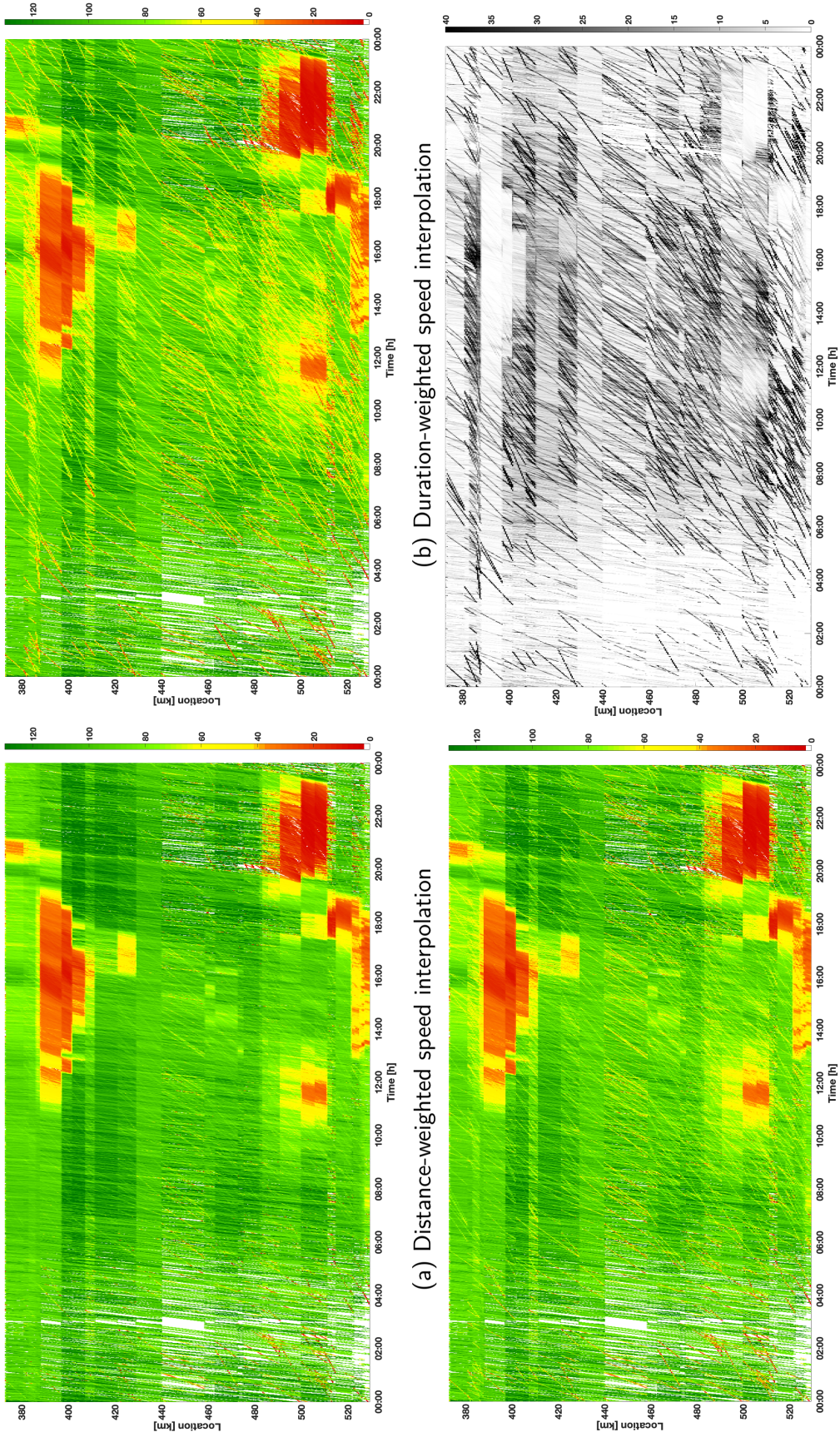
(a) Distance-weighted speed interpolation

(b) Duration-weighted speed interpolation

(c) Distance-duration-weighted speed interpolation

(d) Difference plot of Figures (a) and (b)

Figure 5.1: Average speeds for cells $2\,\mathrm{km} \times 5\,\mathrm{min}$, color corresponds to speed (in [km/h])



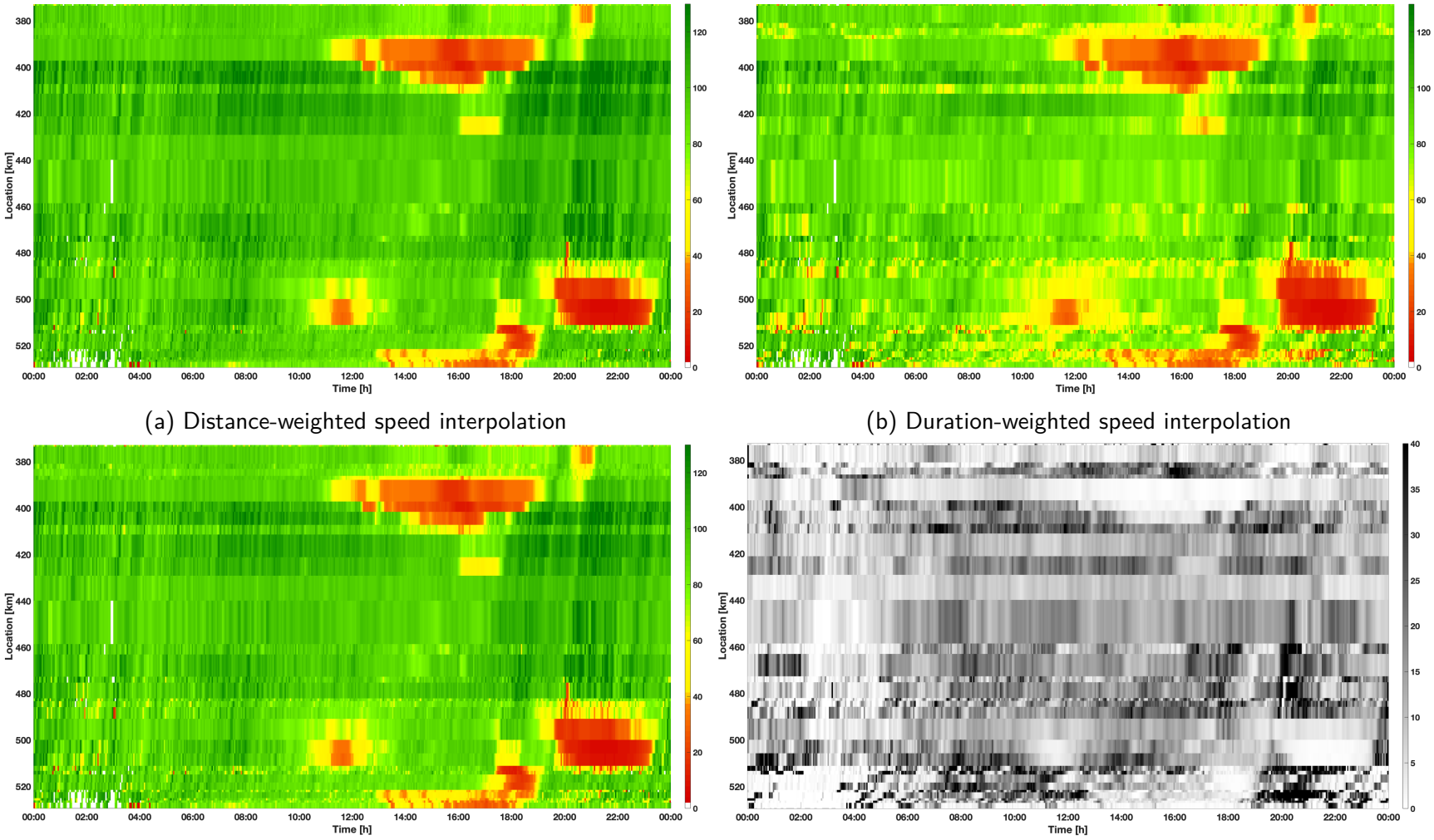
(a) Distance-weighted speed interpolation

(b) Duration-weighted speed interpolation

(c) Distance-duration-weighted speed interpolation

(d) Difference plot of Figures (a) and (b)

Figure 5.2: Average speeds for cells $500\,\mathrm{m} \times 1\,\mathrm{min}$, color corresponds to speed (in [km/h])



(a) Distance-weighted speed interpolation

(b) Duration-weighted speed interpolation

(c) Distance-duration-weighted speed interpolation

(d) Difference plot of Figures (a) and (b)

Figure 5.3: Average speeds for cells *BT spacing* $\times 3\,\mathrm{min}$, color corresponds to speed (in [km/h])

5.2 Multi-Sensor Data Fusion

Section 4.2 introduced several fusion methods for multiple data sources to obtain a superimposed optimized speed distribution. Three different fusion approaches were presented, a weighted average, a reliability-sorted cell fill-up, and an Adaptive Smoothing Method (ASM)-based interpolation. Here, these fusion approaches are applied using the data sets from gantry, radar, BT and FCD presented in section 3.1. As an example, the methodology is evaluated using data recorded on May 29, 2019 in NB direction.

First, each speed distribution is reconstructed per data set solely, which results in three spatio-temporally discretized matrices for Stationary Detector Data (SDD), BT, and FCD, respectively. These are the input to the fusion approaches. Thereafter, the ground truth is given to which resulting speed reconstructions are compared. As a comparison measure, SSIMPE is applied.

SDD, that is local speed measurements from gantry and radar sensors, are interpolated by applying the ASM approach. The parameters used are given in Table 5.1. The result is the speed matrix v^{SDD}, illustrated in Figure 5.4.

Parameter	**Value**
Spatial grid distance	500 m
Temporal grid distance	1 min
Speed in congestion	-18 km/h
Free-flow speed	80 km/h
Crossover from free to congested traffic	80 km/h
Width of the transition region	10 km/h

Table 5.1: ASM parameter values used for smoothing local speed data

BT travel times are interpolated to a speed distribution using the LTSM (section 4.1). The weighting measure is set to the distance-duration-weighted average (section 5.1). The parameters used are as shown in Table 5.2.

Parameter	**Value**
Spatial grid distance	500 m
Temporal grid distance	1 min
Weighting factor	$dist_p \cdot dur_p$

Table 5.2: LTSM parameter values used for smoothing BT travel times

The resulting speed plot is given in Figure 5.5. The jam waves around km 510-530 – easily visible in Figure 5.4 – can hardly be recognized in Figure 5.5. Due to the travel times instead of speed measurements, they cannot clearly be captured. Congestion appearing in Figure 5.4 as fragments (due to the detection location lacks) have a complete spatio-temporal extension in Figure 5.5.

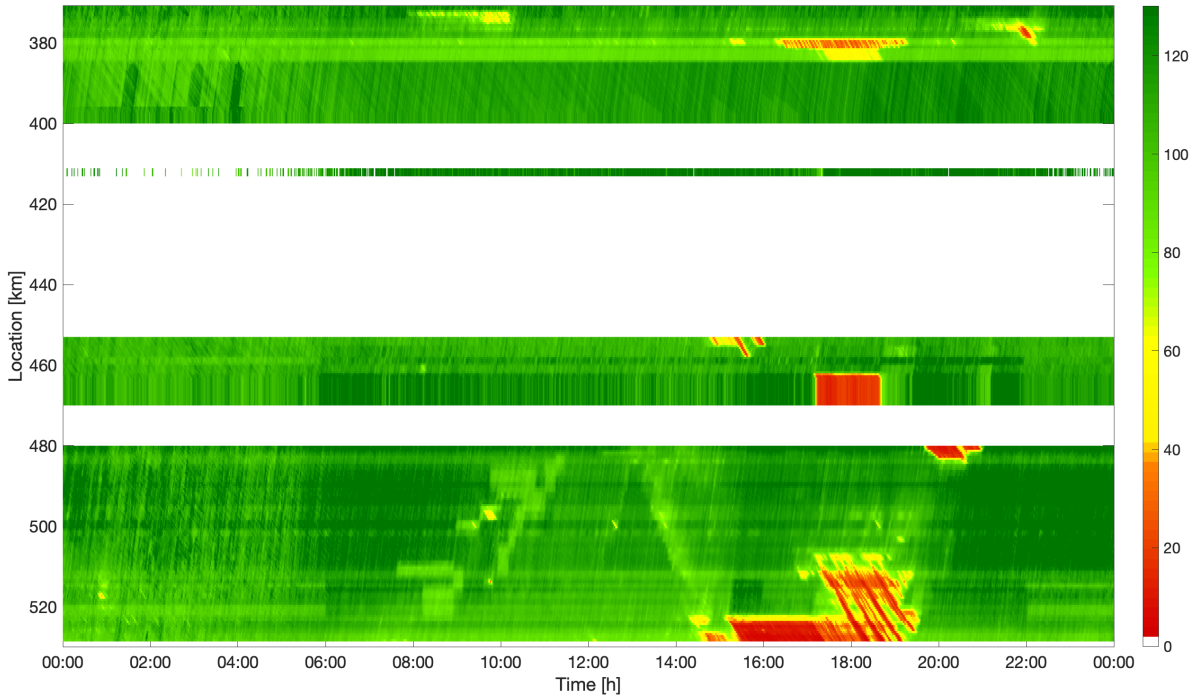


Figure 5.4: v^{SDD} derived from gantry and radar sensor data, interpolated using ASM [Schreiter et al., 2010]

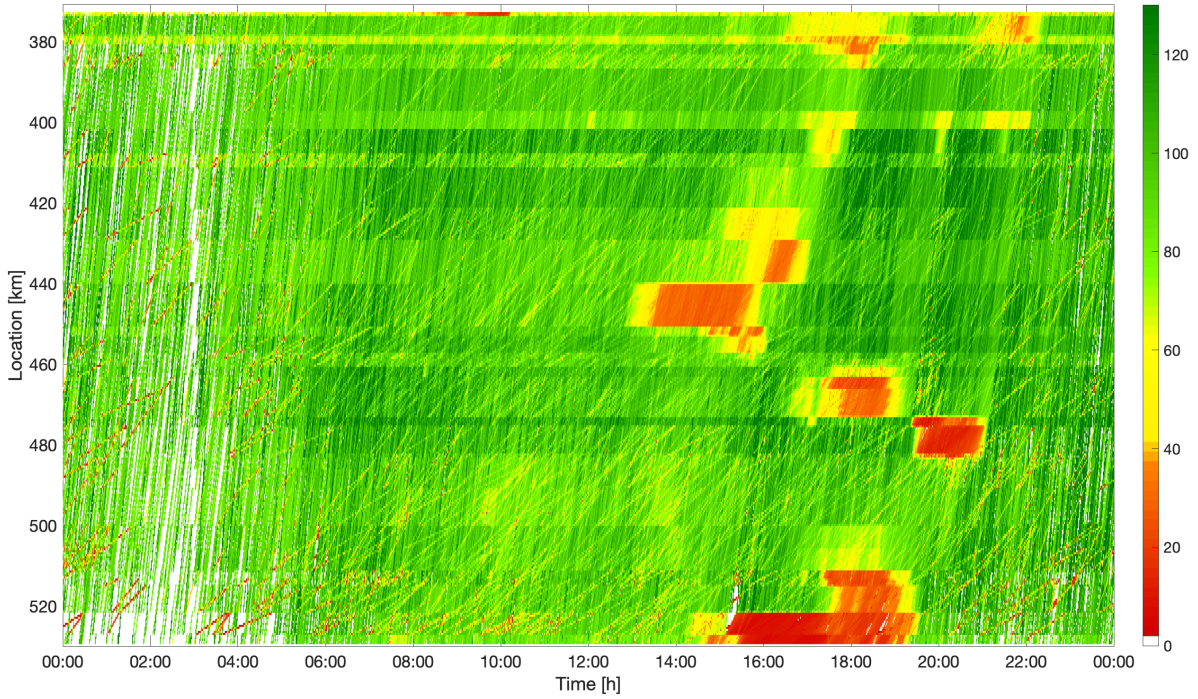


Figure 5.5: v^{BT} derived from BT measurements, interpolated using LTSM

The FCD contour is illustrated in Figure 5.6. For this data set, represented in cells with $|x_i| = 500\,\text{m}$ and $|t_j| = 1\,\text{min}$, interpolation is not necessary because segment-based speed

information is available. Since FCD correspond to a relatively small share of the vehicle collective, many gaps without any measurement exist (undefined cells). Wherever congestion occurs, it can be clearly captured.

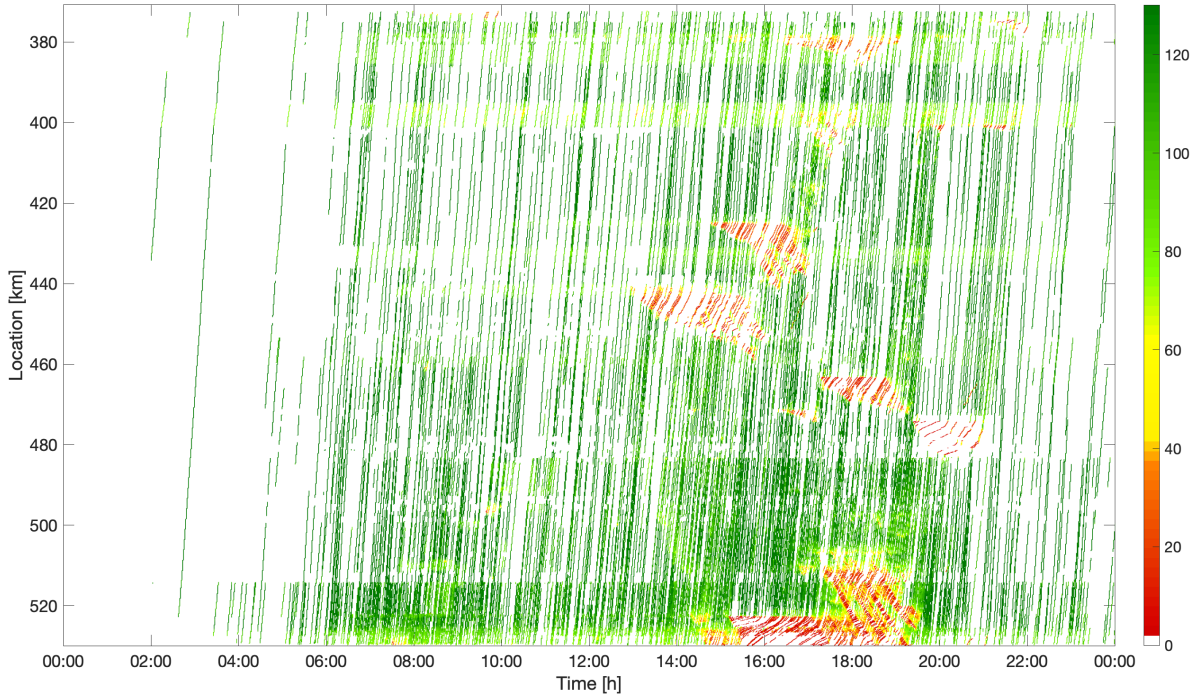


Figure 5.6: v^{FCD} derived from FCD, trajectories averaged in spatio-temporal cells

5.2.1 Evaluation

In general, all three detection technologies recognize similar traffic situations which allows for multi-sensor data fusion for an optimized traffic state reconstruction. In this study, a combination of gantry, radar and FCD speeds are used as approximation of the Ground Truth (GT), using fusion method 1, the weighted average (equation 4.3), to combine the sources into a sparse matrix v^{GT}. This matrix is used as test set against which the training set is compared (section 4.2.3). Figure 5.7 shows the GT if all available information is included in the test set (optimal GT data set).

The results applying fusion method 1, the weighted convex combination of all data sources, are illustrated in Figure 5.8. The algorithm computes the minimum error for all weights in the range from 0.05 to 0.95 in steps of 0.05 for each data set. In each of the 50 runs from the training and test set divisions, the minimum SSIMPE value is considered. Over all, the reconstructed speed returns an average deviation of $\text{SSIMPE}_1 = 0.170$.

The speed reconstruction using method 2, the hierarchical ordering of reliability per data source, results in the contour depicted in Figure 5.9. Each combination of the three sources being the most, the second most and the least reliable data set, is considered and again, minimum SSIMPEs per run are computed. The fused matrix shows an average SSIMPE_2 value of 0.156.

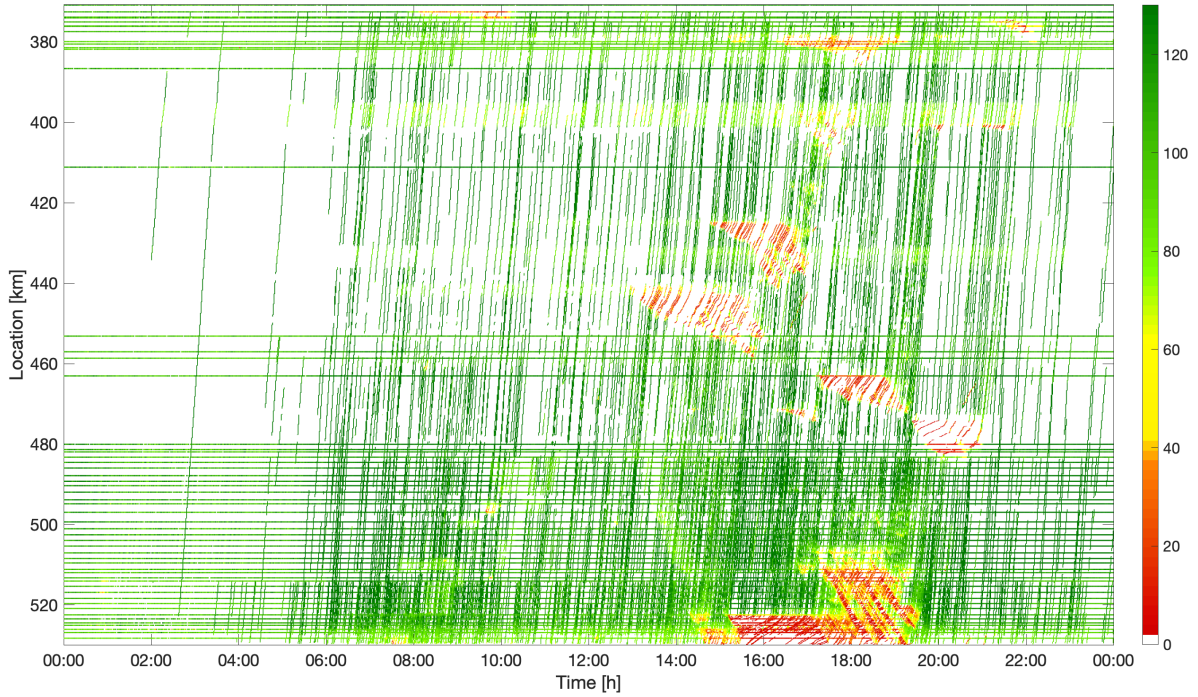


Figure 5.7: v^{GT}, GT derived from non-interpolated SDD and FCD

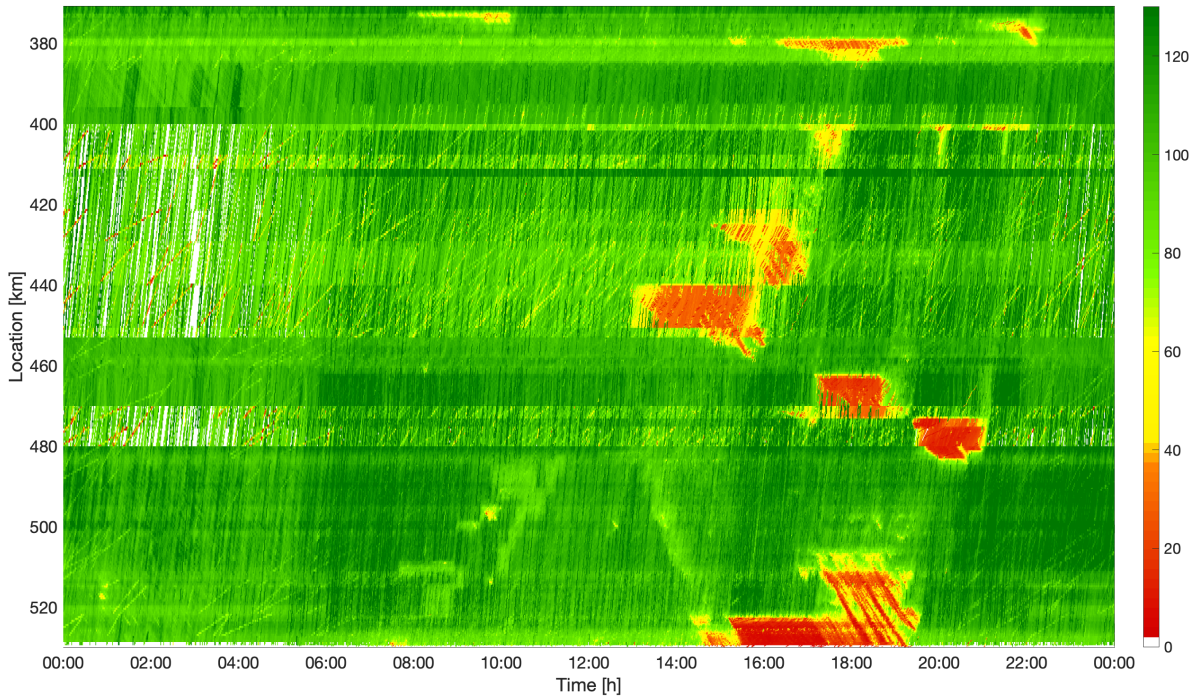


Figure 5.8: Fusion of all data sources according to method 1

Method 3, the interpolation of a fused matrix (both from methods 1 and 2) according to the ASM interpolation method, results in a speed distribution as illustrated in Figure 5.10, applying ASM parameters from Table 5.1. The interpolation of method 1 shows a SSIMPE_{3a}

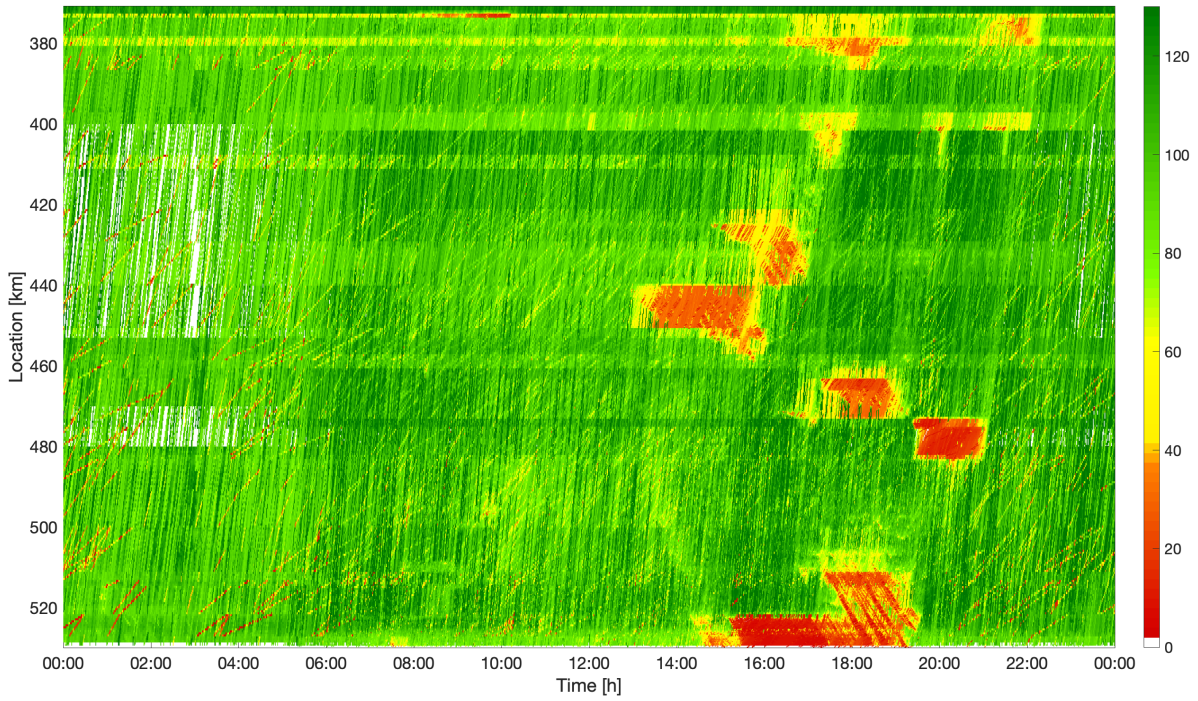


Figure 5.9: Fusion of all data sources according to method 2

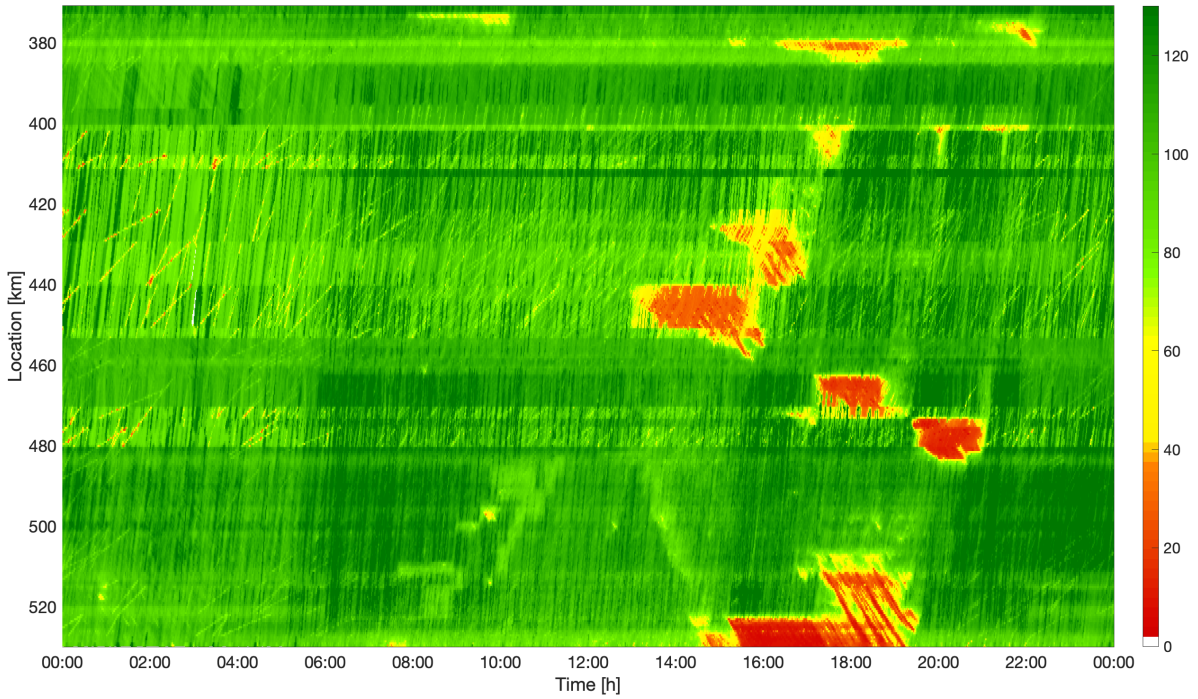


Figure 5.10: Fusion of all data sources according to method 3

value of 0.148; interpolating the fused matrix derived using method 2 with the identical ASM parameters, the resulting SSIMPE_{3b} value amounts to 0.145. Both values are equivalent as they are in the same order of magnitude. Therefore, method 3 does not need a subdivision

and as a whole outperforms both other approaches without ASM-interpolation.

Figure 5.11 shows a synopsis of the resulting average SSIMPEs. Generally, the median error of all methods converges after around 10 iterations to a stable solution. To overcome unsuitable random partitions between training and test data selections, 50 runs are simulated. Using this conservative number of runs, potential outliers do not have an effect on the solutions. With 50 iterations, all variances decrease as expected compared to 10 runs. The variance between the 50 runs is small which indicates that the methodology is not sensitive to the selection of training and test data sets. Methods 1 and 2 have significantly higher median SSIMPE values than method 3 which in turn has a higher variance. Method 3 outperforms methods 1 and 2 as even the worst case results show a better performance.

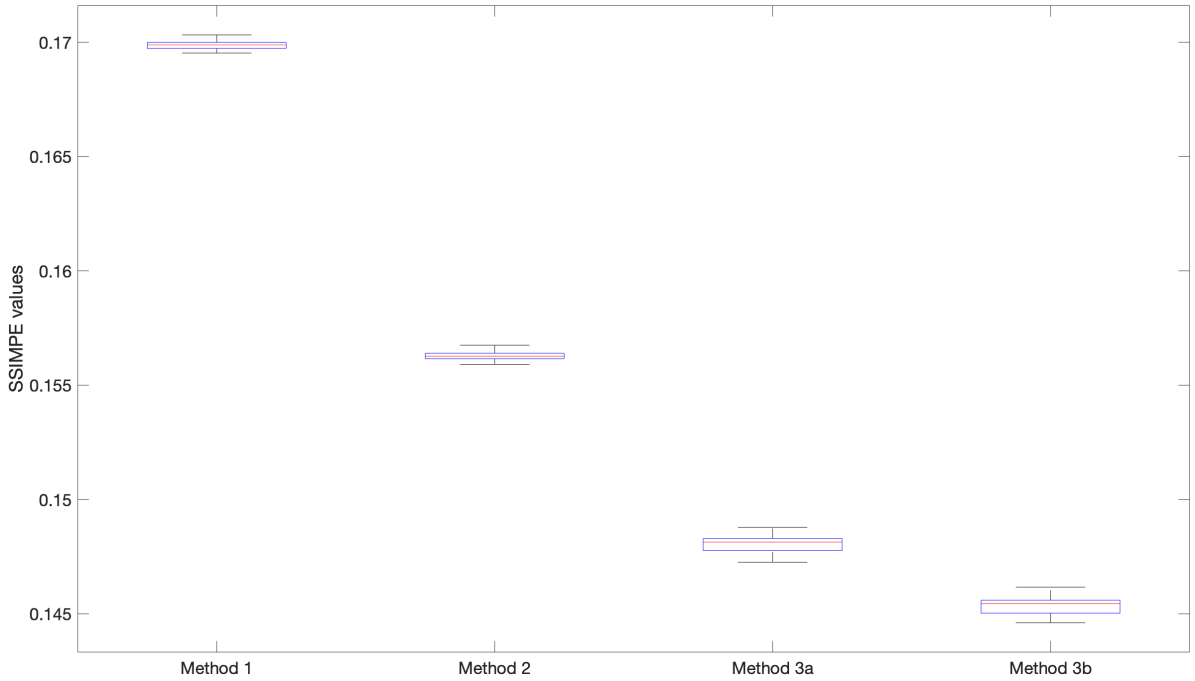


Figure 5.11: Resulting average SSIMPEs (median values in red, 25 and 75 % quartiles in blue)

5.2.2 Conclusion

The results show that each sensor technology is a valuable source of information: While gantry and radar sensors are not available along the complete road stretch, they provide relatively accurate speed information in space and time. FCD have a varying data rate; especially at night, the amount of data is limited. Though, the high spatial and temporal accuracy reveals the structure of congestion with high precision. BT data are available in large quantities allowing to get speed estimates for nearly each cell in time and space. However, the relatively large spacing between the sensors does not allow to detect exact congestion fronts and detailed speed information.

All approaches of information fusion allow for reconstructing the traffic speed using multiple data sources. Compared to each data set solely, the overall accuracy and data completeness

increases. The comparison of the fusion methods shows that method 2 outperforms method 1. Applying method 3 to the fused data sets, both methods 1 and 2 are outperformed according to the resulting SSIMPE values. The main advantage of method 3 is that also undefined cells are filled-up. In this example, 9.5 % of all cells do not have set any speed value, neither in SDD nor in BT nor in FCD data sets. Using neighbored cells and their speed information, these gaps in the grid can be filled and the corresponding cells become defined. The resulting speed contour plot shows fewer noise due to the data interpolation. Specifically, the interpolation of the reliability ordered filling method might return a discontinuous speed distribution. The most accurate result is obtained when using all data sources for reconstruction. The computational complexity of all approaches is similarly low and all fusion methods can be run in real-time.

Concluding, in the following sections, the data fusion according to method 3 is taken as the standard fused data set when comparing data sets from multiple detection technologies.

5.3 Primary Congestion Detection

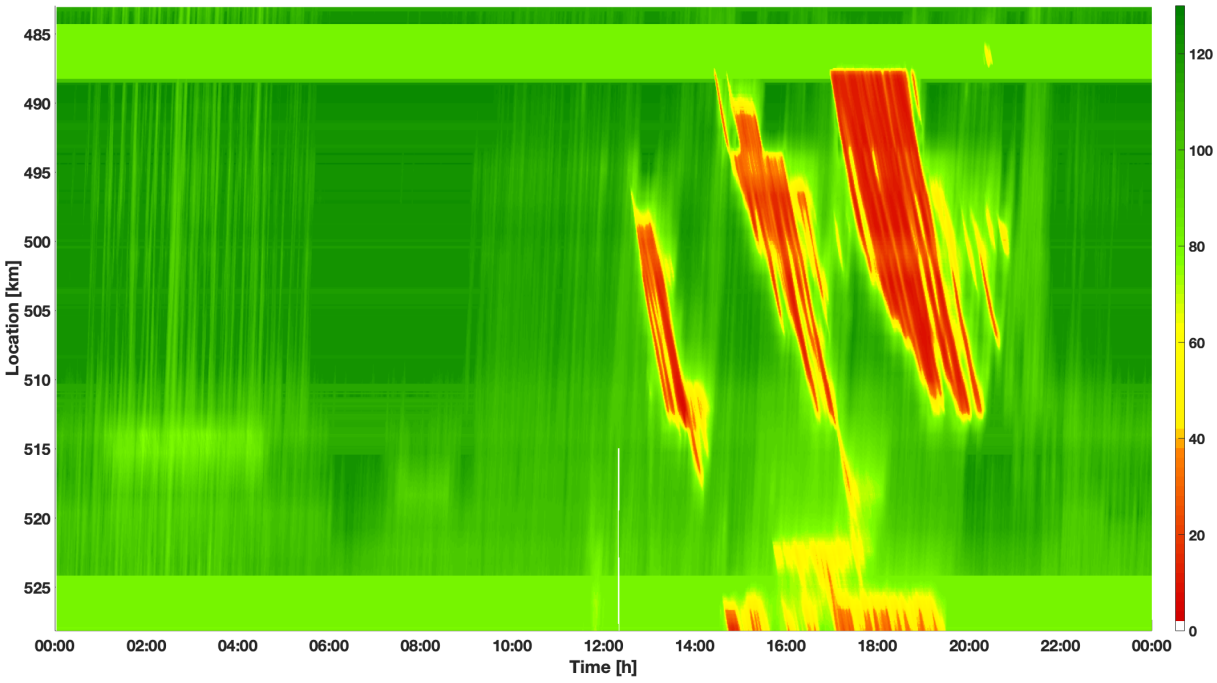
Section 4.5 introduced an algorithm to identify which data source detects congestion earliest and which delay all others show. This way, it answers the question, on which data source a speed estimation should be based on to detect congestion early. On the one hand, there are temporally accurate data derived from stationary sensors such as gantry or radar measurements but limited available in space. These raw data need interpolation first. On the other hand, there exist spatially exact data such as Floating Car Data (FCD) which mirror the traffic state with certain time lags because single Global Navigation Satellite System (GNSS) updates need to be fused by the data provider. BT data need to await their corresponding second signal in order to match both timestamps to a travel time (time consuming, therefore temporally inaccurate) and are spatially inexact due to segment speeds. Both FCD and BT represent only a subset of the traffic collective. In this section, this algorithm is quantitatively applied to data sets from SDD and FCD comprising three months in 2015.

5.3.1 Evaluation

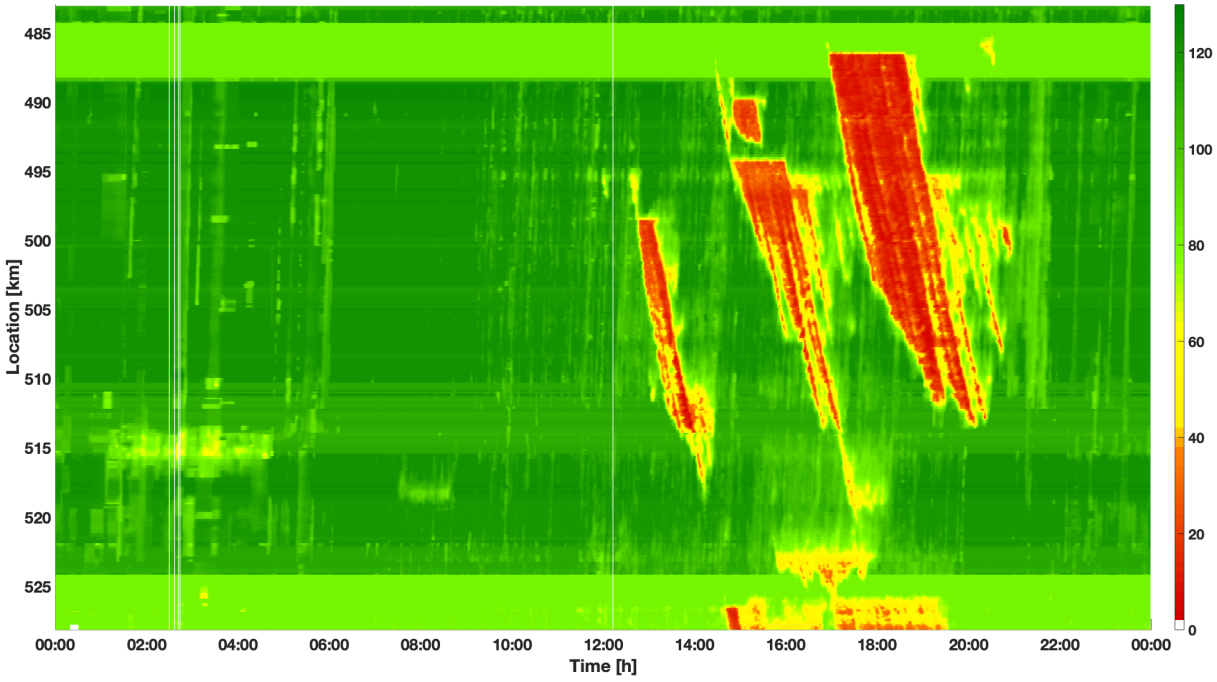
As an example, Figure 5.12 shows the speed distributions from SDD and FCD, respectively, from April 30, 2015 in NB direction. SDD are interpolated using the ASM approach with parameters given in Table 5.3.

Parameter	**Value**
Spatial grid distance	100 m
Temporal grid distance	30 s
Speed in congestion	-18 km/h
Free-flow speed	80 km/h
Crossover from free to congested traffic	70 km/h
Width of the transition region	10 km/h

Table 5.3: ASM parameter values used for smoothing SDD



(a) Speed distribution from interpolated SDD



(b) Speed distribution from FCD

Figure 5.12: Speed distributions from April 30, 2015 in NB direction (in [km/h])

TomTom FCD are available as spatio-temporally discretized speed cells (temporal resolution of 1 min, irregular spatial grid). There is no need for interpolation since the matrices are abundant. The light green horizontal bars denote general speed limits. Due to the interpolation, congestion starting around 15:00 at km 490 cannot clearly be captured by the SDD and appears as one coherent jam in contrast to FCD where two separate jams occur. Besides that, both contours show a similar speed distribution and no general bias.

Figure 5.13 depicts the comparison of congested areas during the afternoon of the same day where the speed threshold v_{crit} amounts to 40 km/h. Black horizontal lines denote stationary sensor locations. Various phases are illustrated: in green, where only interpolated SDD detect congestion; in blue, where only FCD detect congestion. Red areas denote phases where both SDD and FCD detect congestion. The downstream congestion front of the incident starting around 17:00 at km 487 is located farther downstream in FCD compared to SDD since due to the ASM interpolation of SDD, the bottleneck cannot be localized exactly between the two sensors.

SDD recognize the first and the second congestion regime slightly earlier (green stripe on the left of both jams), FCD show a delay of some minutes (red area surrounded by a green stripe on the left and a blue stripe on the right). FCD determine congestion on the third event earlier both temporally and spatially, and overall show a detailed spatial resolution.

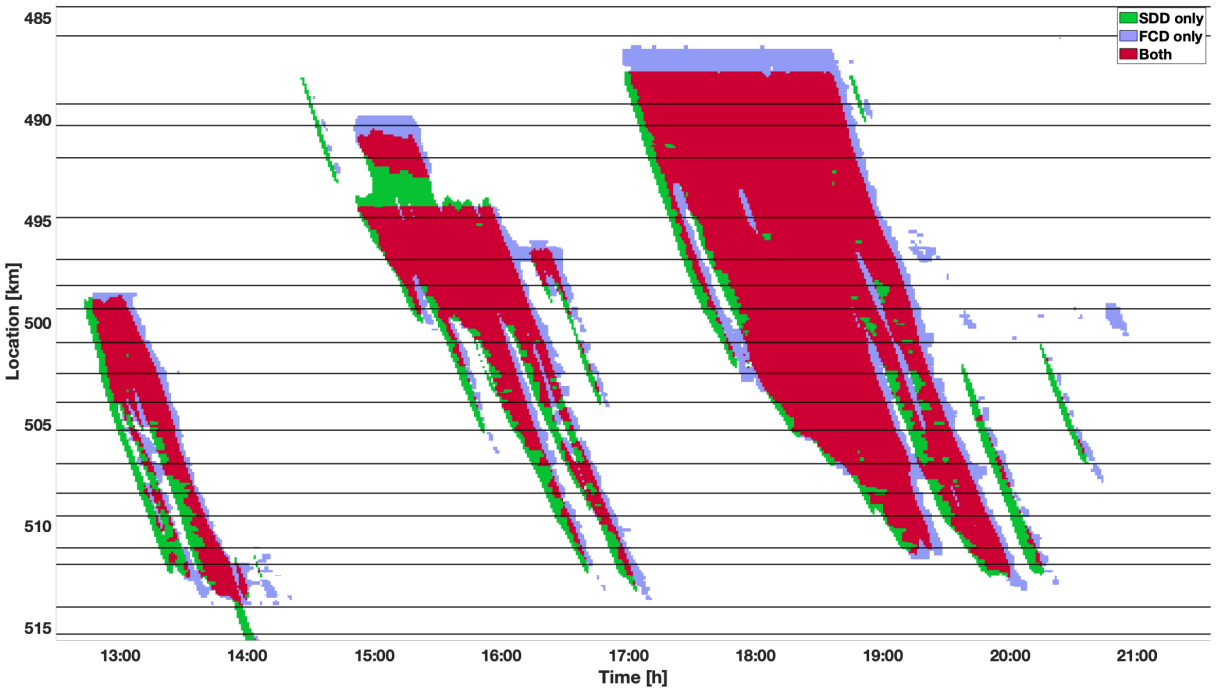
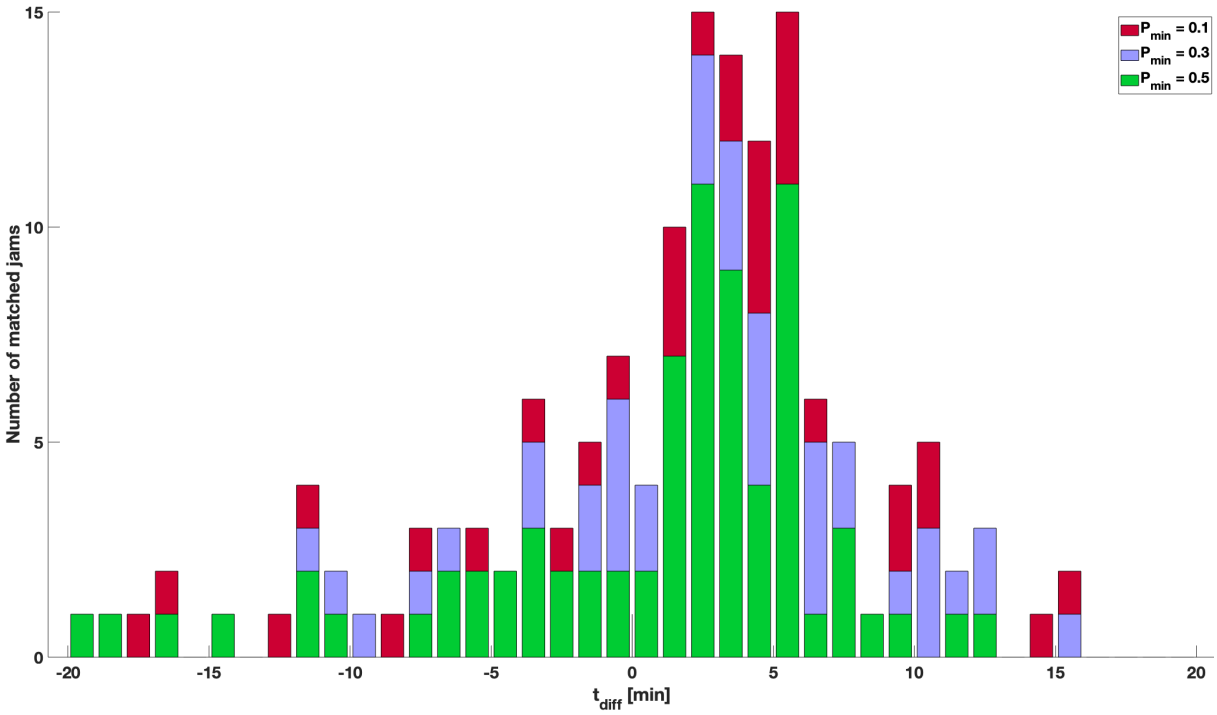


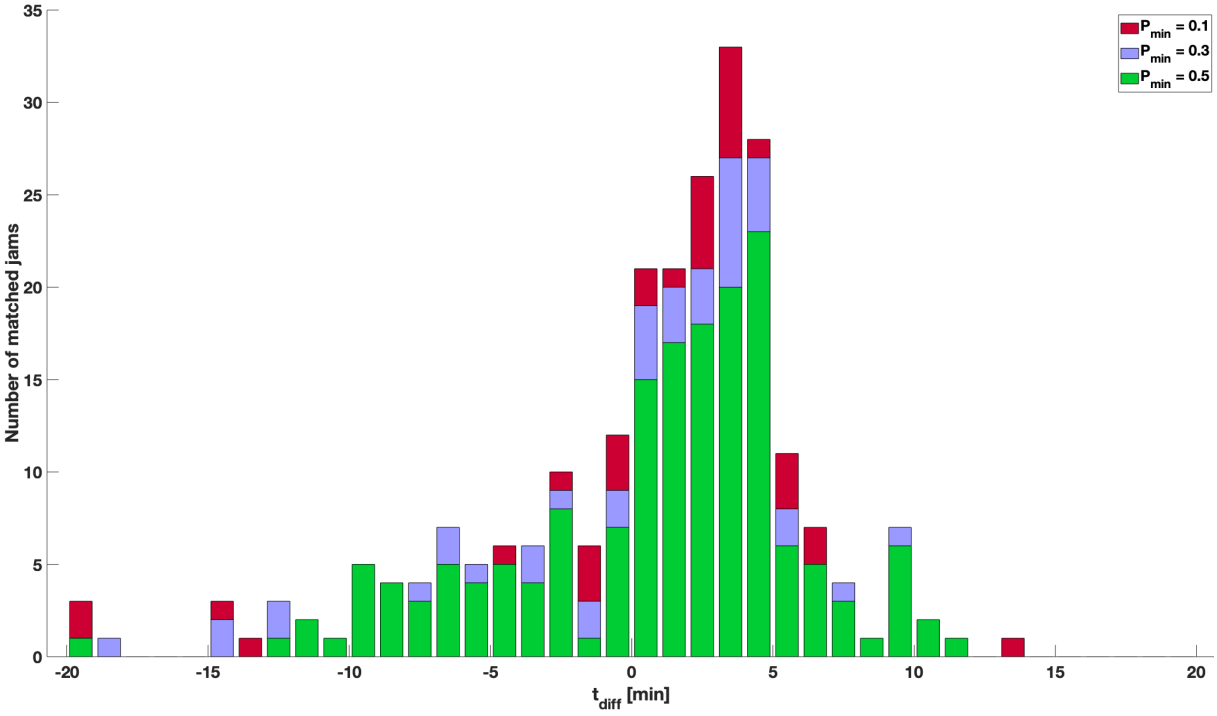
Figure 5.13: Matching of both data sources with a speed threshold of $v_{\text{crit}} = 40\,\text{km/h}$

Congestion clusters are determined for either data source and these clusters are matched between data sets with the parameter set as follows in Table 5.4.

For each match, the deviation t_{diff} (equation 4.19) between the starting times is calculated. Figure 5.14 shows the distributions of t_{diff} for all matchable congestion events in the 80 days period in both driving directions with $v_{\text{crit}} = 40\,\text{km/h}$ and 60 km/h, respectively. In case



(a) Deviation of primary congestion detection with $v_{\text{crit}} = 40\,\text{km/h}$



(b) Deviation of primary congestion detection with $v_{\text{crit}} = 60\,\text{km/h}$

Figure 5.14: Minutes of deviation for different values of v_{crit} and P_{min}

Parameter	Value Set
Velocity threshold v_{crit}	$\{40, 60\}$ km/h
Minimum coverage of associated clusters' intersection P_{min}	$\{0.1, 0.3, 0.5\}$
Minimum free-flow time between congested areas t_{merge}	4 min
Minimum cluster size A_{min}	2 km·min

Table 5.4: Parameter values to identify and match clusters

$t_{\text{diff}} = 0$, both sources detected the associated congested regions at the same time, $t_{\text{diff}} < 0$ means that FCD detected congestion earlier by the indicated minute. Analogously, $t_{\text{diff}} > 0$ implies that SDD registered congestion first. Due to the uniform grid spacing of 1 min, the deviation can only take integer numbers. The minutes of deviation in the plots are limited to $\pm$ 20 min. Matched clusters with $P_{min} = 0.5$ are also included in $P_{min} = 0.3$ bars, clusters with $P_{min} = 0.3$ also in $P_{min} = 0.1$ bars. Average, quartile, and extreme values are given in Table 5.5.

		Matchable	Deviations [min]		
v_{crit}	P_{min}	Jams	Mean/Median	q_{25}/q_{75} Quartiles	Min/Max
40 km/h	0.1	152	0.2/2	$-3/5$	$-31/21$
40 km/h	0.3	116	0.5/2	$-2/5$	$-31/15$
40 km/h	0.5	75	0.3/2	$-2.8/4.8$	$-20/12$
60 km/h	0.1	244	$-0.2/1$	$-3/4$	$-31/13$
60 km/h	0.3	211	$-0.2/1$	$-3/4$	$-31/11$
60 km/h	0.5	168	0.4/2	$-3/4$	$-20/11$

Table 5.5: Deviations for $v_{\text{crit}} \in \{40, 60\}$ km/h and $P_{min} \in \{0.1, 0.3, 0.5\}$

A throughout positive median value indicates that congestion is generally detected earlier by interpolated SDD than by FCD (height of the bars in Figure 5.14). Nevertheless, there are more matchings where FCD recognize a congested situation significantly earlier than SDD (positive skew). The balanced statistical results show that both approaches are equally good in detecting matchable traffic jams, hence, congestion detected by both approaches. The larger the speed limit parameter is chosen, the smaller is the deviation. However, FCD are able to find more relevant congestion clusters: for $v_{\text{crit}} = 40$ km/h, FCD identify 394, SDD 372 clusters; for $v_{\text{crit}} = 60$ km/h, FCD find 472, SDD 407 clusters. Others do not obey equations 4.14 – 4.16 on P_{min} (minimum percentage of congestion regarding both associated clusters) and A_{min} (minimum size of congested area).

This algorithm finds reliable matchings. However, there are examples where the first congestion recognition time differs a lot. In Figure 5.15, a spatio-temporal area from April 21, 2015 in NB direction with a speed threshold of $v_{\text{crit}} = 60$ km/h is depicted. Again, horizontal lines denote the detector positions. The congested region detected by FCD is marked in red, that one identified by SDD is marked in bold on the detector position lines. FCD detected congestion beginning at 16:12, SDD at 16:42. Hence, the difference t_{diff} amounts to -30 min. Obviously, the detector positions are disadvantageous to detect this congestion early due to the large spacing.

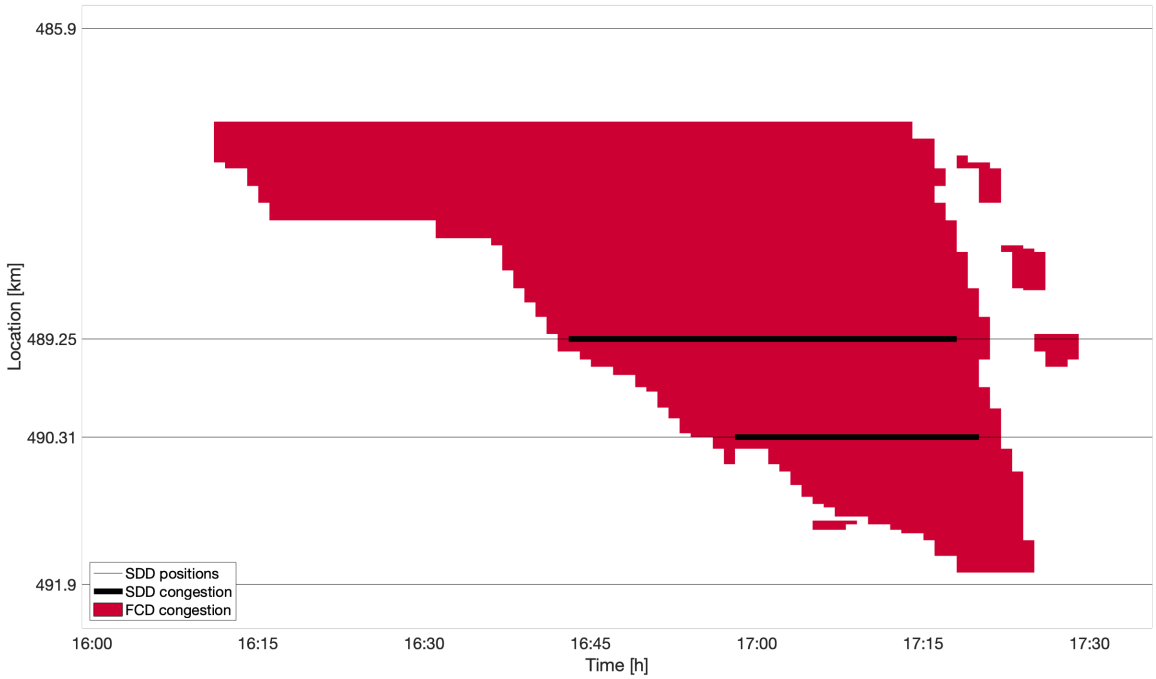


Figure 5.15: Congestion cluster where FCD detected congestion 30 min earlier than SDD

5.3.2 Conclusion

To conclude, neither data set clearly outperforms the other one. The results show that for congestion having its downstream front on a detector position or not far downstream, SDD recognize the it by 2-3 min earlier on average compared to FCD. If a congestion front starts between detector positions or slightly upstream of one, stationary detectors naturally are not suited best to identify congestion early. In these cases, floating cars detect congestion significantly earlier than detectors. Hence, a congestion hot spot analysis (see section 5.5) might help to find optimized locations for stationary sensors. In general, FCD find more (smaller) congestion clusters compared to the smoothed detector data. A slight time lag of few minutes regarding floating cars appears because FCD are processed and fused whereas detector data are derived by direct measurements.

5.4 Detection Rate of Congestion Types Per Data Source

Section 4.4 introduced a 4+1 congestion classification. Four mainline congestion types are *Jam Wave*, *Stop and Go*, *Wide Jam*, and *Mega Jam*. The sensor technologies presented in section 3.1 do not detect each congestion type equally. In the following, the detection rate of these congestion patterns is assessed both per data source individually and on the fused data set. As an example, the same setting as presented in section 5.2 is analyzed, from May 29, 2019 in NB direction. It is investigated which congestion types can be identified by which data set reliably. The speed contour plots of the data sets are shown in Figure 5.4 (gantry

and radar speeds), Figure 5.5 (BT travel times), Figure 5.6 (FCD speeds), and Figure 5.10 (fused speeds).

If a road segment is not equipped with traffic detectors, the speed reconstruction leads to undefined cells. As remarked in section 4.6, these locations without data can be accepted as part of the input speed matrix. To apply the algorithm, all cells representing an unavailable location over the entire time line as well as unavailable times (e.g., time zone offsets because some data sets are in UTC time, others in local time) are set to a free-flow speed v_{freeflow}. During the assignment of a congestion type to an identified congestion cluster, all other areas except the cluster are set to free-flow conditions. Similarly, these locations or times of undefined speed values can be filled with free-flow defaults. Note that assigning a free-flow value is not valid for single undefined speed value cells surrounded by defined speed value cells. This interpolation of particular cells is done using the smoothing algorithms described in section 2.1 and section 4.1.

Speed data are available as a flow-weighted mainline mean. The identification of congestion clusters utilizes the algorithm described in section 4.3.2. The parameter values used to create the clusters are as given in Table 5.6.

Parameter	**Value**
Velocity threshold v_{crit}	40 km/h
Free-flow speed v_{freeflow}	120 km/h
Minimum free-flow time between congested areas t_{merge}	4 min
Minimum size of congested areas A_{min}	12 km·min

Table 5.6: Parameter values to identify separate congestion clusters

Thereafter, a congestion type is assigned to each individual cluster according to the algorithm described in section 4.4.1. Virtual trajectories traverse through the entire congested area and whenever they hit a certain cluster, they change their speeds from free-flow to the speeds defined inside the cluster. The parameter values used for the assignments are given in Table 5.7. In order to minimize the number of assignments to *Mixed* – the undefined congestion type – n_{2types} and n_{3types} are set relatively low in an empirical way. As assessment method, number and size of each recognized congestion type is considered per data source. The size of a congestion cluster is defined as the area of its convex hull.

Parameter	**Value**
Maximum duration of *Jam Wave* $t_{JamWave}$	3 min
Minimum duration of *Mega Jam* $t_{MegaJam}$	30 min
Minimum number of speed drops $n_{StopandGo}$	2
Temporal offset of starting trajectories t_r	5 min
Minimum share of congestion types (2 types) n_{2types}	0.51
Minimum share of congestion types (3 types) n_{3types}	0.41

Table 5.7: Parameter values to assign congestion types

5.4.1 Gantry Data Set

Applied to interpolated gantry speed measurements, the results of the congestion type assignment show four congested areas on this day. Either type *Stop and Go*, *Wide Jam*, and *Mega Jam* is assigned to one cluster each (Figure 5.16). The *Mixed* cluster at km 380 in the late evening is traversed by four trajectories if the temporal offset t_r is set to 5 min. Thereof, two *Jam Wave* and two *Wide Jam* trajectories are identified. Therefore, the *Mixed* cluster remains unclassified. The area without detectors between km 480 and km 400 is set to free-flow conditions such that congestion detection is not false positive.

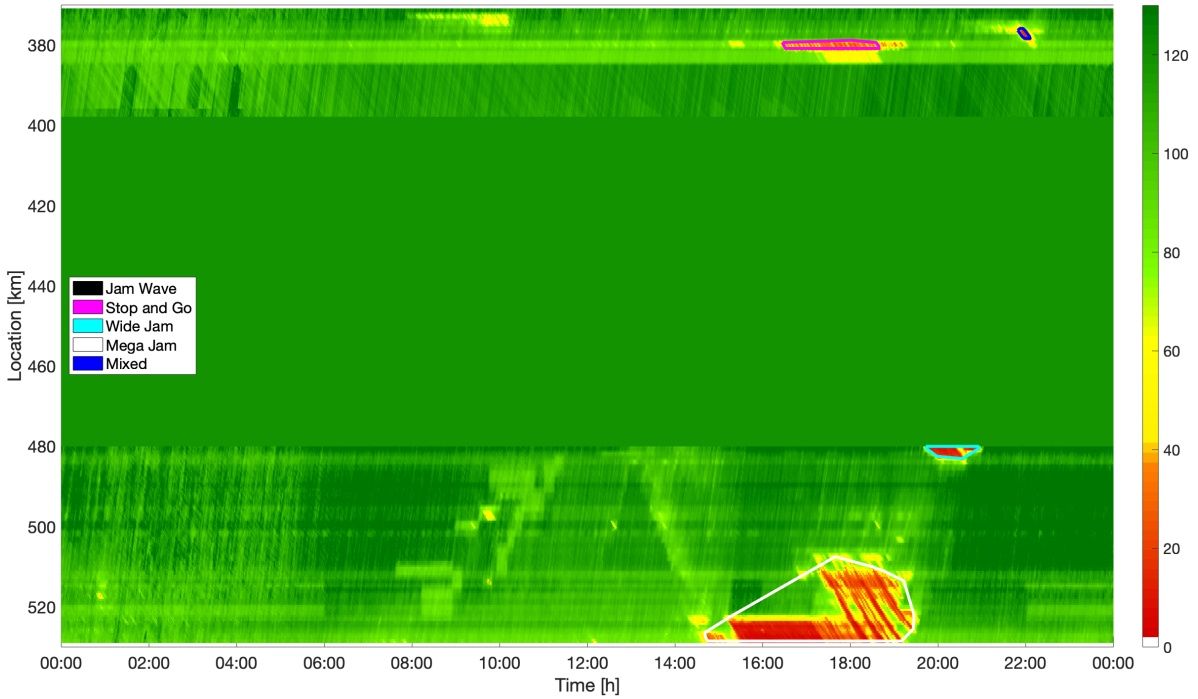


Figure 5.16: Congestion type identification from gantry speed measurements: 1 *Stop and Go*, 1 *Wide Jam*, 1 *Mega Jam*, 1 *Mixed* clusters

5.4.2 Radar Data Set

In contrast to gantry speed measurements, only few radar sensor locations are available, in this example one at km 411 and five around km 460. The identification of congestion types returns one *Jam Wave*, one *Wide Jam*, and one *Mixed* cluster (Figure 5.17). The *Mixed* cluster is crossed by three trajectories, *Stop and Go*, *Wide Jam*, and *Jam Wave* types, respectively. Again, all unequipped road stretches are set to free-flow speeds. The tails of the jam occurring downstream (in the unequipped region) appear as small single jam events instead of one merged cluster. From these sparse radars only, the algorithm cannot assign a reliable congestion type in this case.

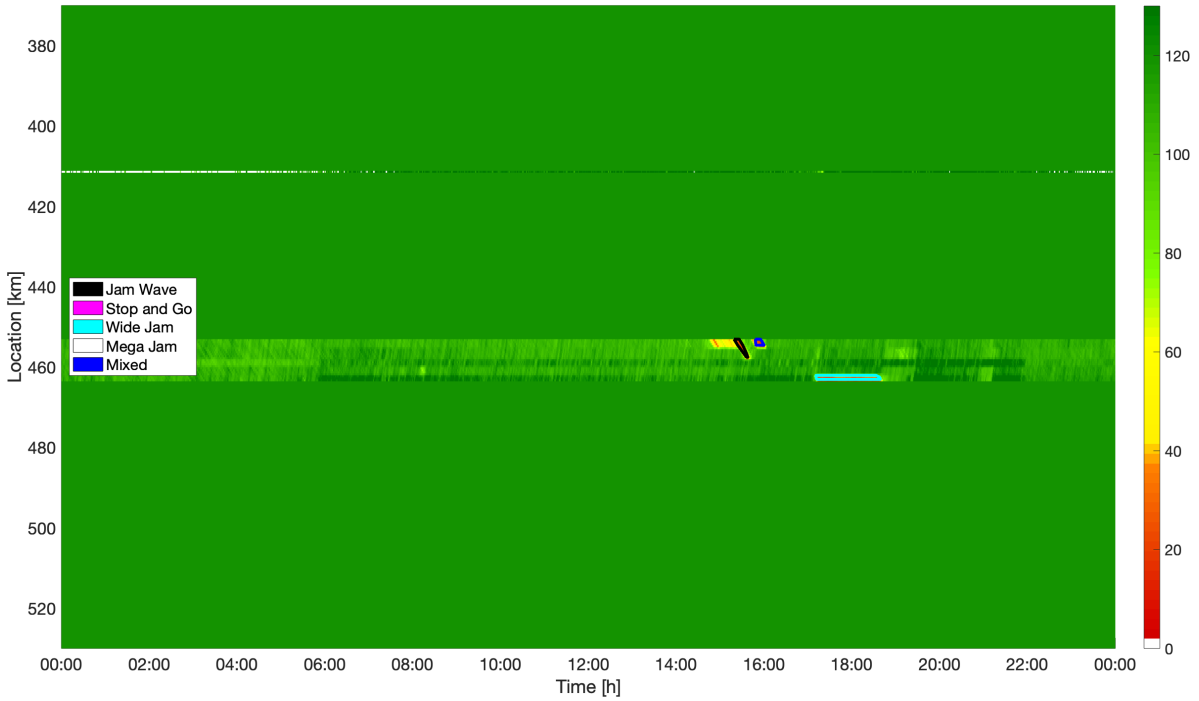


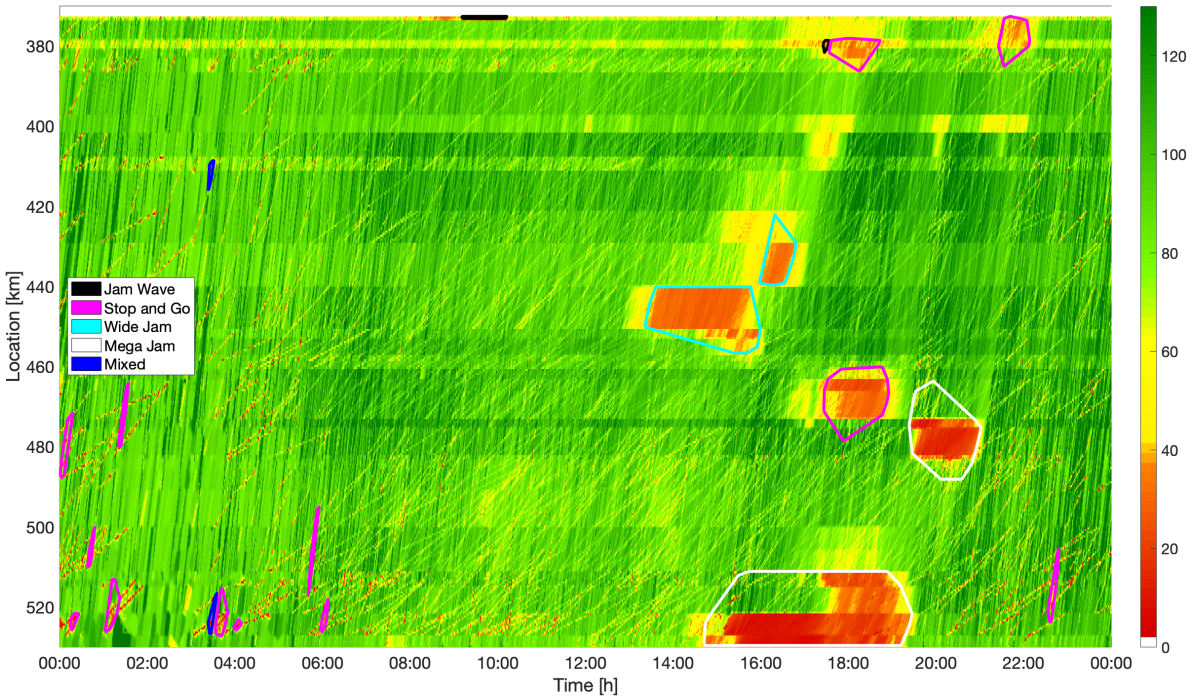
Figure 5.17: Congestion type identification from radar speed measurements: 1 *Jam Wave*, 1 *Wide Jam*, 1 *Mixed* clusters

5.4.3 Bluetooth Data Set

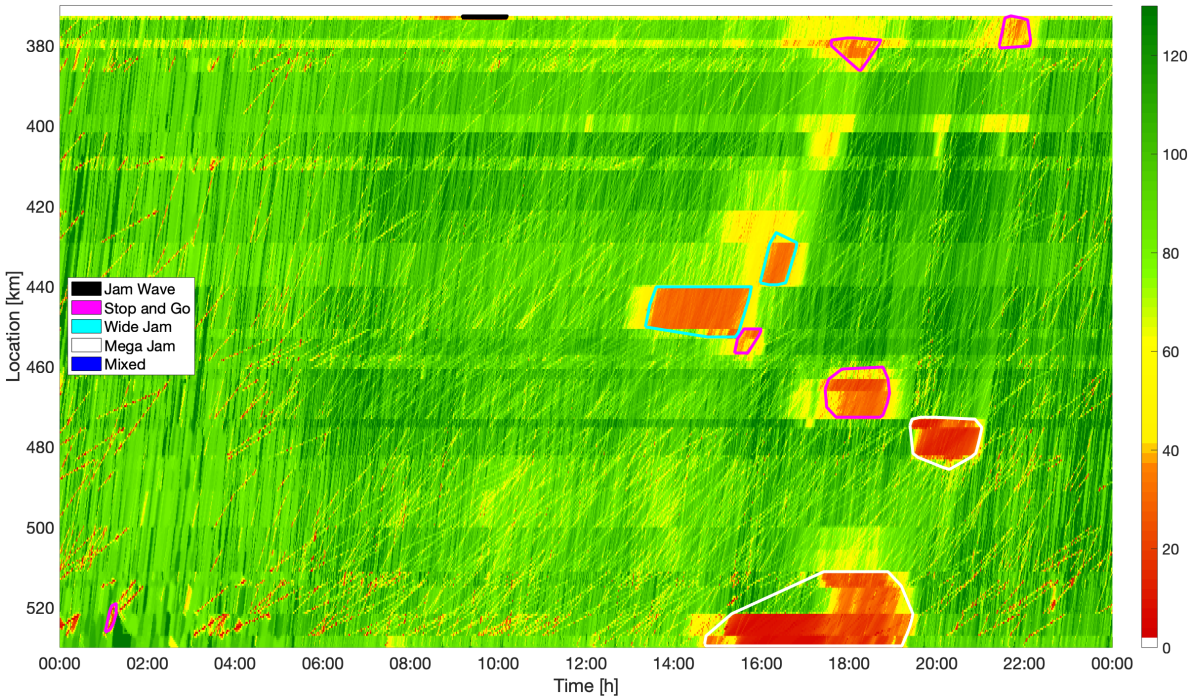
Interpolated travel times derived from BT measurements detect many congestion clusters more than local speed detectors because single slow vehicles which were not eliminated beforehand have a greater influence. Therefore and because the sensors are installed on a longer road stretch than both previously mentioned detection technologies, the results with the same parameterization show 21 congestion clusters whereof *Jam Wave* is assigned to 2, *Stop and Go* to 13, *Wide Jam* to 2, and *Mega Jam* to 2 clusters (Figure 5.18a). Two clusters cannot be assigned to a unique congestion type and are set to the undefined congestion type *Mixed*. Both of them are touched by one *Jam Wave* and one *Stop and Go* trajectory each.

This prompts a need for a sensitivity analysis and some parameter adjustments, specifically the increase of A_{min} and the decrease of t_{merge}. With an average spacing of approximately 5 km between the BT sensors, it is suggested to set $A_{min} = 30$ km·min. This value corresponds to (spatially short) congestion of at least 6 min or spatially more extended congestion which both is reasonable for a congestion definition measured by BT. t_{merge} separates congestion occurrences. Especially during night hours when – supposedly – slow vehicles are moving around, the differentiation should be set sharp. Therefore, $t_{\text{merge}} = 1$ min is recommended. Both parameter changes lead to more precise convex hulls of the clusters and reduce high speeds inside the congested clusters.

After these adjustments, the results for actual congestion (not derived from single vehicles) return 10 clusters whereof 1 *Jam Wave*, 5 *Stop and Go*, 2 *Wide Jam*, and 2 *Mega Jam* clusters are assigned (Figure 5.18b). The parameter adjustments lead to an underestimation of the congestion type *Jam Wave*. However, larger jam events can be detected reliably. One main



(a) Without parameter adjustments: 2 *Jam Wave*, 13 *Stop and Go*, 2 *Wide Jam*, 2 *Mega Jam*, 2 *Mixed* clusters



(b) With parameter adjustments: 1 *Jam Wave*, 5 *Stop and Go*, 2 *Wide Jam*, 2 *Mega Jam* clusters

Figure 5.18: Congestion type identification from BT measurements

finding is that *Jam Wave* is a pattern detectable worst through travel time measurements with large sensor spacings due to an upstream propagation of congestion waves in contrast to downstream travel time measurements.

5.4.4 Floating Car Data Set

FCD travel times are interpolated using the *Phase-Based Smoothing Method (PSM)* described in section 2.1 with the parameter values as proposed in [Rempe, Franeck, et al., 2017] and given in Table 5.8.

Parameter/Phase	Wide Moving Jam	Synchronized Flow	Free-flow Traffic
v_p^{thres}	30 km/h	65 km/h	55 km/h
λ_p	0.5 h/km		
v_p^{dir}	-18 km/h	0 km/h	
τ_p	30 s	250 s	
σ_p	500 m	150 m	
$v_p^{dir,H}$	-18 km/h		70 km/h
τ_p^H	30 s		100 s
σ_p^H	200 m		100 m

Table 5.8: PSM parameter values used for smoothing FCD

Figure 5.19 depicts the reconstructed traffic speeds with identified 3 *Jam Wave*, 5 *Stop and Go*, 1 *Wide Jam*, 2 *Mega Jam*, and 2 *Mixed* clusters. The *Mixed* cluster at 20:00 at km 400 is traversed by two *Jam Wave* and two *Wide Jam* trajectories. The larger cluster at 9:30 around km 372 is passed by 3 *Jam Wave*, 3 *Stop and Go*, and 2 *Wide Jam* trajectories. In both cases, n_{2types} and n_{3types} are not fulfilled, respectively.

5.4.5 Fused Data Set

The data set fused from SDD, BT, and FCD computes 10 congestion clusters. The result is depicted in Figure 5.20. It shows one *Jam Wave* cluster, 6 clusters of congestion type *Stop and Go*, 1 *Wide Jam* cluster, and 2 clusters of *Mega Jam*.

5.4.6 Findings

Figure 5.21 conjointly shows all recognized congestion clusters and their corresponding convex hulls. In general, clusters are detected similarly. The recognized *Mega Jam* at km 480 at 20:00 is classified as a *Wide Jam* by SDD because the downstream sensor equipment ends at km 480. Analogously, the radar-equipped area starts around km 462 and therefore, congestion can only be detected in a narrow stripe. Therefore, SDD do not detect the whole congestion area.

Summarizing, all considered detection technologies are able to identify congested areas with a different level of detail. The basic requirement is a sufficiently small detector spacing.

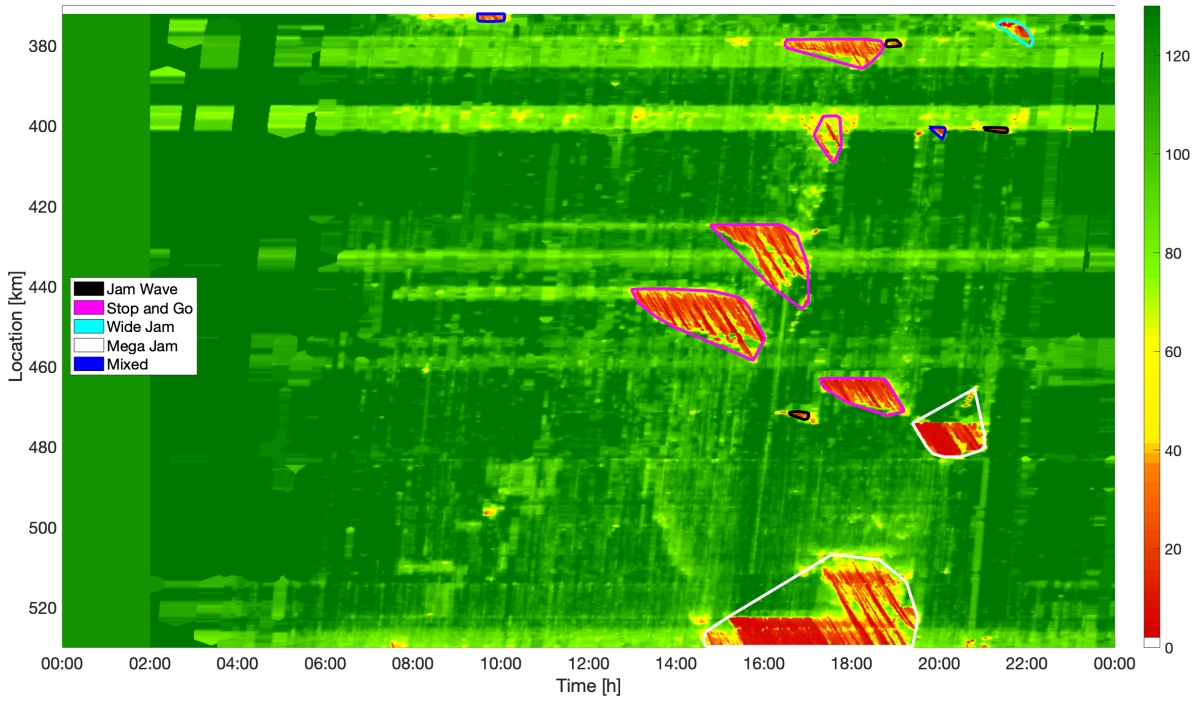


Figure 5.19: Congestion type identification from FCD measurements: 3 *Jam Wave*, 5 *Stop and Go*, 1 *Wide Jam*, 2 *Mega Jam*, 2 *Mixed* clusters

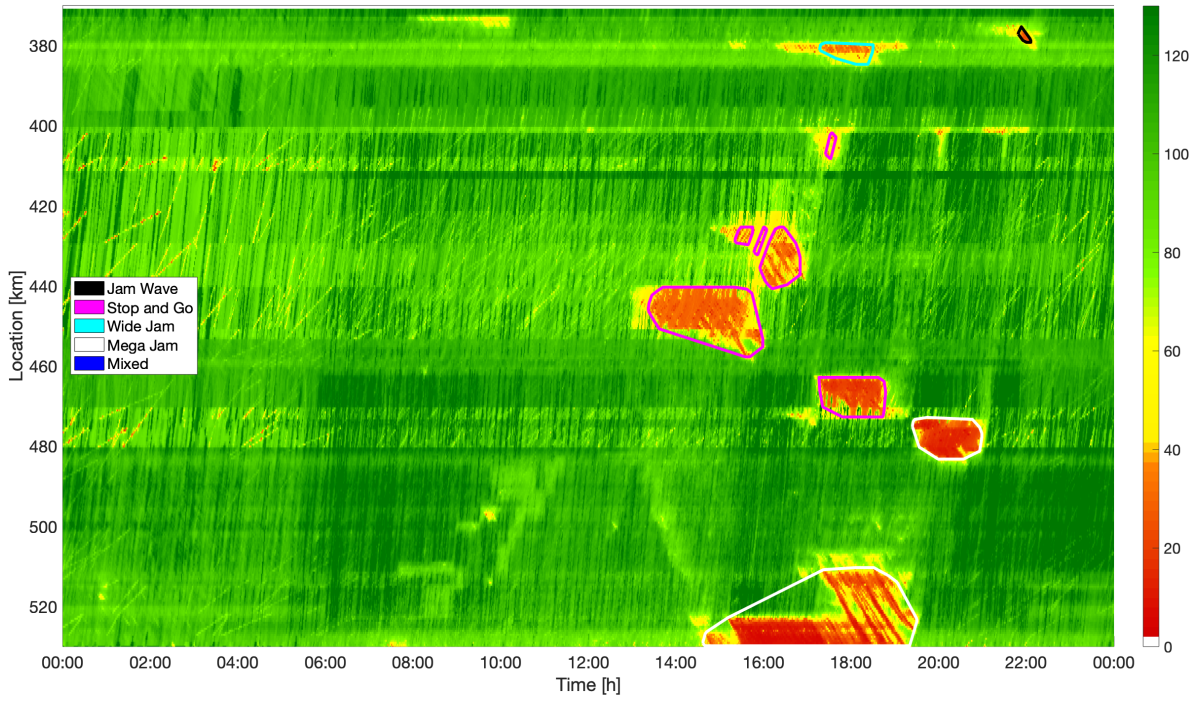


Figure 5.20: Congestion type identification from fused data set: 1 *Jam Wave*, 6 *Stop and Go*, 1 *Wide Jam*, 2 *Mega Jam* clusters

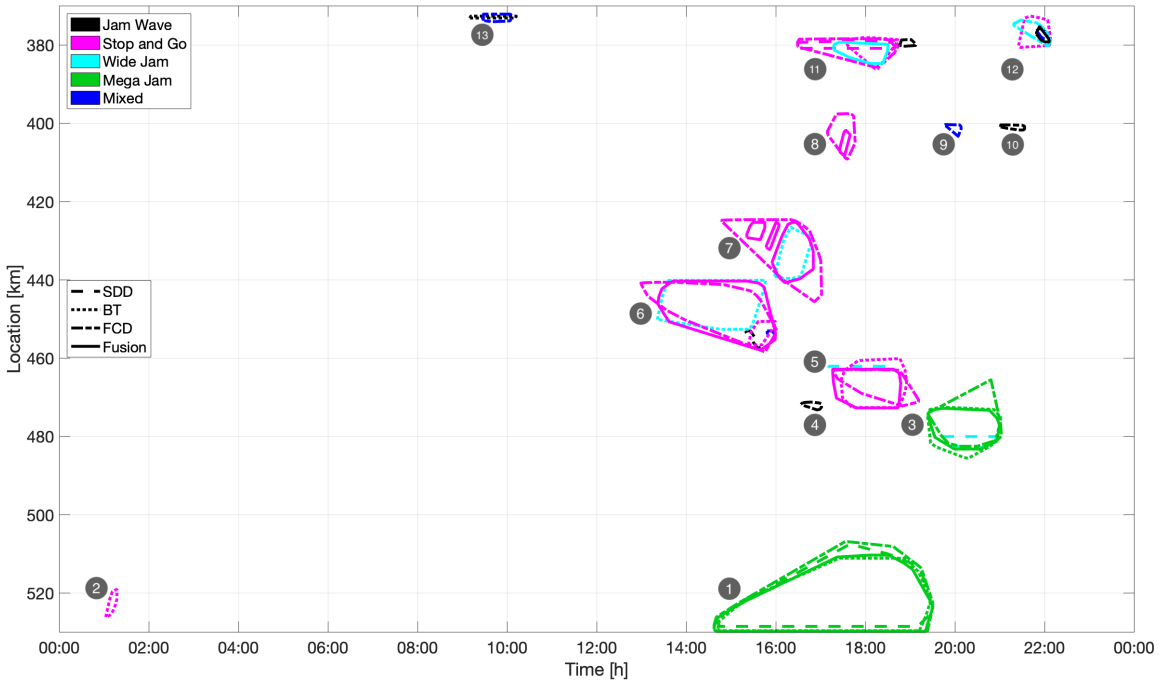


Figure 5.21: Convex hulls of congested areas per data source

Reconstructed BT speeds are rarely suited to detect *Jam Wave* due to their downstream travel time propagation in contrast to upstream congestion propagation. ASM-smoothed gantry speed measurements detect congestion reliably and are able to assign an appropriate type reasonably. FCD prefer congestion type *Stop and Go* over the other types, especially in comparison with BT which is the detection technology with most overlapping locations.

The quality assessment of the recognition strategy also utilizes the area of the convex hull defined by a cluster. Let analogously to section 4.3.2 a cluster k have two properties: a congestion type ct_k and an area A_k. The size of a cluster's area corresponds to the sum of temporal times spatial sizes of all affected cells in the cluster. For example, a congested area of size 150 km·min corresponds to congestion with extensions of 15 min and 10 km or 30 min and 5 km. The size of all identified clusters is summarized in Table 5.9. Additionally, the related region according to Figure 5.21 is indicated.

In regions where data from all detection technologies are available, one *Mega Jam* is identified (region 1). Its size varies between 3,900 (both gantry and BT), 4,600 (FCD), and 4,000 (fusion) km·min. Hence, it is comparable and no systematic bias exists. BT prefers *Wide Jam* over *Stop and Go*, especially visible in the area where neither gantry nor radar sensors are installed. In region 6, the BT *Wide Jam*, conjoint with the small *Stop and Go* cluster directly connected, sums up to an area of size $1528.25 + 141 = 1669.25$, the corresponding FCD *Stop and Go* cluster amounts to 1770.3 km·min. The two jam waves (region 8), recognized as one *Stop and Go* pattern in FCD at 17:00 at km 400, are not detected by gantry sensors (out of their region) and cannot clearly be captured by BT.

To conclude, data sources recognize congestion events in a similar way and therefore are suited to be fused. Detected clusters have a comparable size, however, some sources prefer more and smaller clusters whereas others prefer fewer and larger clusters.

Data Source	Congestion Type	Size [km·min]	Related Region
Gantry	*Stop and Go*	226.75	11
Gantry	*Wide Jam*	158.25	3
Gantry	*Mega Jam*	3933.75	1
Gantry	*Mixed*	23	12
Radar	*Jam Wave*	24.5	6
Radar	*Wide Jam*	85.75	5
Radar	*Mixed*	13.75	6
Bluetooth	*Jam Wave*	30.5	13
Bluetooth	*Stop and Go*	53	2
Bluetooth	*Stop and Go*	960.25	5
Bluetooth	*Stop and Go*	141	6
Bluetooth	*Stop and Go*	317.75	11
Bluetooth	*Stop and Go*	273.25	12
Bluetooth	*Wide Jam*	1528.25	6
Bluetooth	*Wide Jam*	422	7
Bluetooth	*Mega Jam*	3917.25	1
Bluetooth	*Mega Jam*	1001	3
FCD	*Jam Wave*	34.85	4
FCD	*Jam Wave*	31.35	10
FCD	*Jam Wave*	28.6	11
FCD	*Stop and Go*	652.25	5
FCD	*Stop and Go*	1770.3	6
FCD	*Stop and Go*	1462.3	7
FCD	*Stop and Go*	293.2	8
FCD	*Stop and Go*	631.05	11
FCD	*Wide Jam*	150.2	12
FCD	*Mega Jam*	4609.8	1
FCD	*Mega Jam*	1071.65	3
FCD	*Mixed*	34.45	9
FCD	*Mixed*	67.15	13
Fusion	*Jam Wave*	33.3	12
Fusion	*Stop and Go*	824.15	5
Fusion	*Stop and Go*	2007.3	6
Fusion	*Stop and Go*	591.75	7
Fusion	*Stop and Go*	82.35	7
Fusion	*Stop and Go*	40.8	7
Fusion	*Stop and Go*	49.6	8
Fusion	*Wide Jam*	288.2	11
Fusion	*Mega Jam*	4044.95	1
Fusion	*Mega Jam*	818.75	3

Table 5.9: Results of assignment of congestion types and respective cluster sizes, sorted by decreasing location per congestion type, respectively

5.5 Congestion Type Hot Spots

This section describes spatial, temporal, and spatio-temporal hot spots of each particular congestion type. The algorithm explained in section 4.6 is applied quantitatively to the various data sets along the road stretch of autobahn A9. In contrast to section 5.4, gantry and radar measurements are combined and referred to as SDD in the following. All parameter values are identical to those used for the qualitative analysis in the previous section.

5.5.1 Evaluation

Analyzing data from several months in 2019 (see Table 3.3 for an overview), numerous congestion events are identified and classified on NB and Southbound direction (SB). Table 5.10 gives a statistical overview of all identified clusters.

Detection Technology	Driving Direction	Stretch Length (approx.)	Considered Days	Number of Identified Congestion Events
SDD	NB	90 km	150 days	508
BT	NB	160 km	60 days	346
FCD	NB	160 km	41 days	133
Fusion	NB	160 km	3 days	30
SDD	SB	90 km	150 days	841
BT	SB	160 km	60 days	520
FCD	SB	160 km	45 days	271
Fusion	SB	160 km	7 days	61

Table 5.10: Number of identified congestion events

Normalizing the results as a congestion probability (number of congestion events divided by the number of considered days and the stretch length) yields a congestion proneness per day per 10 km (Table 5.11). These values can be interpreted as the probability to experience congestion when driving through a sub-segment stretch of 10 km any time of a day.

Detection Technology	Driving Direction	Normalized Proportion of Congestion Events
SDD	NB	38 %
BT	NB	36 %
FCD	NB	20 %
SDD	SB	62 %
BT	SB	54 %
FCD	SB	38 %

Table 5.11: Congestion proneness per day per 10 km

The SDD values are larger which indicates that gantries are installed in congestion-prone regions of metropolitan areas. Table 5.12 shows the share of each congestion type per data source. Fusion results are not representative because of their collection on too few but congestion-related days.

Detection Technology	Driving Direction	Jam Wave	Stop and Go	Wide Jam	Mega Jam	Mixed
SDD	NB	34.8 %	38.8 %	16.9 %	1.0 %	8.5 %
BT	NB	1.7 %	70.2 %	16.2 %	6.1 %	5.8 %
FCD	NB	14.3 %	57.9 %	12.0 %	6.8 %	9.0 %
Fusion	NB	20.0 %	50.0 %	10.0 %	13.3 %	6.7 %
SDD	SB	41.7 %	32.7 %	17.5 %	1.2 %	6.9 %
BT	SB	2.1 %	76.2 %	8.8 %	5.6 %	7.3 %
FCD	SB	22.9 %	57.9 %	7.0 %	3.3 %	8.9 %
Fusion	SB	19.7 %	49.2 %	14.8 %	6.6 %	9.8 %

Table 5.12: Share of classified events per congestion type

The results show that incident detection via BT systematically underestimates the congestion type *Jam Wave*. The propagation of a jam wave is upstream while in contrast the BT measurement is re-identified downstream because the travel time computation must await its corresponding second scan. False positive *slow* vehicles (especially due to a parking stop between two sensors) create a *virtual congestion*. During night hours, when flow is little, these vehicles falsify the congestion identification because they cannot be averaged by a high share of real-time driving vehicles. Depending on the parameters, two slow vehicles following each other with a certain offset could also cause *Stop and Go* traffic. Hence, the share of *Stop and Go* does not always mirror the real situation. BT data are well suited to recognize *Wide Jam* and *Mega Jam* because of their synchronized speed over all lanes and large extensions of the congested area.

SDD recognize a higher share of *Wide Jam* than the other detection technologies because they are installed in the Greater Munich Area and the Greater Nuremberg Area and thus focus on regions with a large potential of congestion at all. The stretch between the metropolitan areas is less congestion-prone. In case, congestion emerges there, it is more likely to be a larger event such as *Mega Jam*. Therefore, BT and FCD, which both detect along the entire stretch, average the share of their detected *Jam Wave* events over all congestion types. The method ASM, that smoothed SDD, interpolates in its congestion kernel direction and thus prefers *Jam Wave* over the other types. PSM and LTSM-interpolation do not utilize any neighbored cells' speed information which more likely yields either no jam or *Wide Jam* and *Mega Jam* events.

In SB direction, congestion types *Jam Wave* and *Stop and Go* occur more frequently than in the opposite direction. Larger congestion events such as *Wide Jam* and *Mega Jam* are often detected in NB direction. Reasons are the impact of vacation traffic along the main route of the Bavarian highway network and large-scale events such as soccer games. The relationship of a computed congestion type and the cause of the respective congestion is left for future research (section 6.2). Unclassified congestion clusters (with type *Mixed*) have a share of

approximately 6 to 9 % of each data source.

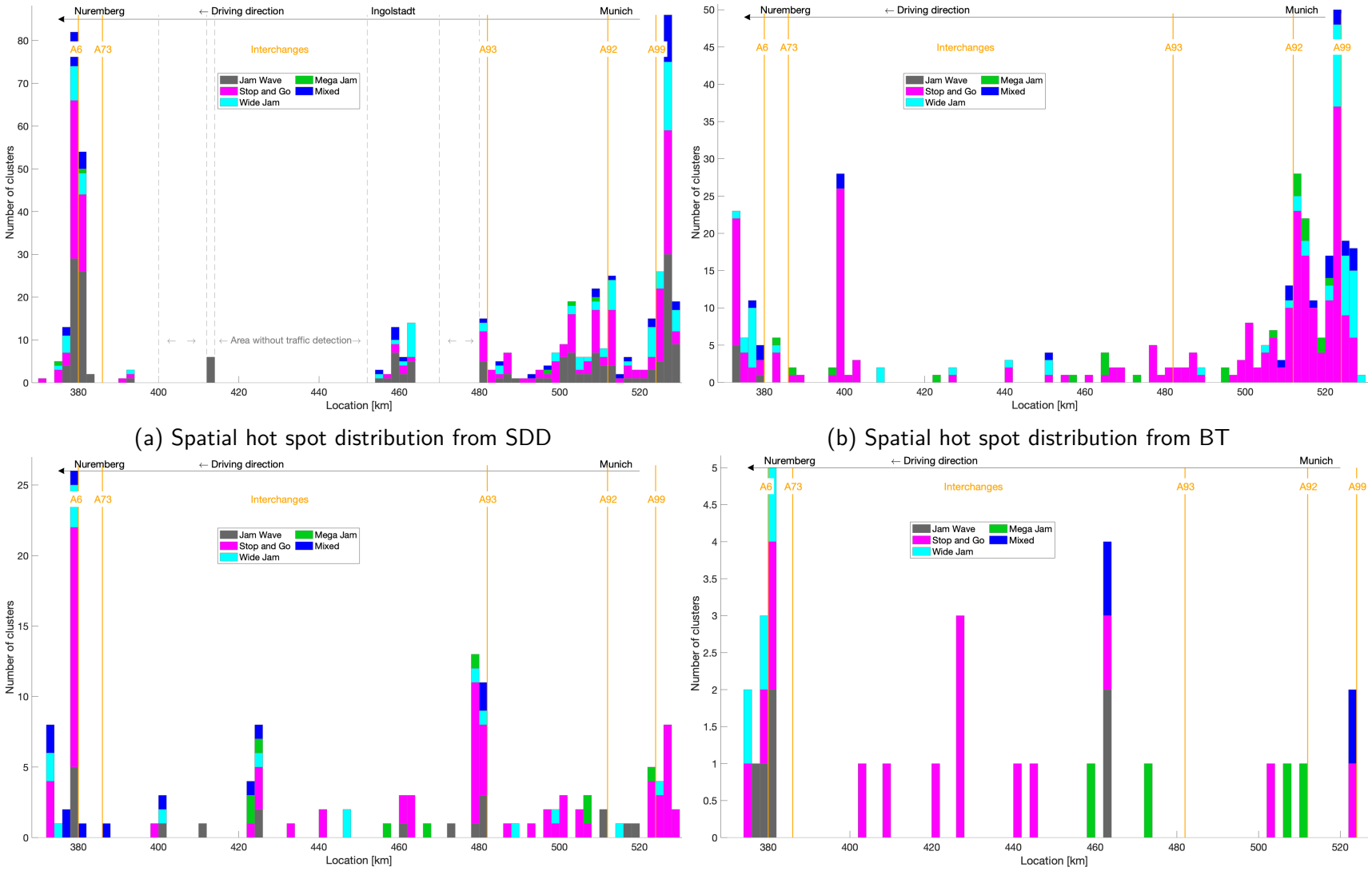
The following Figures 5.22 – 5.27 show the spatial, the temporal, and the spatio-temporal pattern accumulations in NB and SB driving direction, respectively. Each figure is divided into four single plots for SDD, BT, FCD, and the fused data set, respectively. The fusion plot is not taken as a reference because the amount of available data is very low. But at least all identified congestion events are illustrated to discuss the results qualitatively.

Figures 5.22 and 5.23 depict the spatial congestion type hot spots in NB and SB direction, respectively. For a better illustration, all spatial starts are categorized into discrete bins of 2 km. Upstream of large interchanges, congestion occurs frequently. In NB direction, FCD (Figure 5.22c) recognize congestion fronts at particular locations only, focused on stretches in metropolitan areas. In SB direction, spillback caused by both the autobahn A99 (SDD) and the Munich city highway (at km 529, FCD) are often reasons for congestion emergence. In SB direction, the test field starts at slightly different positions over all data sources, therefore SDD recognize another congestion hot spot between km 370 and km 375 (Figure 5.23a), which is not represented in BT and FCD data sets.

BT data recognize a large share of *Stop and Go* upstream of the interchange of autobahn A99 in SB direction with varying starting locations of congestion. Indeed, this is a hidden *Lane Jam* because two lanes turn to A99 (tend to be congested) whereas three lanes continue on A9 (tend to be uncongested). Depending on vehicles on either lanes, which influence the averaged speed of the discretized cell, the velocity is below the threshold v_{crit} or not, yielding *Stop and Go*. The start locations of congestion slightly deviate between SDD, BT, and FCD due to different sensor positions and different data processing techniques such as another spatial grid structure during the interpolations. Moreover, BT sensors are influenced by their entire environment. If placed between a ramp and the main lanes, congestion is registered earlier or later than that measured by induction loops or sensors at gantries.

The temporal hot spots are illustrated in Figures 5.24 and 5.25, again in NB and SB direction, respectively. Similarly, for a better visualization, data are discretized in intervals of 30 min. Two typical peaks occur in SB in all data sets, essentially due to the detected *Stop and Go* traffic. In the opposite direction, mainly one peak is recognizable in SDD and BT data, FCD result in a late morning peak hour. This comes from the fact that the FCD data set detected fewer congestion events at all (due to the smaller time period; note the value range of the vertical axis) and that fleet drivers are supposed to typically drive at certain times only. A more detailed analysis of the temporal distribution of congestion including holidays, days of week, or monthly/quarterly oscillations would help to interpret the results more qualitatively.

The share of *Stop and Go* traffic is significantly increased in the morning between 6:00 and 8:00 in SB while in NB congestion formation generally starts later than 8:00. *Wide Jam* events are evenly distributed throughout the day, only during the rush hours (9:00-10:00 and 16:00-18:00), they are more distinctive in both driving directions. During night hours (late evening and early morning), BT data detect more jams than the others because of fewer flow and greater impact of long travel times than during the day, as already discussed earlier. Additionally, BT data lack of an more advanced data preprocessing step reducing congestion occurring during night hours, especially in the early morning hours before 6:00. Since these times are not relevant for traffic planning or control, the data filtering is skipped in this analysis.



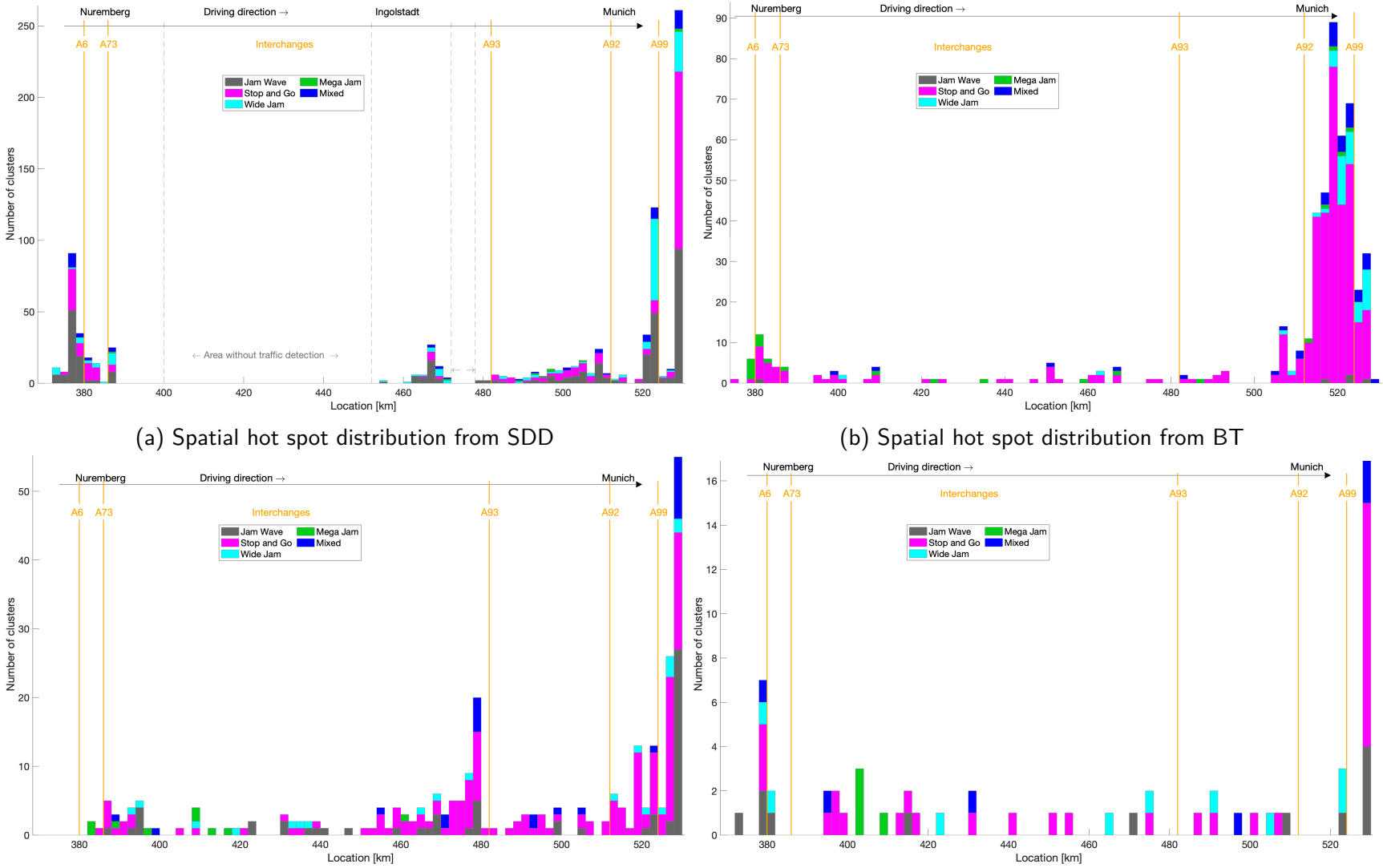
(a) Spatial hot spot distribution from SDD

(b) Spatial hot spot distribution from BT

(c) Spatial hot spot distribution from FCD

(d) Spatial hot spot distribution from fusion

Figure 5.22: Spatial congestion type hot spots in NB direction



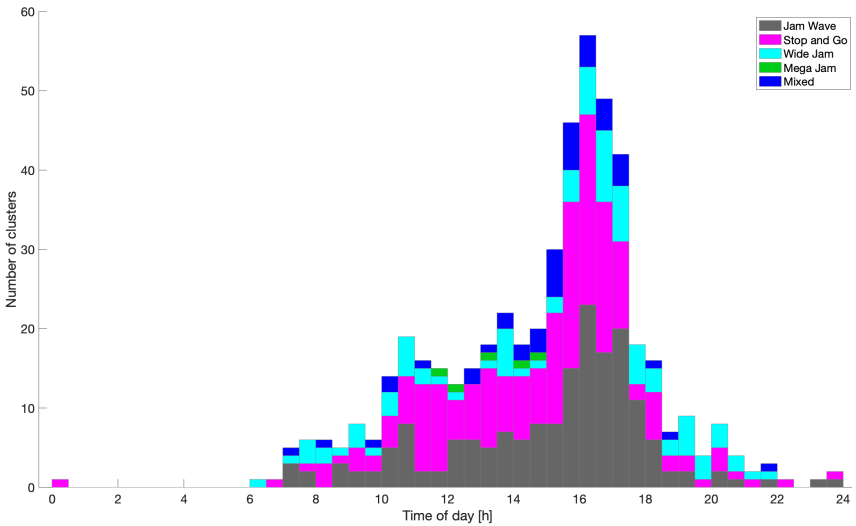
(a) Spatial hot spot distribution from SDD

(b) Spatial hot spot distribution from BT

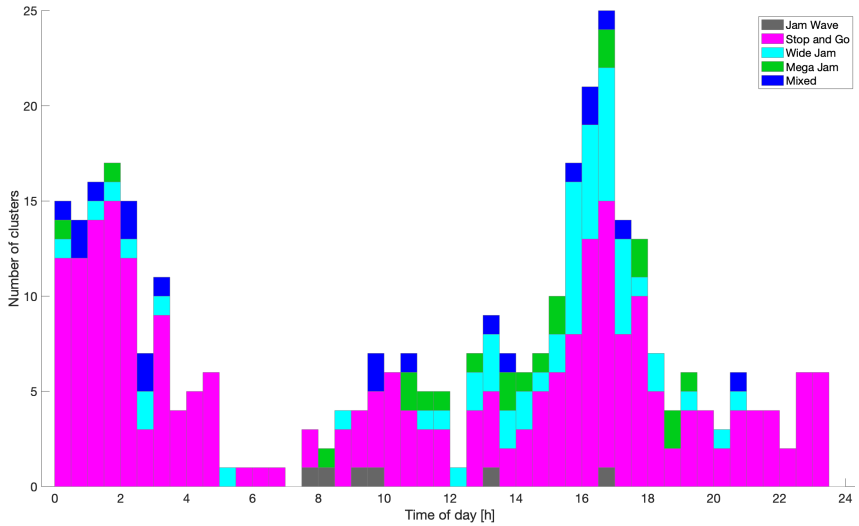
(c) Spatial hot spot distribution from FCD

(d) Spatial hot spot distribution from fusion

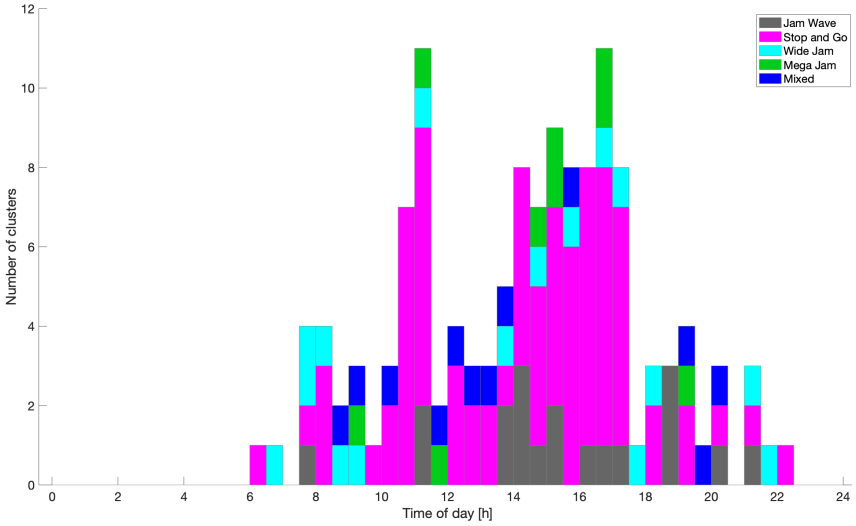
Figure 5.23: Spatial congestion type hot spots in SB direction



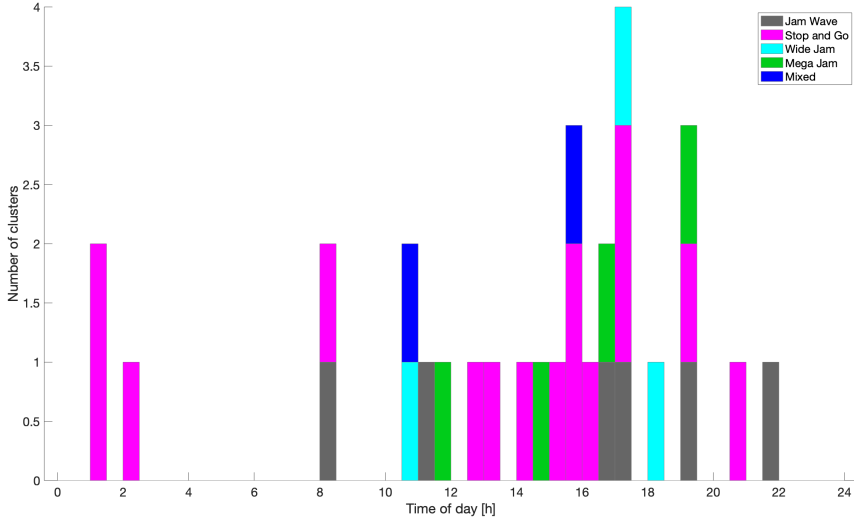
(a) Temporal hot spot distribution from SDD



(b) Temporal hot spot distribution from BT

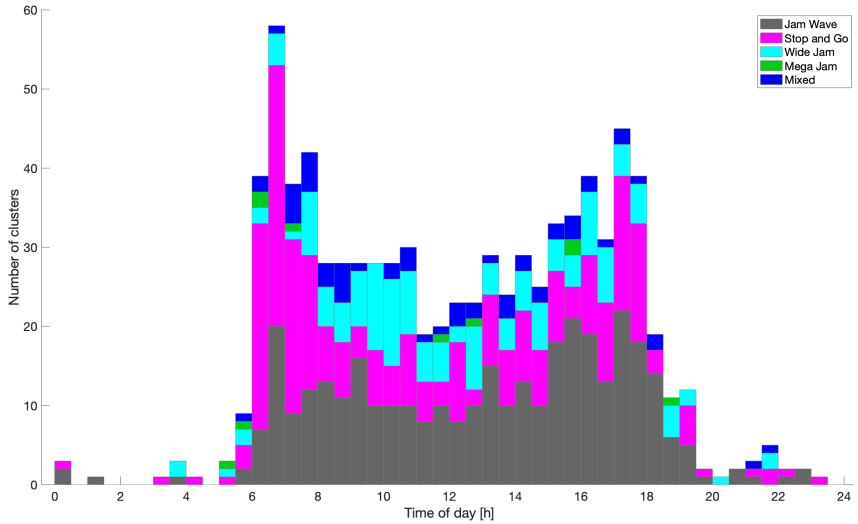


(c) Temporal hot spot distribution from FCD

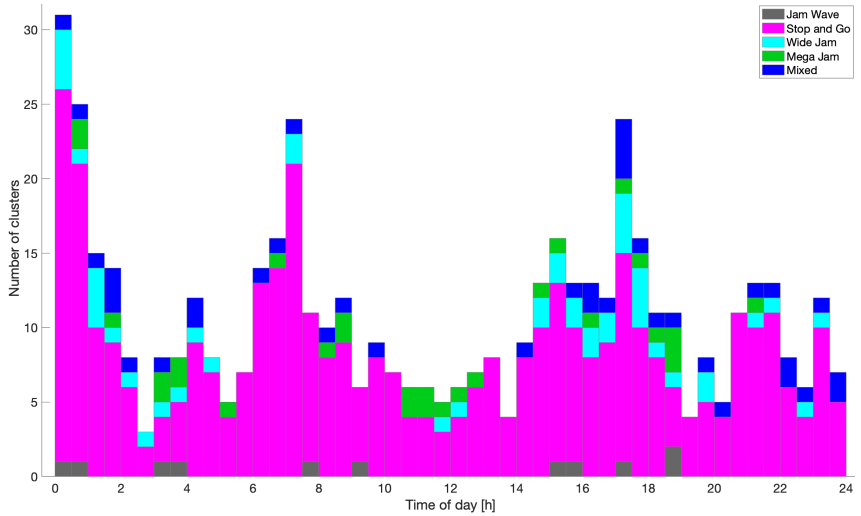


(d) Temporal hot spot distribution from fusion

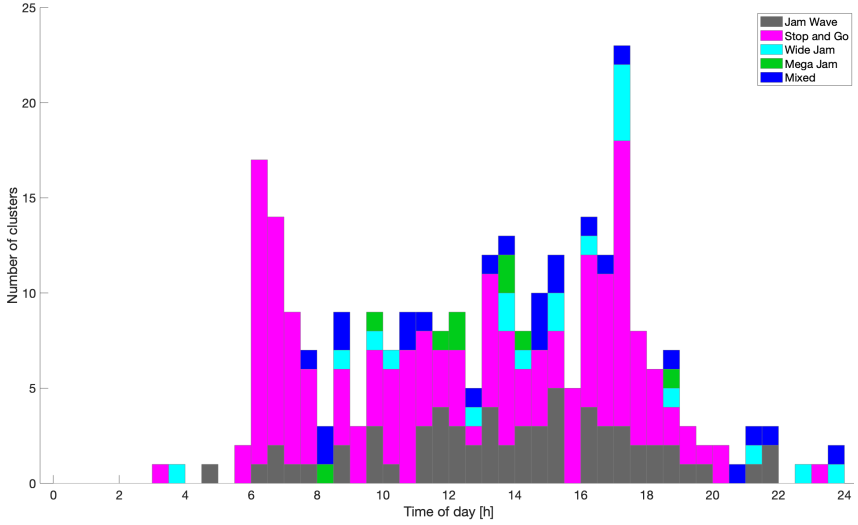
Figure 5.24: Temporal congestion type hot spots in NB direction



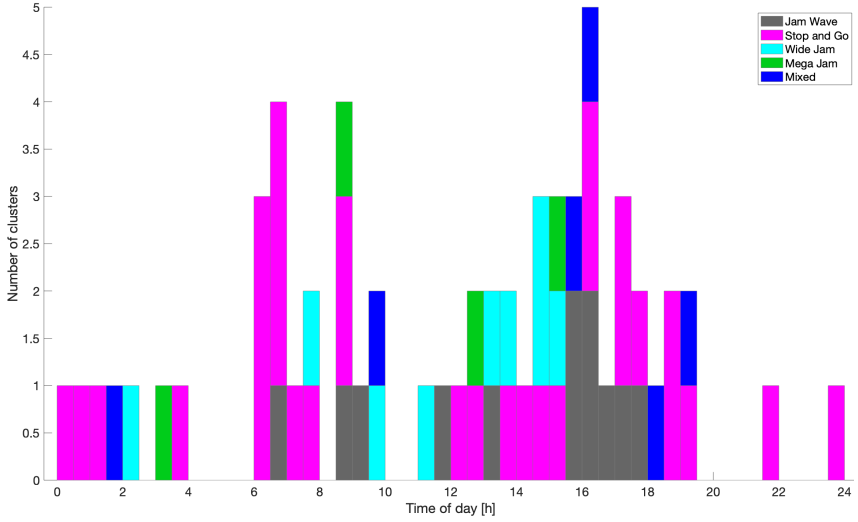
(a) Temporal hot spot distribution from SDD



(b) Temporal hot spot distribution from BT

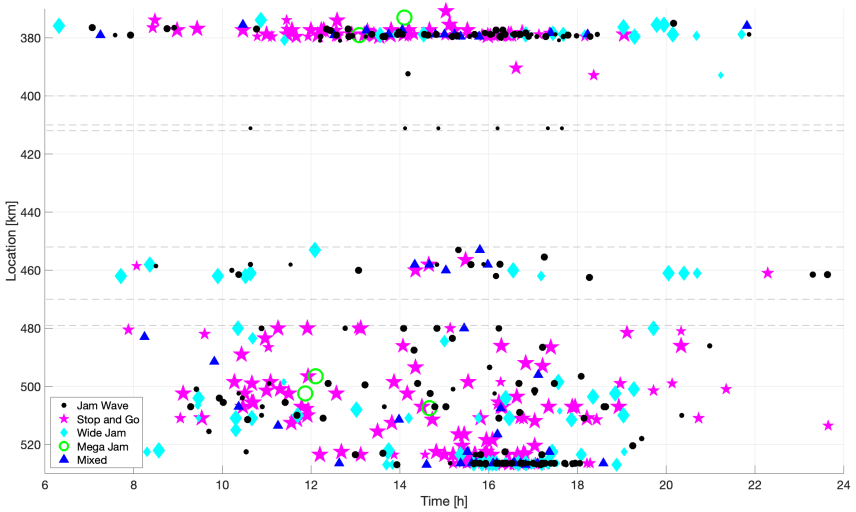


(c) Temporal hot spot distribution from FCD

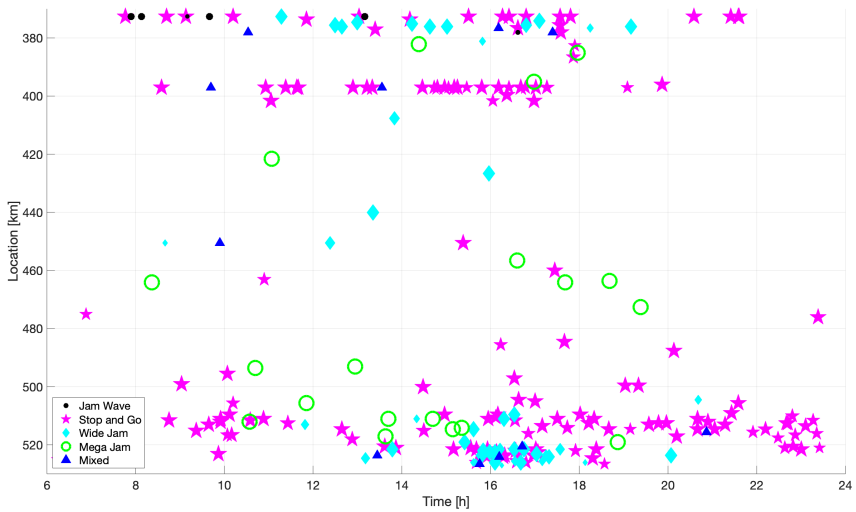


(d) Temporal hot spot distribution from fusion

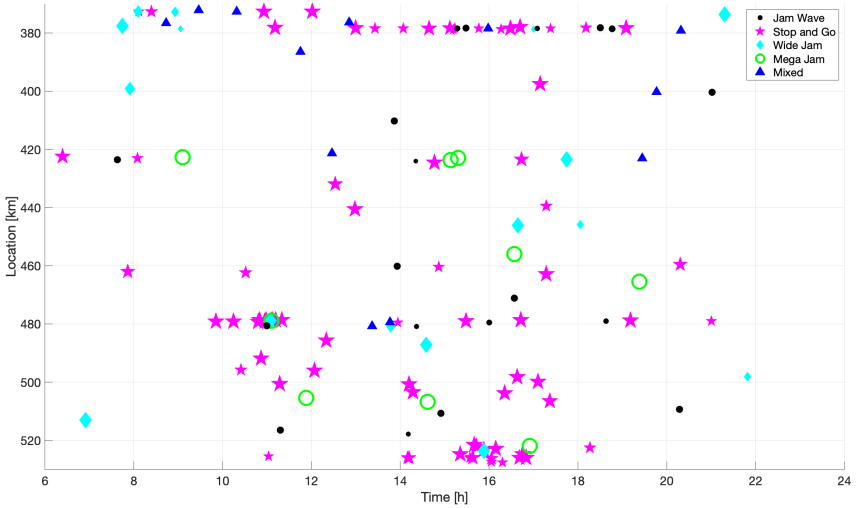
Figure 5.25: Temporal congestion type hot spots in SB direction



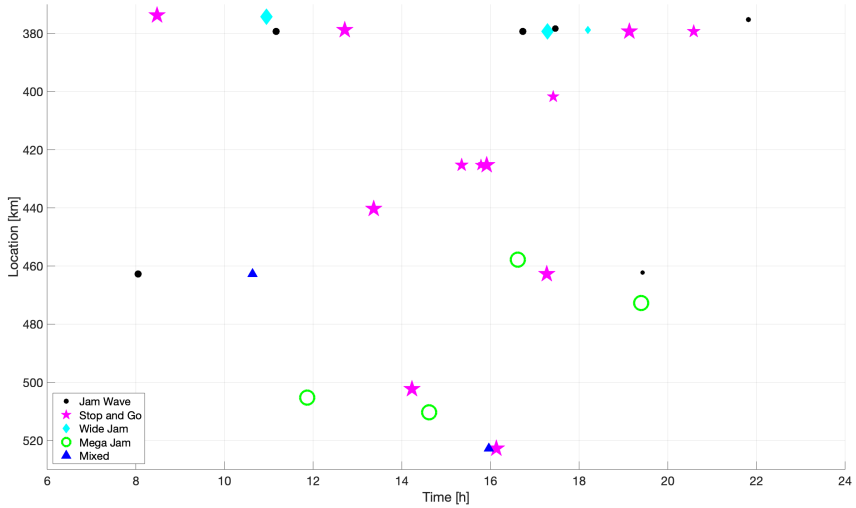
(a) Spatio-temporal hot spot distribution from SDD



(b) Spatio-temporal hot spot distribution from BT

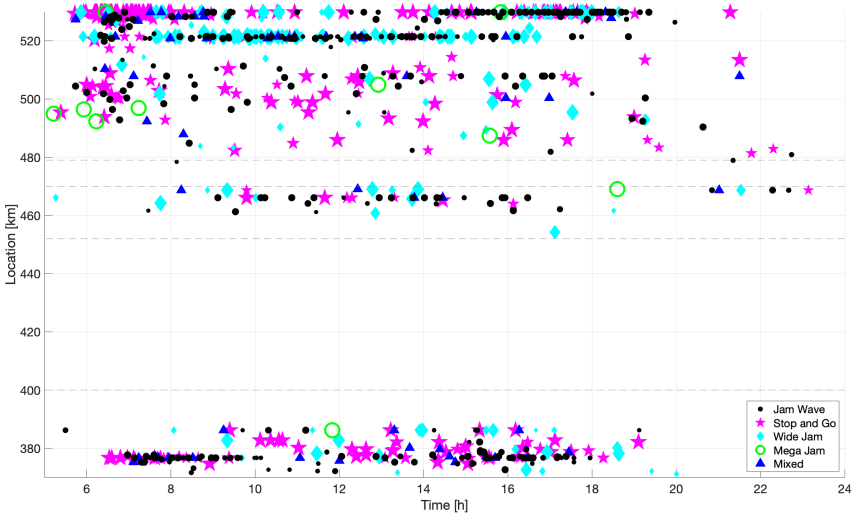


(c) Spatio-temporal hot spot distribution from FCD

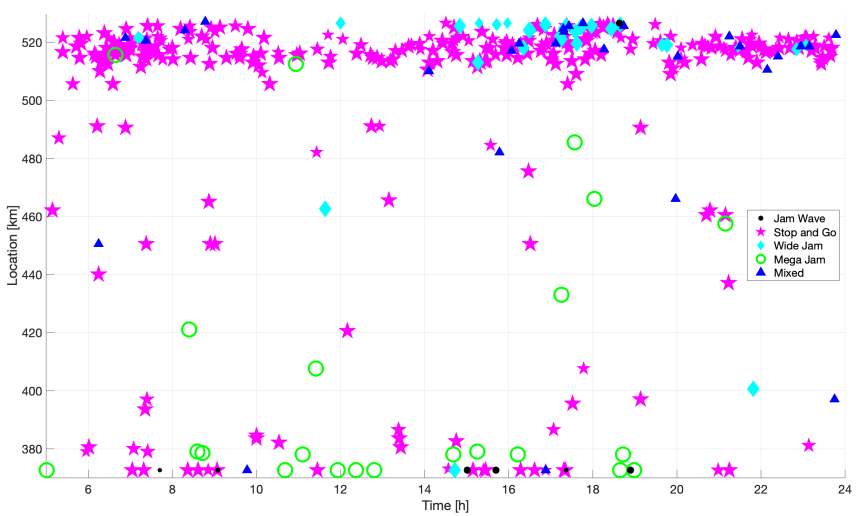


(d) Spatio-temporal hot spot distribution from fusion

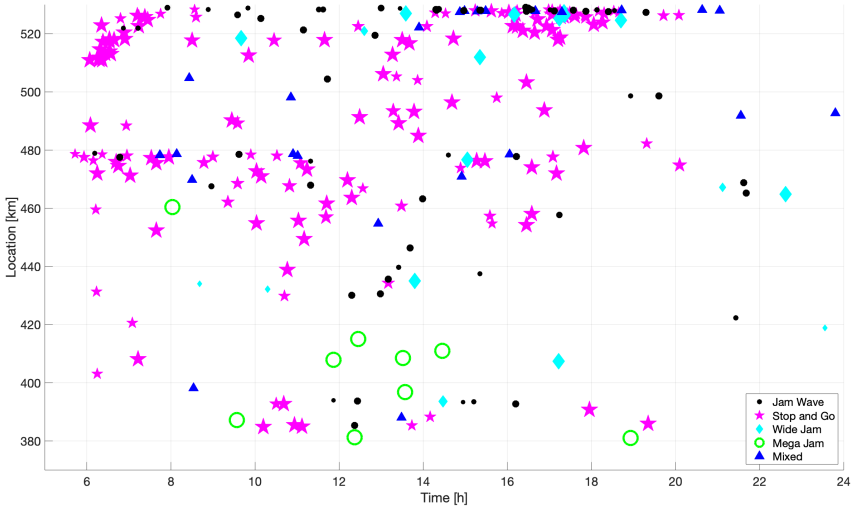
Figure 5.26: Spatio-temporal congestion type hot spots in NB direction



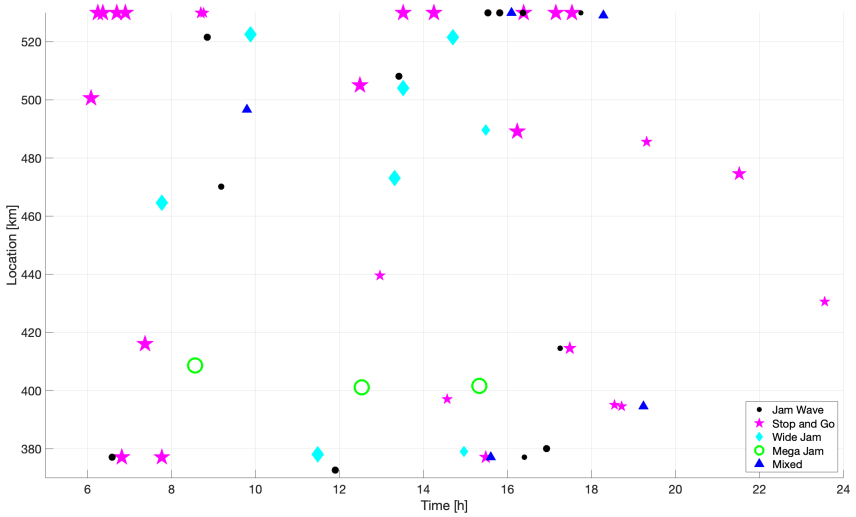
(a) Spatio-temporal hot spot distribution from SDD



(b) Spatio-temporal hot spot distribution from BT



(c) Spatio-temporal hot spot distribution from FCD



(d) Spatio-temporal hot spot distribution from fusion

Figure 5.27: Spatio-temporal congestion type hot spots in SB direction

Both information, spatial and temporal, is united to spatio-temporal hot spots which are illustrated in Figures 5.26 and 5.27 for NB and SB direction, respectively. Hereby, all locations and starting times of occurring congestion events are not rounded but instead taken from the original grid. The plots in NB direction are shown bottom-top, identical to the time-space orientation of the speed contour plots shown before. This way, all congestion events are marked at their temporal and spatial beginning, the upper left corner of the corresponding convex hull. As marker sizes, each congestion size averaged by the median of the respective congestion type sizes is given. It helps to understand where heavy hot spots are located to support local traffic operators.

Stop and Go traffic in the BT data set appears very dominant between km 500 and km 520 in both directions, especially in SB direction. As explained earlier, this comes from the hidden *Lane Jam* and spillback both from A92 and A99, respectively. In NB direction, a *Stop and Go* traffic hot spot exists around km 380 throughout all data sets. Some deviations of the exact location implies either a data calibration inaccuracy or an unequal density of the traffic detection. Furthermore, it is visible that SDD, although not detecting the entire stretch, catch most of occurring congestion because the unequipped region is not congestion-prone. In SB direction, several *Stop and Go* occurrences are detected between 6:00 and 8:00 and later on during the afternoon around km 520. Again, it comes from the spillback of autobahn A99, onto which two out of five lanes exit from autobahn A9.

In SB direction, congestion type *Jam Wave* is not distributed all throughout the grid, but restricted to a few locations (km 375-380 (SDD), km 520 (SDD), and km 525 (SDD and FCD)). In particular, *Jam Wave* clusters detected by SDD and FCD in the afternoon in NB direction at km 525 correspond to a *Wide Jam* hot spot recorded by BT data. Especially in the Nuremberg area (km 375-380), *Jam Wave* events are not limited to peak hours, but occur during all times of day. In contrast to NB direction, the data set of SDD misses several heavy congestion events in SB direction due to the unequipped regions compared to FCD. BT data and FCD recognize a *Mega Jam* accumulation in the Greater Nuremberg Area.

5.5.2 Conclusion

To conclude, congestion hot spots are mainly identified upstream of interchanges and larger exit ramps. SDD recognize *Jam Wave* events most frequently, BT and FCD *Stop and Go* traffic. BT detection underestimates type *Jam Wave* and overestimates *Mega Jam*. Hot spots identified by FCD as *Stop and Go* are frequently matched to a cluster of *Wide Jam* in the BT data set. Especially in SB direction, both SDD and FCD detect more congestion events than BT in the area between Ingolstadt and Munich (km 460 to 500).

5.6 Lane-Specific Congestion

Section 4.4.2 introduced a novel congestion type called *Lane Jam*. It describes situations where congestion does not affect all lanes equally. The novelty in this thesis is to perceive it as congestion clusters per lane. The union of intersecting clusters is called a *region*. *Lane Jam* is defined as region where the share of at least one lane is at most a maximum percentage p_{max}.

This indicates that the union and the intersection of clusters in the region differs. Otherwise, it is said to be mainline congestion.

In this evaluation, *Lane Jam* is detected on lane-by-lane speeds coming from SDD solely, that is gantry and radar measurements. Both BT and FCD do not detect lane-specifically and are therefore not able to generate lane-by-lane speed information. The detection of *Lane Jam* based on non-lane-specific velocities is not tackled within the scope of the dissertation. Thus, comparisons with other data sources or the fusion are not treated here.

In the following, the results from the *Lane Jam* analysis are presented based on data from five months. The parameters are set as given in Table 5.13.

Parameter	**Value**
Velocity threshold v_{crit}	40 km/h
Free-flow speed v_{freeflow}	120 km/h
Minimum free-flow time between congested areas t_{merge}	4 min
Temporal offset of starting trajectories t_r	5 min
Minimum size of regions A_{min}	12 km·min
Maximum percentage of lanes in region p_{max}	0.7

Table 5.13: Parameter values to identify *Lane Jam*

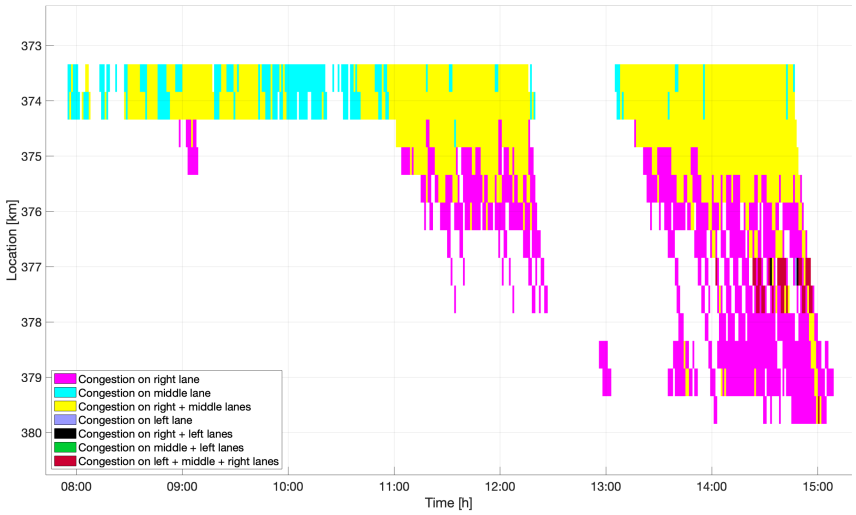
5.6.1 Evaluation

Figure 5.28 depicts three examples of congested regions. In Figure 5.28a, congestion emerges from the middle lane with a stationary bottleneck. It influences the right lane, which then in turn is affected by congestion. Several hours later, congestion moves upstream to the right lane only. The left lane is not affected by congestion in this region except from a few minutes. Figure 5.28b depicts congestion mainly on the middle and left lanes with whiskers on the left lane only. The middle lane is affected upstream one kilometer more (2 segments) than the other lanes. The third example (Figure 5.28c) shows a region starting with congestion on the right lane, extending it to the middle lane and ending up with congestion on all three lanes. On the left lane, the first spatial occurrence is one kilometer upstream than that of the other lanes.

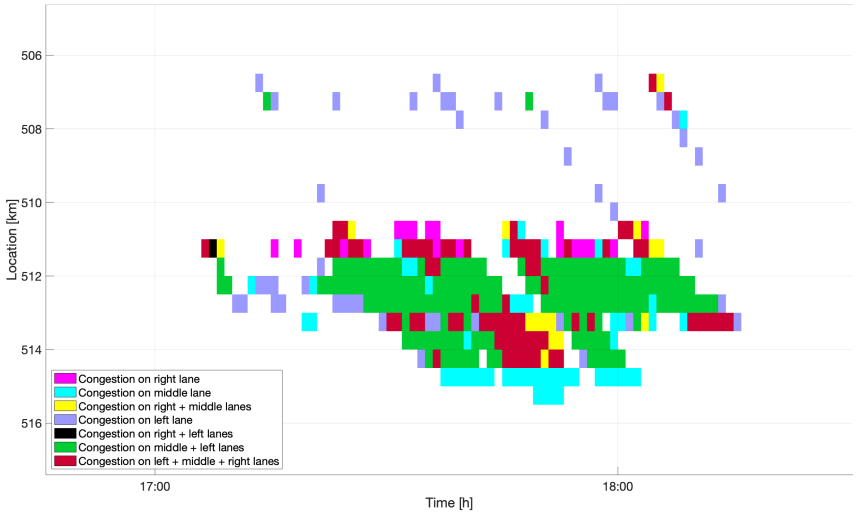
Share of Congestion per Lane

In total, 480 and 842 regions were found in NB and SB direction, respectively. The average share per lane over all regions is computed as given in Table 5.14. The mean and median values indicate the percentage of the area per lane over all regions. Additionally, the share of lane-X-only congestion events is given.

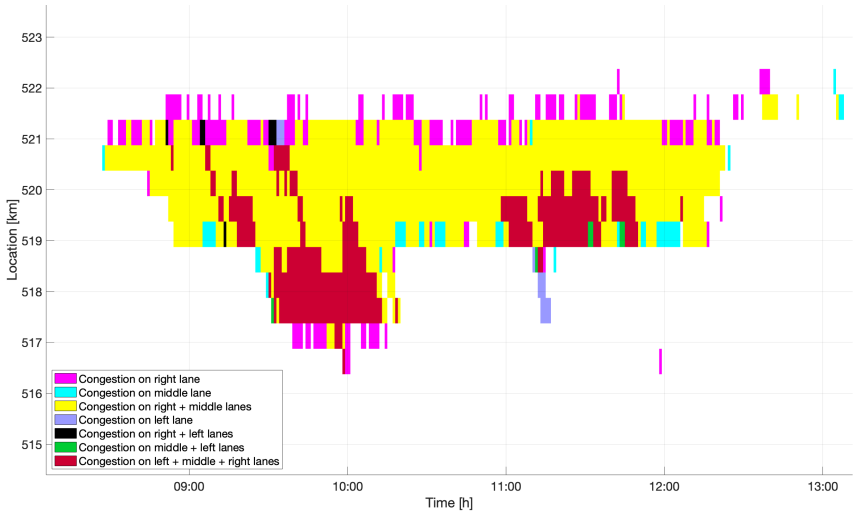
The right lane has a significant proportion on each resulting region. With a median of more than 80 % in each driving direction, it shows that the right lane is frequently concerned by *Lane Jam* and by congestion at all. In NB direction, both other lanes are also often affected by congestion. In SB direction, the emergence of congestion on the right lane dominates over



(a) Congested cells on April 16 in NB direction



(b) Congested cells on May 10 in NB direction



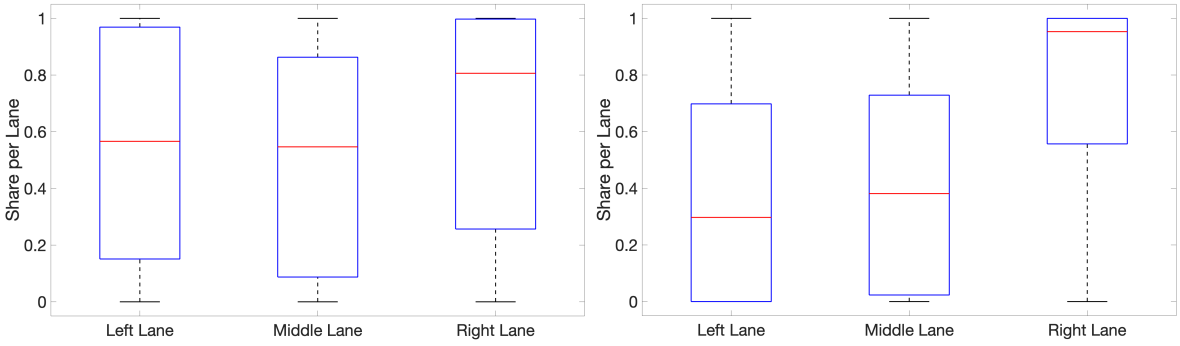
(c) Congested cells on July 14 in SB direction

Figure 5.28: Three examples of *Lane Jam* regions

	Left Lane	Middle Lane	Right Lane
Mean share (NB)	54 %	50 %	63 %
Median share (NB)	57 %	55 %	81 %
Share of lane-X-only congestion events (NB)	10 %	1 %	8 %
Mean share (SB)	39 %	41 %	74 %
Median share (SB)	30 %	38 %	95 %
Share of lane-X-only congestion events (SB)	6 %	1 %	14 %

Table 5.14: Average share per lane in *Lane Jam* regions

the middle and left lanes. The average share of 40 % implies that a SB region describes congestion on mainly the right lane. In contrast, 10 % and 6 % of all congestion events in NB and SB direction, respectively, are left-lane-only congestion occurrences. In total, one fifth of all congestion events (19 % and 21 % in NB and SB driving direction, respectively) concern exactly one lane. Figure 5.29 depicts the detailed statistical distribution as a box plot per lane.



(a) NB direction (b) SB direction

Figure 5.29: Average share per lane in *Lane Jam* regions (median values in red, 25 and 75 % quartiles in blue)

Primary Lane Congestion Detection

In each region, the clusters from either lane have a first temporal and a first spatial occurrence. In order to compare the starting points on each lane, the union of all clusters within the region per lane are taken. Also, only congestion events are considered where at least two lanes are affected. This means that the identified lane-X-only jams – 19 % and 21 % (Table 5.14) of all regions in NB and SB direction, respectively – are not included in the average considerations. The basis form 390 and 665 regions in NB and SB direction, respectively.

The temporal delay t_{diff} (equation 4.12) and the spatial offset x_{diff} (equation 4.13) denote the difference between the lane detecting congestion earliest and all other lanes. Table 5.15 presents an overview of all mean and median values for t_{diff} and x_{diff} per direction. The results show that mostly the right lane detects congestion first. The middle lane experiences a speed breakdown mainly at the same congestion front (kilometer) but few minutes later than the

right lane. On the left lane, congestion also starts at the same location but some more minutes later than on the middle lane. During the emergence of congestion, it appears to be best to drive on the middle or left lane where speeds are supposed to be slightly higher than on the left lane (at least above v_{crit}). In SB direction, the delay times are even larger for the middle and the left lanes than in NB direction. The spatial offset median amounts to zero kilometer in both directions for each lane.

	Left Lane	**Middle Lane**	**Right Lane**
Mean delay (NB)	5.6 min	3.7 min	1.6 min
Median delay (NB)	2 min	1 min	0 min
Mean offset (NB)	0.4 km	0.5 km	0.5 km
Median offset (NB)	0 km	0 km	0 km
Mean delay (SB)	9.3 min	6.0 min	1.7 min
Median delay (SB)	3 min	2 min	0 min
Mean offset (SB)	0.9 km	0.7 km	0.3 km
Median offset (SB)	0 km	0 km	0 km

Table 5.15: Average delay and offset per lane in *Lane Jam* regions

Least Travel Time

The average travel times in each region are determined per lane. This means, each region R is embedded in a free-flow environment and virtual trajectories pass the entire region with their dedicated speeds per lane. To this end, virtual trajectories start with an offset of $t_r = 5\,\text{min}$ (Table 5.13). The average travel time over all trajectories of one lane through the region is computed. The average travel times per lanes over all regions are given in Table 5.16. To compare travel times from different regions with varying sizes, the considerations always include the entire road stretch. The uncongested travel times amount to 78.83 min and 79.84 min in NB and SB direction, respectively. The deviation comes from different detector locations per direction, in total 2 km.

	Left Lane	**Middle Lane**	**Right Lane**
Mean travel time (NB)	80.6 min	80.7 min	81.2 min
Median travel time (NB)	80.0 min	80.2 min	80.6 min
Mean travel time (SB)	81.5 min	81.7 min	82.7 min
Median travel time (SB)	80.8 min	80.9 min	81.9 min

Table 5.16: Average travel time per lane in *Lane Jam* regions

The results show that the right lane is usually the slowest lane. Compared to the uncongested situation, the average travel times are only slightly increased. The reason for this is that virtual trajectories are created for every region. These trajectories cross the congestion cluster(s) of

each lane at congested speed but also areas where only the other lanes are congested and the currently considered lane is not in congestion. For these trajectories, free-flow speeds are assumed, which causes a high share of uncongested travel times influencing the averages in Table 5.16. Still, the tendency of travel time increase compared to the free-flow situation is recognizable.

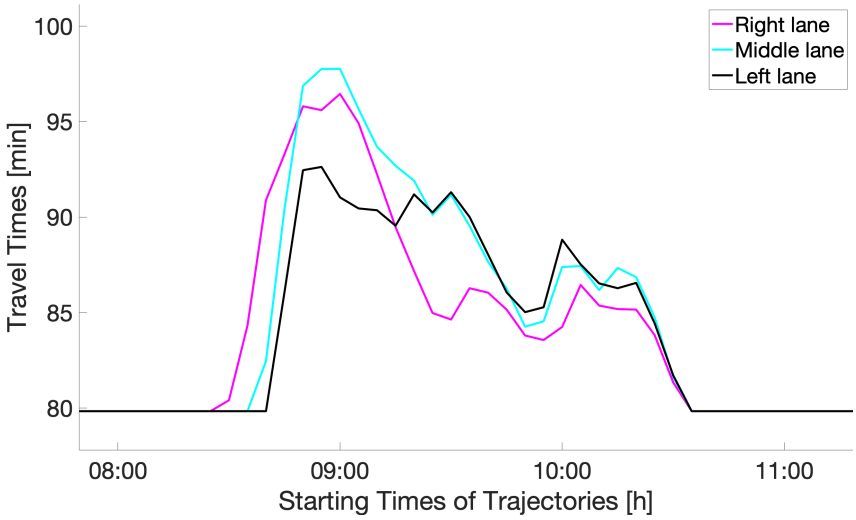
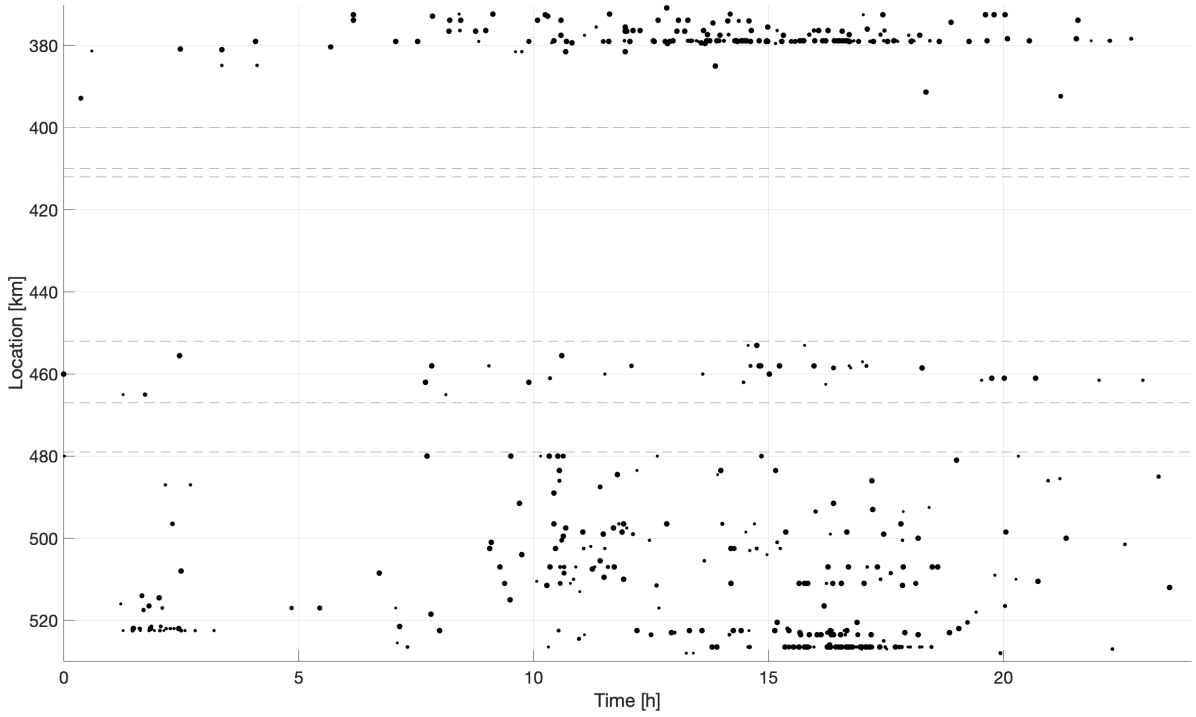


Figure 5.30: Typical travel time increase per lane during emergence of congestion

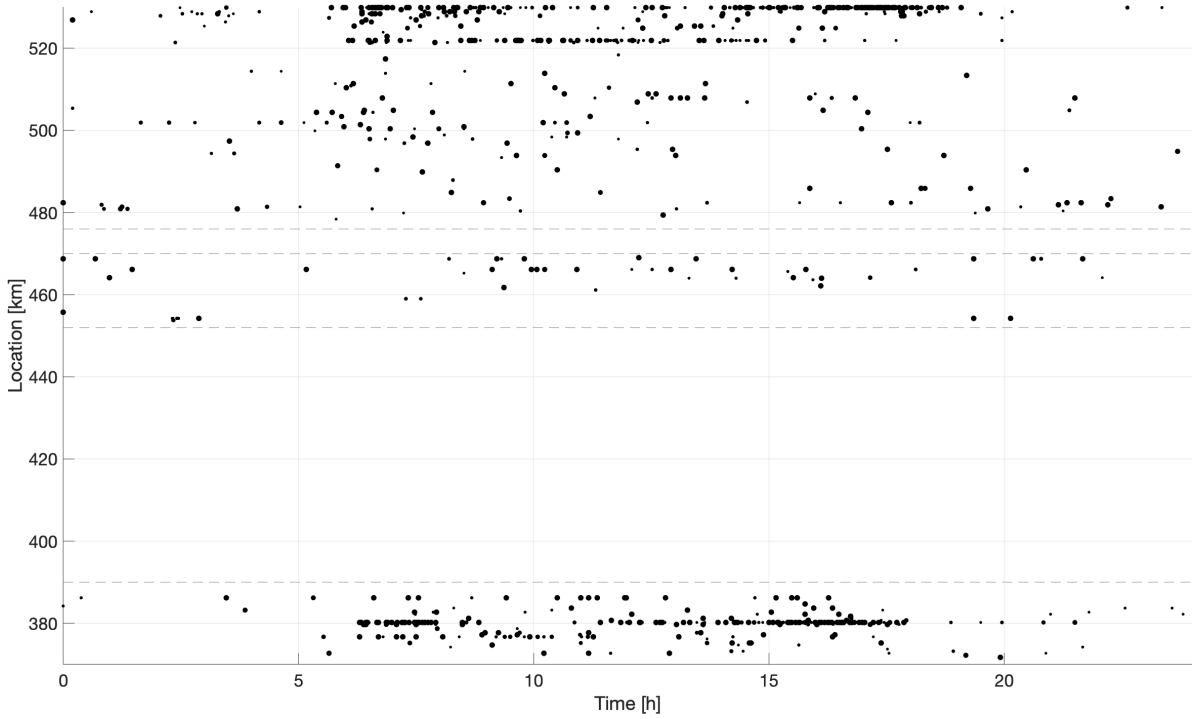
A typical process of a congestion propagation behaves like illustrated in Figure 5.30, an example from June 23, 2019 in SB direction. This plot shows the travel times of each trajectory starting on that day. The velocity threshold v_{crit} is first reached on the right lane, some minutes later also on the middle lane and again some minutes later on the left lane. When all lanes are congested, the travel times are highest on the middle lane. On the right lane, congestion resolves relatively fast after its emergence and the travel time increase keeps a low level. Middle and left lane are more likely affected by congestion after the first third. When it resolves, all lanes are affected equally in a relatively short time period. In this example, the middle lane is mostly affected by congestion at all, considering the integral over the line curve. It appears that if known in which phase of congestion one is, it was most efficient to take the left lane for the first third of congestion and then to move to the right lane to lose minimum travel time. Surely, the congestion phase is not known a priori, each congestion behaves differently, and lane hopping involves other challenges than just a consideration of travel time loss minimization ex-post, e.g. traffic safety issues. Also, negative effects for the other users with regard to the system optimum should be taken into account.

Hot Spots

All starting points of regions are compared in a spatio-temporal area. Again, the symbol size is correlated to the size of the region. *Lane Jam* hot spots are illustrated in Figure 5.31. The results in NB direction show an accumulation at the exit ramps of interchange A99 (km 522) during the afternoon. In the remaining Munich area, no special concentration can be found except from a higher share during late morning and afternoon.



(a) NB direction



(b) SB direction

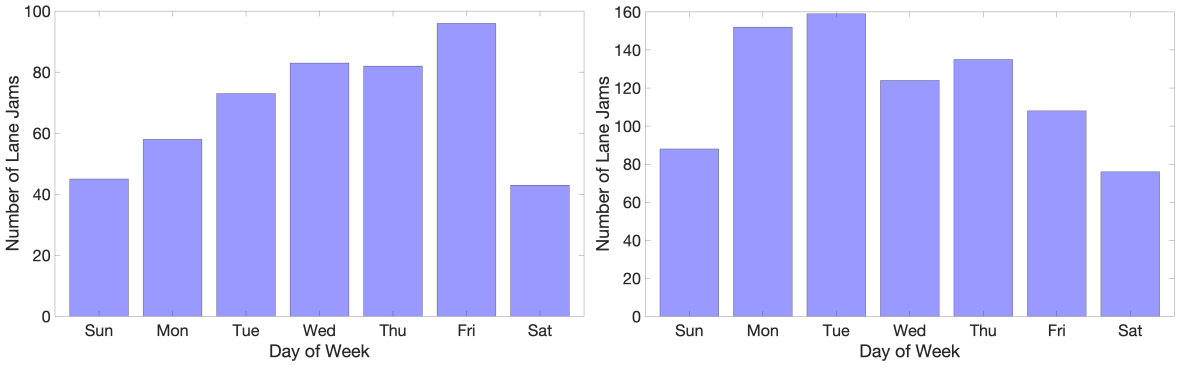
Figure 5.31: Spatio-temporal hot spots of *Lane Jam* regions

In the Ingolstadt area (around km 460), predominantly smaller sizes of region areas occur, spatially due to the smaller equipment area but temporally not limited. In the Nuremberg area (km < 400), a hot spot occurs at the interchange of A6 (km 380) and further on towards the interchange of A3 (end of the test field at km 372). There, the right lanes continue the A9 whereas the left lanes merge onto the A3 causing heavy lane changes. Both hot spots, at A99 and A6, appear also on mainline congestion detection, see Figure 5.22.

In SB direction, data from the Nuremberg area show the identical hot spot at km 380 (interchange of A6). Albeit the equipped area is smaller compared to NB, more congestion regions are found. In the Ingolstadt and the Munich area, some *Lane Jam* events occur during night hours. These come from a construction site and their preparations such as road markings or the installation of traffic cones, executed while traffic flow is little.

In the Munich area, two significant hot spots occur. One is again the interchange of A99 at km 522. There, two out of five lanes leave the A9 and merge onto the A99. During the data collection period, a large construction site was installed on this exit ramp. The lane number on the ramp decreased from two to one and caused a wide spillback. This causes different speeds on the left/middle and the right lane and the emergence of *Lane Jam*. The condensation occurs between 6:00 and 15:00, and hence, does not arise on late afternoon and evening. This hot spot was also discussed in the mainline hot spot section 5.5 where BT data interpreted these congestion events as *Stop and Go* traffic which was said to be indeed a hidden *Lane Jam*.

The second hot spot is located at the end of the autobahn, at km 529. At the very end of the freeway stretch, two lanes are split up onto the Munich city highway: one goes to the Westbound, the other to the Eastbound direction. Depending on the traffic volume on that highway, either the Westbound or the Eastbound lane of A9 are congested and cause a *Lane Jam*. This hot spot is divided into two phases: one during the morning peak between 6:00 and 10:00, the other one during the afternoon peak between 14:00 and 19:00. In SB direction, *Lane Jam* events occur on different locations than mainline congestion (see Figure 5.23). The hot spot times are similar, showing a morning and an afternoon peak.



(a) NB direction (b) SB direction

Figure 5.32: Day of week distribution of *Lane Jam* occurrences

The distribution of *Lane Jam* occurrences on the days of week is depicted in Figure 5.32. In NB direction, the most concerned day is Friday whereas during the weekend, fewer *Lane*

Jam events occur. In the opposite direction, mainly weekdays are affected. Both is aligned with mainline congestion, which mostly occurs on Wednesdays and Thursdays [Karl and Bogenberger, 2019]. A reason for that might lie in the causes of *Lane Jam*. While construction sites are preferably prepared during off-times, spillback emergence at ramps due to overload occurs surrounding the weekend.

5.6.2 Conclusion

Summarizing the key findings of this section, *Lane Jam* frequently occurs starting from the right lane spreading out to the middle and left lanes. Congestion starts at the same location for all lanes. However, the middle and left lane have an average delay of few minutes to recognize it. In total, the middle lane is most affected by low travel times seen over an entire emerging incident. Like hot spots of the other congestion types, also *Lane Jam* often occurs upstream of interchanges and exit ramps, basically due to spillback. Correspondingly, it mainly appears during weekdays, especially before and after the weekend in NB and SB direction, respectively.

5.7 Discussion

This section gives a short overview of all congestion types and their detection rate, and discusses the suitability of the detection technologies. The recognition of 4+1 congestion types *Jam Wave*, *Stop and Go*, *Wide Jam*, *Mega Jam*, and *Lane Jam* are compared using data from four detection technologies gantry, radar, BT, and FCD on a freeway stretch. The evaluation shows that each congestion type has a different level of recognition and results in different hot spots. Table 5.17 gives an overview of the suitability.

	Gantries	**Radars**	**Bluetooth**	**FCD**
Jam Wave	✓	✓	✗	✓
Stop and Go	✓	✓	✗	✓
Wide Jam	✓	✓	✓	✓
Mega Jam	✓	✗[1]	✓	✓
Lane Jam	✓	✓	✗[2]	✗[2]

[1] due to the short detection area here

[2] left for future research

Table 5.17: Suitability of detection technologies and congestion types

In general, gantry sensors are able to detect all kinds of congestion both mainline and lane-by-lane. Their temporally high accuracy and their small spacings in this study lead to an exact traffic state. Radars are basically also suited to detect all kinds of congestion. However, in this data set, only a short stretch of the road is equipped (approximately 10 km). Generally, a *Mega Jam* has greater extensions, which challenges the detection evidence based on radars. Only if a *Mega Jam* lies ideally in this spatio-temporal area, it can be detected reliably. Therefore,

based on the data set in this study, it is not recommended to equip radars to detect *Mega Jam*.

BT measurements are not suited to detect a single or several jam waves (*Stop and Go*). The propagation direction of jam waves is upstream with a constant velocity of approximately 15 km/h, in contrast to the travel time measurements, which are downstream awaiting the second detection at the end of the segment. Especially, when the speed falls below v_{crit} for a short period in a segment, BT speeds result in two options: Either congestion is not detected at all because the segment speed averages to a value greater than v_{crit}. Or a congestion event supposed to be *Jam Wave* decreases the segment speed to a value below v_{crit} which then increases the probability of detecting *Wide Jam* because of the large spacing of several kilometers. The same holds for *Stop and Go* traffic. Multiple jam waves after another are not to be determined reliably by travel time measurements. *Wide Jam* and *Mega Jam* are suited best to be detected by BT.

FCD are able to detect all mainline congestion types well as long as the penetration rate of probe vehicles is sufficiently high. During night hours, no or little speed information can be gathered. If the penetration rate is low, few vehicles cause the speed information and false velocities such as a wrong map-matching cannot be averaged. If the vehicle fleet is large enough, a reliable speed reconstruction along with the determination of congestion types is possible. A drawback here is that the data set comes from a car manufacturer which provides its customers with in-car traffic information. In advance of congestion, drivers receive an incident warning and are more likely to leave the freeway or use an alternative route. Therefore, the fleet size decreases upstream of congestion and increases again downstream of it, which reduces the number of vehicles driving through the congested area.

Lane-specific traffic information based on both BT and FCD are left for future research. *Stop and Go* traffic is characterized by oscillating speeds. Its recognition requires dense detection and harmonization. If a section-based travel time detection system uses small sensor spaces, it is likely to reliably recognize *Jam Wave* and *Stop and Go* events though it could not be proven with this data set.

Takeaways from Chapter 5: Application to Various Traffic Detection Technologies

Evaluation of Fusion Approaches

Low-Resolution Travel Time Smoothing Method (LTSM)	Fusion method that is generally applicable, easy to implement, and ready for practice
Multi-sensor data fusion	Computationally efficient superimposition of multiple data sources for an accurate speed reconstruction

Congestion Detection and Classification

General findings	All data sources comparable, no systematic bias; smoothed data sets detect fewer and larger congestion clusters, raw data more and smaller; size of cluster for estimating and assessing congestion severity appropriately
Earliest congestion detection	Appropriate methodology to identify first congestion occurrences ex-post; SDD faster than FCD by 2-3 min if sensor spacing is small, otherwise FCD significantly earlier
Detection rate per data source	Small differences in resulting clusters regarding types and numbers; FCD and BT data overestimate *Stop and Go*, BT data underestimate *Jam Wave*
Mainline congestion type hot spots	Comprehensive big data analysis, ready for traffic planning; hot spots mainly upstream of interchanges, *Jam Wave* occurrences restricted to few locations only
Lane-by-lane congestion	Novel congestion type derived via congestion clusters; right lane affected most by emerging congestion, middle lane highest travel times on average; hot spots at interchanges of A6 and A99
Suitability of detection technologies	Gantry and radar sensors suitable for all congestion patterns, FCD and BT data for mainline types only; BT data not suited to *Jam Wave* and *Stop and Go* detection if sensor spacing is large

Chapter 6

Conclusion and Future Research

This chapter concludes the dissertation. It first summarizes all findings and discusses the results. Second, topics for further research are given. In the end, a vision is laid out for handling congestion in future transportation systems.

6.1 Summary

This dissertation deals with strategies and comparisons of the performance of sensor technologies while identifying congestion patterns on freeways. A formal 4+1 congestion pattern definition is given, involving four mainline types and one lane-based type: *Jam Wave*, *Stop and Go*, *Wide Jam*, *Mega Jam*, and *Lane Jam*. Each congestion pattern is detected based on an assessment of trajectories resulting from the simulation of virtual vehicles. The proposed algorithm classifies occurring congestion into one of these categories. For each congestion type, metrics are computed such as a hot spot distribution, a primary congestion detection, and – on lane-specific speed information – an estimation of the least travel time per lane.

Each of these congestion detection and comparison metrics is applied to data sets from various detection technologies: gantry speed measurements, radar speed measurements, Bluetooth (BT) travel time measurements, and Floating Car Data (FCD), and the results are compared. The comparison is based on a spatio-temporally discretized speed distribution per data source. For section-based data measurement approaches that deliver travel times, such as BT, a novel algorithm called Low-Resolution Travel Time Smoothing Method (LTSM) is introduced. As a second step, a data fusion approach is presented, which leads to a fused speed distribution based on all data sources. For each of the speed distributions, a methodology is invented, which extracts congested regimes. Each so-called *congestion cluster* is assigned a particular shape – a convex hull of relevant congested cells of the discretized speed matrix –, a particular congestion type, and a size as the area of the congested region. This enables quantitative, machine-interpretable insights and statistics on freeway traffic incidents.

In a comparison of all data sources, the suitability of each detection technology with each congestion type is discussed. Among other findings, it could be observed that BT travel times are not suited to detect short congestion or highly oscillating speeds such as *Jam Wave* or *Stop and Go* traffic. Gantry and radar speed measurements are able to detect each congestion type well. However, radars are installed only on a short stretch of the considered freeway, which challenges the identification of a *Mega Jam*. FCD are likely to detect *Stop and Go* clusters compared to the other congestion types and yield reliable results with an appropriate penetration rate of some percent of probe vehicles.

Considering the autobahn stretch of A9 between Munich and Nuremberg, congestion hot spots are mainly detected upstream of interchanges or exit ramps. The result shows a significantly high share of *Stop and Go* traffic in the Greater Munich Area in the morning hours. In particular, spillback caused by the autobahn A99 and by the Munich city highway are often reasons for congestion emergence. Also, the analysis determined that *Jam Wave* events are not distributed all over the space-time domain, but instead restricted to a few locations. In the Nuremberg area, these *Jam Wave* events occur throughout the day, as opposed to other areas, in which they are limited to only certain times of day. BT data and FCD recognize a *Mega Jam* accumulation in the Greater Nuremberg Area. The reasons for these hot spot distributions are left for future research. Lane-specific congestion is mainly detected upstream of a large construction site together with a lane reduction on an exit ramp during the data collection phase.

This dissertation aims at providing a set of practical tools which are ready to use and are implemented in a computationally efficient way. With these results, traffic authorities have the necessary tools to specifically optimize traffic flow at particular critical locations or times. The comprehensible approaches give support to every road planner, traffic controller, and freeway operator.

6.2 Further Research Topics

Some aspects touched upon in this dissertation are left for future research. Due to limitations of the data collection phase, the amount of common dates from all data sets is small. When gathering traffic measurements during a longer, less error-prone time period, the multi-sensor data fusion should be executed on a large-scale data basis so that results are more comparable, especially regarding the mainline congestion type hot spot analysis and the equipment recommendation with regard to the detection capability for each congestion pattern.

A second point is the investigation of the reasons for the emergence of a particular congestion type, as there are many different possible causes, such as construction sites, accidents, slow heavy-duty vehicles, etc. In a comparison, a matching between congestion causes and occurring patterns should be conducted. This will be helpful to forecast congestion behavior and congestion fronts based on their causes.

A third topic is the detection of *Lane Jam* based on non-lane-specific traffic information such as BT or FCD speeds. During the discretization into spatio-temporal cells, the trajectory speeds of individual vehicles could be clustered into two or more groups. If the speeds between these clusters differ distinctly – indicating a slow and a fast subset of speeds –, it may be an indicator of lane-specifically varying congestion. Possibly, a new definition of *Lane Jam* is needed, not derived via the share of each lane's cluster in the united region, but instead using different segment travel times. Also, for *Lane Jam*, a thorough derivation of possible causes should be conducted.

A further research question is the separation of two conjoint congestion clusters. If the algorithm to identify isolated congestion clusters combines what is actually two bottlenecks into one cluster, in general, one of them gets lost. In a study, the impact of this phenomenon, the misidentifying of the lost congestion, should be investigated. To this end, the share of congested versus uncongested cells in the cluster could be analyzed since it is likely that the

cluster contains many free-flow speed cells.

As a last point, the online applicability of the currently offline available methodology should be investigated. All algorithms can handle live data and act in an online context, however this has not been tested. It is challenging to determine the appropriate congestion type based on the first few trajectories solely. Depending on the congestion type, traffic managers can then take action in real-time leading to an optimized traffic management. Still, the feedback time to inform travelers before the last exit needs consideration.

The online handling of congestion types could be implemented in a macroscopic traffic flow simulation to test various traffic control scenarios. One example of a small adjustment that could be made to the methodologies is a filtering of slow BT section-based speeds during night hours. Also, a detailed consideration of the congestion type of each crossing trajectory should be examined. For example, if the first and the last trajectory traversing the cluster result in a different congestion type than the rest (mainly *Jam Wave* because the intersection with the cluster is only short), these trajectories should get a smaller weight in the frequency-based determination of a cluster's congestion type. With the help of traffic control optimization, if the upstream congestion front can be predicted, alternative routes or bypasses can be recommended for safety reasons.

6.3 Outlook

The results of this thesis help to optimize traffic sensor placement on freeways. *Stop and Go* traffic needs more dense detection in order that free-flow sections, which should be considered part of the *Stop and Go* congestion, are not misidentified. *Jam Wave* events require small-scale detection, for example via local measurements, otherwise the averaged speeds (along a larger road section) could be above the velocity threshold, in which case congestion is not recognized at all. Traffic flow with *Lane Jam* can be optimized if the lanes are handled dynamically and controlled individually. The detection of hot spots leads to more detailed traffic information in advance or even to the ability to make predictions, in which case warnings can be sent, for example that drivers should change their current lanes earlier to increase traffic safety.

In the future, an economical traffic state reconstruction will mainly utilize FCD, that is probe vehicle and mobile device data. The installation and maintenance costs of both induction loops and gantries are generally high. BT measurements are a cost-effective alternative, but they miss traffic flow since they detect a sub-sample of the traffic collective. Radar measurements detect flow and are often suited for an energy self-sufficient operation. In addition to mobile device detection, which also lacks traffic flow information, radars are recommended as mesh points to count the traffic volume. Still, a large amount of traffic detection will be gathered via mobile apps or probe vehicles.

This moves the focus of traffic state reconstruction and specifically congestion detection from a macroscopic view towards a mesoscopic view, as a combination of macroscopic flows and microscopic vehicle speeds. To this end, in the future, congestion could be represented in an event-based way. A speed breakdown then corresponds to an event at a certain time, at a certain place, and on a certain lane, instead of a cell-based spatio-temporal grid. In such a scenario, congestion can be detected by single vehicles. The precondition of a real-time traffic state estimation is a high penetration rate of probe vehicles and a dense mobile net coverage,

which might improve with 5G and beyond.

In a world with 100 % automated vehicles, traffic will change from lane-based to lane-free. In such a world, *Lane Jam* events are obsolete. Detailed congestion type information is even easier to gather because the vehicles are connected. Each vehicle can experience a congestion type individually and determine spatial and temporal extents of the current congestion. An information fusion of all connected vehicles then gives the opportunity to estimate the cluster with its congestion type appropriately. There is no need to reconstruct the entire space-time domain. Also, destinations are known, and therefore, micro-routing can be optimized in real-time for a larger subset of traffic participants. In this way, emerging congestion can be nipped in the bud. Until such a world is realized, however, in mixed traffic conditions with automated and human-driven vehicles, lane-by-lane speed information combined with the approaches introduced in this work, allows traffic planners to improve traffic flow.

Use Cases of the Developed Methodologies

- Traffic planning: By deriving congestion type hot spots and by matching detection technologies to congestion types, the ideal locations of sensors can be determined. For example, at *Wide Jam* hot spots, BT equipment as a cheap alternative suffices, whereas *Jam Wave* hot spots should be equipped with other technologies.
- Traffic control: An online control which knows about the actual congestion pattern can act in an optimized way. Further, it can estimate which sensor technology is applied in which cases to receive a most accurate traffic state. The provided tools support the controllers in Intelligent Transportation Systems (ITS) in making decisions, such as:
  - where and when to use the shoulder lane
  - where and when to employ bi-directional usage of lanes
  - recommendations on lane changes
  - recommendations of alternative routes
- Traffic prediction: Knowing the type of the currently occurring congestion, the upstream and downstream congestion fronts can be predicted. The behavior (e.g. speed profiles) in the interior of this congestion are estimated reliably.

Acknowledgments

The present dissertation concludes the results of my studies conducted to gain new insights into traffic flow optimization and congestion detection. In this acknowledgment, I would like to express my gratitude to all the great people who supported me in reaching this goal.

First of all, I want to thank my supervisor Klaus. He gave me the opportunity to take part in his great team, both at the Bundeswehr University and the Technical University of Munich. His valuable input into a challenging topic was always inspiring to me. I very much appreciate his trust in my research and his continuous support. Danke, Klaus!

Also, I am very grateful to Axel for accepting the part of my second supervisor. His thoughtful comments and remarks improved and sharpened my research a lot. Danke, Axel!

Furthermore, I would like to express my sincere thanks to Costas who served as a third supervisor for this dissertation. His acceptance was a very spontaneous decision from him which I honestly appreciate. Σε ευχαριστώ Κώστα!

I would like to thank my mentor and co-group leader Matthias. Even though he accompanied my dissertation for a rather short period towards its end, it was an intensive last phase and he always supported me and was open for discussions on any topic. Danke, Matthias!

A deep gratitude goes to all my colleagues and to the entire team from the Chair of Traffic Engineering and Control at the TUM for smaller and larger discussions, valuable feedback, and proofreading. Especially, I would like to gratefully thank my co-authors Felix, Barbara, and Gerhard for their contributions to my research and specifically to this thesis. Also our natives Patrick and Allister supported me with their fabulous English. Danke, Leute!

Sincere thanks go to the local authority *Zentralstelle Verkehrsmanagement*, to *BMW*, and to *TomTom* for providing their data for this scientific research. In particular, I would like to thank Johannes and Arne who patiently explained their data for my studies.

Last, I would like to thank my family and friends who helped me realizing every day what work-life balance means in reality. Danke euch allen!

List of Figures

List of Tables

List of Terms and Abbreviations

AID Automatic Incident Detection 3, 21

ANPR Automatic Number Plate Recognition 14

ASM Adaptive Smoothing Method 3, 9, 15, 16, 18, 22, 30, 36, 37, 41, 46, 57, 75, 76, 78–81, 83, 93, 96

BT Bluetooth 2–4, 12, 13, 17, 18, 21, 23, 28, 31, 33, 37–43, 45–47, 62, 69–71, 75, 76, 80, 81, 87, 89–91, 93, 95–105, 111–117, 119

EV Electric Vehicle 1, 4

FCD Floating Car Data 2–4, 13–18, 21–23, 29–33, 38–41, 45, 47, 62, 64, 71, 75–78, 80–83, 85–87, 91–93, 95–105, 112–117, 123

GNSS Global Navigation Satellite System 2, 13, 14, 16, 18, 20, 29, 30, 38, 81

GT Ground Truth 47, 52, 77, 78

ITS Intelligent Transportation Systems 17, 119

LOS Level of Service 10

LTSM Low-Resolution Travel Time Smoothing Method 41, 67, 69, 75, 76, 96, 114, 115

NB Northbound direction 24, 26, 27, 30, 43, 69, 75, 81, 82, 85, 86, 95–98, 100, 102, 104–112

PSM Phase-Based Smoothing Method 3, 9, 15, 16, 18, 22, 38, 41, 91, 96

SB Southbound direction 24, 26, 27, 30, 56, 60, 61, 95–97, 99, 101, 103–112

SDD Stationary Detector Data 3, 21, 22, 32, 33, 64, 75, 78, 81–83, 85, 86, 91, 95–105, 114

SSIMPE Symmetric Square Inverse Mean Percentage Error 45, 47, 69, 75, 77–81

TMC Traffic Message Channel 1, 17

Publications

The content of the following peer reviewed papers are entirely or partly described in this dissertation. For each paper, the addressing sections are indicated. The publications are ordered by year.

- [Rempe, Kessler, et al., 2017]
  - Section 3.1 Data
- [Kessler, Huber, et al., 2018a] and [Kessler, Huber, et al., 2018b]
  - Section 3.1 Data
  - Section 2.2.1 Congestion Definition
  - Section 4.5 Primary Congestion Detection (Theory)
  - Section 5.3 Primary Congestion Detection (Evaluation)
- [Kessler, Karl, et al., 2019]
  - Section 3.1 Data
  - Section 4.1 Low-Resolution Travel Time Smoothing Method (Theory)
  - Section 5.1 Low-Resolution Travel Time Smoothing Method (Evaluation)
- [Karl, Kessler, et al., 2019] and [Karl, Kessler, et al., 2020]
  - Section 3.1 Data
  - Section 2.2.1 Congestion Definition
  - Section 4.4 Congestion Types
- [Kessler, Karl, et al., 2020]
  - Section 3.1 Data
  - Section 2.2.1 Congestion Definition
  - Section 4.4 Congestion Types
  - Section 4.5 Primary Congestion Detection
  - Section 4.6 Congestion Type Hot Spots (Theory)
  - Section 5.5 Congestion Type Hot Spots (Evaluation)
- [Kessler, Rempe, et al., 2021]
  - Section 3.1 Data

The research was conducted while working on the project *iRoute2*, therefore the final project report contains parts of this dissertation [Kessler and Bogenberger, 2021].

- Section 3.1 Data
- Section 3.2 Quality of all Data Sets
- Section 5.4 Detection Rate of Congestion Types Per Data Source
- Section 5.5 Congestion Type Hot Spots

Bibliography

Ambühl, Lukas; Monica Menendez (2016). "Data fusion algorithm for macroscopic fundamental diagram estimation". In: *Transportation Research Part C: Emerging Technologies* 71, pp. 184–197. DOI: 10.1016/j.trc.2016.07.013.

Antoniou, Constantinos; Haris N. Koutsopoulos; George Yannis (2013). "Dynamic data-driven local traffic state estimation and prediction". In: *Transportation Research Part C: Emerging Technologies* 34, pp. 89–107. DOI: https://doi.org/10.1016/j.trc.2013.05.012.

ASTRA, Bundesamt für Strassen (2021). *Verkehrsfluss und Stauaufkommen – Definitionen*. URL: https://www.astra.admin.ch/astra/de/home/themen/nationalstrassen/verkehrsfluss-stauaufkommen/definitionen.html.

Autobahndirektion Südbayern (2021). *Zahlen und Fakten*. URL: http://www.abdsb.bayern.de/zahlen/.

Bachmann, Christian (2011). "Multi-Sensor Data Fusion for Traffic Speed and Travel Time Estimation". MA thesis. Department of Civil Engineering, University of Toronto.

Bachmann, Christian; Baher Abdulhai; Matthew J. Roorda; Behzad Moshiri (2013). "A comparative assessment of multi-sensor data fusion techniques for freeway traffic speed estimation using microsimulation modeling". In: *Transportation Research Part C: Emerging Technologies* 26, pp. 33–48. DOI: 10.1016/j.trc.2012.07.003.

Bachmann, Christian; Matthew J. Roorda; Baher Abdulhai; Behzad Moshiri (2013). "Fusing a Bluetooth Traffic Monitoring System With Loop Detector Data for Improved Freeway Traffic Speed Estimation". In: *Journal of Intelligent Transportation Systems* 17.2, pp. 152–164. DOI: 10.1080/15472450.2012.696449.

Barcelo, Jaume; Lídia Montero; Laura Marques; Carlos Carmona (2010). "Travel Time Forecasting and Dynamic Origin-Destination Estimation for Freeways Based on Bluetooth Traffic Monitoring". In: *Journal of the Transportation Research Board* 2175, pp. 19–27. DOI: 10.3141/2175-03.

Barreau, Matthieu; Anton Selivanov; Karl Henrik Johansson (2020). "Dynamic Traffic Reconstruction using Probe Vehicles". In: pp. 233–238. DOI: 10.1109/CDC42340.2020.9304446.

BASt, Federal Highway Research Institute (2018). *Merkblatt für die Ausstattung von Verkehrsrechnerzentralen und Unterzentralen (MARZ)*.

BayernInfo (2021). *BayernInfo*. Bayerisches Staatsministerium für Wohnen, Bau und Verkehr. URL: https://www.bayerninfo.de/en/.

Bekiaris-Liberis, Nikolaos; Claudio Roncoli; Markos Papageorgiou (2016). "Highway Traffic State Estimation With Mixed Connected and Conventional Vehicles". In: *IEEE Transactions on Intelligent Transportation Systems* 17.12, pp. 3484–3497. DOI: 10.1109/TITS.2016.2552639.

Bickel, Peter J.; Chao Chen; Jaimyoung Kwon; John Rice; Erik van Zwet; Pravin Varaiya (2007). "Measuring Traffic". In: *Statistical Science* 22.4, pp. 581–597. doi: 10.1214/07-STS238.

Blumthaler, Wolfgang; Bartosz Bursa; Markus Mailer (2020). "Influence of floating car data quality on congestion identification". In: *European Journal of Transport and Infrastructure Research* 20.4, pp. 22–37. doi: 10.18757/ejtir.2020.20.4.5304.

Bogenberger, Klaus (2003). "Qualität von Verkehrsinformationen". In: *Straßenverkehrstechnik* 47, pp. 518–526.

Bogenberger, Klaus; Martin Hauschild (2009). "QFCD – A Microscopic Model for Measuring the Individual Quality of Traffic Information". In: *ITS World Congress, Stockholm, Sweden*.

Bogenberger, Klaus; Simone Weikl (2012). "Quality Management Methods for Real-Time Traffic Information". In: *Procedia - Social and Behavioral Sciences* 54, pp. 936–945. doi: 10.1016/j.sbspro.2012.09.809.

Bursa, Bartosz; Nemanja Gajic; Markus Mailer (2018). "Classification of traffic jams on alpine motorways". In: *Proceedings of 7th Transport Research Arena TRA 2018*. doi: 10.5281/zenodo.1441017.

Bursa, Bartosz; Nemanja Gajic; Markus Mailer (2019). "Insights into the congestion patterns on alpine motorways based on separate traffic lane analysis". In: *Transportation Research Procedia* 37. 21st EURO Working Group on Transportation Meeting (EWGT), pp. 441–448. doi: 10.1016/j.trpro.2018.12.220.

Chang, Gang-Len; Yao Cheng; Yen-Yu Chen; Yen-Hsiang Chen (2020). *Ingegration of Ramp Metering and Off-Ramp Progression*. Tech. rep. University of Maryland. url: https://www.roads.maryland.gov/OPR_Research/MD-20-SHAUM514_RampMetering_Report.pdf.

Chang, Tang-Hsien; Albert Y. Chen; Yu-Ting Hsu; Chien-Li Yang (2016). "Freeway Travel Time Prediction Based on Seamless Spatio-temporal Data Fusion: Case Study of the Freeway in Taiwan". In: *Transportation Research Procedia* 17. International Conference on Transportation Planning and Implementation Methodologies for Developing Countries (12th TPMDC) Selected Proceedings, pp. 452–459. doi: 10.1016/j.trpro.2016.11.087.

Chen, Chao; Alexander Skabardonis; Pravin Varaiya (2004). "Systematic Identification of Freeway Bottlenecks". In: *Transportation Research Record* 1867, pp. 46–52. doi: 10.3141/1867-06.

Chen, Yanyan; Cong Chen; Qiong Wu; Jianming Ma; Guohui Zhang; John Milton (2020). "Spatial-temporal traffic congestion identification and correlation extraction using floating car data". In: *Journal of Intelligent Transportation Systems* 25.3, pp. 263–280. doi: 10.1080/15472450.2020.1790364.

Chiabaut, Nicolas; Rémi Faitout (2021). "Traffic congestion and travel time prediction based on historical congestion maps and identification of consensual days". In: *Transportation Research Part C: Emerging Technologies* 124, p. 102920. doi: 10.1016/j.trc.2020.102920.

Cohen, Simon; Zoi Christoforou (2015). "Travel Time Estimation Between Loop Detectors and FCD: A Compatibility Study on the Lille Network, France". In: *Transportation*

Research Procedia 10. 18th EURO Working Group on Transportation Meeting (EWGT), pp. 245–255. DOI: 10.1016/j.trpro.2015.09.074.

Coifman, Benjamin; Balaji Ponnu (2020). "Adjacent lane dependencies modulating wave velocity on congested freeways – An empirical study". In: *Transportation Research Part B: Methodological* 142, pp. 84–99. DOI: 10.1016/j.trb.2020.10.005.

Corsi, N.; Alessio Capitanelli (2011). "Multi-sensor Data Fusion for Traffic Planning and Control". In: *WIT Transactions on The Built Environment* 116. DOI: 10.2495/UT110161.

Cotten, Drew; Julius Codjoe; Matthew Loker (2020). "Evaluating Advancements in Bluetooth Technology for Travel Time and Segment Speed Studies". In: *Transportation Research Record* 2674.4, pp. 193–204. DOI: 10.1177/0361198120911931.

Dakic, Igor; Monica Menendez (2018). "On the use of Lagrangian observations from public transport and probe vehicles to estimate car space-mean speeds in bi-modal urban networks". In: *Transportation Research Part C: Emerging Technologies* 91, pp. 317–334. DOI: 10.1016/j.trc.2018.04.004.

Day, Christopher; Toushik Niloy (2021). "Influence of Data Aggregation Processes on the Estimation of Travel Time Distributions from Probe Vehicle Data for Assessing Travel Time Reliability". In: *100th Annual Meeting of the Transportation Research Board (TRB)*.

Denaes, Samuel; Andrea Schieferstein; Stefanie Riess; Peter Ermer (2009a). "Neue Methoden zur Steuerung von Streckenbeeinflussungsanlagen – Teil 1: Zielfunktion zur Wirkungsmodellierung von Harmonisierungssteuerprogrammen und Stauwarnungen". In: *Straßenverkehrstechnik* 2009/3, pp. 133–140.

Denaes, Samuel; Andrea Schieferstein; Stefanie Riess; Peter Ermer (2009b). "Neue Methoden zur Steuerung von Streckenbeeinflussungsanlagen – Teil 2: Das neue Steuerungsverfahren INCA". In: *Straßenverkehrstechnik* 2009/4, pp. 233–240.

Dion, Francois; Hesham Rakha (2006). "Estimating Dynamic Roadway Travel Times using Automatic Vehicle Identification Data for Low Sampling Rates". In: *Transportation Research Part B: Methodological* 40.9, pp. 745–766. DOI: 10.1016/j.trb.2005.10.002.

Faouzi, Nour-Eddin El; Lawrence A. Klein (2016). "Data Fusion for ITS: Techniques and Research Needs". In: *Transportation Research Procedia* 15. International Symposium on Enhancing Highway Performance (ISEHP), pp. 495–512. DOI: 10.1016/j.trpro.2016.06.042.

Faouzi, Nour-Eddin El; Lawrence A. Klein; Olivier de Mouzon (2009). "Improving Travel Time Estimates from Inductive Loop and Toll Collection Data with Dempster–Shafer Data Fusion". In: *Transportation Research Record* 2129.1, pp. 73–80. DOI: 10.3141/2129-09.

Fulari, Shrikant; Lelitha Vanajakshi; Shankar Subramanian; T Ajitha (2015). "Application of Multisensor Data Fusion for Traffic Congestion Analysis". In: *Multisensor Data Fusion – From Algorithms and Architectural Design to Applications*. Hassen Fourati. Chap. 33, pp. 596–613. DOI: 10.1201/b18851-33.

Greenshields, Bruce D.; J. R. Bibbins; W. S. Channing; H. H. Miller (1935). "A Study of Traffic Capacity". In: *Highway Research Board* 14, pp. 448–477.

Haghani, Ali; Masoud Hamedi; Kaveh Farokhi Sadabadi; Stanley Young; Philip Tarnoff (2010). "Data Collection of Freeway Travel Time Ground Truth with

Bluetooth Sensors". In: *Transportation Research Record* 2160.1, pp. 60–68. DOI: 10.3141/2160-07.

He, Shanglu; Jian Zhang; Yang Cheng; Xia Wan; Bin Ran (2016). "Freeway Multisensor Data Fusion Approach Integrating Data from Cellphone Probes and Fixed Sensors". In: *Journal of Sensors* 2016.7269382. DOI: 10.1155/2016/7269382.

He, Zhengbing; Ying Lv; Lili Lu; Wei Guan (2017). "Constructing spatiotemporal speed contour diagrams: using rectangular or non-rectangular parallelogram cells?" In: *Transportmetrica B: Transport Dynamics* 7.1, pp. 44–60. DOI: 10.1080/21680566.2017.1320774.

Hegyi, Andreas; Bart Netten; Meng Wang; Wouter J. Schakel; Thomas Schreiter; Yufei Yuan; Bart van Arem; Tom Alkim (2013). "A cooperative system based variable speed limit control algorithm against jam waves – an extension of the SPECIALIST algorithm". In: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 973–978. DOI: 10.1109/itsc.2013.6728358.

Helbing, Dirk; Martin Treiber; Arne Kesting; Martin Schönhof (2009). "Theoretical vs. Empirical Classification and Prediction of Congested Traffic States". In: *The European Physical Journal B* 69, pp. 583–598. DOI: 10.1140/epjb/e2009-00140-5.

Heshami, Seiran; Lina Kattan (2021). "A queue length estimation and prediction model for long freeway off-ramps". In: *Journal of Intelligent Transportation Systems* 25.1, pp. 122–134. DOI: 10.1080/15472450.2020.1846125.

Huber, Gerhard; Klaus Bogenberger (2013). "A quality evaluation model for real-time-traffic-information". In: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 2126–2131. DOI: 10.1109/ITSC.2013.6728543.

Huber, Gerhard; Klaus Bogenberger; Robert L. Bertini (2014). "New Methods for Quality Assessment of Real Time Traffic Information". In: *93rd Annual Meeting of the Transportation Research Board (TRB)* 14.2918.

INRIX (2021). *INRIX 2020 Global Traffic Scorecard*. URL: https://inrix.com/scorecard/.

Jong, Joost de (2012). "Quality of Real-Time Travel Time Information". University of Twente. MA thesis.

Karantanos, Michail; Andy H. F. Chow (2015). "Incident Detection for Congested Motorways". In: *94th Annual Meeting of the Transportation Research Board (TRB)*.

Karl, Barbara; Klaus Bogenberger (2019). "Staubericht Bayern 2018 im Auftrag der Autobahndirektion Südbayern, Zentralstelle für Verkehrsmanagement". (unpublished).

Karl, Barbara; Lisa Kessler; Klaus Bogenberger (2019). "Automated Classification of Different Congestion Types". In: *22nd International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 2312–2317. DOI: 10.1109/ITSC.2019.8917410.

Karl, Barbara; Lisa Kessler; Klaus Bogenberger (2020). "Automatisierte Klassifikation verschiedener Stautypen". In: *Tagungsband HEUREKA 2021* FGSV 002/127.

Kerner, Boris S. (1999). "The physics of traffic". In: *Physics World* 12.8, pp. 25–30. DOI: 10.1088/2058-7058/12/8/30.

Kerner, Boris S. (2002). "Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks". In: *Phys. Rev. E* 65 (4), p. 046138. DOI: 10.1103/PhysRevE.65.046138.

Kerner, Boris S. (2004). *The Physics of Traffic. Empirical Freeway Pattern Features, Engineering Applications, and Theory*. Vol. 1860-0832. Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-540-40986-1.

Kerner, Boris S. (2008). "A theory of traffic congestion at heavy bottlenecks". In: *Journal of Physics A: Mathematical and Theoretical* 41.21, p. 215101.

Kerner, Boris S. (2009). *Introduction to Modern Traffic Flow Theory and Control. The Long Road to Three-Phase Traffic Theory*. Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-642-02605-8.

Kerner, Boris S.; Hubert Rehborn (1996). "Experimental properties of complexity in traffic flow". In: *Phys. Rev. E* 53 (5), R4275–R4278. doi: 10.1103/PhysRevE.53.R4275.

Kerner, Boris S.; Hubert Rehborn; Mario Aleksic; Andreas Haug (2004). "Recognition and tracking of spatial–temporal congested traffic patterns on freeways". In: *Transportation Research Part C: Emerging Technologies* 12.5, pp. 369–400. doi: 10.1016/j.trc.2004.07.015.

Kessler, Lisa; Klaus Bogenberger (2015). "Forecast of the Energy Consumption of BEV Based on Dynamic Traffic Information". In: *Mobil.TUM 2015*.

Kessler, Lisa; Klaus Bogenberger (2016). "Mobility patterns and charging behavior of BMW i3 customers". In: *19th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1994–1999. doi: 10.1109/ITSC.2016.7795878.

Kessler, Lisa; Klaus Bogenberger (2019). "Dynamic traffic information for electric vehicles as a basis for energy-efficient routing". In: *21st EURO Working Group on Transportation Meeting (EWGT)*. Vol. 37, pp. 457–464. doi: 10.1016/j.trpro.2018.12.218.

Kessler, Lisa; Klaus Bogenberger (2021). "Empfehlungen für die Ausstattung von lokaler, mobiler und streckenbezogener Detektion für Störungs- und Verkehrslageerkennung – Endbericht iRoute2". In: *Digitales Testfeld A9 – Endbericht*. Federal Highway Research Institute.

Kessler, Lisa; Gerhard Huber; Arne Kesting; Klaus Bogenberger (2018a). "Comparing Speed Data from Stationary Detectors Against Floating-Car Data". In: *IFAC-PapersOnLine, 15th IFAC Symposium on Control in Transportation Systems (CTS)* 51.9, pp. 299–304. doi: 10.1016/j.ifacol.2018.07.049.

Kessler, Lisa; Gerhard Huber; Arne Kesting; Klaus Bogenberger (2018b). "Comparison of floating-car based speed data with stationary detector data". In: *97th Annual Meeting of the Transportation Research Board (TRB)*.

Kessler, Lisa; Barbara Karl; Klaus Bogenberger (2019). "Spatiotemporal Traffic Speed Reconstruction from Travel Time Measurements Using Bluetooth Detection". In: *22nd International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 4275–4280. doi: 10.1109/ITSC.2019.8917084.

Kessler, Lisa; Barbara Karl; Klaus Bogenberger (2020). "Congestion Hot Spot Identification using Automated Pattern Recognition". In: *23rd International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–7. doi: 10.1109/ITSC45102.2020.9294598.

Kessler, Lisa; Felix Rempe; Klaus Bogenberger (2021). "Multi-Sensor Data Fusion for Accurate Traffic Speed and Travel Time Reconstruction". In: *Frontiers in Future Transportation*. doi: 10.3389/ffutr.2021.766951.

Kim, Seoungbum; Benjamin Coifman (2014). "Comparing INRIX speed data against concurrent loop detector stations over several months". In: *Transportation Research Part C: Emerging Technologies* 49, pp. 59–72.

Klein, Lawrence A. (2019). *Sensor and data fusion for intelligent transportation systems*. Vol. PM305. SPIE Press. isbn: 9781510627642.

Klein, Lawrence A. (2020). *Traffic flow sensors: technologies, operating principles, and archetypes*. Vol. SL58. SPIE Press. isbn: 9781510636682.

Leonhardt, Axel (2008a). "Ein Instanzbasiertes Lernverfahren zur Prognose von Verkehrskenngrößen unter Nutzung Räumlich-Zeitlicher Verkehrsmuster". Dissertation. München: Technische Universität München.

Leonhardt, Axel (2008b). "Method for Estimating and Predicting Traveltime based on Probe Vehicles and Local Detector Data using Regression Techniques". In: *ITS World Congress, New York*.

Leonhardt, Axel; Albert Steiner (2012). "Instance Based Learning for Estimating and Predicting Traffic State Variables using Spatio-Temporal Traffic Patterns". In: *91th Annual Meeting of the Transportation Research Board (TRB)*.

Lesani, Asad; Taras Romancyshyn; Luis Miranda-Moreno (2016). "Arterial Traffic Monitoring Using Integrated Wi-Fi-Bluetooth System". In: *95th Annual Meeting of the Transportation Research Board (TRB)*.

Li, Jia; H. Michael Zhang (2011). "Fundamental Diagram of Traffic Flow: New Identification Scheme and Further Evidence from Empirical Data". In: *Transportation Research Record* 2260.1, pp. 50–59. doi: 10.3141/2260-06.

Li, Yuguang; Qingquan Li (2010). "A fast algorithm for identifying candidate links for floating car map-matching: a vector to raster map conversion approach". In: *Annals of GIS* 16.3, pp. 177–184. doi: 10.1080/19475683.2010.513153.

Lint, Hans van (2010). "Empirical Evaluation of New Robust Travel Time Estimation Algorithms". In: *Transportation Research Record* 2160.1, pp. 50–59. doi: 10.3141/2160-06.

Lint, Hans van; Serge Hoogendoorn (2010). "A Robust and Efficient Method for Fusing Heterogeneous Data from Traffic Sensors on Freeways". In: *Computer-Aided Civil and Infrastructure Engineering* 25.8, pp. 596–612. doi: 10.1111/j.1467-8667.2009.00617.x.

Listl, Gerhard (2003). "Anwendung neuer Technologien zur Erfassung des Verkehrsablaufs". Dissertation. Schriftenreihe Verkehr: Universität Kassel.

Listl, Gerhard; Hans-Peter Beck; Frank Bölling; Franz Custodis; Anja Estel; Michael Feldges; Johannes Grötsch; Gaby Gurczik; Jakob Gundolf; Matthias Marohn; Ralf Meschede; Christoph Roth; Ralf Träger (2019). *Hinweise zu Detektionstechnologien im Straßenverkehr*. url: https://www.fgsv-verlag.de/h-detektoren-fgsv-reader.

Liu, Yingshun; Shanglu He; Bin Ran; Yang Cheng (2018). "A Progressive Extended Kalman Filter Method for Freeway Traffic State Estimation Integrating Multisource Data". In: *Wireless Communications and Mobile Computing* 2018.6745726. doi: 10.1155/2018/6745726.

Lux, Carsten (2011). "QBench – Evaluation of Traffic Flow Quality". In: *Quality of on-trip road traffic information*. Ed. by Christine Lotz; Malte Luks. Vol. 82. Berichte der Bundesanstalt für Straßenwesen: F, Fahrzeugtechnik. Wirtschaftsverlag NW, pp. 56–63. URL: https://bast.opus.hbz-nrw.de/opus45-bast/frontdoor/deliver/index/docId/525/file/F82.pdf.

Margreiter, Martin (2016a). "Automatic Incident Detection Based on Bluetooth Detection in Northern Bavaria". In: *Transportation Research Procedia* 15. International Symposium on Enhancing Highway Performance (ISEHP), pp. 525–536. DOI: 10.1016/j.trpro.2016.06.044.

Margreiter, Martin (2016b). "Fast and Reliable Determination of the Traffic State Using Bluetooth Detection on German Freeways". In: *14th World Conference on Transport Research*. Shanghai, China.

Margreiter, Martin; Fritz Busch; Christian Carstensen (2019). "The Evolution of Bluetooth Detection Rates". In: *98th Annual Meeting of the Transportation Research Board (TRB)*.

Margreiter, Martin; Matthias Spangler; Thomas Zeh; Christian Carstensen (2015). "Bluetooth-Measured Travel Times for Dynamic Re-Routing". In: *Annual International Conference on ACE 2015 Singapore*. Ed. by Global Science; Technology Forum. DOI: 10.13140/RG.2.1.4981.6081.

Martchouk, Maria; Fred Mannering; Darcy Bullock (2011). "Analysis of Freeway Travel Time Variability Using Bluetooth Detection". In: *Journal of Transportation Engineering* 137.10, pp. 697–704. DOI: 10.1061/(ASCE)TE.1943-5436.0000253.

Masters, Philip; Joseph Lam; Kam Wong (1991). "Incident detection algorithms for COMPASS - An advanced traffic management system". In: *Vehicle Navigation and Information Systems Conference, 1991*. Vol. 2, pp. 295–310. DOI: 10.1109/VNIS.1991.205776.

Mercader, Pedro; Jack Haddad (2020). "Automatic incident detection on freeways based on Bluetooth traffic monitoring". In: *Accident Analysis & Prevention* 146, p. 105703. DOI: 10.1016/j.aap.2020.105703.

Nedoma, Jan; Marcel Fajkus; Radek Martinek; Jan Vaňuš; Stanislav Kepak; Radana Kahankova; Rene Jaros; Daniel Cvejn; Michal Prauzek (2018). "Analysis of the use of fiber-optic sensors in the road traffic". In: *IFAC-PapersOnLine* 51.6. 15th IFAC Conference on Programmable Devices and Embedded Systems PDeS 2018, pp. 420–425. DOI: 10.1016/j.ifacol.2018.07.117.

Ou, Qing; Hans van Lint; Serge Hoogendoorn (2010a). "An integrated algorithm for fusing travel times, local speed and flow". In: *2010 13th International Conference on Information Fusion*, pp. 1–8. DOI: 10.1109/ICIF.2010.5711912.

Ou, Qing; Hans van Lint; Serge Hoogendoorn (2010b). "TravRes: A method for high-resolution traffic speed reconstruction using GPS-based travel-times". In: *13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1195–1201. DOI: 10.1109/ITSC.2010.5625129.

Palmer, Jochen; Hubert Rehborn; Ivan Gruttadauria (2011). "Reconstruction Quality of Congested Freeway Traffic Patterns Based on Kerner's Three-Phase Traffic Theory". In: *International Journal on Advances in Systems and Measurements* 4, pp. 168–181.

Rainer, Bernd; Martin Müllner (2011). "TMCplus – Improving the TMC information chain". In: *Quality of on-trip road traffic information*. Ed. by Christine Lotz; Malte Luks. Vol. 82. Berichte der Bundesanstalt für Straßenwesen: F, Fahrzeugtechnik. Wirtschaftsverlag NW, pp. 25–31. URL: https://bast.opus.hbz-nrw.de/opus45-bast/frontdoor/deliver/index/docId/525/file/F82.pdf.

Rakha, Hesham; Hao Chen; Ali Haghani; Kaveh Farokhi Sadabadi (2013). "Assessment of Data Quality Needs for use in Transportation Applications". In: *Report MAUTC-2011-01*.

Rehborn, Hubert; Boris S. Kerner; Jochen Palmer (2011). "How can we determine the quality of traffic information?" In: *Quality of on-trip road traffic information*. Ed. by Christine Lotz; Malte Luks. Vol. 82. Berichte der Bundesanstalt für Straßenwesen: F, Fahrzeugtechnik. Wirtschaftsverlag NW, pp. 46–55. URL: https://bast.opus.hbz-nrw.de/opus45-bast/frontdoor/deliver/index/docId/525/file/F82.pdf.

Rempe, Felix; Philipp Franeck; Ulrich Fastenrath; Klaus Bogenberger (2016). "Online freeway traffic estimation with real floating car data". In: *19th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1838–1843.

Rempe, Felix; Philipp Franeck; Ulrich Fastenrath; Klaus Bogenberger (2017). "A phase-based smoothing method for accurate traffic speed estimation with floating car data". In: *Transportation Research Part C: Emerging Technologies* 85, pp. 644–663. DOI: 10.1016/j.trc.2017.10.015.

Rempe, Felix; Lisa Kessler; Klaus Bogenberger (2017). "Fusing probe speed and flow data for robust short-term congestion front forecasts". In: *5th International IEEE Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pp. 31–36. DOI: 10.1109/MTITS.2017.8005695.

Rostami-Shahrbabaki, Majid; Ali Akbar Safavi; Markos Papageorgiou; Ioannis Papamichail (2018). "A data fusion approach for real-time traffic state estimation in urban signalized links". In: *Transportation Research Part C: Emerging Technologies* 92. DOI: 10.1016/j.trc.2018.05.020.

Sasahara, Fabio; Luan Carvalho; Tanay Datta Chowdhury; Zachary Jerome; Lily Elefteriadou; Alexander Skabardonis (2020). "Predicting Lane-by-Lane Flows and Speeds for Freeway Segments". In: *Transportation Research Record* 2674.9, pp. 1052–1068. DOI: 10.1177/0361198120933634.

Sasahara, Fabio; Lily Elefteriadou (2018). "Lane-by-lane analysis of congestion effects due to queue spillback at an off-ramp". In: *97th Annual Meeting of the Transportation Research Board (TRB)*.

Schönhof, Martin; Dirk Helbing (2007). "Empirical features of congested traffic states and their implications for traffic modeling". In: *Transportation Science* 41 (2), pp. 1–32.

Schöpplein, Elke (2013). "Integration of lane-based traffic parameters and data of sidewise detection into a macroscopic traffic flow model for three-lane carriageways". Dissertation. München: Technische Universität München.

Schreiter, Thomas; Hans van Lint; Martin Treiber; Serge Hoogendoorn (2010). "Two fast implementations of the Adaptive Smoothing Method used in highway traffic state estimation". In: *13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1202–1208. DOI: 10.1109/ITSC.2010.5625139.

Sheikh, Muhammad Sameer; Jun Liang; Wensong Wang (2020). "An Improved Automatic Traffic Incident Detection Technique Using a Vehicle to Infrastructure Communication". In: *Journal of Advanced Transportation* 2020. doi: 10.1155/2020/9139074.

Sunderrajan, Abhinav; Vaisagh Viswanathan; Wentong Cai; Alois Knoll (2016). "Traffic State Estimation Using Floating Car Data". In: *Procedia Computer Science* 80. International Conference on Computational Science (ICCS), pp. 2008–2018. doi: 10.1016/j.procs.2016.05.521.

TomTom (2021). *TomTom Traffic Index*. url: https://www.tomtom.com/en_gb/traffic-index/ranking/.

Transver GmbH (2010). "Stauklassifizierung und Untersuchung des Zusammenhangs Verkehrsstärke und Verkehrszusammenbruch – Abschlussbericht". (unpublished).

Treiber, Martin; Dirk Helbing (2002). "Reconstructing the Spatio-Temporal Traffic Dynamics from Stationary Detector Data". In: *Cooper@tive Tr@nsport@tion Dyn@mics* 1. (Internet Journal, www.TrafficForum.org/journal), pp. 3.1–3.24.

Treiber, Martin; Dirk Helbing (2003). "An adaptive smoothing method for traffic state identification from incomplete information". In: *Interface and Transport Dynamics*. Springer, Berlin, Heidelberg, pp. 343–360. doi: 10.1007/978-3-662-07969-0_33.

Treiber, Martin; Arne Kesting (2011). "Evidence of convective instability in congested traffic flow: A systematic empirical and theoretical investigation". In: *Transportation Research Part B: Methodological*. doi: 10.1016/j.trb.2011.05.011.

Treiber, Martin; Arne Kesting; Dirk Helbing (2010). "Three-phase traffic theory and two-phase models with a fundamental diagram in the light of empirical stylized facts". In: *Transportation Research Part B: Methodological* 44.8-9, pp. 983–1000. doi: 10.1016/j.trb.2010.03.004.

Treiber, Martin; Arne Kesting; R. Eddie Wilson (2011). "Reconstructing the Traffic State by Fusion of Heterogeneous Data". In: *Computer-Aided Civil and Infrastructure Engineering* 26, pp. 408–419. doi: 10.1111/j.1467-8667.2010.00698.x.

Trinh, Xuan Sy; Dong Ngoduy; Mehdi Keyvan-Ekbatani; Blair Robertson (2019). "A comparative study on filtering methods for online freeway traffic estimation using heterogeneous data". In: *22nd International IEEE Conference on Intelligent Transportation Systems (ITSC)*. doi: 10.1109/ITSC.2019.8917363.

Tsogas, Manolis; Nikos Floudas; Panagiotis Lytrivis; Angelos Amditis; Aris Polychronopoulos (2011). "Combined lane and road attributes extraction by fusing data from digital map, laser scanner and camera". In: *Information Fusion* 12.1. Special Issue on Intelligent Transportation Systems, pp. 28–36. doi: 10.1016/j.inffus.2010.01.005.

Vukanovic, Svetlana (2006). "Intelligent link control framework with empirical objective function: INCA". Dissertation. München: Technische Universität München.

Wang, Yinhai; Yegor Malinovskiy; Yao-Jan Wu; Un Kun Lee (2011). *Error Modeling and Analysis for Travel Time Data Obtained from Bluetooth MAC Address Matching*. Tech. rep. University of Washington.

Work, Daniel B.; Sébastien Blandin; Olli-Pekka Tossavainen; Benedetto Piccoli; Alexandre M. Bayen (2010). "A Traffic Model for Velocity Data Assimilation". In: *Applied Mathematics Research eXpress* 2010.1, pp. 1–35. doi: 10.1093/amrx/abq002.

Xie, Kun; Kaan Ozbay; Di Yang; Hong Yang; Yuan Zhu (2021). "Modeling Lane-specific Breakdown Probabilities at Freeway Diverge Sections". In: *Physica A: Statistical Mechanics and its Applications* 561, p. 125231. DOI: 10.1016/j.physa.2020.125231.

Yildirimoglu, Mehmet (2019). "Joint estimation of paths and travel times with Bluetooth traffic monitoring". In: *98th Annual Meeting of the Transportation Research Board (TRB)*.

Yuan, Yufei; Hans van Lint; Femke van Wageningen-Kessels; Serge Hoogendoorn (2014). "Network-Wide Traffic State Estimation Using Loop Detector and Floating Car Data". In: *Journal of Intelligent Transportation Systems* 18.1, pp. 41–50. DOI: 10.1080/15472450.2013.773225.

Zeng, Dehuai; Jianmin Xu; Gang Xu (2008). "Data Fusion for Traffic Incident Detection Using D-S Evidence Theory with Probabilistic SVMs". In: *Journal of Computers* 3.10.

Zhao, Xue; Dongbo Zhang (2018). "A Review of Multi-sensor Data Fusion for Traffic". In: *9th International Symposium ISICA 2017, Revised Selected Papers, Part I*, pp. 432–444. DOI: 10.1007/978-981-13-1648-7_37.

Zhou, Zhuoyang; Pitu Mirchandani (2015). "A Multi-Sensor Data Fusion Framework for Real-Time Multi-Lane Traffic State Estimation". In: *94th Annual Meeting of the Transportation Research Board (TRB)* 15-0186.