TUDelft

Delft University of Technology

Multimodal Transportation Simulation for Emergencies using the Link Transmission Model

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Jeroen P.T. van der Gun

Multimodal Transportation Simulation for Emergencies
using the Link Transmission Model

Multimodal Transportation Simulation for Emergencies using the Link Transmission Model

Jeroen P.T. van der Gun

Delft University of Technology

Multimodal Transportation Simulation for Emergencies using the Link Transmission Model

Proefschrift

ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft,

op gezag van de Rector Magnificus prof.dr.ir. T.H.J.J. van der Hagen,

voorzitter van het College voor Promoties,

in het openbaar te verdedigen op donderdag 24 mei 2018 om 12:30 uur

door

Jeroen Peter Tjebbo VAN DER GUN

Master of Science in Civil Engineering,

Technische Universiteit Delft,

geboren te Utrecht

Dit proefschrift is goedgekeurd door de promotor en copromotor.

Samenstelling van de promotiecommissie:
Rector Magnificus   voorzitter
Prof.dr.ir. B. van Arem   Technische Universiteit Delft, promotor
Dr.ir. A.J. Pel   Technische Universiteit Delft, copromotor

Onafhankelijke leden:
Prof.dr.ir. C.M.J. Tampère   Katholieke Universiteit Leuven, België
Prof.dr. H.S. Mahmassani   Northwestern University, Verenigde Staten
Prof.dr. H. Tu   Tongji University, China
Prof.dr. T.A. Arentze   Technische Universiteit Eindhoven
Prof.dr.ir. S.P. Hoogendoorn   Technische Universiteit Delft
Prof.dr.ir. J.W.C. van Lint   Technische Universiteit Delft, reservelid

Printed in the Netherlands

"The first criterion is quite a likely one in practice, since it might be assumed that traffic will tend to settle down into an equilibrium situation in which no driver can reduce his journey time by choosing a new route."

J.G. Wardrop (1952, p. 345)

Preface

According to article 7.18 of the Dutch law on higher education and scientific research, a doctoral thesis must demonstrate the capability of independently practicing science. However, such a capability can only surface and flourish in the right environment. I would like to greatly thank my supervisors Adam Pel and Bart van Arem for giving me the freedom to do my research as I saw fit, allowing me to select my own scientific challenges to tackle, and to adjust my focus as my research progressed. As a result, this thesis has a seemingly diverse table of contents, at some points rather loosely related to the original research proposal, that nevertheless resulted in a coherent thesis that should be useful both inside and outside the specific domain of emergencies. Adam and Bart, thank you for your trust in this end result and your helpful constructive feedback to get there. It is great to be able to focus on the contents of your work without a need to worry whether your supervisors support what you're doing. I look forward to continue working with you, both inside and outside the university.

Yet, independently practicing science does not mean one does not get any help. Adam, thank you for the in-depth discussions we had on the topics of this thesis and your assistance while preparing and revising the papers it is based upon. Wei Gu, thank you for developing the bus bridging model that formed the basis of Chapter 6 and was essential to its case study. Erik-Sander Smits and Luuk Brederode, thank you for provoking me to dive into the wonderful world of the Link Transmission Model, which greatly contributed to the selection of methodology applied in this thesis. Yu Han and Kai Yuan, thank you for the valuable and enjoyable discussions on our various approaches of implementing a capacity drop in first-order traffic flow theory. Theo Arentze, thank you for providing the Albatross data that was crucial input to two of my three case studies. Chris Tampère and Willem Himpe, thank you for the thoughtful discussions of my contributions to the Link Transmission Model. Furthermore, I'd like to thank my various roommates over time and other colleagues at the Transport & Planning department for our enjoyable interactions and insightful discussions, also those not related to my own work, that I hope will continue in the future.

Finally, I'd like to thank all members of the doctoral committee for examining my thesis. I look forward to an interesting academic debate during the defence ceremony.

Jeroen van der Gun

Delft, 23 April 2018

Contents

Chapter 1

Introduction

The capacity of transportation systems of urban regions is finite. This is evidenced by the significant congestion and delays that travellers by car or public transport incur every working day. Much effort is being put into relieving these problems, both in practice and in theoretical research. However, the situation goes from bad to worse in case of major emergencies affecting the transportation system. From a transportation market perspective, such unplanned events can be characterised by a strong decrease in transportation supply and/or a strong increase in transportation demand, easily overloading the system. Compared to daily congestion, the extent of this overload is more severe, and due to the urgency of the situation, potentially also its consequences.

This thesis will look into new computer simulation models that authorities can use to assess the transportation consequences of emergencies and to test plans to manage the associated transportation problems. Within this introduction chapter, Section 1.1 gives an overview of emergencies and their impacts on transportation. Section 1.2 describes the transportation management problems posed by such emergencies. Using this, Section 1.3 states the purpose of this thesis and lists its scientific contributions. Finally, Section 1.4 explains the outline of this thesis.

1.1 Characteristics of emergencies

An overview of emergencies that may affect the transportation system is provided in Table 1.1. Although many structural characteristics may be used to classify disasters causing emergencies (Leach, 1994), this thesis focuses on the impacts on the transportation system, and thus we will classify emergencies here according to their types of impact on transportation supply and demand, that are discussed further below.

Firstly, there are emergencies that reduce or eliminate the capacity of parts of the transportation system, i.e. a reduction of transportation supply, and hence cause delays to everyday travel. This includes adverse weather conditions like heavy precipitation or snowfall, damage to infrastructure, and failure of public transport services.

Table 1.1. Emergency categories and example emergencies, including example studies assessing impacts on the transportation system.

Capacity reduction	Emergency services	Evacuation traffic	Examples
✓	✗	✗	Extreme weather (Chung, 2012; Zhao and Sadek, 2013) Infrastructure damage Public transport failures (Tahmasseby, 2009)
✓	✓	✗	Traffic accidents (Knoop, 2009; Zhao and Sadek, 2013) Train crashes
✗	✓	✗	Aeroplane crashes (Rosenthal et al., 2013) Building fires/collapses
✓	✓	✓	Floods (Tu et al., 2010; Pel, 2011; Kolen, 2013) Tsunamis (Lämmel et al., 2008) Hurricanes (Litman, 2006; Lindell and Prater, 2007; Lindell, 2008) Sediment disasters (Chiba et al., 2017) Wildfires (Wolshon and Marchive, 2007) Volcanic eruptions (Baxter and Ancia, 2002) Chemical spills (Brachman and Dragicevic, 2014)
✗	✓	✓	Nuclear disasters (Urbanik, 2000) Industrial accidents

While these can result in major disruptions, they are not emergency situations that endanger people. This is different for severe traffic accidents and train crashes, which therefore, in addition to (partial) temporary infrastructure unavailability in the road and public transport network respectively, require dispatching of emergency services, generating additional high-priority traffic demand primarily directed towards the disaster site. For major emergencies, the incoming emergency services may be accompanied by volunteers, spectators and friends and relatives of suspected victims (Rosenthal et al., 2013). Near the disaster site, rubbernecking effects may further reduce road capacity (Knoop, 2009). There are also emergencies that require emergency services to be deployed but do not necessarily affect the urban transportation system capacity, such as aeroplane crashes and building fires/collapses, provided they occur not too close to significant transportation corridors.

Impacts on the transportation system are even more severe for emergencies that generate additional demand consisting of evacuation traffic fleeing to safety. People may be trying to reach a hotel or motel, the home of friends or family, a public shelter (Murray-Tuite and Wolshon, 2013; Deka and Carnegie, 2010), or their own home (Trainor et al., 2013; Kolen, 2013). Heading for shelter within the endangered area is often referred to as vertical evacuation (Kolen, 2013). Prior to travelling to the final evacuation destination, people may also need to make several intermediate trips in preparation (Trainor et al., 2013; Murray-Tuite and Mahmassani, 2003; Yin et al., 2014). In addition to rescue operations, emergency services are needed for delivery of water, food and utility maintenance, helping stranded vehicles, protecting and evacuating non-self-reliant people (Litman, 2006) and traffic regulation (Tu et al., 2010).

For evacuations to occur prior to a disaster – which is assumed in this thesis – the disaster must to a sufficient extent be observed or predicted in advance with sufficient likelihood and accuracy, so that the public can be warned and has sufficient time to evacuate. If floods are predicted, the warning time varies from a couple of hours to multiple days (Younis et al., 2008; Kolen, 2013). Tsunami warning times range from several minutes to more than half an hour (Koshimura et al., 2006; Lämmel et al., 2008). Hurricane warnings can be issued multiple days in advance (Regnier, 2008). Although affecting smaller, more suburban areas, wildfires typically move much faster than hurricanes, leading to considerably lower warning times in the order of hours (Wolshon and Marchive, 2007). Warning time varies for sediment disasters like landslides (Michoud et al., 2013). While volcanic eruptions are difficult to predict reliably (McNutt, 1996), there may also be time between the eruption and the impact on an urban area. The latter also applies to chemical spills. The threat of a nuclear disaster is generally imminent, not immediate (Urbanik, 2000), and thus allows for preventive evacuation.

In many cases, the emergency warranting evacuation reduces the capacity or availability of the transportation system itself, e.g. due to adverse weather, infrastructure damage and public transport failures. However, even in possible exceptions like industrial accidents and threats of nuclear disasters, road capacity may change due to different driving behaviour in evacuation conditions. In a driving simulator, Hoogendoorn (2012) observed on average an increase in speed, acceleration and deceleration rates and a decrease in headway, collectively increasing road capacity. Urbanik (2000) mentions that households may select their best driver for the evacuation trip, increasing road capacity. Capacity may also decrease due to driver confusion (Litman, 2006), because of heavily loaded vehicles (Litman, 2006; Lindell and Prater, 2007), and because lane changes might be more difficult (Tu et al., 2010), although the latter might be compensated by changes in lateral driving behaviour (Hamdar, 2004). Particularly for large-scale evacuations, traffic incidents may also play a role (Robinson et al., 2009), and the previously mentioned more aggressive driving behaviour and adverse weather may increase crash rates and reduce the expected road capacity. While the overall effect is hard to estimate, it is thus usually wise to account for the possibility of a reduction of road capacity in evacuation.

1.2 Transportation management problem

Public administration, the police, fire brigades, emergency medical services, traffic control centres, public transport companies can typically prepare themselves for emergencies in advance by formulating potential emergency scenarios and testing their response in a computer simulation model, in order to minimise the adverse impacts on society. Collectively,

we will refer to them as authorities, and assume cooperation. All sorts of control actions that are to be undertaken by authorities in case of an emergency can be described by an emergency plan. In order to evaluate an emergency plan with a computer simulation model, the model must capture both the behavioural and physical processes inside the transportation system, as well as the effects of the control actions. If this is the case, the simulation model can help authorities to understand the transportation consequences of a disaster, to analyse and improve their emergency plan, and to understand whether, why and how particular actions contribute to a successful plan.

The emergency management problem from the perspective of authorities can be formulated as an optimisation problem, consisting of decision variables, an objective, and a set of constraints. The value of the objective is obtained by simulating the transportation system and assessing the results of this simulation with respect to a particular objective, yielding practical insights into how the decision variables affect the objective.

1.2.1   Optimisation objectives and constraints

For evacuations, one common objective is to maximise the number of evacuated people within a predefined time limit (Abdelgawad et al., 2010). To incorporate uncertainty in the time available for evacuation, this objective can be extended with weights such that early-arriving evacuees are preferred over evacuees arriving later on (Pel, 2011; Huibregtse, 2013). Alternatively, one may minimise the evacuation time or network clearance time, i.e. the time required to evacuate the population (Murray-Tuite and Mahmassani, 2004; Van Zuilekom et al., 2005; Tu et al., 2010; Pel, 2011; Bish et al., 2014) or some specified percentage thereof (Abdelgawad et al., 2010). One may also minimise total travel time, possibly supplemented with total pre-departure waiting time so that the total "exposure time" is minimised (Abdelgawad et al., 2010). For emergencies not involving mass evacuations, typical objectives one could think of are to minimise total delay for the travellers, minimise road congestion, or minimise deviation from the original public transport schedule. For given values of the decision variables, all these indicators can be estimated quantitatively with a computer simulation model.

There can also be constraints. For emergencies involving dispatching of emergency services, one constraint is that the quality and safety of their rescue operations must meet particular thresholds: although the performance of the transportation system should be optimised, the resulting decisions should not hinder rescue operations. In case of evacuations, it is also possible to treat this as a trade-off in the objective, since both the evacuation and the rescue processes relate to the safety of the public. In case the emergency plan provides instructions to the public on what to do during the emergency, one may include constraints on how much worse off some individuals who follow the official advice may be (for sake of the greater good) compared to when they would not; if trade-offs are deemed acceptable, these effects may instead be added to the objective function. Aside from ethical considerations, this helps ensure a high compliance rate (Abdelgawad et al., 2010), improving the controllability of the current and possible similar future situations.

1.2.2   Decision variables

Finally, decision variables represent   possible actions by authorities to improve the transportation system performance according to the chosen objective. Before discussing them, it is important to realise that even in day-to-day traffic management, successful control actions often need to be based on real-time monitoring of the traffic system. Since emergency

situations are typically accompanied with much greater uncertainty, it is advisable for authorities to update decisions during the emergency based on new relevant information that becomes available to authorities as the event progresses, resulting in an adaptive emergency plan that is likely to be more successful. Acknowledging this, we can separate decision variables into several categories based on their adaptiveness, each of which will be described below.

First of all, there are *strategic decision variables*. These decisions must remain constant or can only change infrequently since any change would require non-local communication with the public, and are therefore fixed in the emergency plan. They can affect both the supply side and the demand side of the problem. Demand-oriented examples are general recommendations to the population on whether, when and how to travel, such as departure advice (Pel, 2011; Huang et al., 2012; Dixit et al., 2012b; Huibregtse, 2013), mode advice, destination advice (Huibregtse, 2013), and route advice (Pel, 2011; Huibregtse, 2013). Supply-oriented examples are deployment of roadblocks (Tu et al., 2010; Huibregtse, 2013), the usage of contraflow roads (Litman, 2006; Huibregtse, 2013), the construction of temporary roads, the availability and schedule of public transport (Litman, 2006; Abdelgawad et al., 2010; Shiwakoti et al., 2013) and the provision of public shelter facilities (Litman, 2006).

Next, there are *tactical decision variables*. These would otherwise be strategic variables, except that a set of tactics or conditions is defined to apply infrequent changes to them semi-automatically while the emergency situation progresses. Unlike strategic variables, the emergency plan thus contains no pre-specified decisions for tactical variables, but quantified rules to make decisions during the emergency situation, based on the development of the situation so far. These semi-automatic changes are simulated in the computer simulation model by applying the selected tactics. How variables are divided over the strategic and tactical categories, i.e. what is fixed and what is flexible, may differ per emergency situation that is considered. Examples include varying departure and route advice during an evacuation depending on the observed congestion or modification of public transport schedules in response to the observed demand.

Finally, there are *operational decision variables*. These decisions can be frequently changed real-time based on the prevailing conditions, in some cases even autonomously by a computer varying from simple local rule-based control to advanced network-wide model-predictive control, and thus focus primarily on the supply side of the transportation problem. Examples are traffic light and ramp metering settings, variable speed limits, the operation of peak-hour and contraflow lanes, information shown on dynamic route information panels, and information provided through announcements at public transport stops and in stations. Manual traffic control, such as by traffic regulators, and the deployment of emergency services also fall into this category.

1.3   Research goal and contributions

The discussion of the transportation management problem above poses a lot of requirements for the to be used transportation simulation model. As will be shown in the next chapter, the model needs to be dynamic, to describe relevant choice behaviour of the population, to deal with initial conditions at the start of the emergency, to include interactions between individual travellers, to be multimodal, to take travellers who are not directly affected into account, and to consider emergency services. Additionally, there is a need for the propagation of car traffic

to be detailed, simple to calibrate, and computationally efficient. While existing models address subsets of these requirements, they have not yet been addressed in an integrated manner, limiting what emergencies can be analysed and with what level of detail.

The goal of the research in this thesis is to identify requirements for, to develop, and to test a simulation model and its components to analyse how the aforementioned emergencies affect transportation in urban areas and evaluate candidate plans for transportation management problems posed by them. They should allow calculation of the value of aforementioned objective functions and respond to changes in the aforementioned decision variables. Ideally, the simulation model and components are also computationally efficient so that enough different candidate plans can be compared and enough different emergency scenarios can be simulated to account for the uncertain nature of the problem.

1.3.1   Methodological contributions

This thesis makes three main scientific contributions to this research goal, which are primarily methodological:

- We formulate a novel *general simulation modelling framework for emergencies*. This framework features agent-based modelling of transportation demand, permitting advanced choice models including interactions between individuals, background traffic, and multimodal transportation. By using an activity-based choice model for normal travel conditions as a reference, it furthermore enables direct comparison of the simulation results to what the same simulation model would predict if there were no emergency, or if the emergency would occur at a different time-of-day with corresponding initial conditions. By using a macroscopic or mesoscopic dynamic network loading model to simulate road traffic over time, traffic dynamics are simulated in a computationally-efficient way without the need for a very detailed specification of driving behaviour. The impact of road congestion on public transport operation is also included. Overall, the formulated modelling framework can therefore be applied to a wide range of emergency conditions.

- We present new explicit *considerations for integrating agent-based demand models into macroscopic traffic simulation models*. These ensure that both model components work well together, allowing agents to dynamically make en-route choices while retaining the computational efficiency and physical accuracy of the traffic propagation model in case of a large number of agents. This provides the necessary wiring between the components of the above emergency simulation framework, satisfying the requirements of the framework.

- We develop new and improved *extensions to enhance the level of detail of the Link Transmission Model* (Yperman, 2007), which is a computationally-efficient first-order traffic simulation model based on the traffic flow theory by Lighthill and Whitham (1955) and Richards (1956) that is very suitable for our emergency modelling framework. These extensions relate to
  - o   the speed on a road gradually decreasing as it gets loaded up to capacity,
  - o   the dispersion of platoons,
  - o   the capacity of a road dropping once congestion has formed,
  - o   within-simulation adaptation of driving behaviour on road links,
  - o   starting the simulation with traffic already in the network, and
  - o   the application of signalling schemes at signalised intersections.

These extensions are especially beneficial for simulating emergency plans involving traffic management measures and emergencies with changing environmental conditions, but are also useful for more general usage of transportation models.

These contributions provide important advancements relative to previous literature. In particular,

- no previous modelling framework satisfies all requirements posed by the wide range of emergencies in an integrated manner;

- no previous study comprehensively addresses or satisfactorily deals with the consistency challenges of integrated agent-based demand modelling and macroscopic or mesoscopic dynamic network loading;

- no previous first-order traffic flow theory with capacity drop allows the existence of moving jams;

- no previous first-order traffic simulation model supports the features of our Link Transmission Model extensions while matching their efficiency and accuracy.

1.3.2   Societal contributions

Through three case studies, all situated in the Netherlands, this thesis also shows how the model behaves in various practical emergency situations. These cases vary in

- geographic scale of the network (corridor, city, metropole),

- type of emergency (public transport failure, traffic accident, evacuation),

- simulation purpose and types of decision variables (reproduction of previous emergency with operational decisions, analysis of possible emergency with implications for strategic and operational decisions, test of tactical decisions in possible emergency), and

- which model components are used (full framework, traffic propagation only).

The case studies result in societal contributions as they illustrate how the methodology can be used in diverse settings, but also theoretical contributions as they provide general insight into the features of the methodology. The cases are:

- A hypothetical *multimodal evacuation of the city of Delft*. For this case, equilibrium route choices are computed for normal conditions, and the results are subsequently compared with what happens in the emergency scenario where people in Delft gradually begin facing choices regarding how to gather family members and how to evacuate. The study highlights the importance of interactions between modes, interactions between inbound, outbound, and background traffic, and the potential sub-optimality of user mode and route choice from a system perspective.

- A true *vehicle crash on the A13 motorway corridor*, with traffic management systems changing the number of available lanes and speed limit over space and time. This case analyses the propagation of car traffic in detail and compares it with the measured

data, demonstrating the strengths and limitations of our Link Transmission Model improvements. Additionally, we present simulations of several scenarios of the same corridor without any accident, showing the impact of the capacity drop on recurring congestion, again with comparisons with measurements.

- A hypothetical *disruption of metro and tram operations in the city of Rotterdam*, where substitute bus services are deployed to bring the stranded passengers to their destinations through an also disrupted road network with signalised intersections. These bus bridging services are optimised in real time, offering a specific example of how our modelling framework can be used to test the tactics of an adaptive emergency plan in simulation.

As an indication of the scale of the transportation networks used in these case studies, Table 1.2 lists some basic network statistics.

Table 1.2. Network statistics for the case studies. Household counts exclude weights.

Network	Delft	A13 motorway	Rotterdam
Image			
Junctions	481	11 – 36	6,752
Links	2,298	23 – 48	28,924
Centroids	24	13	167
PT stops	44	0	867
PT lines	20	0	465
Households	30,155	N/A	648,823
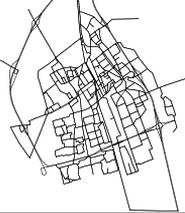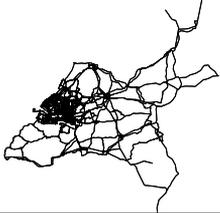
1.4   Thesis outline

An overview of the structure of this thesis is depicted in Figure 1.1. Chapter 1, this introduction, has described and classified the wide range of emergencies that may affect an urban transportation system. The resulting management problem from the perspective of transportation authorities has been discussed, including a description and classification of decision variables representing all kinds of transport and traffic management measures. This chapter furthermore stated the purpose of this research, namely the creation of a simulation model that permit analysis of the management problem and evaluation of candidate management plans.

Chapter 2 develops a simulation framework to address this problem. After identifying model requirements from literature, it proposes a methodological framework that includes an escalation model to represent the choice behaviour of individuals in the network based on

their normal-day activity-travel patterns, making them respond in different ways depending on the severity of the impact the emergency has on them, and a macroscopic or mesoscopic multimodal dynamic network loading model to simulate traffic propagation for all modes. Chapter 2 also implements the framework and tests it using the case study of the evacuation of Delft, using the Yperman (2007) version of the Link Transmission Model. The case shows that various complexities introduced by the methodology are indeed important for the model outcomes.

The following three chapters serve to discuss and further improve specific aspects of this simulation model. Chapter 3 looks more closely into the issue of integrating an agent-based traffic demand model and a non-agent-based traffic propagation model. Chapter 3 discusses how such an integration can be realised in a way that is both efficient from a computational perspective and accurate from a traffic flow perspective, identifying several pitfalls that should be avoided. It notes that the Link Transmission Model can be a suitable traffic propagation model for this purpose.

The Link Transmission Model is further extended in Chapters 4 and 5. The extension proposed by Chapter 4, which is particularly relevant for the evaluation of traffic management measures, is realised in two steps, both improving the shape of the fundamental diagram that describes the driving behaviour on road links. First, it extends the link model from a simple triangular fundamental diagram to a general continuous fundamental diagram – particularly improving the simulation of light traffic – and points out inaccuracies in previous such extensions. Second, after an extension of the underlying traffic flow theory, the chapter extends both the link model and the node model to account for a capacity drop in case of congestion, introducing stop-and-go waves in addition to standing queues. These improvements are tested using the A13 motorway case study without accidents.

Chapter 5 continues the extension of the Link Transmission Model, and is also useful for emergencies involving changes in environmental conditions and other changes in driving behaviour, in addition to emergencies involving traffic management. This chapter namely derives algorithms for working with the traffic states within the link interior, permitting initialisation of the link model with an initial condition and computation of the traffic densities within the link during a simulation. This enables capabilities such as starting the simulation with initially non-empty roads and changing the driving behaviour parameters of the fundamental diagram during a simulation, which are demonstrated using again the A13 motorway case study, this time including an accident and various traffic management measures applied in response to it.

After formulating the simulation framework in Chapter 2 and detailing the simulation model in Chapters 3-5, Chapter 6 offers an example of how the framework can be applied in an optimisation context. It proposes a real-time bus bridging scheduling method to address the problem of dispatching bus bridging services for stranded passengers in case of a severe disruption of public transport. The scheme serves as an adaptive emergency plan where the real-time optimisation method is the applied tactic. It is tested in the large-scale Rotterdam case study, featuring a road network with signalised intersections that is also affected by the disruption. The chapter also compares the results from different variations of the simulation model.

Finally, Chapter 7 synthesises the findings and conclusions from the research in this thesis, discusses its practical implications, and recommends directions for future research.

1. Introduction

↓

Simulation framework

2. A general activity-based methodology for simulating multimodal transportation networks during emergencies

Case study: city of Delft

↓

Model improvements

3. Propagating agents with macroscopic dynamic network loading: challenges and possible solutions	*4. Extending the Link Transmission Model with non-triangular fundamental diagrams and capacity drops*
	5. The Link Transmission Model with variable fundamental diagrams and initial conditions
	Case study: A13 motorway corridor

↓

Optimisation application

6. Real-time bus-bridging scheduling and dynamic simulation for multimodal urban network disruptions

Case study: city of Rotterdam

↓

7. Conclusions and recommendations

Figure 1.1. Overview of the structure of this thesis.

Chapter 2

A general activity-based methodology for simulating multimodal transportation networks during emergencies

Abstract

Many possible emergency conditions, including evacuations, negatively affect the urban transportation system by substantially increasing the travel demand and/or reducing the supplied capacity. A transportation model can be used to quantify and understand the impact of the underlying disasters and corresponding management strategies. To this end, we develop an efficient methodology suitable for simulating multimodal transportation systems affected by emergencies, based on the novel integration of an activity-based choice model with both pre-trip and en-route choices, and a macroscopic or mesoscopic dynamic network loading model. The model structure first estimates the daily equilibrium and then uses that result as a starting point to simulate the emergency situation without further iterations. Unlike previous efforts, our methodology satisfies all requirements identified from literature regarding transportation modelling for emergencies, and is sufficiently general to investigate a wide range of emergency situations and management strategies. An evacuation case study for Delft shows the feasibility of applying the methodology. Furthermore, it yields practical insights for urban evacuation planning that stem from complex system dynamics, such as important interactions among travel directions and among modes. This supports the need for a comprehensive modelling methodology such as the one we present in this chapter.

This chapter is a version of the following publication:

2.1   Introduction

In today's world, many types of disasters can pose significant challenges to the transportation systems of urban areas. Ample studies have hence been undertaken to understand the impact of these disruptions and disasters, ranging from extreme weather to large-scale accidents. Notwithstanding the specifics of the consequent emergency situations, from a transport perspective we can also discern three ways how such emergencies commonly differ from a normal situation. First of all, there may be a reduced capacity for daily traffic. That is, disruptions and disasters tend to reduce the capacity of the road infrastructure and public transport network, e.g. because of adverse or extreme weather conditions (Hoogendoorn, 2012; Litman, 2006; Lindell, 2008), traffic accidents (Knoop et al., 2010), damaged infrastructure (Brachman and Dragicevic, 2014) and public transport failures (Tahmasseby, 2009). These capacity reductions can lead to substantial delays. Second of all, there may be evacuation traffic leaving the affected area or sheltering in-place. That is, in case of notice, evacuation may precede the disaster event, with evacuees trying to return home (Trainor et al., 2013; Kolen, 2013), or to reach a hotel, family or public shelter (Murray-Tuite and Wolshon, 2013; Deka and Carnegie, 2010), possibly including intermediate trips (Trainor et al., 2013; Murray-Tuite and Mahmassani, 2003; Yin et al., 2014), where heavy road congestion is likely to occur (Litman, 2006). And third of all, there may be emergency services trying to reach the disaster site. That is, many disasters require the transportation of a significant amount of emergency services personnel, e.g. for delivery of medical care, water, food and utility maintenance (Litman, 2006), traffic regulation (Tu et al., 2010), helping stranded vehicles and protect and evacuate non-self-reliant people (Litman, 2006) and rescue operations (Dombroski et al., 2006).

As urban transportation systems typically have a modest capacity compared to the local population and workforce, they can easily become overloaded by the surge in travel demand and/or reduction of supplied capacity caused by the characteristics of the emergency situation listed above. This poses a problem for the resilience of the system and necessitates authorities to have proper transportation management strategies. In this regard, the added value of a transportation model is twofold. Firstly, such a model enables quantifying the effects of disruptions as well as management strategies, where the former is often used to identify the most critical emergency conditions and vulnerable parts of the network, and the latter assists in ranking alternative strategies and measures. Secondly, such a model predicts the manner in which these effects take place with respect to travelers' decisions and traffic flow operations, which is helpful in understanding the underlying causes why certain effects occur and certain measures are successful, or not. Evidently, for a transportation model to be of value, its predictive validity is essential.

This brings us to what requirements such a transportation model should satisfy. Based on literature and the previously discerned commonalities of emergencies, we can identify seven main model features that are needed to capture the transport-related characteristics of an emergency. Each affects the choice model for the behaviour of the affected people, the network loading model for the propagation of traffic, or both. We find that a transportation model for emergencies should:

- *Be dynamic*. An emergency situation and the emerging traffic conditions are time-varying (Fu and Wilmot, 2004), and consequently people's choices also have a time dimension. It is important that such time dynamics are taken into account (Lin et al., 2009). The time range depends on the type and severity of the disaster, but would

typically vary from about one hour for no-notice and short-notice disasters to a few days for hurricane evacuations.

- *Describe the relevant choice behaviour*. An emergency is an unusual situation and as mentioned above, people need to make choices dynamically over time, instead of planning the whole day in advance. They may even need to adjust their choices en-route based on the information then available to them at that moment (Pel et al., 2012; Robinson and Khattak, 2010). The emergency can also put people in entirely new choice situations, such as evacuation-related decisions, resulting in unusual behaviour.

- *Predict the initial conditions of the emergency* or otherwise allow specifying these starting conditions if determined exogenously. The initial locations of people evidently affect the travel demand pattern, as it determines where people depart from (Noh et al., 2009) or where people need to be picked up (Murray-Tuite and Mahmassani, 2003), while the initial traffic pattern affects the network performance, together with aspects of possible capacity reductions and induced emergency traffic flows.

- *Include interactions between individuals*. Particularly for evacuations, households tend to act as a unit (Murray-Tuite and Mahmassani, 2003) due to activities at the household level, such as the pick-up of household members (Murray-Tuite and Mahmassani, 2003; Trainor et al., 2013) and necessary purchases (Yin et al., 2014). It has been shown that these activities may have important implications for traffic flows (Murray-Tuite and Mahmassani, 2004; Lin et al., 2009). We presume that interactions within or across households may also play a role in other disasters that severely disrupt the transportation system.

- *Be multimodal*. Emergencies may directly affect various transport systems, or may cause severe spillover effects especially in urban regions with interacting transport modes, as, for example, public transport and pedestrian traffic play an important role as fallback alternatives for people without a car or in case of severe congestion (Shiwakoti et al., 2013).

- *Include travelers who are not directly affected* and their behaviour. This may pertain to, for example, background traffic that itself is not affected by the emergency but does affect the situation as they co-consume road and public transport capacity, or is affected indirectly by changes in the traffic situation or the availability of destinations (Murray-Tuite and Wolshon, 2013).

- *Include emergency services*. This can be either to enable evaluating the deployment of emergency services (as decision variable) or evaluating how these traffic flows affect the situation similar to the previous requirement.

Despite the existence of models that address subsets of this set of challenges, we lack methodology and tools to satisfy all requirements in an integrated manner, hampering disaster planning. In this chapter, we address this knowledge gap by presenting a generic modelling methodology to simulate the impacts of emergencies on urban transportation networks. To adequately incorporate the choice behaviour of the affected people, we propose using an activity-based escalation model for travel choices, which we show to connect well to the existing literature on travel choices during emergencies. Additionally, the chapter contributes a new and computationally efficient methodology to couple such a choice model with macroscopic or mesoscopic dynamic network loading models for the simulation of

evacuations as well as other emergency conditions, in a way that satisfies the listed requirements. We retain a high amount of flexibility in the specification of the choice model that can even include en-route choices, and we show that with the escalation-based formulation we propose, this is sufficiently flexible to incorporate insights from earlier studies on choice behaviour in a wide range of emergencies.

We present this methodology in Section 2.2. Through a case study for a hypothetical evacuation of the city of Delft, we discover a number of important modelling issues, such as to capture interactions between transportation modes and between inbound, outbound and background traffic, as these show to potentially cause failure mechanisms that may be overlooked with a less comprehensive model. This model application is presented in Section 2.3. In Section 2.4 we conclude with a discussion on the model structure, its performance, and the modelling issues highlighted by the case study.

2.2   A general methodology to model travel choices and traffic propagation

As mentioned in the introduction, emergency conditions may both affect the choice behaviour of the affected people and the propagation of traffic. Therefore, we derive specific structures of the choice model in Subsection 2.2.1 and the network loading model in Subsection 2.2.2 that satisfy the indicated requirements. We couple them in Subsection 2.2.3 resulting in a general modelling methodology for the complete transportation simulation of an emergency.

2.2.1   Travel choice modelling

To describe travel choices, we start by acknowledging that the demand to travel is derived from the demand to undertake activities at different locations (Bowman, 2009; Ortúzar and Willumsen, 2011). By modelling complete activity patterns, rather than individual trips, one can consider resource (e.g. vehicle) and task allocation within households, whether people travel together, how activities are dynamically rescheduled, and the consequences for the load on the transportation system. When applied in the context of emergencies, this means that an activity-based model can predict the locations and activities of individuals and the vehicles they use at the time of an emergency event (i.e., the initial conditions), as well as it can simulate background traffic that is unaffected by the event, including transition effects from normal conditions to emergency conditions.

Our activity-based approach implies microscopic agent-based choice models. The modelled agents represent individuals, resources (e.g. vehicles), or groups of communicating individuals that make decisions together (e.g. households). Depending on their type, agents can have constant and mutable attributes, including preferences, memory, and physical locations, and can obtain information, think, and make choices.

Our approach generalises earlier methods where a normal day model forms the basis for an emergency model. Noh et al. (2009) use normal day demand matrices to estimate evacuation demand matrices per time-of-day. Lin et al. (2009) use an activity-based model for determining evacuation demand, and later Yin et al. (2014) use a more advanced one, but they cannot include time dynamics in the choice process. However, unlike these previous attempts, the framework we present here should ensure sufficient flexibility to specify how people dynamically respond to the emergency, e.g. by rerouting, rescheduling activities or evacuating.

To this end, we define an escalation model to categorise behavioural responses to emergencies, consisting of three possible behavioural states of individuals at any moment in time. These are an initial state, for those who are not or not yet affected by the emergency, an adaptation state for those responding to the disruption of the transportation system and an evacuation state for those directly threatened by the emergency. As individuals become increasingly affected, their choice behaviour escalates, causing shifts in preferences and resulting in changes compared to the original activity-travel patterns.

Relying on existing literature, let us now summarise the most important choice behaviour associated with the elements of the activity-based escalation model we propose:

1. In the initial, *normal state* the individual performs its (equilibrium) travel and activity plans as usual, that is consistent with a normal day.

2. In the *adaptation state* the individual responds to the disruption and may adapt its activity and travel plans accordingly by e.g. switching routes or rescheduling activities. For example, Kitamura and Fujii (1998) and Joh et al. (2006) propose models that, given an initial schedule, evaluates the utility of possible changes to the activities and their durations, sequencing, locations and modes, to see if a significant improvement can be found that outweighs the (mental) effort of the re-evaluation process. However, these models are not yet specifically targeted to within-day re-planning in response to unforeseen events, that additionally requires an estimate of perceived future travel times. Illenberger et al. (2007) do propose a model for this, focusing on the time and route choices in the schedule and comparing various assumptions on the availability of travel time information. Knapen et al. (2014) reschedule begin and end times of planned activities in response to unforeseen events and use an explicit model to dynamically estimate the perceived travel times from a combination of normal day travel times and incident characteristics. Analysing empirical data, Knoop et al. (2010) find that the presence of a traffic incident is an even stronger encouragement to switch routes than travel time differences alone, and that the response of travelers is delayed.

3. In the *evacuation state* the individual either evacuates or seeks shelter. Obviously, this third state is only relevant when imminent danger is present, as well as acknowledged and acted upon (Leach, 1994; Vorst, 2010), which, for example, Dixit et al. (2012b) incorporate by modelling risk attitudes. Besides this choice on whether and when to engage in evacuations, which can be captured with e.g. a sequential binary logit model, research on evacuation behaviour has traditionally focused on two other choices related to the evacuation trip itself, namely the accommodation type and destination choice, with family and friends as the most favored and public shelters as the least favored accommodation types, and the mode choice, where, if possible, evacuation by car is the most preferred option. Murray-Tuite and Wolshon (2013) give a comprehensive overview of knowledge and models resulting from this. Nonetheless, there is increasing attention to activity-based aspects of evacuation modelling (Trainor et al., 2013). In particular, returning home for pick-up activities within households is an important aspect (Murray-Tuite and Mahmassani, 2003; Murray-Tuite and Mahmassani, 2004), which was recently re-emphasised by the Great East Japan Earthquake (Hara and Kuwahara, 2015). Yin et al. (2014) formulate and estimate a detailed activity-based model for evacuation behaviour, including child pick-up activities, shopping activities to make necessary purchases and joint travelling with other households. Regarding route choice, Sadri et al. (2014) have found that people

tend to choose familiar routes during evacuations. Despite this, Robinson and Khattak (2010) find that more people are willing to make en-route choices than in normal circumstances. Pel et al. (2012) recommend to model choices both pre-trip and en-route.

From this overview, we see that in addition to satisfying the requirements listed in the introduction, our proposal of an activity-based escalation model relates well to existing literature on both adaptation and evacuation behaviour. This allows existing choice models to be embedded in our framework: the agent-based models can be directly incorporated whereas aggregate-level models can be easily translated to the agent level. Our setup requires a description of under what conditions the behaviour of an agent may escalate as well as of dynamic choices within each state. In general, these choice models can be a function of the plans and experiences on a normal day, the characteristics and attitudes of the considered agent, the on-going emergency event, and the information available to the agent. Interaction with other agents can be included here as well, in the sense of either joint decision-making or responding to perceived previous choices of others. Overall, our model structure thus provides a high level of flexibility with respect to the specification of the choice behaviour, which, given a particular emergency situation, can be filled in accordance with the cited literature.

For clarity, note that, unlike the activity-rescheduling model by Knapen et al. (2014) and the evacuation model by Lin et al. (2009), the above choice modelling includes route choice which is hence also performed at agent-level. This ensures that at any time, an explicit location is defined for each individual, so that not only pre-trip but also en-route choices can be modelled, which in turn increases the realism of local traffic dynamics in the model (Knapen et al., 2014). Furthermore, this allows to explicitly incorporate observed heterogeneity in route choice, e.g. the tendency of people to choose familiar routes during evacuations, allowing to further increase the predictive validity of the model.

2.2.2   Transportation network loading

Following our list of requirements, here we prescribe a multimodal dynamic network loading model that can include emergency services. A key characteristic to decide upon is whether the traffic simulation will be microscopic, mesoscopic or macroscopic, i.e. whether traffic is represented as individual vehicles, vehicle packets, or aggregated flows (Hoogendoorn and Bovy, 2001). Note that our microscopic (agent-based) representation of travelers in the previous choice models does not necessitate also having a microscopic traffic representation for the network loading model. At the same time, mesoscopic and macroscopic traffic flow simulation models typically are computationally more efficient, which is beneficial for applications that are large-scale or require iterative optimisation, as well as suffice with aggregated data for calibration and validation purposes, which is beneficial since current empirical knowledge is limited with respect to individual driving behaviour, e.g. during emergency conditions (Hoogendoorn, 2012; Tu et al., 2010) and adverse weather conditions and due to heavy vehicle loads (Litman, 2006; Lindell and Prater, 2007) and traffic incidents during evacuations (Robinson et al., 2009; Fonseca et al., 2013). For these reasons we propose a network loading model with mesoscopic or macroscopic traffic representation.

Note that such a traffic representation relates to all vehicles on the road network, potentially distinguishing separate user classes with specific properties for describing affected individuals, background traffic, emergency services and public transport vehicles. Further details are discussed in the next subsection. Furthermore, this approach can be extended to the pedestrian network, e.g. using a uni- or bidirectional pedestrian fundamental diagram

(Flötteröd and Lämmel, 2015), allowing a relatively simple extension of the model to multimodal networks. If the amount of pedestrian traffic is significant, one could also model the interaction between pedestrians and cars macroscopically (Meschini and Gentile, 2009).

2.2.3   Integration of the choice model and network loading model

In combining the dynamic network loading model and the choice model, attention needs to be paid to their interaction, which is challenging in this case because the latter operates at a microscopic level of detail while the former does not. To this end we develop a new method to tightly integrate these models, considering both equilibrium and emergency conditions, with the inclusion of en-route choices, public transport and emergency services. This method relies on both a serial procedure and a parallel procedure.

The *serial procedure* solves the user equilibrium assignment, representing a normal day. One runs the choice model to yield dynamic route demand, that is then input to the network loading model to yield dynamic travel times, that are then input to the choice model, and so on (Lin et al., 2009). The method of successive averages (Ortúzar and Willumsen, 2011), which is usually applied to macroscopic flows, can be adapted to find the equilibrium of agents: to simulate flow averaging, one can fix the choices of a random share of agents that increases in size over iterations, eventually yielding agent choices that reproduce the equilibrium situation. This may be extended further into, e.g., the more comprehensive approach by Raney and Nagel (2006).

The *parallel procedure* solves the non-equilibrium assignment, representing the emergency situation. During execution of the network loading model, one can already determine dynamic travel times up till the current time. While tracking individuals throughout the network, we repeatedly alternate between the network loading model and the choice model as time progresses: after each network loading time step, the choice model receives the locations of travelers and the current travel times, and provides the (possibly adapted) departure and route choices for the next time step.

The results for the normal day equilibrium, to be found using the serial procedure, serve as input to the emergency choice model described previously, so that the system remains in equilibrium until the emergency situation causes a disturbance, to be simulated via the parallel procedure. The overall process is illustrated in Figure 2.1. The parallel procedure does require the emergency choice model to be causal, which is a realistic assumption as people can only base their (expectations and) choices on the conditions up till now (Pel et al., 2012; Qian and Zhang, 2013).

Although a serial procedure can simulate the emergency scenario as well (Lin et al., 2009), the parallel procedure skips the construction of intermediate infeasible solutions where some people depart for their next trip before they arrived from their previous trip. The absence of iterations makes this method also much more efficient. For these reasons, we propose the parallel procedure in our methodology. Of course, the parallel procedure does require the software of the choice model and the network loading model to be tightly integrated so that the overall model can rapidly alternate between them.

Figure 2.2 shows a more detailed flow chart for the parallel procedure for simulating the emergency scenario. Here, we added an optional control component that represents the actions undertaken by authorities during the emergency situation, which affect the
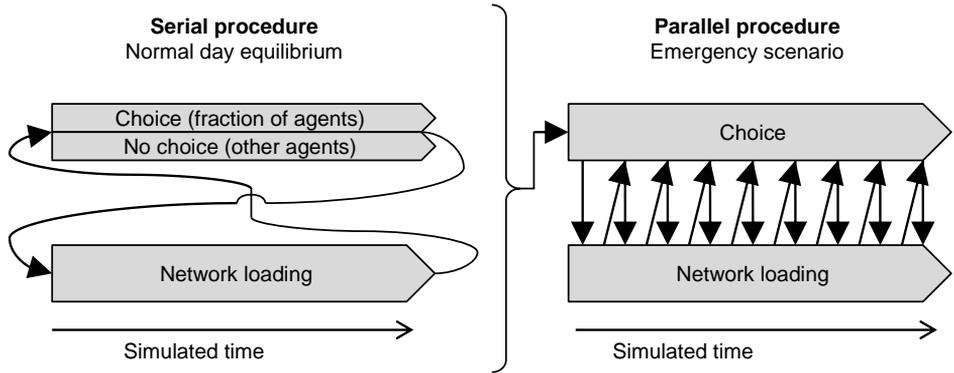
Figure 2.1. Flow charts of procedures for the overall model.
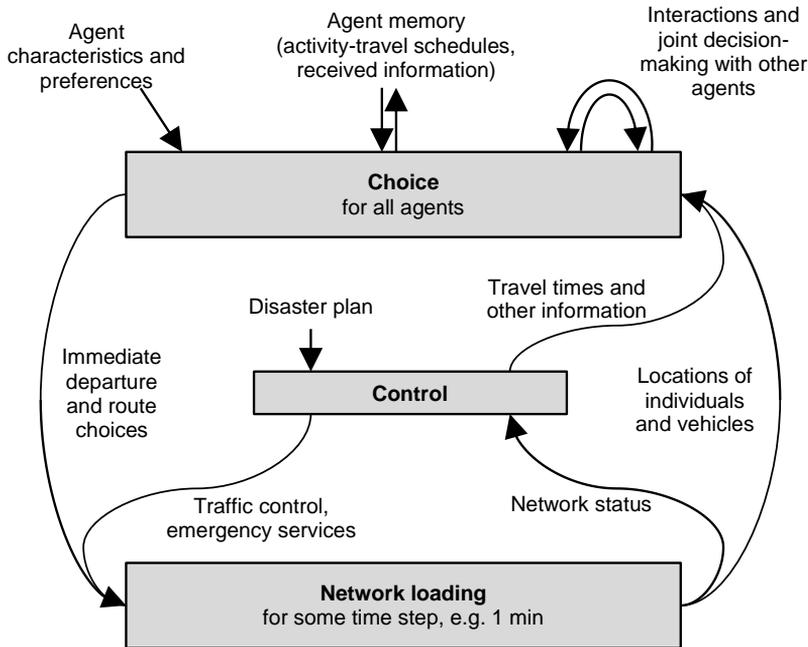
Figure 2.2. Detailed flow chart of the parallel procedure for the emergency scenario.

transportation system either directly via traffic control and deployment of emergency services or indirectly by influencing the choice behaviour of the population.

Tracking individuals and handling en-route choices

One requirement for the parallel procedure is that the network loading model continuously informs the choice model of the location of individual travelers. In a microscopic network

loading process, this is straightforward, and it is also easy to immediately apply, for example, a route change in response to an en-route choice. However, we will show that we can also do this in a macroscopic or mesoscopic network loading process by accurately tracking the location of individuals while the aggregated traffic flow is propagated through the network.

Dynamic network loading models propagate traffic at intersections based on turn fractions. However in this application these are not constant over time and the time dynamics depend on the delays the traffic encounters before reaching the intersection as well as the en-route choices that people make. As explained below, we solve this by disaggregating the traffic flow into various commodities with different routing behaviour, so that the turn fractions are specified per commodity rather than for the total flow (Daganzo, 1995a; Papageorgiou, 1990; Yperman, 2007). We thereby choose the disaggregation such that we can both track the location of individual travelers and delay the definitive route choice of a traveler until she/he passes the relevant divergence point. More specifically, we define that each commodity is associated with exactly one spatial position, i.e. a link, a cell of a link, or a packet of vehicles in a link, depending on how space is discretised. At the next intersection, the commodity is converted into other commodities associated with downstream links, according to particular splitting rates (i.e. turn fractions). Furthermore, we define that if an individual microscopic vehicle is en route, it is associated with exactly one of the commodities in the network. This setup is schematised in the class diagram in Figure 2.3.

The currently associated commodity of a vehicle defines its current location. As soon as the first link outflow of a commodity is registered, its splitting rates are fixed and all associated vehicles are immediately transferred to the corresponding commodities on the outgoing links
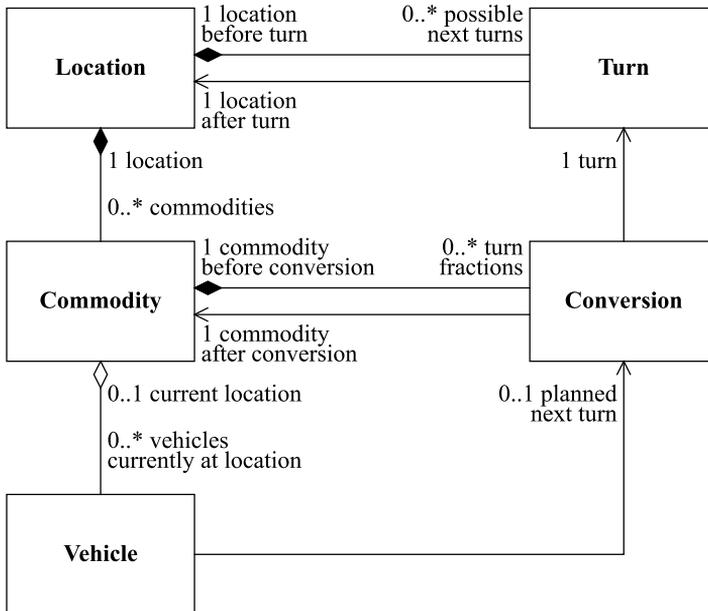
Figure 2.3. Unified Modelling Language class diagram of commodities.

of the intersection, even if not all corresponding flow passes the intersection within that time step. If a vehicle needs to switch to a different route due to an en-route choice, we simply modify the splitting rates of the currently associated commodity in response to its new planned next turn. New commodity objects are constructed on the fly as needed.

At an origin, all vehicles with the same departure time and first link can be grouped into a commodity, and this traffic will then split up into an increasing number of distinct commodities upon passing downstream intersections as their routes diverge. This aggregation of multiple vehicles into commodities saves us computation time compared to creating separate commodities for every vehicle.

Incorporation of public transport

The previous definition of commodities considers only private transport. To incorporate public transport, we define an additional commodity for each public transport vehicle for each location along its route, such that the road traffic propagation model naturally simulates any interactions, including, e.g., occupancy of road space and adaptation of the speed of the vehicle subject to the traffic conditions. Of course, if the network loading model can differentiate user classes, these public transport commodities may have different dynamics than normal cars.

The location of the currently associated commodity of a vehicle, as defined for the unimodal case, will be the key to modelling how (individual) passengers propagate throughout the public transport network. For this, we require that network nodes containing public transport stops must have sink and source capabilities in the network loading model. We can then define the following system:

- Each public transport stop node has a queue of passengers waiting to board any line.

- Each public transport vehicle has a list of passengers who are on board.

- Passengers arriving at a public transport stop via incoming links, e.g. of the pedestrian network, use the sink capability and are placed into the queue. If they arrive by private vehicle, this vehicle must be parked at the node, and this must fit within the parking capacity.

- Public transport vehicles arriving at a public transport stop via incoming links use the sink capability to stop.

- Passengers who want to alight from a vehicle, use the source capability to leave via outgoing links, e.g. of the pedestrian network, and are removed from the vehicle's passenger list. Alternatively, if they want to transfer to another public transport line at the same stop, they are removed from the vehicle and added to the queue of the public transport stop. If they continue by private vehicle, this vehicle is retrieved from the parking capacity of the node.

- Passengers in the queue, who want to board a vehicle, are added to the vehicle's passenger list and removed from the queue, until the vehicle's capacity is reached.

- Public transport vehicles, when ready, depart from stops using the source capability via outgoing links.

This fully defines the propagation of both public transport vehicles and passengers, and the transfers of passengers from and to the system of private modes. The next two rules complete the system with support for en-route choices of public transport passengers:

- Passengers in a vehicle may choose to switch to any other public transport route containing the part of the public transport line where they currently are.

- Passengers in the queue may choose to switch to any other route from the corresponding node, with or without public transport. If they do not need to board anymore, they use the source capability to leave via outgoing links, e.g. of the pedestrian network, and are removed from the queue.

Incorporation of emergency services

Finally, any emergency services may be modelled similarly to public transport, but without intermediate stops and passengers, and with the possibility of having multiple vehicles belonging to a commodity. Like public transport, it is up to the network loading model to correctly model the interaction with other traffic by treating the emergency services as a separate user class, potentially with dedicated infrastructure (Litman, 2006; Maassen, 2012), different speeds (Petzäll et al., 2011) and priority at intersections (Teng et al., 2010). The "timetable" of emergency services may be derived from the disaster plan. If emergency services are deployed to regulate traffic (Tu et al., 2010), then the model may activate the regulation once they arrive at their destination.

Summary

Let us briefly summarise the proposed framework. The travel choices follow from an activity-based model. On a normal day the activity-travel patterns of agents are in equilibrium. We can find this equilibrium by adapting the method of successive averages to discrete agents, which iterates until the choices resulting from the choice model are consistent with the travel times resulting from the network loading model. In an emergency scenario agents begin executing their equilibrium activity-travel patterns, but may show adaptation behaviour after noticing a disruption and evacuation behaviour after being confronted with a threat, possibly interacting with other agents or influenced by information provision by authorities. Consequently, the traffic situation will gradually start deviating from that of a normal day. For the propagation, we use a macroscopic or mesoscopic dynamic network loading model, possibly including traffic control measures, in which we represent agents using commodities. The commodities allow us to track the locations of agents and handle their en-route choices. We define additional commodities to represent public transport vehicles and vehicles of emergency services, and define mechanisms for agents to board and alight at public transport stops.

2.3   Delft evacuation model application

To demonstrate the feasibility and investigate the properties of our newly proposed methodology, we here present a case study application. The application describes the hypothetical multimodal evacuation of the city of Delft with the following setup. In our scenario, authorities take no action other than informing the public, and the public transport system operates according to the normal timetable. This means we are essentially investigating a do-nothing variant, whose results can be used as a starting point to develop more proper control strategies and as a reference point when evaluating such strategies.

Households living in external zones follow their normal activity-travel patterns, and households living in Delft start evacuating between 16:00 and 17:30, thus interacting with initially substantial background traffic due to the evening peak. The Delft transport network is plotted in Figure 2.4 and consists of motorways and provincial roads, main arterials, and urban streets for cars and busses as well as rail infrastructure. In total it contains 24 centroids (7 are external), 4 train stations (2 are external), 40 nodes with bus/tram stops and 437 other nodes. There are 1206 uni-directional vehicular links and 1092 uni-directional pedestrian links. Counting each direction separately, the public transport network contains 4 train lines, 2 tram lines and 14 bus lines. We based our network on the default network included with the OmniTrans modelling software that we also use for generating artificial public transport timetables and visualising the results of our model.

For the evacuation choice model, we assume that people go home prior to evacuating, depart from home once all household members have arrived, use their car if possible, use public transport or walk if not, and choose their routes according to instantaneous travel times (i.e., based on current conditions) and household-level personal preferences. If multiple members of a household want to make the same trip, we assume they share a single car. Preferred evacuation destinations are assumed to be randomly distributed equally between the northern and southern safe destinations (see Figure 2.4). The southern destination is reachable only by the A13 motorway, by train and by bus, whereas the northern destination is also reachable by the A4 motorway, by a small local road and by tram.
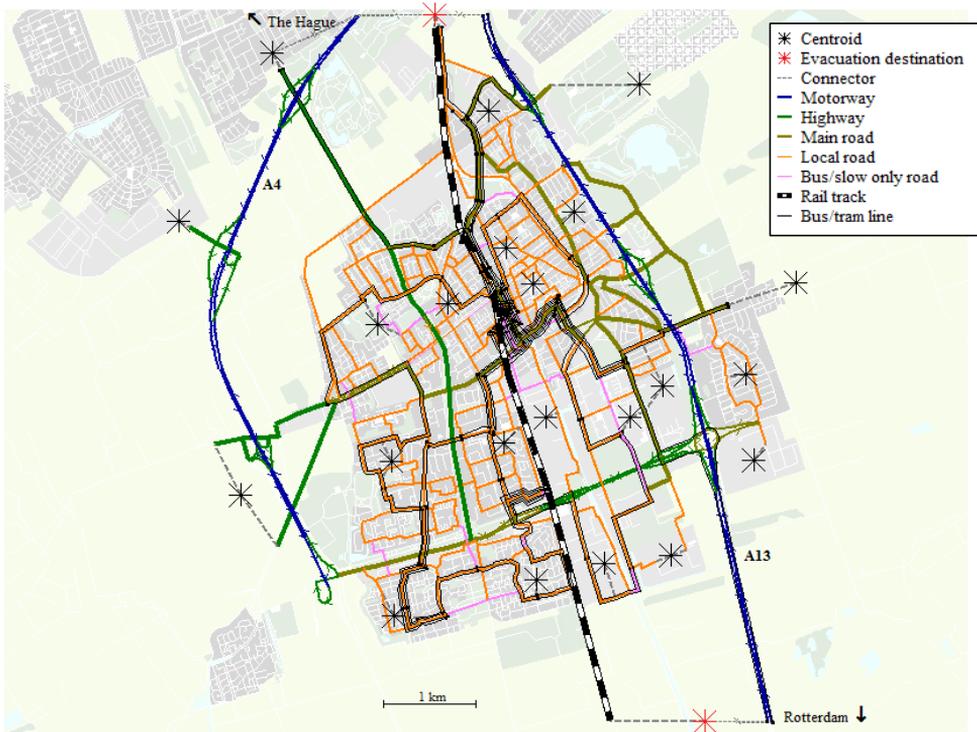
Figure 2.4. Network of the city of Delft used in our case study.

- Household members, vehicles, home location.
- Normal-day route preferences of household.
- Route choice sets.

- Normal-day activity-travel patterns.
- Indicator of evacuation notice reception by household.
- Chosen evacuation destination of household.

- Wait for other household members before starting final evacuation trip together.
- Share cars with household members.
- Share route preferences within household.

Choice for all agents. Execute normal-day activity-travel patterns until notified and home is in evacuation area. In the latter case:
- Departure choice: go home once notified, then evacuate.
- Destination choice: first home, then fifty-fifty north/south.
- Mode choice: car if available, public transport or walk otherwise.
- Route choice: random utility maximisation using instantaneous travel times and normal-day route preferences.
These choices are also evaluated en-route.

- Gradual notifications for households living in Delft.
- Instantaneous link travel times and predicted arrival times of public transport vehicles.

Disaster plan: information provision and normal public transport operation.

Immediate departure and route choices.

Locations of household members and vehicles.

Control: provide information and try to operate public transport according to normal timetable.

Attempted execution of normal public transport timetable.

Traffic information for all modes.

Multimodal **network loading** for 1 minute using the Link Transmission Model. Propagates cars, pedestrians and public transport vehicles.

Figure 2.5. Overview of how the evacuation scenario implements the general framework of Figure 2.2.

Before we present the simulation results, in the next subsection we first discuss model implementation issues relating to the choice set generation, the activity-travel pattern generation for a synthetic population, and the multimodal dynamic network loading. Figure 2.5 summarises the main points of the case-specific assumptions and methodology for the emergency scenario.

2.3.1   Case-specific methodology

Route choice sets

As stated previously, route choice, including en-route choices where individuals adapt to the prevailing conditions, is simulated microscopically. For sake of computational efficiency we generate explicit route choice sets beforehand, rather than repeatedly searching for a new fastest path for each individual. To include en-route choices, we need to generate a choice set for each location in the multimodal network where an individual may update his/her route. Furthermore, each choice set should include all destinations reachable from the corresponding location, with one or more alternative routes for each possible mode.

For the private modes, we generate choice sets at all road network links, including connector links. Here an accelerated Monte Carlo approach is used (Fiorenzo-Catalano and Van der Zijpp, 2001), where we repeatedly draw link travel times τ from normal distributions with mean equal to the free flow time τ_0:

$$\tau = \max\left(\tau_0 + \varepsilon\sqrt{\theta\tau_0}, 0\right) \text{ with } \varepsilon \sim N(0,1) \tag{2.1}$$

We repeat this for 500 iterations, linearly increasing the dispersion coefficient θ from zero to 0.0025 h. In each iteration the shortest paths are extracted using Dijkstra's (1959) algorithm and merged into the choice sets of the links if new. Note that we execute Dijkstra's algorithm backwards to efficiently generate routes from each destination centroid to each link. Turn prohibitions are taken into account by operating at link level rather than node level. For efficient storage in computer memory, routes are defined recursively as a combination of the first link and, if this is not the only link, a reference to the route after that link.

The inclusion of public transport is complicated by the common lines problem where travelers can choose among multiple lines for a single public transport leg (Kurauchi et al., 2003). One possible solution is to merge all lines between the boarding and alighting stops into a single public transport link representing the effective level of service (Cominetti and Correa, 2001). This simplifies route choice as individuals then board the first departing vehicle along the lines. However, to include en-route choices, the location of an individual needs to be known in more detail to determine at which intermediate or later stops a traveler can choose to alight. In this application, we therefore define public transport links between each pair of consecutive stops, merging any lines serving that pair. We extend the previous route definition for the unimodal case to include a set public transport line numbers that may be used to traverse the link, which is constant over a leg.

Combining these public transport links with the links of the pedestrian network, we use a version of the branch-and-bound algorithm (Friedrich et al., 2001) for the generation of public transport routes, again working backwards to construct routes from each destination centroid to each link. In the algorithm, we concatenate public transport legs and pedestrian legs, and save the generated route for each link in each leg. As input, we generate pedestrian legs for access, egress and transfers with Dijkstra's algorithm, considering only the shortest routes for simplicity. One public transport leg is generated for each possible combination of a boarding and alighting stop along each line, with common lines merged. While concatenating legs, we only use the following logical constraints:

- a route may not contain two consecutive pedestrian legs;

- a consecutive pedestrian leg and public transport leg (in any order) may not have any node in common, except for the boarding/alighting node, to prevent walking between stops of a used public transport line;

- routes may not be cyclic, except if the cycle is at the start of the route, since these circular routes could be chosen by travelers en-route to return to their origin;

- no individual pedestrian leg may exceed a specified maximum walking time (30 minutes), where the maximum is lower for transfer legs than for access/egress legs (5 minutes);

- a specified maximum number of public transport legs per route may not be exceeded (2).

Synthetic population, activity-travel patterns and route choice

For practical reasons, we use the existing Albatross model (Arentze and Timmermans, 2008) to generate synthetic households with normal-working-day activity-travel patterns, for its base year 2004. This model has been calibrated for the Netherlands using annual census data and travel diaries. From the Albatross output we extract all households that conduct any of their activities inside the study area, as well as households generating through traffic. For the latter, we extract households with one or more trips between any (near) location north and south of Delft, as Delft is situated in a corridor. The size of our synthetic population equals 20% of the representative population and hence all households are assigned a weight of 5 in the network loading simulation to get correct traffic volumes. Finally, note that Albatross constructs the activity-travel schedules for the adults of a household, and the synthetic households only indicate the age of the youngest child, but not the number of children. For simplicity, we assume that such households have one child, who needs to be picked up at home in case of an evacuation.

Albatross distinguishes a car driver mode, a car passenger mode, a public transport mode and a slow mode (Arentze and Timmermans, 2004). Note that this does not include park-and-ride separately, hence we assume the slow mode is the access and egress mode for public transport. To avoid explicit modelling of car sharing within and between households on a normal day – constraints that are principally handled within Albatross – we always create a car for users of the car driver mode, even if the number of cars owned by the household is lower according to Albatross. For trips not related to the evacuation, we "teleport" car passengers directly to their destination for simplicity, avoiding the need to explicitly couple every car passenger to a car driver as this has no influence on the traffic conditions. This also happens for public transport and slow mode trips without any route available from their origin, which may occur for some external zones in our case. The slow mode covers walking and cycling, and these users are also simulated jointly in our case study.

Until the household joins the evacuation, household members must execute their activity-travel patterns as specified, leaving only routes to be chosen. Route choice on a normal day is assumed to follow the (multimodal) dynamic stochastic user equilibrium, approximated by the serial procedure in Subsection 2.2.3. The weights used in the method of successive averages are set as $\alpha_i = i^{-2/3}$, where i is the iteration number (Polyak, 1990), instead of the usual $\alpha_i = i^{-1}$, to emphasise the later iterations. Thus, in each iteration we randomly select $\lfloor \alpha_i N \rfloor = \lfloor i^{-2/3} N \rfloor$ households to update their choices, where N is the number of households. The sampling is stratified with respect to the number of iterations ago each household was last updated. Furthermore, we prioritise households who can gain more than 45 minutes of utility by changing the route of a single trip.

Households' route choice follows Random Utility Maximisation, where route disutilities U^r are based on to-be-experienced travel times for choices on a normal day and instantaneous travel times for choices during the emergency conditions:

$$U^r = \sum_{a \in A^r} \left(\tau^a + \max\left(\varepsilon^a \sqrt{\theta \tau_0^a}, -\tau_0^a \right) \right) + \beta_{\text{PT}} \max\left(\varepsilon^{\text{PT}}, 0 \right) \left| P^r \right| + \sum_{p \in P^r} \left(t_{\text{wait}}^p + t_{\text{IV}}^p + \sum_{a \in A_p} \max\left(\varepsilon^a \sqrt{\theta \tau_0^a}, -\tau_0^a \right) \right) \quad (2.2)$$

with $\varepsilon^\bullet \sim N(0,1)$. Equation (2.2) sums up the travel times τ^a of private mode links $a \in A^r$, and the in-vehicle time t_{IV}^p, waiting time t_{wait}^p and a boarding penalty β_{PT} of public transport legs $p \in P^r$, where each component includes an error term to include heterogeneous preferences. For the computation of public transport in-vehicle time and waiting time, we use the departure and arrival times of all individual public transport vehicles as realised in the previous iteration, or as in the timetable for the first iteration; for simplicity we do not include the delay due to vehicles that cannot be boarded due to capacity constraints. In our application, we set β_{PT} to 8 minutes and θ to 0.0005 h. Stochastic error terms $\varepsilon^\bullet$ are generated and stored per household.

For choices during the evacuation, the same error terms are used as on the normal day, such that households' route preferences are fully correlated between these conditions. This effectively yields an increased likelihood that people choose familiar routes to evacuate. We use instantaneous travel times to predict the arrival times of each public transport vehicle at each of its stops, which in turn are used in the utility calculation of public transport routes in the emergency scenario.

Multimodal dynamic network loading

For the macroscopic dynamic network loading model, we use the Link Transmission Model (Yperman, 2007), derived from kinematic wave theory (Lighthill and Whitham, 1955; Richards, 1956). This model discretises time at every node, but, unlike e.g. the Cell Transmission Model, does not need to discretise space within links, both yielding a high computational efficiency and a small numerical error. Here, we maximise the time step of each node, within the constraints set by the adjacent links, and process nodes in parallel when possible in order to minimise computation time. In order to prevent small amounts of traffic from travelling faster than the free flow speed due to numerical diffusion – this would give unrealistically small travel times on the microscopic level and negatively affect the possibilities for en-route choice behaviour – we use a modified method to interpolate cumulative curves in our model, the details of which are explained in Appendix 3.A. The maximum time step of any node equals 1 minute, since travel demand at origins is aggregated per minute and, in the evacuation scenario, the choices of households are updated every minute as well. To avoid numerical errors in the arrival and departure times of public transport vehicles, all nodes with public transport stops are assigned a small time step of 1 second.

We define links and fundamental diagrams separately for vehicular and pedestrian traffic. The fundamental diagram for vehicular traffic is piecewise linear with three pieces, based on the link free speed, link critical speed, link capacity, and a jam density of 180 vehicles per kilometer per lane. For dedicated public transport links, we use a triangular fundamental diagram based on the public transport speed. For pedestrian traffic, lacking specific data, we use a rather arbitrary fixed triangular fundamental diagram with a free speed of 5 km/h.

Our node model is based on the Tampère et al. (2011) node model for unsignalised intersections. Note that this model does not include the interaction of crossing flows on intersections, e.g. via red phases at traffic lights, but only the diverging and merging of flows. As described in Subsection 2.2.3, we need to extend our node model with various source and sink capabilities to handle public transport, which we implement as follows:

- For alighting pedestrians and departing public transport vehicles, we define a source immediately downstream of the node. We give this source traffic absolute priority over traffic coming from the intersection by subtracting the source traffic from the receiving flow provided to the node model.

- Pedestrians arriving at the node to board public transport and public transport vehicles arriving at their final stop, are part of the sending flow provided to the node model, but are destined to a virtual sink turn. Thus, they are constrained by the conservation of turn fractions, but do not show up on any outgoing link. For public transport vehicles, this is consistent with an assumption that the final stop is situated just upstream of the intersection.

- Public transport vehicles arriving at intermediate stops are part of the sending flow, use a regular turn and are also part of the transition flow of that turn, but are removed from the flow just before the transition flow is added to the inflow of the outgoing link. These vehicles thus go through the node model before stopping, consistent with an assumption that the public transport stop is situated just downstream of the intersection.

2.3.2   Simulation results

Below we present the case study results. First of all, Figure 2.6 indicates the trip departures over the day for a normal day, thus following directly from the Albatross output. The "teleportation" trips correspond to all trips not explicitly assigned to the network, of which almost all relate to car passenger trips (and almost none are public transport and slow trips for which no route exists in the case study network). The strange peaks in the figure exist because in Albatross, the time-of-day constraints on activity types are deterministic rather than stochastic, leading to many identical starting and ending times of activities among the synthetic population.

The next step is to perform route choices for the activity-travel patterns of Albatross. Let U_i^r denote the disutility of route r, as in Equation (2.2), using the travel times resulting from iteration i. Let r_i then indicate the chosen route in iteration i, based on the travel times resulting from iteration $i-1$. This allows us to define a duality gap DG_i and a maximum utility difference MUD_i for iteration $i>1$ as follows:

$$r_i = \arg \min_r U_{i-1}^r$$

$$DG_i = \frac{\sum \left(U_{i-1}^{r_{i-1}} - U_{i-1}^{r_i} \right)}{\sum U_{i-1}^{r_i}} = \frac{\sum U_{i-1}^{r_{i-1}}}{\sum U_{i-1}^{r_i}} - 1 \tag{2.3}$$

$$MUD_i = \max \left(U_{i-1}^{r_{i-1}} - U_{i-1}^{r_i} \right)$$

In a perfect stochastic Wardrop equilibrium, both convergence indicators would be zero since the included error terms are drawn explicitly and remain constant over iterations. From Figure 2.7, we observe it is difficult to obtain reasonable convergence using our current method, even though there is almost no road congestion outside the evening peak hours. This might partly be explained by the fact that the travel times allocated by Albatross may deviate from the travel times in our case study network, so that encountered delays in one trip affect later trips as well because departures are delayed. Additionally, in the public transport system, very
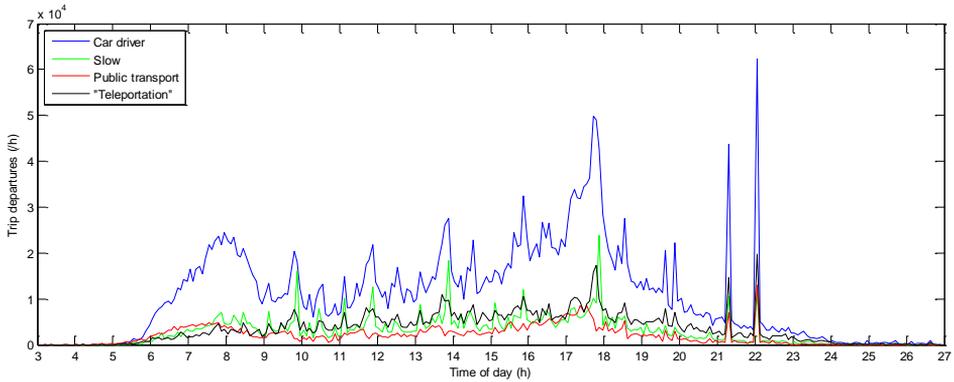
Figure 2.6. Trip departures per mode in Albatross activity-travel patterns for the case study.
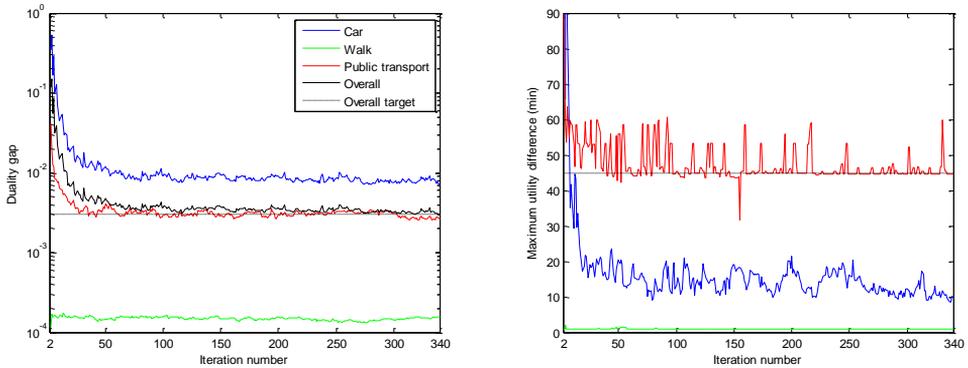
Figure 2.7. Convergence of the duality gap (left) and the maximum utility difference (right).

subtle changes in traffic conditions and hence in running times may have a big impact for individual travelers, as this may force them to board a different vehicle, causing a very unequal distribution of the duality gap over the trips (Figure 2.7, right). The indivisibility of agents (with a weight of 5 each) may also play a role (Bekhor et al., 2014). Of course, in this study we focus on simulating an emergency, but better convergence for the normal day would be desirable. Here, we stop once the overall duality gap is lower than 0.003, which takes 340 iterations.

Proceeding with the simulation of the emergency conditions, the average status of the transportation system between 17:30 and 18:00 is visualised in Figure 2.8 for vehicular traffic and in Figure 2.9 for the public transport network (the pedestrian traffic is omitted here for brevity). For the vehicular traffic (Figure 2.8), we notice many traffic jams on roads inside Delft. Because of the combination of inbound, outbound and background traffic with various destinations each, we observe no clear pattern regarding the directions of traffic jams on the

urban roads – at various places, both directions of the same road are heavily congested simultaneously. Looking at access points to the motorways, the traffic conditions are especially bad for the people that try to evacuate to the south.

One main origin of congestion is the southernmost A13 motorway on-ramp directed towards the south. This bottleneck is also active in the evening peak of a normal day, but in the evacuation scenario we can see it spill back over the urban arterial, via one of its intersections back in the other direction of the same arterial, and then onto the opposite direction of the motorway. The route of this shockwave is indicated with an arrow in Figure 2.8 (right). This is a major conflict between outbound, inbound and background traffic, causing very large delays for through-traffic travelling on the A13 northbound *towards* Delft – this motorway queue does not resolve until 21:30. If the disaster plan would require emergency services to access Delft from the south, they would also be severely delayed.

Due to the same southbound on-ramp, the southbound A13 also suffers from heavier congestion than on a normal day, with much more spillback into the urban network at the other two on-ramps as well. In response, we see many people evacuating to the north using the A4 motorway instead. Looking at the A4, at the circled off-ramps in Figure 2.8 (right), we see traffic heading into Delft blocking the through-traffic on the motorway. Although to some extent this also occurs on a normal day for southbound traffic, this problem starts earlier due to the inbound evacuation traffic and now also blocks outbound evacuation traffic heading north. Hence, at this moment in time, all three motorway routes towards Delft are simultaneously congested, which again could be problematic for emergency services.

For the public transport network (Figure 2.9), all outbound trains are fully loaded until 19:30. The bus and tram lines are however almost not used, due to road congestion as indicated in Figure 2.8. Both tram and bus in the northern evacuation direction show zero passenger flow out of the network between 17:00 and 18:30, as traffic jams cause delays so large that they are too unattractive for evacuating travelers. The next half an hour, the road congestion between Delft central station and The Hague reduces a bit and the tram line is boarded up till capacity before leaving Delft. Later, the passenger flow of the tram line reduces again as for most people taking the train is faster, while the demand from residential zones near the tram line is already decreasing. The bus line remains practically unused by passengers, but it is also barely used on a normal day. Overall there appears to be a difference between user-optimal and system-optimal evacuation in public transport, although this may be overestimated because the *choice* model does not take into account that the train services are overloaded with passengers.

Also, we note that since the last traffic jam resolves after 23:30, there is a large period of time of four hours in which the train lines have spare capacity, but the roads are still congested. This suggests that a true multimodal evacuation plan, where authorities try to convince car owners to use public transport, may significantly reduce the time required to evacuate the city, in this case especially for people with a destination south of Delft. A public transport schedule specifically designed for evacuations may further increase the network production.

We plot the overall progress of the evacuation in Figure 2.10. Here, we see that not all, but a large majority of households can gather at home quickly, yet the final evacuation trip from home to the safe destination takes substantially longer. This supports our earlier expectation that early outbound traffic together with background traffic will activate prominent network bottlenecks, and due to these traffic jams then also late inbound traffic gets delayed.

Figure 2.8. Average vehicle traffic flows (h⁻¹) between 17:30 and 18:00 for the Delft case study on a normal day (left) and in the evacuation scenario (right). Bar widths are proportional to flow, colours indicate traffic speed.
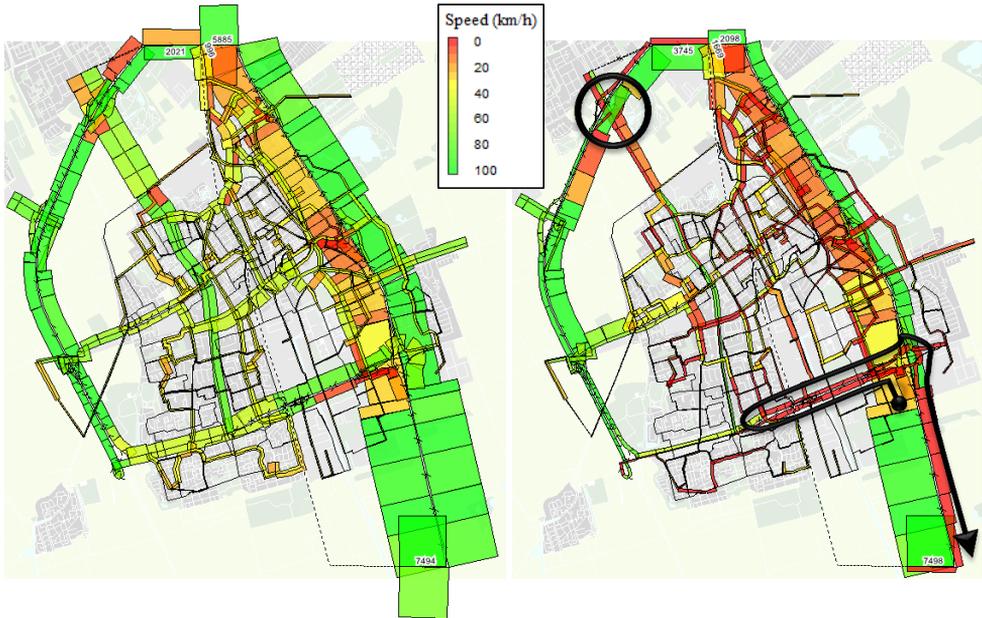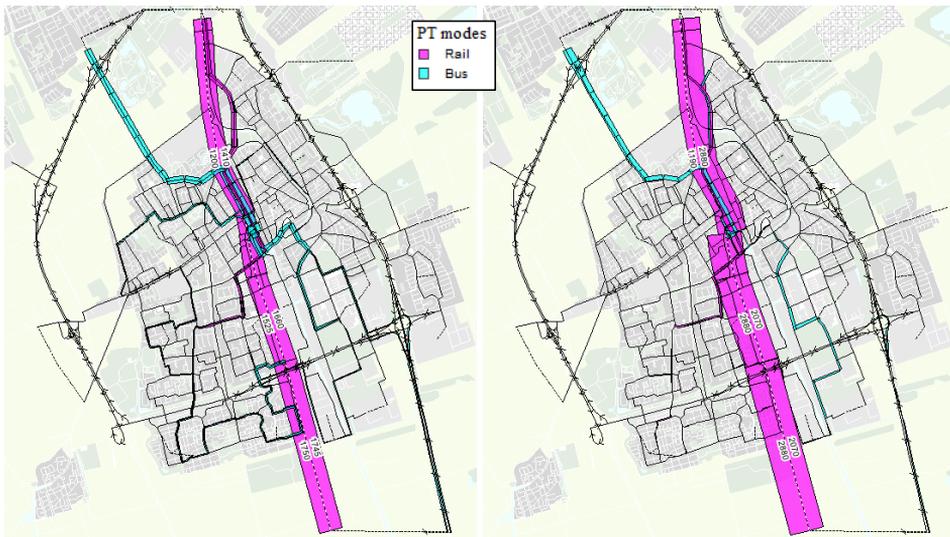

Figure 2.9. Average public transport flows (h⁻¹) between 17:30 and 18:00 for the Delft case study on a normal day (left) and in the evacuation scenario (right). Bar widths are proportional to flow, colours indicate public transport mode.
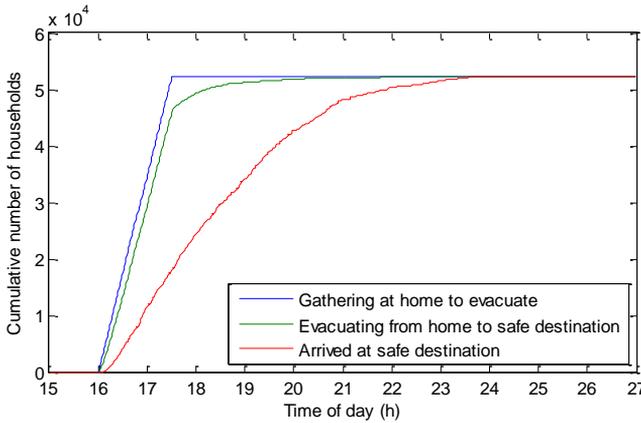
Figure 2.10. Evacuation progress over time for the Delft case study.

Table 2.1. Computation time per model component, measured with a C++ implementation on a Dell Precision T3600 PC with 12 logical processors.

Model component	Computation time
Network and population import	62.1 s
Car/walk route choice set generation	8.9 s
Public transport route choice set generation	11.7 s
Normal day equilibrium simulation	158.8 min (on average 28.0 s per iteration)
Emergency simulation	154.0 s

Finally, we look at computation time. From Table 2.1, we see that although a lot of time is spent finding the normal day equilibrium – simply due to the very many iterations – the actual evacuation simulation is very fast: it takes only about two and a half minutes on our PC. This makes the model suitable for efficiently investigating a large number of possible emergency scenarios and candidate disaster plans. Note that if the overall model is ran with increased accuracy by enforcing a maximum time step of 1 second for all nodes, the convergence of the normal day equilibrium takes a lot longer, but the increased accuracy does not significantly affect the results reported above, indicating that the use of larger time steps was adequate.

2.4 Conclusions

In this chapter, we propose a methodology to simulate a multimodal transportation network during emergency conditions. Starting with model requirements identified from literature, we included an activity-based escalation model for the choice behaviour of individuals in the network and a macroscopic or mesoscopic multimodal dynamic network loading model. We show that such an escalation-based choice model captures the choice behaviour found in literature, and subsequently presented a new method to integrate it with the network loading model for efficient simulation of emergency conditions. As a result, the overall methodology supports arbitrary pre-trip and en-route choice models with interactions among individuals

and is agnostic regarding the modelling of traffic propagation, so that it is flexible and thus general enough to study a wide range of emergency situations, and corresponding management strategies of authorities. Our case study application showed that the resulting model structure is indeed very efficient for the simulation of emergencies. This for example allows model users to compare various alternative management strategies and to check the robustness of a strategy by varying simulation inputs.

The Delft case study application yielded the following insights, which both provide directions for the development of urban evacuation plans, as well as emphasise essential model requirements and thus the need for our comprehensive modelling methodology that includes these complex system dynamics:

- One should consider the interaction of inbound, outbound and background traffic. Within the endangered area, congestion may occur in multiple directions simultaneously. Due to spillback, important roads *towards* the endangered area may become seriously congested. This may have important consequences for emergency services trying to reach this area.

- One should also consider the interaction between modes. Urban public transport may fail due to road congestion. Consequently, public transport users may need to adapt their route.

- User-optimal route choice may not be system-optimal. This is well-known for car traffic, but during evacuations this can also be the case within the public transport network, as this network has its own capacity restrictions.

- Spare capacity may become available in public transport while the road network is still congested. Encouraging car owners to use public transport may hence reduce the total evacuation time.

Once candidate strategies have been developed, preferably with these insights in mind, these can again be assessed with the same framework and model. Important aspects of the case study that need further verification are the equilibrium traffic pattern for a normal day, the emergency choice model, and the emergency driving behaviour in the network loading model. At the same time, relying on an activity-based choice model for a normal day for all members of the population and the need for an accurate model that predicts all travel choices of people during an emergency are likely important practical limitations of our methodology. To remedy the latter, we offer recommendations to develop and calibrate an emergency choice model through a stated-preference experiment in Section 7.3. The methodology presented in this chapter is flexible enough to handle any outcome of such further research.

Acknowledgements

Chapter 3

Propagating agents with macroscopic dynamic network loading: challenges and possible solutions

Abstract

While agent-based modelling of traffic demand is gaining attention, a macroscopic dynamic network loading model may be beneficial, particularly in large-scale applications. We investigate the implications of coupling such models, with inclusion of en-route choices, for the modelling of links and the determination of turning fractions, yielding useful recommendations to help select an appropriate solution scheme of the macroscopic traffic flow theory and overcome other practical challenges specifically associated with the coupling of agent-based traffic demand and macroscopic traffic propagation.

Keywords: agent-based modelling, dynamic network loading, macroscopic traffic model, en-route choice.

3.1   Introduction

Dynamic network loading models represent the roads and intersections of a transportation network as a graph consisting of links and nodes, and use this graph to simulate the propagation of traffic over time. The traffic demand is specified as a dynamic route vector, defining the amount of traffic departing on each route for each moment of time. With this routing information, the dynamic network loading model is able to determine the dynamic turning fractions at each node endogenously. The dynamic route vector itself may also be determined endogenously from a dynamic origin-destination matrix and a route choice model, in which case the overall model becomes a dynamic traffic assignment model.

Within dynamic network loading, the propagation of traffic can be simulated by means of agents, each representing individual vehicles, such that the overall traffic pattern emerges from the interactions between vehicles. Helbing and Balietti (2012) show that complex traffic phenomena that can be reproduced this way. The dynamic network loading model is said to be microscopic in this case (Hoogendoorn and Bovy, 2001). On the other hand, macroscopic models describe traffic at an aggregated level. As discussed by Van Wageningen-Kessels et al. (2015), macroscopic models tend to have less parameters that are more easily observable, while are also getting better at reproducing complex traffic phenomena. Furthermore, compared to microscopic approaches, the computation time of a macroscopic model can be much shorter because one does not need to move all vehicles in the network individually, every small time step. Therefore a macroscopic model can be advantageous, especially for large-scale applications.

Notwithstanding, for modelling purposes other than driving behaviour, agent-based modelling in transportation is becoming increasingly attractive. The primary reason for this is the rise of activity-based modelling, which calculates traffic demand by generating a synthetic population of agents with activity-travel patterns (Bowman, 2009). The coupling of agent-based demand modelling with microscopic dynamic network loading is relatively straightforward: an agent of the population simply maps to a vehicle agent on the road, as illustrated by e.g. Illenberger et al. (2007). Although one could use macroscopic dynamic network loading here by aggregating the activity-travel patterns to a dynamic origin-destination matrix or route vector, the agent is then unable to exercise any choice behaviour while en-route.

This en-route choice behaviour of agents is important in various applications dealing with disruptions and resilience of the transportation system, which can be divided in three categories. Firstly, rescheduling of activity-travel patterns in case of unexpected delays is a good example, where it is particularly important for the traffic dynamics near the location of the disruption (Knapen et al., 2014). A related example is the simulation of transportation systems during emergency conditions, including regional evacuations, as these also benefit from an activity-based approach (Chapter 2). Secondly, outside of activity-based modelling, agents with en-route choices are also needed in other cases where parameters of choice behaviour are sampled. In route choice under either uncertainty or risk, agent-based choice modelling allows to ensure consistent choice behaviour of drivers over time. This applies regardless of whether route choice is modelled as en-route route choice or, as proposed by Gao et al. (2008), as en-route implementation of a pre-trip route policy choice. Thirdly, in multimodal transportation systems, public transport vehicles, e.g. buses, can be modelled as agents in the traffic network, operating the public transport service en-route as efficiently as possible. Public transport passengers can be represented as agents as well and simulated

including access, egress and transfers (Subsection 2.2.3). This allows studying the interactions between modes, and the consequent impacts on reliability and resiliency, in detail.

In this chapter, we therefore explore the fundamental challenges and opportunities that need to be addressed when using agents to represent traffic demand in a macroscopic dynamic network loading model while retaining the option for agents to effectuate individual choice behaviour while en-route. By doing so, we seek to combine the advantages of agent-based modelling of choice behaviour and non-agent-based modelling of driving behaviour.

3.2   Overview of the problem

Let us first start with an overview of the problem at hand, and the general principles behind a solution. In macroscopic dynamic network loading models, flows and densities are continuous variables, which can be disaggregated into so-called commodities, each of which employing specific routing behaviour (Daganzo, 1995a; Yperman, 2007). Hence, with our agent-based traffic demand, we will need to use commodities to link parts of the aggregate traffic flow to individual agents. However, since all flows and densities are continuous variables, an agent does not have an unambiguous location in the network. Its vehicle, represented by disaggregate continuous densities, may consist of multiple vehicle parts that are spread out over multiple links, possibly even with a gap in between. For the purposes of en-route choices, we will use the front of the vehicle, i.e. the first vehicle part, as the location of the agent, and we require that any remainder of the vehicle always follows the same route as the front. This ensures en-route choices are possible and yield a unique route for the agent.

We now proceed to define the interface between the choice modelling and the dynamic network loading. For en-route choices, the current location of the agent, as defined above, will always be a link. Here, the agent needs to choose which turning movement to take at the node downstream of this link, or choose to leave the network there. In other words, the agent needs to have an intended next turn, where leaving the network can be included in the list of possible turns. The intended next turn may be derived from a previously chosen route, where the route choice is periodically updated, or chosen ad hoc whenever the dynamic network loading model requests to use it. Note that an intended next turn must be known before an agent actually leaves the link and the intention is made final, since, depending on the downstream congestion, the intention itself may affect the ability of an agent to leave the link.

Finally, let us look at the possible structures of the macroscopic dynamic network loading model itself. We will here focus on the Lighthill-Whitham-Richards theory of traffic flow (Lighthill and Whitham, 1955; Richards, 1956), which, if needed, can be extended to include more complex higher-order phenomena (Van Wageningen-Kessels et al., 2015). Then, any traffic pattern can be defined as the relation between three continuous variables: time t, location x and cumulative number of vehicles n (Laval and Leclercq, 2013). Given such a relation, the flow $\partial n / \partial t$ and density $-\partial n / \partial x$ are obtained by differentiation. We can now categorise the dynamic network loading models based on which of these three continuous variables are discretised. Furthermore, we can distinguish models based on whether they serve as solution methods to the Lighthill-Whitham-Richards equations or whether they were formulated separately. Figure 3.1 lists the resulting six discretisation categories with example models, indicating for each of them whether or not it solves the network loading according to the Lighthill-Whitham-Richards theory. We have included several models that advertise themselves as mesoscopic as these are based on macroscopic traffic propagation rules. In the
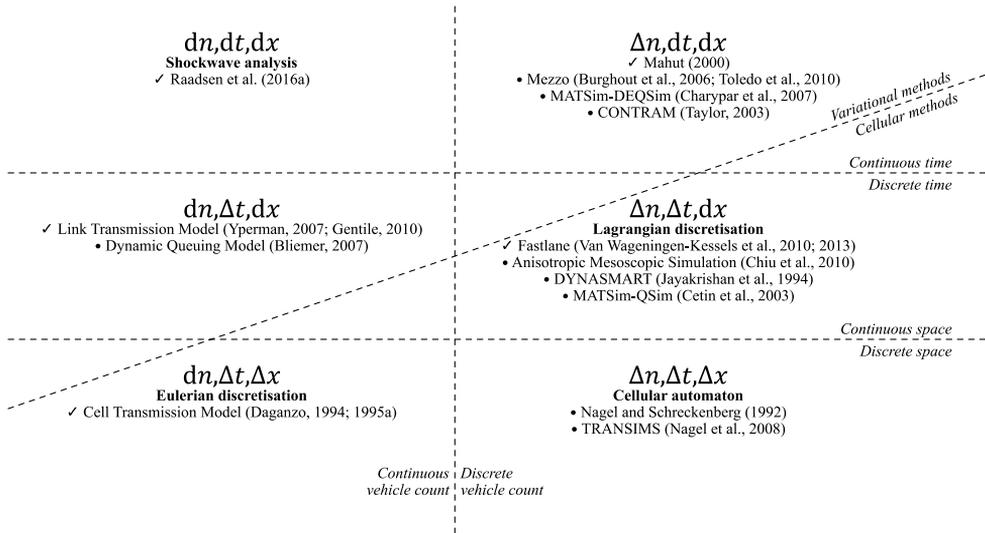
Figure 3.1. Categorisation of macroscopic dynamic network loading models by discretisations. Models indicated with a checkmark serve as solution methods for Lighthill-Whitham-Richards theory.

next section, we will see that each discretisation has important implications for our use of agents.

3.3   Specific challenges and possible solutions

We will now list specific challenges encountered when coupling the agent-based traffic demand and the macroscopic dynamic network loading, and identify possible solutions to these challenges. The challenges below reflect the most fundamental issues we experienced ourselves while constructing such a coupled model. They relate to modelling links and determining turning fractions at nodes. Both categories are discussed below.

3.3.1   Link modelling

The first challenge for link modelling is to *avoid systematic errors* causing deviations from the exact solution according to traffic flow theory. A large number of models in Figure 3.1 do not provide a solution method to Lighthill-Whitham-Richards theory, but are merely inspired by it. Hence, they will have systematic errors no matter how fine the discretisations are. Among the listed models with continuous n, this applies only the Dynamic Queuing Model, which splits the link into a running part and a congested part, the latter having a constant density that does not depend on flow (Bliemer, 2007). The same simplification is commonly encountered among the models with discrete n, e.g. Mezzo (Burghout et al., 2006), CONTRAM (Taylor, 2003) and MATSim-QSim (Cetin et al., 2003). Models like Mezzo, CONTRAM, Anisotropic Mesoscopic Simulation and DYNASMART furthermore compute the vehicle speed as a function of some average link density rather than the current local density (Burghout et al., 2006; Taylor, 2003; Chiu et al., 2010; Jayakrishan et al., 1994), with

CONTRAM even lacking any explicit vehicle interaction during queuing (Taylor, 2003). The accuracy of the traffic propagation in a cellular automaton is limited due to its stringent discretisation. Thus, while one may intuitively prefer models with discrete n for agent-based applications, these are often limited in their accuracy of traffic dynamics. In some cases, the simplifications causing the systematic errors do reduce the computational complexity of the model.

Even if one selects a model without such systematic errors, a follow-up challenge is to *limit numerical diffusion* that causes the numerical solution to deviate from the exact one. Although this is an issue with all applications of macroscopic dynamic network loading, it is of particular interest in agent-based applications as the vehicles of agents should diffuse as little as possible. For Eulerian-discretised models, i.e. the Cell Transmission Model, this is minimal if the ratio $\Delta x / \Delta t$ matches the free speed, while for Lagrangian-discretised models, i.e. Fastlane, it is minimal if the ratio $\Delta n / \Delta t$ matches the congested wave speed, outperforming the Eulerian discretisation (Van Wageningen-Kessels, 2013). The variational methods have even less numerical diffusion because the traffic dynamics within the interior of links are not discretised. The continuous event-based model of Raadsen et al. (2016a) is an extreme example that can eliminate numerical diffusion completely, albeit potentially at a high computational cost. Therefore using variational methods for link modelling that are free of systematic errors can be recommended to tackle these first two challenges.

For completeness, we remark that if one seeks to find an equilibrium in the choices of agents, numerical diffusion may also be somewhat helpful in achieving this as it softens the changes in travel times, which can be particularly sharp when using first-order Lighthill-Whitham-Richards traffic flow theory – hence one may not want to go at great lengths to completely eliminate numerical diffusion in such applications. On the other hand, it is theoretically better to calculate the traffic propagation without numerical diffusion and apply some kind of moving average filter when resulting travel times are used in agent choice models.

A next challenge is to *prevent agents from traversing links faster than the free speed* u_F, which is one consequence of numerical diffusion that is particularly troublesome, because we do not want agents to experience impossibly low travel times. For this assessment, we use the location of an agent as defined previously, hence we investigate to what extent numerical diffusion can speed up the front of the vehicle. Naturally, this problem does not apply to models with continuous t. An Eulerian-discretised model makes a speed error of up to $\Delta x / (\lfloor \Delta x / (u_F \Delta t) \rfloor \Delta t) - u_F$, which vanishes at $\Delta x / \Delta t = u_F$ when the numerical diffusion in general is minimal. In general, models with continuous x and discrete t have an identical formula for the error, with the link length substituted for Δx. However, for the Link Transmission Model, the overall numerical diffusion is smaller and it is computationally less costly to increase the accuracy by reducing Δt, since one does not need to simultaneously reduce Δx too. Appendix 3.A further mitigates the issue by proposing to modify the interpolation of the cumulative inflow curve such that the speed error becomes unbiased. Some of the models with discrete n keep track of the exact time within a time step a vehicle packet passes a node, making the error unbiased as well or even eliminating it, e.g. Fastlane (Van Wageningen-Kessels et al., 2013), Anisotropic Mesoscopic Simulation (Chiu et al., 2010) and DYNASMART (Jayakrishan et al., 1994).

A final, but not unimportant link modelling challenge is to *accept agents to flow into the link*. This may seem trivial, but it is actually problematic for models with discrete n, as they need to collect Δn vehicles before these can enter the link as a vehicle packet. Unless $\Delta n \leq 1$, an

individual agent is therefore unable enter a link on its own. Using such a small Δn is problematic for those models that discretise t in addition to n, as their computational complexity then becomes similar to that of a microscopic model. An exception is MATSim-QSim which avoids this by calculating the travel speed of a vehicle packet only when it enters a link (Cetin et al., 2003), loosening adherence to Lighthill-Whitham-Richards theory. While CONTRAM reduces its computational burden by varying Δn over origin-destination pairs depending on the size of their demand (Taylor, 2003), this approach is not suitable if en-route choices need to be possible (Jayakrishan et al., 1994).

3.3.2   Turning fractions

At nodes, turning fractions of an incoming link are determined based on the order in which vehicles entered the incoming link, using the either the intended next turn or a past finalised turn of each corresponding agent, depending on whether or not the agent location is still on the incoming link. An important challenge is to *specify this ordering of vehicles on a link*. One can choose between an order of agents and an order of vehicle parts that correspond to agents. Moreover, one can choose between a strict weak order and a strict total order, that is, it may or may not be possible for multiple elements of the order to have an equivalent priority for leaving the link. For example, microscopic models use a strict total order of agents per lane, whereas macroscopic models typically use a strict weak order of vehicle parts of which the size and commodity are known. All four possible combinations are illustrated in Figure 2.3.

In our application, it seems attractive to use an order of agents, since it prevents the sequence of vehicles from becoming fuzzy due to numerical diffusion. It is also relatively easy to implement, especially in case of a strict total order of agents for models with discrete n. However, a strict weak order of agents results in highly fluctuating discrete turning fractions, and a strict total order of either agents or vehicle parts even in binary turning fractions. This can cause congestion even if there is sufficient road capacity. For example, consider the situation of a motorway with an off-ramp. Highly fluctuating turning fractions may cause the demand for the off-ramp to temporarily exceed its capacity, even though there would be sufficient capacity on average. The fluctuations will decrease the average throughput of the node, and will have even larger consequences if the dynamic network loading includes a capacity drop. In principle, such fluctuations should hence be avoided, although in models with discrete t the issue may be obscured by a large Δt. The issue may also not appear if the capacity constraint is not applied to the total link outflow, as in CONTRAM where vehicle packets never block each other at all (Taylor, 2003), but also as in e.g. Mezzo where different outgoing turns from a link process queued vehicles independently within a specified "look-back limit" (Burghout et al., 2006). Either way, this makes the queuing process difficult to interpret physically.

Therefore, we should prefer a strict weak order of vehicle parts, as common in macroscopic network loading. This means that multiple vehicle parts can enter a link simultaneously, and that they also leave the link simultaneously. For models with discrete n, this implies that the composition of vehicle packets must be modified at nodes, rather than simply forwarding vehicle packets one-by-one. Unfortunately, this is unusual (Van Wageningen-Kessels et al., 2013). The effects of a strict total order may be mitigated by explicitly considering individual lanes like Mahut (2000) does, at the cost of having a more complex model.

Strict total order of agents:

Strict weak order of agents:

Strict total order of vehicle parts:

Strict weak order of vehicle parts:

Figure 3.2. Types of orderings of vehicles, showing traffic composition as function of n. In all examples, the cumulative link inflow is $N(x_0,t)$ and the cumulative link outflow is $N(x_L,t)$, implying that vehicles of agents A and B fully left the link, and vehicles of agents C-G are fully inside the link.

Another challenge is to *choose a node model of an appropriate form*, that is, the dynamic network loading component that computes the flows through nodes. One can choose between an incremental node model, that starts its computation with zero flow and gradually increases the flows, or a squeezing node model, that starts with the maximum outflows permitted by the incoming links and gradually decreases these flows. A good example is the comparison of Flötteröd and Rohde (2011) and Tampère et al. (2011), which, from the same assumptions on traffic behaviour, develop an incremental algorithm and a squeezing algorithm respectively. While these produce the same results if turning fractions are constant, it is possible to adapt the turning fractions in the middle of the computation of an incremental node model. Consequently, an incremental model can ensure first-in-first-out behaviour when the outflow of a link is congested, whereas a squeezing model cannot. This reduces numerical diffusion at nodes and hence an incremental model is to be preferred. Note that for efficiency, the order of vehicle parts on a link should be stored as an explicit queue in computer memory, so that the vehicle parts can be incrementally extracted from the front of the queue, unlike e.g. the formulation of Yperman (2007) that needs to iterate over all commodities in a link to disaggregate its outflow – which can be many in an agent-based application.

A final challenge related to turning fractions is to *prevent "agent-based gridlock"*. As stated before, parts of the vehicle of an agent may lag behind the location of the agent, and are

required to follow the agent. Hence, this part of the flow cannot exercise en-route choices, potentially leading to a gridlock-like situation where agents are stuck on a circle of network links. Despite that the agents may choose to leave such a circle in the en-route choice process, none of the agents are able to effectuate such a choice as they are not located exactly at the downstream end of their link. One possible solution, similar to what is proposed by e.g. Charypar et al. (2007), is to always accept some small percentage of the capacity as inflow into a link, even if that means exceeding the jam density. Although this method, which partially disables spillback, can also speed up convergence in cases where one seeks an equilibrium between choices of agents, it is of course not ideal for the accuracy of traffic propagation. Hence, we believe a better solution is to relax the first-in-first-out condition such that in cases where the outflow of a link would otherwise be jammed, vehicle parts belonging to agents that have already left the link can be overtaken. This allows agents to drive out of a circle of congested links while maintaining the first-in-first-out property on the level of agents. This seems to be a balanced solution to the problem, although it does increase the computation time because it precludes grouping multiple agents into a single commodity at departure as proposed in Subsection 2.2.3.

3.4   Conclusions

In this chapter, we investigated whether and how macroscopic dynamic network loading models may be coupled with an agent-based representation of traffic demand, whose agents are explicitly propagated throughout the network such that they can exercise en-route choices. We conclude that this is indeed possible, also with less computational complexity than microscopic dynamic network loading models, but there are several of pitfalls that a modeler must consider. Many existing traffic flow models that are commonly used in agent-based applications turn out to have systematic inconsistencies in the modelling of links and nodes if we compare them to macroscopic traffic flow theory. Perhaps surprisingly, models which discretise the number of vehicles do not seem ideal for the task at hand, due to limitations of current node models and the need for complete vehicle packets. In particular, we find that models that discretise both time and the number of vehicles are unable to outperform microscopic models in this context. Based on the above, we recommend using a dynamic network loading model with a continuous number of vehicles. Here, the modeler can reduce numerical diffusion by avoiding discretisations of space and time. Currently, discretisation of time will be necessary as Raadsen et al. (2016a) have not yet formulated a multi-commodity version of their continuous-time model, while their model is also restricted to triangular fundamental diagrams and can thus not handle complex traffic phenomena. The Link Transmission Model with small time steps hence seems a suitable recommendation for now. We further recommend to use a strict weak order of vehicle parts and an incremental node model to determine the turning fractions, with a relaxed first-in-first-out constraint to prevent unnatural gridlock-like situations.

Acknowledgement

3.A Enhanced interpolation for the Link Transmission Model

In this appendix we formulate an alternative interpolation procedure for the cumulative inflow curves of the Link Transmission Model that prevents the maximum speed violations reported in Subsection 3.3.1 from accumulating over links. Instead of linear interpolation within the time step, it uses linear extrapolation of both surrounding time steps if these extrapolated lines intersect within the surrounded time step. This reduces diffusion of the shape of the curve while translated to downstream nodes. The interpolation formula becomes:

$$t^*(t) = \frac{(\lceil t \rceil - 2\Delta t_{x_0})N(x_0,\lceil t \rceil - \Delta t_{x_0}) - (\lceil t \rceil - \Delta t_{x_0})N(x_0,\lceil t \rceil - 2\Delta t_{x_0}) + (\lceil t \rceil + \Delta t_{x_0})N(x_0,\lceil t \rceil) - \lceil t \rceil N(x_0,\lceil t \rceil + \Delta t_{x_0})}{N(x_0,\lceil t \rceil - \Delta t_{x_0}) - N(x_0,\lceil t \rceil - 2\Delta t_{x_0}) + N(x_0,\lceil t \rceil) - N(x_0,\lceil t \rceil + \Delta t_{x_0})}$$

$$N(x_0,t) = \begin{cases} \frac{\lceil t \rceil - \Delta t_{x_0} - t}{\Delta t_{x_0}} N(x_0,\lceil t \rceil - 2\Delta t_{x_0}) + \frac{t - \lceil t \rceil + 2\Delta t_{x_0}}{\Delta t_{x_0}} N(x_0,\lceil t \rceil - \Delta t_{x_0}) & \text{if } \lceil t \rceil - \Delta t_{x_0} < t \le t^*(t) \le \lceil t \rceil \\ \frac{\lceil t \rceil + \Delta t_{x_0} - t}{\Delta t_{x_0}} N(x_0,\lceil t \rceil) + \frac{t - \lceil t \rceil}{\Delta t_{x_0}} N(x_0,\lceil t \rceil + \Delta t_0) & \text{if } \lceil t \rceil - \Delta t_{x_0} \le t^*(t) \le t \le \lceil t \rceil \\ \frac{\lceil t \rceil - t}{\Delta t_{x_0}} N(x_0,\lceil t \rceil - \Delta t_{x_0}) + \frac{t - \lceil t \rceil + \Delta t_{x_0}}{\Delta t_{x_0}} N(x_0,\lceil t \rceil) & \text{if } t^*(t) \le \lceil t \rceil - \Delta t_{x_0} \vee \lceil t \rceil \le t^*(t) \end{cases}$$

(3.1)

Here $N(x_0,\cdot)$ denotes the cumulative inflow curve of a link, Δt_{x_0} denotes the time step of the node upstream of the link and $\lceil \cdot \rceil$ denotes ceiling to a multiple of Δt_{x_0}. Now, let L be the link length, u_F the maximum speed, Δt_{x_L} the time step of the downstream node and $N(x_L,\cdot)$ the cumulative outflow curve. Since the determination of $N(x_L,t+\Delta t_{x_L})$ requires the value of $N(x_0,t+\Delta t_{x_L} - L/u_F)$ as input, $N(x_0,\lceil t+\Delta t_{x_L} - L/u_F \rceil + \Delta t_{x_0})$ must already be known. In general, this is the case if $\Delta t_{x_0} + \Delta t_{x_L} \le L/u_F$. The enhanced interpolation thus poses more stringent requirements on the time steps. An easy method to determine satisfactory maximum time steps for all nodes in a network is requiring $\Delta t_{x_0} \le L/2u_F$ and $\Delta t_{x_L} \le L/2u_F$.

Chapter 4

Extending the Link Transmission Model with non-triangular fundamental diagrams and capacity drops

Abstract

The original Link Transmission Model as formulated by Yperman et al. (2006) simulates traffic according to Lighthill-Whitham-Richards theory with a very small numerical error, yet only supports triangular fundamental diagrams. This chapter relaxes that restriction in two steps. Firstly, we extend the model to handle any continuous concave fundamental diagram, and prove that this extension is still consistent with Lighthill-Whitham-Richards theory. Secondly, we extend the theory and model to handle a capacity drop, explicitly looking into the handling of both the onset and release of congestion. The final model is still first-order and suitable for general networks. Numerical examples show that it qualitatively improves on the original model due to uniquely featuring complex traffic patterns including stop-and-go waves, with crisp shockwaves between traffic states, as well as acceleration fans.

Keywords: Link Transmission Model, Lighthill-Whitham-Richards theory, first-order model, capacity drop, node model, stop-and-go wave.

4.1 Introduction

Lighthill-Whitham-Richards (LWR) theory or kinematic wave theory, introduced by Lighthill and Whitham (1955) and Richards (1956), consists of two main equations: the conservation of vehicles and the equilibrium flow-density relationship. Assuming that traffic is always in an equilibrium state, these combine into a single partial differential equation for the propagation of traffic along a network link. Traditionally, this partial differential equation has often been solved by the Cell Transmission Model (CTM) (Daganzo, 1994), that discretises roads into small cells according to the Godunov (1959) scheme. The Lagged Cell Transmission Model (LCTM) (Daganzo, 1999) and its later enhancement (Szeto, 2008) are variants of this method, reducing the numerical error.

Newell (1993) proposed a very different solution scheme, using cumulative numbers of vehicles as the primary variable. Later, this idea led to the development of the Link Transmission Model (LTM) (Yperman et al., 2006; Yperman, 2007), which does not discretise space and consequently leads to substantially smaller numerical errors (or computation time) than both the CTM and the LCTM. Daganzo (2005a; 2005b) and Jin (2015) show implicitly and Han et al. (2016) show explicitly that this numerical procedure indeed solves the partial differential equation as the time step tends to zero, but only for triangular fundamental diagrams (FDs).

However, the requirement of triangular FDs is restrictive in multiple ways. Edie (1961) already identified that the speed in subcritical traffic decreases as traffic density increases. It is important to capture this relationship for the modelling of travel times in light traffic, yet a triangular FD does not do so. A non-linear free-flow branch in the FD furthermore captures platoon dispersion (Geroliminis and Skabardonis, 2005). Edie also recognised that there may be a discontinuity between the free-flow capacity and the queue discharge rate. This is commonly referred to as the capacity drop, i.e. the effect that the presence of congestion reduces the maximum flow. Papageorgiou (Papageorgiou, 1998) mentioned this as an important aspect of traffic flow that models should be able to reproduce, particularly when considering traffic control. This is especially relevant when testing or optimising traffic control measures aimed at preventing or postponing the occurrence of the capacity drop, like many ramp metering installations, or at dissolving stop-and-go waves or wide moving jams, like SPECIALIST (Hegyi et al., 2008). While Hajiahmadi et al. (2013) propose an extension to the LTM for variable speed limits and ramp metering, the lack of a capacity drop in triangular FDs thus significantly restricts its usability, e.g. in assessing control strategies. Separate modelling of a free-flow capacity and a queue discharge rate can furthermore be expected to benefit strategic assessments of intelligent in-vehicle systems designed to intervene specifically in case of congestion, such as the Congestion Assistant (Van Driel and Van Arem, 2010).

Unlike the continuous-space LTM, these issues have mostly been addressed for the discrete-space models. General continuous FDs can be handled by the CTM and LCTM with proven convergence to LWR theory (Daganzo, 1995b; Daganzo, 1999; Szeto, 2008) and have been incorporated into CTM-based optimisation problems (Nie, 2011; Carey and Watling, 2012). Multiple different modifications of the CTM have been proposed to deal with a capacity drop (see Section 4.4). None of this has so far been the case for the LTM.

Hence, the purpose of this chapter is to overcome the aforementioned limitations of the shape of the FD in the LTM. More specifically, we extend the LTM to handle general concave FDs, optionally including capacity drops. The resulting model, which we show to converge to

LWR theory if there is no capacity drop, is applicable to general networks and features both standing queues, with a head initially fixed at the bottleneck that may move upstream later, and stop-and-go waves that can both grow and dissolve. Qualitative properties of this model are demonstrated with numerical examples.

This chapter is structured as follows. First, Section 4.2 will briefly introduce the structure of the LTM, consisting of a link model and a node model. Then, in Section 4.3, we derive a link model algorithm for the general case of a continuous concave FD, proving its consistency with LWR theory, and compare it to other link model formulations previously proposed in literature. Next, in Sections 4.4-4.6, we review previous work extending first-order models with capacity drops and subsequently extend LWR theory, the previous link model and a node model to allow for a capacity drop. Section 4.7 demonstrates the final model with two numerical examples. Finally, we list our conclusions in Section 4.8.

4.2 Structure of the LTM

The starting point of this chapter is the LTM for dynamic network loading, as formulated by Yperman et al. (2006) and Yperman (2007). Its primary components are a link model and a node model, that together are used to update in time steps of size Δt the cumulative number of vehicles $N(x,t)$ at the entrance $x_{i,0}$ and exit $x_{i,L}$ of each link i.

The link model is used to determine the sending and receiving flows, which for each time step indicate the number of vehicles that could potentially exit and enter the link respectively. For link i, these quantities are denoted $S_i(t)$ and $R_i(t)$ respectively. The procedures for determining them are similar and will be discussed in more detail below.

The node model then considers the interactions of traffic at intersections to derive transition flows $G_{ij}(t)$, the number of vehicles that succeed in crossing the intersection, where different node models can be used to represent different types of intersections. They indicate how much of each turn-specific sending flow $S_{ij}(t)$ will actually pass the node during the time step.

The node model also needs to know the turning fractions $S_{ij}(t)/S_i(t)$ as input, which can be either specified exogenously or modelled endogenously by splitting the traffic into multiple so-called commodities with different routing behaviour. Yperman et al. (2006) and Yperman (Yperman, 2007) assumed the latter option in their model formulation. Although we do not consider the specification of turning fractions in this chapter, the results of this chapter can be used in both these cases.

Algorithm 4.1 summarises the overall process, showing how the link model and the node model together specify the traffic flow propagation.

The LTM discretises only time, not space. Because due to the Courant-Friedrichs-Lewy (1928) condition, the maximum possible time step of a node depends on the length of the attached links, Yperman (2007, p. 25) suggested that different nodes may be operated with different time step sizes to retain a high computational efficiency without being restricted by the smallest link in a (large) network. In this chapter we explicitly incorporate this suggestion

Algorithm 4.1. Link Transmission Model.

➢ For each time step t for each node:
  ➢ Using the link model, determine sending flow $S_i(t)$ for each incoming link i.
  ➢ Using the link model, determine receiving flow $R_j(t)$ for each outgoing link j.
  ➢ Determine turning fractions $S_{ij}(t)/S_i(t)$ for each turn ij.
  ➢ Using a node model, determine transition flows $G_{ij}(t)$ for each turn ij.
  ➢ $N(x_{i,L}, t+\Delta t) := N(x_{i,L}, t) + \sum_j G_{ij}(t)$ for each incoming link i.
  ➢ $N(x_{j,0}, t+\Delta t) := N(x_{j,0}, t) + \sum_i G_{ij}(t)$ for each outgoing link j.

by writing Δt_{x_0} and Δt_{x_L} for the time step sizes of the upstream and downstream nodes of a link respectively and only using Δt without subscript if the node under consideration is obvious.

Note that two close variants of the explicit forward simulation scheme of Algorithm 4.1 have been proposed recently. Himpe et al. (2016) turned it into an implicit iterative scheme, permitting the use of larger time steps. At least Section 4.3 of this chapter regarding continuous concave FDs is also compatible with this variant. Alternatively, Hajiahmadi et al. (2016), Van de Weg et al. (2016) and Long et al. (2016) turned Algorithm 4.1 into optimisation problems, with points of the cumulative curves as decision variables and the link and node models as constraints. Because for non-triangular FDs we find constraints that only apply conditionally, this chapter is likely not compatible with that variant. Further discussion of these two variants is outside the scope of this chapter.

4.3 Link model for continuous concave FDs

In this section, we will define a link model for the case of a continuous concave FD $Q(k)$, i.e. without capacity drop. First, Subsection 4.3.1 introduces our notation and axioms. Next, we derive a method to compute sending and receiving flows in Subsection 4.3.2 resulting in the algorithms listed in Subsection 4.3.3. Finally, Subsection 4.3.4 compares this newly proposed model with other models in literature.

4.3.1 Notation and axioms

A general continuous concave FD $Q(k)$ is depicted in Figure 4.1. Here (k_c, q_c) denotes the capacity point and k_J denotes jam density. Note that we omit link indices on all variables for brevity.

As an alternative to a function $Q(k)$, the FD may be written as two functions $K(q)$ and $K'(q)$, describing the free-flow branch and the congested branch respectively. For each branch, we define a set of relevant wave speeds, indicating the speeds at which traffic states propagate, as

$$
\begin{aligned}
Z &= \left[\inf \operatorname{im} \frac{dq}{dK}, \sup \operatorname{im} \frac{dq}{dK} \right] \\
Z' &= \left[\inf \operatorname{im} \frac{dq}{dK'}, \sup \operatorname{im} \frac{dq}{dK'} \right]
\end{aligned}.
$$

(4.1)
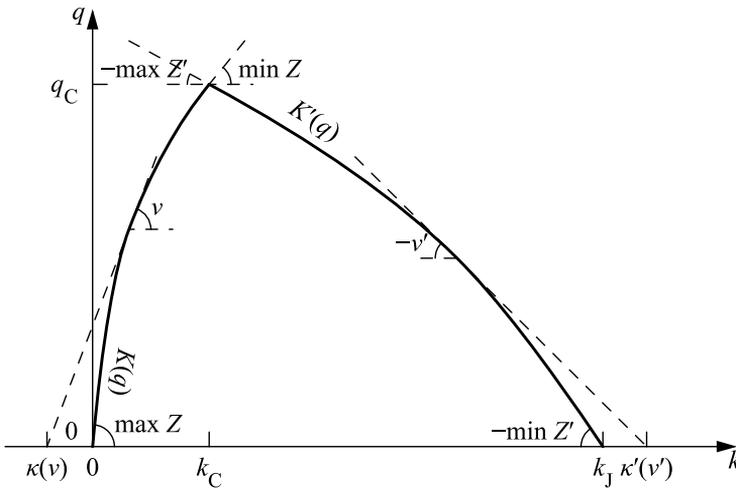
Figure 4.1. Example fundamental diagram satisfying Equation (4.2).

The prime notation indicates the congested branch rather than a derivative and im is used to denote the image of a function. We impose the following restriction on the shape of the FD:

$$\{\min Z, \max Z, \min Z', \max Z'\} \cap \{-\infty, 0, \infty\} = \varnothing. \tag{4.2}$$

Note that contrary to this restriction, from a traffic flow theory point of view, $\min Z = 0$ and $\max Z' = 0$ could be permitted, which would allow the FD to be horizontal at the capacity point. However, we will later see that the full restriction allows to compute sending and receiving flows efficiently and exactly on the link level. Note that we do not require the branches of the FD to be continuously differentiable, e.g. piecewise-linear FDs can be used if desired.

Let us now define function $\kappa(v)$ that indicates where a tangent line of slope v to the free-flow branch would intersect the density axis. Likewise, we define function $\kappa'(v')$ for the congested branch. The corresponding formulas, domains and images are

$$
\begin{aligned}
\kappa(v) &= \min_{q \in \text{dom}\,K}\left(K(q) - \frac{q}{v}\right), & \text{dom}\,\kappa &= (0, \max Z], & \text{im}\,\kappa &= (-\infty, 0] \\
\kappa'(v') &= \max_{q \in \text{dom}\,K'}\left(K'(q) - \frac{q}{v'}\right), & \text{dom}\,\kappa' &= [\min Z', 0), & \text{im}\,\kappa' &= [k_J, \infty)
\end{aligned}
\tag{4.3}
$$

Conversely, by appropriately taking the upper and lower envelopes of these tangent lines, we get back the FD:

$$
\begin{aligned}
&\forall q \in \operatorname{dom} K: \quad K(q) = \max_{v \in Z}\left(\frac{q}{v} + \kappa(v)\right) \\
&\forall q \in \operatorname{dom} K': \quad K'(q) = \min_{v' \in Z'}\left(\frac{q}{v'} + \kappa'(v')\right) \\
&\forall k \in [0, k_J]: \quad Q(k) = \min\left(\min_{v \in Z} v(k - \kappa(v)), \min_{v' \in Z'} v'(k - \kappa'(v'))\right)
\end{aligned}
\tag{4.4}
$$

We furthermore define $V(q)$ and $V'(q)$ as the sets of positive and negative tangent line slopes compatible with the FD at flow q:

$$
\begin{aligned}
V(q) &= \arg\max_{v \in Z}\left(\frac{q}{v} + \kappa(v)\right) = \left\{v \in Z \,\middle|\, \frac{q}{v} + \kappa(v) = K(q)\right\} \\
V'(q) &= \arg\min_{v' \in Z'}\left(\frac{q}{v'} + \kappa'(v')\right) = \left\{v' \in Z' \,\middle|\, \frac{q}{v'} + \kappa'(v') = K'(q)\right\}
\end{aligned}
\tag{4.5}
$$

The theoretical basis for traffic propagation along the link is formed by LWR theory. Traditionally, this takes the form of the following scalar conservation law, which combines conservation of vehicles with the FD:

$$
\frac{\partial k}{\partial t} + \frac{dQ(k)}{dk}\frac{\partial k}{\partial x} = 0 .
\tag{4.6}
$$

However, it is more convenient to replace this differential equation with the following Hamilton-Jacobi equation, that also works if $Q(k)$ is not continuously differentiable:

$$
\frac{\partial N}{\partial t} - Q\left(-\frac{\partial N}{\partial x}\right) = 0, \quad q = \frac{\partial N}{\partial t}, \quad k = -\frac{\partial N}{\partial x} .
\tag{4.7}
$$

This differential equation states the FD in a way that implicitly guarantees conservation of vehicles (Newell, 1993):

$$
\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = -\frac{\partial^2 N}{\partial x \partial t} + \frac{\partial^2 N}{\partial t \partial x} = 0 .
\tag{4.8}
$$

The differential equation is combined with the Lax (1957) shock admissibility or entropy condition to get a unique weak solution, eliminating alternative solutions where acceleration fans are replaced with shocks. This is achieved by prohibiting shocks from emanating waves, thus allowing discontinuities in flow or density only when they absorb waves or run parallel to them[*]. Then, because Hamiltonian $-Q(-\partial N / \partial x)$ is convex in $\partial N / \partial x$, the Hamilton-Jacobi equation, and thus the propagation of traffic along a link, can be solved with relative ease using variational theory (Evans, 2002; Daganzo, 2005a). We use this in the next subsection to determine the sending and receiving flows, assuming that the link is initially empty.

---

[*] If a discontinuity runs parallel to waves on both sides, it is a contact discontinuity. If it absorbs waves from one or both sides, it is a shockwave.

4.3.2   Computing sending and receiving flows: solution networks

The sending and receiving flows will be solved in terms of cumulative numbers of vehicles. More precisely, our algorithm relies on finding the maximum possible $N(x, t + \Delta t_x)$ at the considered end of the link $x \in \{x_0, x_L\}$ at the end of the time step under consideration, so that $N(x, t + \Delta t_x) - N(x, t)$ is the maximum number of vehicles exiting or entering the link during the time step, which are the sending flow $S(t)$ and the receiving flow $R(t)$ respectively.

We thus rephrased the traffic propagation problem into finding the maximum possible value of $N(x, t + \Delta t_x)$ for $x \in \{x_0, x_L\}$. To do so, we apply the variational theory developed by Daganzo (2005a; 2005b). The boundary condition for this application is formed by the values of the cumulative curves in previous time steps at both link ends. We build a solution network, as defined by Daganzo (2005a), that indicates how each boundary point may constrain the cumulative number of vehicles N at our solution point $P = (x_P, t + \Delta t_{x_P})$. This applies to the determination of both the sending and receiving flow, hence we derive both solution networks simultaneously in the ensuing.

Figure 4.2 illustrates the concept of using solution networks. After each time step, the boundary condition is extended with the newly found solution and the solution network is shifted to compute the next time step. Since we assume the link is initially empty, we can disregard the initial condition by moving it sufficiently far to the left while extending the boundary conditions into the past with $N = 0$.
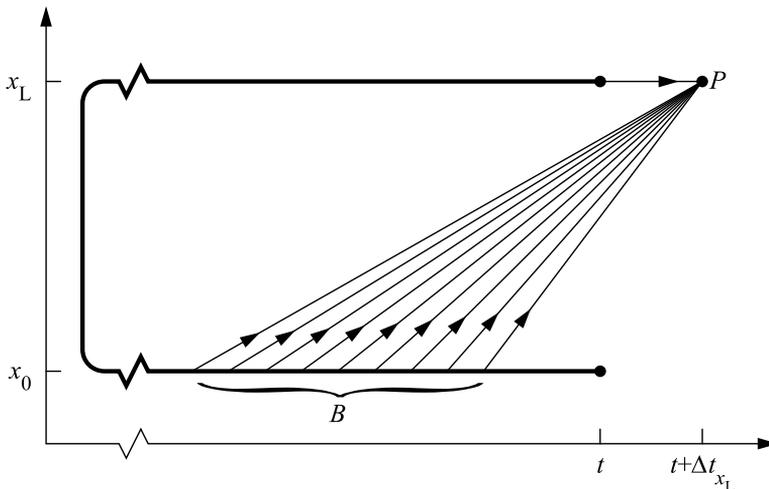
Figure 4.2. Concept of a solution network for the sending flow, showing its boundary condition and space-time paths.

Considering what space-time paths to include in our solution network, Fact 3 of Daganzo (2005b) states that we only need to consider straight space-time paths from the boundary to the solution point, reducing the solution network construction problem to selecting which boundary points to include. For the part of the boundary on the same link end as the solution point, the only possibly relevant path originates from the beginning of the time step, namely $(x_P, t) \rightarrow (x_P, t_P) = (x_P, t + \Delta t_{x_P})$. Its constraint is $N(x_P, t_P) \leq N(x_P, t) + q_C \Delta t_{x_P}$.

Next, let $L = x_L - x_0$ denote the link length, and let us consider the part of the boundary at the opposite link end. First of all, space-time paths originating from any point B for which

$$
t_B > t_P - \begin{cases} \dfrac{L}{\max Z} & \text{if } x_P = x_L \\[2ex] \dfrac{-L}{\min Z'} & \text{if } x_P = x_0 \end{cases} \tag{4.9}
$$

should not be included because, by Daganzo (2005a), these paths are not valid as the necessary slope exceeds the range of possible wave speeds $[\min Z', \max Z]$. For the other boundary points, the path to the solution point has a slope of $(x_P - x_B)/(t_P - t_B)$, and, using Daganzo (2005a), its constraint imposed on $N(x_P, t_P)$ thus equals

$$
\begin{aligned}
N(x_P, t_P) &\leq N(x_B, t_B) + \int_{t_B}^{t_P} \max_{k \in [0, k_J]} \left(Q(k) - k \frac{x_P - x_B}{t_P - t_B} \right) dt \\[2ex]
&= N(x_B, t_B) + (t_P - t_B) \max_{k \in [0, k_J]} \left(Q(k) - k \frac{x_P - x_B}{t_P - t_B} \right) \\[2ex]
&= N(x_B, t_B) + \begin{cases} (x_P - x_B) \max_{k \in [0, k_C]} \left(Q(k) \frac{t_P - t_B}{x_P - x_B} - k \right) & \text{if } x_P > x_B \\[2ex] (x_P - x_B) \min_{k \in [k_C, k_J]} \left(Q(k) \frac{t_P - t_B}{x_P - x_B} - k \right) & \text{if } x_P < x_B \end{cases} \\[2ex]
&= N(x_B, t_B) + \begin{cases} (x_P - x_B) \max_{q \in \operatorname{dom} K} \left(q \frac{t_P - t_B}{x_P - x_B} - K(q) \right) & \text{if } x_P > x_B \\[2ex] (x_P - x_B) \min_{q \in \operatorname{dom} K'} \left(q \frac{t_P - t_B}{x_P - x_B} - K'(q) \right) & \text{if } x_P < x_B \end{cases} \\[2ex]
&= N(x_B, t_B) - \begin{cases} (x_P - x_B) \kappa \left(\frac{x_P - x_B}{t_P - t_B} \right) & \text{if } x_P > x_B \\[2ex] (x_P - x_B) \kappa' \left(\frac{x_P - x_B}{t_P - t_B} \right) & \text{if } x_P < x_B \end{cases} \\[2ex]
&= \begin{cases} N(x_B, t_B) - \kappa \left(\frac{L}{t + \Delta t_{x_L} - t_B} \right) L & \text{if } x_P = x_L \\[2ex] N(x_B, t_B) + \kappa' \left(\frac{-L}{t + \Delta t_{x_0} - t_B} \right) L & \text{if } x_P = x_0 \end{cases} \quad . \tag{4.10}
\end{aligned}
$$

Due to Theorem 4.1 below, we only need to include paths from the opposite boundary with wave speeds within Z (sending flow) or Z' (receiving flow). This is illustrated in Figure 4.3.
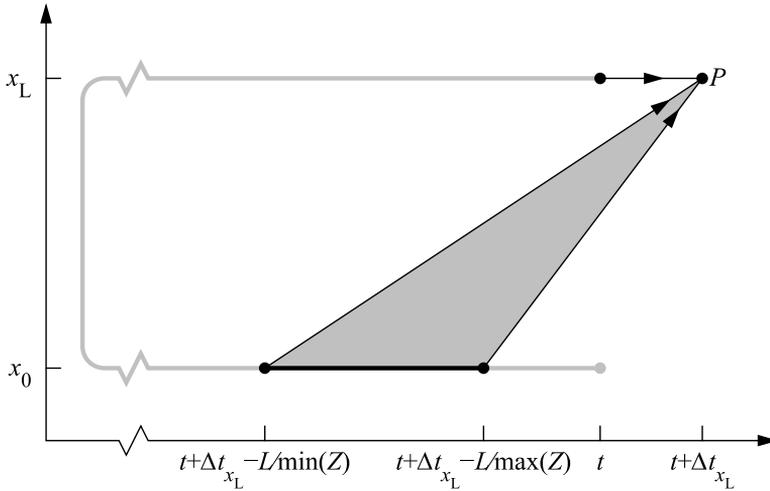
Figure 4.3. Solution network for the sending flow, highlighting the relevant part of the boundary condition and the space-time paths corresponding to these origins.

Theorem 4.1. Paths originating from points (x_B, t_B) **on the opposite part of the boundary with** $t_B < t_P - L/\min Z$ **(sending flow) or** $t_B < t_P + L/\max Z'$ **(receiving flow) can be excluded from the solution network.**

Proof. Consider all points (x_B, t_B) with $t_B \leq t_P - L/\min Z$ (sending flow) or $t_B \leq t_P + L/\max Z'$ (receiving flow), imposing the constraints

$$N(x_P, t_P) \leq \begin{cases} N(x_B, t_B) - \kappa\left(\dfrac{L}{t_P - t_B}\right)L & \text{if } x_P = x_L \\ N(x_B, t_B) + \kappa'\left(\dfrac{-L}{t_P - t_B}\right)L & \text{if } x_P = x_0 \end{cases}$$

$$= \begin{cases} N(x_B, t_B) - \left(k_C - q_C\dfrac{t_P - t_B}{L}\right)L & \text{if } x_P = x_L \\ N(x_B, t_B) + \left(k_C - q_C\dfrac{t_P - t_B}{-L}\right)L & \text{if } x_P = x_0 \end{cases} \tag{4.11}$$

$$= \begin{cases} N(x_B, t_B) - k_C L + q_C(t_P - t_B) & \text{if } x_P = x_L \\ N(x_B, t_B) + k_C L + q_C(t_P - t_B) & \text{if } x_P = x_0 \end{cases}.$$

Because $\partial N(x_B, t)/\partial t \leq q_C$, the point with $t_B = t_P - L/\min Z$ (sending flow) or $t_B = t_P + L/\max Z'$ (receiving flow) is always the most constraining one. ∎
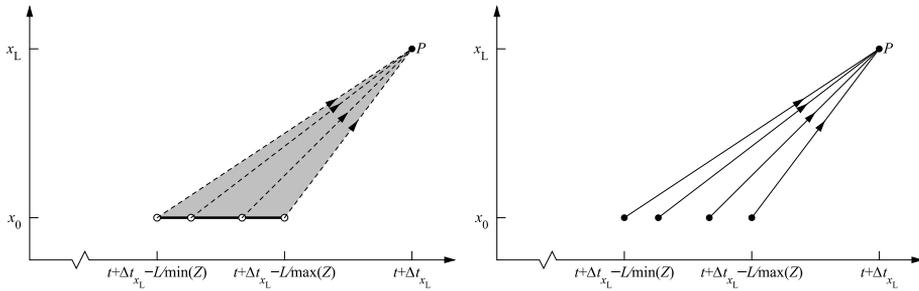
Figure 4.4. Opposite boundary of solution network split into open intervals and points.

Finally, we observe that, due to the time discretisation of the LTM, N is piecewise-linear along the boundary. Therefore we can split the relevant part of the boundary at the opposite link end into a finite set of open intervals with defined, constant flows $\partial N / \partial t$ and a finite set of points. This is illustrated in Figure 4.4.

For the open intervals, instead of investigating all valid paths originating from it, we need only to look at those valid paths with a wave speed compatible with the known boundary flow. More specifically, this means that the wave speed corresponding to the slope from the boundary point to the solution point can be used to construct a tangent line to the FD passing through the traffic state corresponding to the known flow at the boundary. The reason for this is that in order to be constraining, a valid path must be a wave emanated from the boundary (Daganzo, 2005a), and in order for that wave to be emanated, its traffic state must match the traffic state at the boundary. Theorem 4.2 below shows that this results in only one extra closed-form constraint per open interval, that, together with the above formulas and considerations, allow us to compute $N(x_P, t_P)$ exactly if the boundary conditions are exact.

Theorem 4.2. If, along the opposite boundary x_B**, an open interval** (t_1, t_2) **with flow** $q = \left(N(x_B, t_2) - N(x_B, t_1)\right) / (t_2 - t_1)$ **constrains** $N(x_P, t_P)$ **unlike its infimum** t_1 **and its supremum** t_2 **, then**

$$\begin{cases} t_1 + \frac{L}{\min V(q)} < t_P < t_2 + \frac{L}{\max V(q)} & \text{if } x_P = x_L \\ t_1 - \frac{L}{\max V'(q)} < t_P < t_2 - \frac{L}{\min V'(q)} & \text{if } x_P = x_0 \end{cases} \tag{4.12}$$

and

$$N(x_P, t_P) = \begin{cases} N(x_0, t_1) + q \cdot (t_P - t_1) - K(q)L & \text{if } x_P = x_L \\ N(x_L, t_1) + q \cdot (t_P - t_1) + K'(q)L & \text{if } x_P = x_0 \end{cases}. \tag{4.13}$$

Proof. The constraint imposed by the interval equals

$$N\left(x_P,t_P\right) \leq \begin{cases} \displaystyle\inf_{t_B \in (t_1,t_2)}\left(N\left(x_0,t_1\right)+q\cdot\left(t_B-t_1\right)-\kappa\left(\dfrac{L}{t_P-t_B}\right)L\right) & \text{if } x_P = x_L \\[2ex] \displaystyle\inf_{t_B \in (t_1,t_2)}\left(N\left(x_L,t_1\right)+q\cdot\left(t_B-t_1\right)+\kappa'\left(\dfrac{-L}{t_P-t_B}\right)L\right) & \text{if } x_P = x_0 \end{cases}$$

$$\underset{\substack{\text{if constraining} \\ \leq \text{ otherwise}}}{=}\begin{cases} \displaystyle\inf_{\substack{t_B \in (t_1,t_2) \\ \left|\frac{L}{t_P-t_B}\right| \in V(q)}}\left(N\left(x_0,t_1\right)+q\cdot\left(t_B-t_1\right)-\left(K(q)-q\dfrac{t_P-t_B}{L}\right)L\right) & \text{if } x_P = x_L \\[3ex] \displaystyle\inf_{\substack{t_B \in (t_1,t_2) \\ \left|\frac{-L}{t_P-t_B}\right| \in V'(q)}}\left(N\left(x_L,t_1\right)+q\cdot\left(t_B-t_1\right)+\left(K'(q)-q\dfrac{t_P-t_B}{-L}\right)L\right) & \text{if } x_P = x_0 \end{cases} \tag{4.14}$$

$$=\begin{cases} \displaystyle\inf_{\substack{t_B \in (t_1,t_2) \\ \left|\frac{L}{t_P-t_B}\right| \in V(q)}}\left(N\left(x_0,t_1\right)+q\cdot\left(t_P-t_1\right)-K(q)L\right) & \text{if } x_P = x_L \\[3ex] \displaystyle\inf_{\substack{t_B \in (t_1,t_2) \\ \left|\frac{-L}{t_P-t_B}\right| \in V'(q)}}\left(N\left(x_L,t_1\right)+q\cdot\left(t_P-t_1\right)+K'(q)L\right) & \text{if } x_P = x_0 \end{cases}$$

Since the argument of the infimum is independent of t_B, this reduces to the constraint

$$N\left(x_P,t_P\right) \leq \begin{cases} \begin{cases} N\left(x_0,t_1\right)+q\cdot\left(t_P-t_1\right)-K(q)L & \text{if } \exists t_B \in (t_1,t_2): \dfrac{L}{t_P-t_B}\in V(q) \\ \infty & \text{otherwise} \end{cases} & \text{if } x_P = x_L \\[4ex] \begin{cases} N\left(x_L,t_1\right)+q\cdot\left(t_P-t_1\right)+K'(q)L & \text{if } \exists t_B \in (t_1,t_2): \dfrac{-L}{t_P-t_B}\in V'(q) \\ \infty & \text{otherwise} \end{cases} & \text{if } x_P = x_0 \end{cases} \tag{4.15}$$

Instead of looking whether there exists a point $\left(x_B,t_B\right)$ in the interval that can reach $\left(x_P,t_P\right)$, we can look whether $\left(x_P,t_P\right)$ lies within the total area reached by the interval. This turns the constraint into

$$N\left(x_P,t_P\right) \leq \begin{cases} \begin{cases} N\left(x_0,t_1\right)+q\cdot\left(t_P-t_1\right)-K(q)L & \text{if } t_1+\frac{L}{\max V(q)} < t_P < t_2+\frac{L}{\min V(q)} \\ \infty & \text{otherwise} \end{cases} & \text{if } x_P = x_L \\[4ex] \begin{cases} N\left(x_L,t_1\right)+q\cdot\left(t_P-t_1\right)+K'(q)L & \text{if } t_1-\frac{L}{\min V'(q)} < t_P < t_2-\frac{L}{\max V'(q)} \\ \infty & \text{otherwise} \end{cases} & \text{if } x_P = x_0 \end{cases} \tag{4.16}$$

Finally, if $L/\max V(q) < t_P-t_1 \leq L/\min V(q)$ (sending flow) or $-L/\min V'(q) < t_P-t_1 \leq -L/\max V'(q)$ (receiving flow), then this constraint of the open interval (t_1,t_2) is equal to the constraint of its infimum t_1. Likewise, if $L/\max V(q) \leq t_P-t_2 < L/\min V(q)$ (sending flow) or $-L/\min V'(q) \leq t_P-t_2 < -L/\max V'(q)$ (receiving flow), the constraint is equal to that of its supremum t_2. This completes the proof. ∎

4.3.3   Computing sending and receiving flows: algorithms

Algorithms 4.2 and 4.3 implement the described solution networks for the sending and receiving flow respectively. For brevity of notation, we use a compound assignment operator $a \lesssim b$ meaning $a := \min(a,b)$ and a floor-to-multiple-of operator $\lfloor a \rfloor_b$ meaning $\lfloor a/b \rfloor \cdot b$.

Algorithm 4.2. Sending flow.

➢ $N\left(x_L, t + \Delta t_{x_L}\right) := N\left(x_L, t\right) + q_C \Delta t_{x_L}$.   Apply outflow capacity constraint.

➢ $t_1 := t + \Delta t_{x_L} - L / \min Z$.

➢ $N\left(x_L, t + \Delta t_{x_L}\right) \lesssim N\left(x_0, t_1\right) - \kappa\left(\min Z\right) L$ Apply constraint of point t_1.

➢ $t_2 := \lfloor t_1 \rfloor_{\Delta t_{x_0}}$.

➢ Loop:

    ➢ $t_2 := \min\left(t_2 + \Delta t_{x_0}, t + \Delta t_{x_L} - L / \max Z\right)$.   Find next t_2.

    ➢ If $t_1 = t_2$:

        ➢ Exit the loop.

    ➢ $q := \frac{N(x_0, t_2) - N(x_0, t_1)}{t_2 - t_1}$.

    ➢ If $t_1 + \frac{L}{\min V(q)} < t + \Delta t_{x_L} < t_2 + \frac{L}{\max V(q)}$:

        ➢ $N\left(x_L, t + \Delta t_{x_L}\right) \lesssim N\left(x_0, t_1\right) + q \cdot \left(t + \Delta t_{x_L} - t_1\right) - K(q) L$.   Apply constraint for interval (t_1, t_2).

    ➢ $N\left(x_L, t + \Delta t_{x_L}\right) \lesssim N\left(x_0, t_2\right) - \kappa\left(\frac{L}{t + \Delta t_{x_L} - t_2}\right) L$.   Apply constraint of point t_2.

    ➢ $t_1 := t_2$.   Move t_1 to t_2.

➢ $S(t) := N\left(x_L, t + \Delta t_{x_L}\right) - N\left(x_L, t\right)$.   Set sending flow.

Algorithm 4.3. Receiving flow.

➢ $N\left(x_0, t + \Delta t_{x_0}\right) := N\left(x_0, t\right) + q_C \Delta t_{x_0}$.   Apply inflow capacity constraint.

➢ $t_1 := t + \Delta t_{x_0} + L / \max Z'$.

➢ $N\left(x_0, t + \Delta t_{x_0}\right) \lesssim N\left(x_L, t_1\right) + \kappa\left(\max Z\right) L$.   Apply constraint of point t_1.

➢ $t_2 := \lfloor t_1 \rfloor_{\Delta t_{x_L}}$.

➢ Loop:

    ➢ $t_2 := \min\left(t_2 + \Delta t_{x_L}, t + \Delta t_{x_0} + L / \min Z\right)$.   Find next t_2.

    ➢ If $t_1 = t_2$:

        ➢ Exit the loop.

    ➢ $q := \frac{N(x_L, t_2) - N(x_L, t_1)}{t_2 - t_1}$.

    ➢ If $t_1 - \frac{L}{\max V'(q)} < t + \Delta t_{x_0} < t_2 - \frac{L}{\min V'(q)}$:

        ➢ $N\left(x_0, t + \Delta t_{x_0}\right) \lesssim N\left(x_L, t_1\right) + q \cdot \left(t + \Delta t_{x_0} - t_1\right) + K'(q) L$.   Apply constraint of interval (t_1, t_2).

    ➢ $N\left(x_0, t + \Delta t_{x_0}\right) \lesssim N\left(x_L, t_2\right) + \kappa\left(\frac{L}{t + \Delta t_{x_L} - t_2}\right) L$.   Apply constraint of point t_2.

    ➢ $t_1 := t_2$.   Move t_1 to t_2.

➢ $R(t) := N\left(x_0, t + \Delta t_{x_0}\right) - N\left(x_0, t\right)$.   Set receiving flow.

For Algorithm 4.1, the Courant-Friedrichs-Lewy (1928) condition requires that the time step sizes of nodes are chosen such that for each link,

$$\Delta t_{x_0} \leq \frac{-L}{\min Z'} \wedge \Delta t_{x_L} \leq \frac{L}{\max Z} \tag{4.17}$$

Note that these algorithms thus give us exact results according to LWR theory, insofar the input boundary condition is represented exactly as series of time steps with constant flows each.

4.3.4   Comparison with literature

Now, let us compare this result with discrete-time algorithms previously reported in literature:

- In case of a triangular FD, our algorithms reduce to the algorithms given by Yperman et al. (2006) and Yperman (2007).

- In case of a piecewise-linear concave FD that is not triangular, our algorithms include more constraints on the sending and receiving flows than those proposed by Yperman (2007): we include more paths originating from the boundary at the opposite link end. This difference becomes visible in the model output in case of acceleration fans or rarefaction waves, which the Yperman (2007) formulation cannot correctly reproduce. Figure 4.5 demonstrates this graphically, where a sudden increase in link inflow results in an outflow spike rather than a monotonically increasing outflow. Bliemer et al. (2016) show a similar example.

- In case of a continuously differentiable concave FD, we can compare our algorithms with the LTM formulation of Gentile (2010). Actually, our algorithms require a discontinuity in the derivative at the capacity point. While this might appear to be a difference with the Gentile formulation, we see upon closer inspection of his algorithms that he constrains the travel times of waves on a link by limiting his look-ahead window. Our restriction on the shape of the FD makes this implicit constraint explicit. The models are hence similar, e.g. Gentile (2010, pp. 161-162) proofs that the constraints resulting from our Theorem 4.2 represent valid paths (without using that terminology). A difference however is that he uses linear interpolation of cumulative curve in acceleration fans producing constant flows at the opposite link end during the fan. Because each branch of the FD has more than two slopes, this is not correct, as noted by Gentile (2010). Instead, our algorithms interpolate only within time steps rather than over the entire fan, yielding a more accurate representation of acceleration fans with piecewise-constant flows. A final difference is that our algorithms do not require all nodes to have the same time step size.

- For any concave FD, Mazaré et al. (2011) have previously studied the boundary value problem of LWR theory with N piecewise-linear along the boundary. The main difference with our model is that we do not know the traffic states at the link ends a-priori but compute these endogenously for a network, according to Algorithm 4.1. Another difference is that we added the previously mentioned restriction to the shape of the FD near the capacity point, enabling us, via Theorem 4.1, to prevent the computation time from increasing more than linearly with the time horizon of the simulation. Note that our algorithms, derived from variational theory (2005a) with a
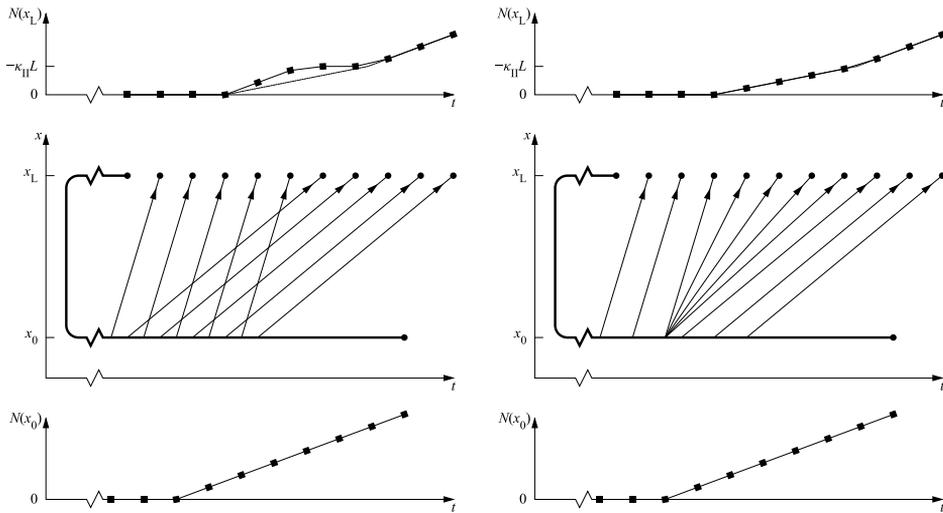
Figure 4.5. Reproduction of acceleration fans by the Yperman (2007) model formulation (left) and our model formulation (right), when both are used with the same piecewise-linear FD. The middle images indicate the most restrictive paths of each solution network. The analytical solution is indicated with a dashed line in the top images.

> solution network with an infinite number of direct paths from the boundary to the solution point, qualify as Lax-Hopf algorithms rather than dynamic programming in the terminology of Mazaré et al. (2011). Hence the proposed formulas are both grid-free within the link. The result of our Theorem 4.2 matches with their Equations (24) and (27).

We conclude that according to LWR theory, our proposed algorithms compute the sending and receiving flows exactly on the link level, proven using variational theory. Compared to our formulation, all variants of the discrete-time LTM previously proposed in literature have either additional restrictions on the shape of the FD, or errors or limitations resulting in inexact solutions, or both.

On the network level, this means that when using our algorithms, the only source of error is the simplification that node flows be constant within time steps, i.e. the time discretisation of the boundary conditions that are input to the algorithms above. Since the exactness of the algorithms is naturally subject to the exactness of their input, the network-level solution is generally not exact but converges to the exact solution as the time steps tend to zero, creating a trade-off between numerical accuracy and computational efficiency. For a single invocation of either algorithm, the computational complexity is $O(\Delta t^{-1})$. With respect to Δt, this yields a complexity of $O(\Delta t^{-2})$ for the dynamic network loading process as a whole: the same as Gentile (2010). These complexities reduce to $O(1)$ and $O(\Delta t^{-1})$ respectively in case the FDs are triangular, which are the same as Yperman et al. (2006).

Finally, we remark that in the specific case of a triangular FD, the discrete-time algorithms can be replaced with event-based algorithms as proposed by Raadsen et al. (2016a) which are able to produce exact solutions on a network level. However, the current extension of that approach to general concave FDs (Raadsen et al., 2014) incorrectly replaces acceleration fans with shocks, yielding an inadmissible weak solution to the differential equations. Thus, while our discrete-time approach cannot provide the exact network-level solution, it does have the advantage of being able to approximate it by using small time steps. Note that in the future such an approximation may become possible for the event-based approach as well if discretisation of FDs into piecewise-linear ones (Bliemer et al., 2016) turns out to be feasible at high resolutions.

4.4   Capacity drop theory for first-order models

We proceed to extend our model with support for capacity drops. Before we do so, let us review previous work extending first-order models with a capacity drop. As observed by Chung et al. (2007), the capacity drop can be related to the traffic density. Many previously proposed models indeed modify the capacity based on the current density, e.g. in cell transmission models by having the demand function decrease past the critical density (Monamy et al., 2012; Muralidharan et al., 2012; Roncoli et al., 2015). Alvarez-Icaza and Islas (2013) proposed to select the capacity based on the sign of the time derivative of density. In case the capacity can only attain two values, i.e. a free-flow capacity and a queue discharge rate, Srivastava and Geroliminis (2013) proposed to use two threshold densities, where the capacity retains its previous value between the threshold densities and only switches when both thresholds are exceeded. Torné et al. (2014) instead look at whether downstream traffic conditions restricted the flow for a particular location in the previous time step, and set the capacity to the queue discharge rate if so. Jin et al. (2015) specify a kinematic wave theory with capacity drop that constrains the flow over a one-to-one node to the queue discharge rate if congestion is unavoidable.

The previously mentioned models turn out to be difficult to formulate correctly for inhomogeneous roads and general networks. The reason is that the queue discharge rate is defined as the outflow of an active bottleneck, whereas a standing queue is located in front of this bottleneck, on a road with a potentially different free-flow capacity or queue discharge rate itself. For example, if the queue discharge rate is larger upstream of a bottleneck than downstream, the flow through the bottleneck will be too large if the model is not specified carefully. One could work around this problem by inserting a special transition cell at discontinuities, that could have the free-flow capacity of the upstream cells and the queue discharge rates from the downstream cells. Monamy et al. (2012) indeed propose inserting a special cell to handle the capacity drop at merges. Alternatively, Torné et al. (2014) have cells modify the capacity of neighbors, in addition to using special cells for merges. Such approaches, even if done correctly, are clearly not very useful for extending the LTM, as it does neither discretise links nor nodes into small cells. Jin et al. (2015) are not affected by this issue, but their model formulation can only apply a capacity drop at a pre-specified one-to-one node and neither elsewhere within links nor at more general nodes.

There are however more problems with the driving behaviour implied by existing models. For the models that use the average density within each cell to select the cell capacity, the cell size will now influence how the capacity of the road changes over time. This is difficult to understand in terms of driving behaviour, which now depends on the discretisation of the road, and it is impossible to use these approaches in the LTM due to the lack of such a

discretisation. Finally, all of the models imply some kind of memory effect in the capacity of a road segment, which may or may not exist in reality. In particular, many models effectively yield a trapezoidal FD when the capacity drop is active, including the extended kinematic wave theory proposed by Jin et al. (2015). As a consequence, traffic at a congested road segment cannot recover to a free-flow state unless the demand for entering that road segment drops and the queue dissolves from its tail. The head of a queue is thus unable to move upstream, while this is observed in reality for so-called wide moving jams or stop-and-go waves. To solve this, one must ensure that the queue discharge state is a point on the free-flow branch. Fortunately, an inverted-lambda style FD (Koshi et al., 1983) can be employed to achieve this, as demonstrated by Hegyi et al. (2008) and Schreiter et al. (2010). However, as pointed out by Torné et al. (2014) and Schreiter et al. (2010), a severe difficulty lies in the possibility of infinite-speed backward shockwaves that can then occur during the onset of congestion. Instead of investigating an inverted-lambda style FD, Lu et al. (2009) mathematically analysed the solutions of unmodified LWR theory with a simpler jump discontinuity in the FD, which did not result in a capacity drop but did result in the infinite-shockwave-speed problem.

Because of the previous considerations, mere modification of the node model as proposed by Jin et al. (2015) will not be sufficient to describe the capacity drop: the link model must be modified as well, with an inverted-lambda style FD. By extending LWR theory this way, we can correctly account for the capacity drop both when a queue is standing in front of a bottleneck node and when a queue is moving upstream within a link. We will do so in this section.

Firstly, Subsection 4.4.1 formulates the dynamics for an infinite link in free-flow and in congestion, and Subsection 4.4.2 defines how the interface between these areas behaves. Next, Subsection 4.4.3 extends this theory to handle finite links. Subsection 4.4.4 provides an example of the link dynamics. Next, nodes are discussed in Subsection 4.4.5. Finally, Subsection 4.4.6 compares the newly developed traffic flow theory with some of the papers cited above. The results from this section will be added to our model in Sections 4.5 and 4.6.

4.4.1   Link dynamics in free-flow and in congestion

An inverted-lambda style FD is shown in Figure 4.6. Clearly, $Q(k)$ ceases to exist as a function. Consequently, we cannot directly use the differential equations from Section 4.3 to describe traffic flow on a link. However, for space-time areas in free-flow and space-time areas in congestion, we still have the following scalar conservation laws respectively:

$$
\begin{aligned}
\frac{\partial q}{\partial x} + \frac{dK(q)}{dq}\frac{\partial q}{\partial t} &= 0, \quad q \in [0, q_C], \quad K(q) \in [0, k_C] \\
\frac{\partial q'}{\partial x} + \frac{dK'(q')}{dq'}\frac{\partial q'}{\partial t} &= 0, \quad q' \in [0, q_D], \quad K'(q') \in [k_D, k_J]
\end{aligned}
. \tag{4.18}
$$

Note the time and space axes have swapped roles compared to Equation (4.6). The equivalent Hamilton-Jacobi equations are

$$
\begin{aligned}
\frac{\partial N}{\partial x} + K\left(\frac{\partial N}{\partial t}\right) &= 0, \quad q = \frac{\partial N}{\partial t} \in [0, q_C], \quad k = -\frac{\partial N}{\partial x} \in [0, k_C] \\
\frac{\partial N}{\partial (-x)} - K'\left(\frac{\partial N}{\partial t}\right) &= 0, \quad q' = \frac{\partial N}{\partial t} \in [0, q_D], \quad k' = \frac{\partial N}{\partial (-x)} \in [k_D, k_J]
\end{aligned}
. \tag{4.19}
$$
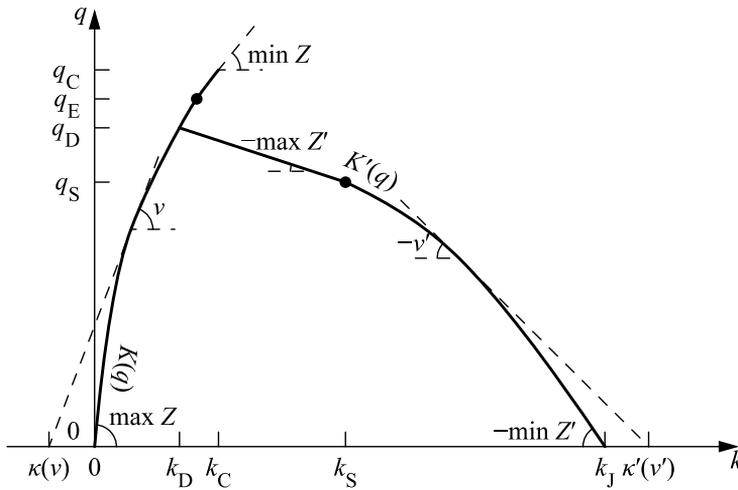
Figure 4.6. Example fundamental diagram including a capacity drop. (q_E will be introduced in Subsection 4.4.5.)

Because Hamiltonians $K(\partial N / \partial t)$ and $-K'(\partial N / \partial t)$ are convex in $\partial N / \partial t$, the weak solutions within a free-flow space-time area and within a congested space-time area can be found the same way as before. A definition of the boundary between free-flow and congestion, which provides a boundary condition for both Hamilton-Jacobi equations, now completes our traffic flow theory with capacity drop for the link.

4.4.2 The interface between free-flow and congestion

To describe the separation between free-flow and congestion, we can use the concept of a separating shock $x_s(t)$, which is a generalised characteristic that divides space-time in a free-flow area upstream and a congested area downstream (Han et al., 2016). Let $N_s(t) = N(x_s(t), t)$ denote the cumulative number of vehicles at the separating shock, and let us describe an infinite link with initial conditions satisfying

$$N(x, 0) = N_s(0) + \begin{cases} \int_x^{x_s(0)} K(f(x)) dx & \text{for } x < x_s(0) \\ -\int_{x_s(0)}^x K'(f(x)) dx & \text{for } x > x_s(0) \end{cases} \tag{4.20}$$

for some $x_s(0)$, $N_s(0)$ and initial flows $f(x)$. This means we have the queue discharge traffic state (k_D, q_D) between congested traffic states and downstream of the last congested state. To ensure that moving observers on both sides of the shock pass vehicles at the same rate, the Rankine-Hugoniot condition (Evans, 2002) requires

$$\frac{dN_S}{dt} = q - k \frac{dx_S}{dt} = q' - k' \frac{dx_S}{dt}$$

$$\Rightarrow q - q' = (k - k') \frac{dx_S}{dt} \tag{4.21}$$

where (k, q) and (k', q') are the traffic states upstream and downstream of the shock respectively. Let us now state the following differential equations for the separating shock:

$$\lim_{h \downarrow 0} \frac{x_S(t+h) - x_S(t)}{h} = \begin{cases} \dfrac{q - q'}{k - k'} & \text{if } k' > k_D \\ \max Z & \text{if } k' = k_D \end{cases}.$$

$$\lim_{h \downarrow 0} \frac{N_S(t+h) - N_S(t)}{h} = q' - k' \lim_{h \downarrow 0} \frac{x_S(t+h) - x_S(t)}{h} \tag{4.22}$$

In the first case, the separating shock is actually a shock and its speed follows from the Rankine-Hugoniot condition. In the second case, the traffic state (k_D, q_D) behind the separating shock lies on the intersection of the free-flow branch and congested branch of the FD, and a speed of $\max Z$ is chosen so that the separating shock is too fast to be an actual shock. Traffic state (k_D, q_D) can then be emanated from the boundary into the free-flow space-time area. Any interaction between traffic state (k_D, q_D) and other free-flow states will thus be handled within the free-flow space-time area by its Hamilton-Jacobi equation, rather than by the separating shock itself.

In the above differential equations, the following definitions ensure that the traffic states upstream and downstream of a space-time point are defined at the shock location:

$$k(x,t) = \lim_{h \downarrow 0} \frac{N(x-h,t) - N(x,t)}{h}, \quad q(x,t) = K^{-1}(k(x,t))$$

$$k'(x,t) = \lim_{h \downarrow 0} \frac{N(x,t) - N(x+h,t)}{h}, \quad q'(x,t) = K'^{-1}(k'(x,t)) \tag{4.23}$$

With the semi-derivatives chosen as above, this forms a causal description of the separating shock and hence of the complete link. While this works fine for FDs with $(k_D, q_D) = (k_C, q_C)$, this leads to two problems if the upstream flow q is larger than the queue discharge rate q_D. The first problem is the possibility that the separating shock violates the Lax (1957) shock admissibility condition, as it may travel upstream too fast for kinematic waves corresponding to downstream state (k', q') to keep up. Hence there exists no solution. The second problem is the infinite-speed shockwave problem described earlier if the downstream flow is also relatively high.

We solve these problems by selecting some fixed point (k_S, q_S) on the congested branch of the FD with $k_C < k_S \leq k_J$ and modifying the shock admissibility criterion for backward shocks whenever an upstream state (k_1, q_1) and a downstream state (k_2, q_2) make contact in space-time, to allow and require shocks emanating traffic state (k_S, q_S) from the shock into the downstream space-time area, provided that

1. this would not violate the Rankine-Hugoniot condition (conservation of vehicles), and

2. there is no solution without a backward shock (flow maximisation).

To ensure that this always defines a solution everywhere, it is necessary that the congested branch of the FD is linear between (k_D, q_D) and (k_S, q_S), as in Figure 4.6. We then have the following implications for our previous coupling of the two Hamilton-Jacobi systems:

- within the free-flow space-time area, no backward shocks are needed and hence there is no change;

- within the congested space-time area, backward shocks are only needed when $k_1 < k_2$, in which case shocks emanating (k_S, q_S) are not possible without violating the Rankine-Hugoniot condition, so there is no change either;

- at the separating shock, backward shocks are only needed when $q_2 < \min(q_D, q_1)$, and a shock emanating (k_S, q_S) will occur if, subject to the Rankine-Hugoniot condition, the shock between (k_1, q_1) and (k_S, q_S) would diverge from the shock or contact discontinuity between (k_S, q_S) and (k_2, q_2), i.e. if $(q_1 - q_S)/(k_1 - k_S) < (q_2 - q_S)/(k_2 - k_S)$ – otherwise there is no change.

Since the modification of the shock admissibility criterion can only affect the separating shock, let us modify the definition of (k', q') in Equation (4.23) as follows, making it dependent on (k, q):

$$k'(x,t) = \begin{cases} k_S & \text{if } x = x_S(t) \wedge \lim_{h \downarrow 0} \frac{N(x,t)-N(x+h,t)}{h} > k_D \wedge \frac{q(x,t)-q_S}{k(x,t)-k_S} < \frac{K'^{-1}\left(\lim_{h \downarrow 0} \frac{N(x,t)-N(x+h,t)}{h}\right) - q_S}{\lim_{h \downarrow 0} \frac{N(x,t)-N(x+h,t)}{h} - k_S} \\ \lim_{h \downarrow 0} \frac{N(x,t)-N(x+h,t)}{h} & \text{otherwise} \end{cases}$$

$$q'(x,t) = K'^{-1}\big(k'(x,t)\big)$$

(4.24)

Here, we eliminated the redundant $q_2 < q_1$ check using $(q_2 - q_S)/(k_2 - k_S) \le (q_D - q_S)/(k_D - k_S)$ and substituted $(k_1, q_1) = (k(x,t), q(x,t))$ and $(k_2, q_2) = \left(\lim_{h \downarrow 0} \frac{N(x,t)-N(x+h,t)}{h}, K'^{-1}\left(\lim_{h \downarrow 0} \frac{N(x,t)-N(x+h,t)}{h}\right)\right)$ into the previous condition for emanating (k_S, q_S). When the outcome differs from the previous definition, this changes the propagation of the separating shock per Equation (4.22), forming a boundary condition for the congested Hamilton-Jacobi system that emanates (k_S, q_S) from there on where needed in accordance with the above. Note that $k(x,t) > k_D$ is necessary but not sufficient for the second inequality inside Equation (4.24) to be true.

One may wonder how to choose the values of (k_S, q_S). This traffic state manifests itself when high-density free-flow traffic breaks down due to downstream congestion. Hence, (k_S, q_S) is the traffic state that occurs in a newly created stop-and-go wave in our model, which can be helpful in choosing appropriate values. A larger k_S means stop-and-go waves become shorter and denser. If one desires to derive (k_S, q_S) empirically, one should consider that our model has infinite deceleration at the back and infinite acceleration at the head of a stop-and-go wave, so one would need to fit reality into this schematic representation of the wave.

Additionally, one will wish to take into account that the choice of these values also influences the greatest possible backward speed of information $w < 0$ according to

$$w = \min\left(\frac{q_C - q_S}{k_C - k_S}, \frac{q_C}{k_C - k_J}, \min Z' \right). \tag{4.25}$$

4.4.3   Imposing boundary conditions

The previous theory for an infinite link can be extended to a finite link with prescribed boundary values $N(x_0, t)$ and $N(x_L, t)$ by constraining x_S to $[x_0, x_L]$ and adding the following definitions of q and k at x_0 and q' and k' at x_L:

$$q(x_0, t) = \lim_{h \downarrow 0} \frac{N(x_0, t+h) - N(x_0, t)}{h}$$

$$k(x_0, t) = K(q(x_0, t))$$

$$q'(x_L, t) = \begin{cases} q_S & \text{if } x_L = x_S(t) \wedge \lim_{h \downarrow 0} \frac{N(x_L, t+h) - N(x_L, t)}{h} < q(x_L, t) \wedge \frac{q(x_L, t) - q_S}{k(x_L, t) - k_S} < \frac{\lim_{h \downarrow 0} \frac{N(x_L, t+h) - N(x_L, t)}{h} - q_S}{K'\left(\lim_{h \downarrow 0} \frac{N(x_L, t+h) - N(x_L, t)}{h}\right) - k_S} \\ \lim_{h \downarrow 0} \frac{N(x_L, t+h) - N(x_L, t)}{h} & \text{otherwise} \end{cases} \tag{4.26}$$

$$k'(x_L, t) = \begin{cases} K'(q'(x_L, t)) & \text{if } q'(x_L, t) \le q_D \\ k_D & \text{if } q'(x_L, t) \ge q_D \end{cases}$$

Similar to before, the dynamics of the separating shock are modified. The main difference is that the necessity of backward shocks is now given by $\lim_{h \downarrow 0} \frac{N(x_L, t+h) - N(x_L, t)}{h} < q(x_L, t)$. Furthermore, note that (k', q') may violate the congested branch of the FD when $x_S = x_L$. This does not matter since at such moments in time, the entire link is and remains in free-flow and the Hamilton-Jacobi equation for the congested space-time area is not needed.

4.4.4   Example

Figure 4.7 illustrates our extension of LWR theory with an example initial-boundary value problem whose solution can be constructed by hand. It has multiple possible choices for $x_S(0)$ that satisfy Equation (4.20), but they all result in the same solution for $N(x,t)$. The example has three time periods during which the flow upstream of $x_S(t)$ exceeds q_D, i.e. is equal to q_C. The first two times, (k_S, q_S) is emanated from the separating shock according to Equation (4.24). The third time, Equation (4.22) initially ensures that the $x_S(t)$ moves downstream. It is constrained to x_L once it reaches the downstream link end, until the link outflow reduces to q_D. The latter causes $x_S(t)$ to start moving upstream, emanating (k_S, q_S), according to Equation (4.26).
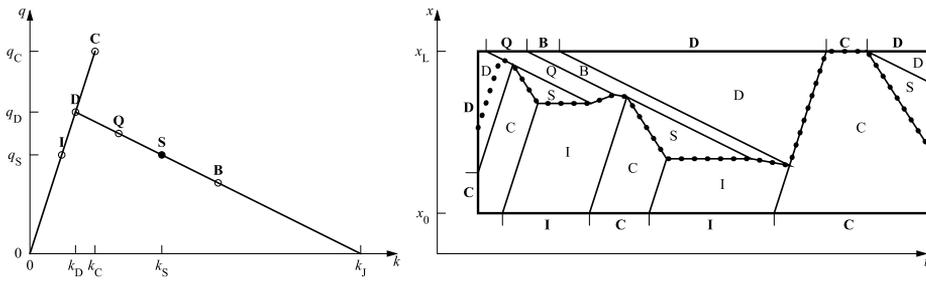
Figure 4.7. Example initial-boundary value problem for a link with capacity drop. Letters indicate traffic states in the FD (left) and solution (right). In the right image, the solution of $x_s(t)$ is dotted for one possible $x_s(0)$.

4.4.5   Node model requirements

Various first-order node models exist in literature. Our extension of LWR theory with capacity drop does not prescribe a specific one to be used. Nevertheless, the fact that each link may be subject to a capacity drop, adds more constraints to the solution of the node model, so that any chosen node model requires some modifications. These constraints, considering capacity drops for both incoming and outgoing links, are discussed in this subsection.

Firstly, each incoming link cannot transmit a flow between its queue discharge rate and its sending flow, since no appropriate congested traffic state exists to represent such a solution in the inverted-lambda FD of the incoming link. Thus, the transition flow over the node either facilitates the sending flow or is restricted by the queue discharge rate. We will refer to this as the "discharge rate for accelerating" q_D, as it refers to the discharge rate for queues standing or moving on the incoming link, as formulated earlier in this section.

Secondly, it would be illogical if an overloaded outgoing link, i.e. an outgoing link with insufficient receiving flow to accommodate all incoming flows directed at it, would accept more flow than some queue discharge rate, even if permitted by its receiving flow: the overloading implies congestion at its entrance and congestion implies a restriction to a queue discharge rate, which must be enforced as well. We will refer to this as the "discharge rate for merging" q_E, as it refers to the discharge rate for the entirety of queues standing before the outgoing link with traffic trying to enter it. Similar reasoning may be applied to internal node capacity constraints if one wishes to include those.

Note that these discharge rates q_D and q_E may differ for the same link. This allows to account for different mechanisms behind, and hence different magnitudes of, the capacity drop in both situations. To avoid inconsistencies at redundant one-to-one nodes, we must have $q_D \leq q_E \leq q_C$ for each link. This is plausible, as there is generally more maneuvering space available at the entrance of a link compared to the interior of the link. Yuan et al. (2017b) provide empirical confirmation of $q_D < q_E$ for a motorway link downstream of a lane drop.

Finally, following our considerations in the introduction of this section, we opt to not include memory effects on nodes, thus avoiding that capacity drops in previous time steps affect the maximum node flows in the current time step. Nevertheless, the physical queuing process on an incoming link will always result in the sending flow not exceeding the queue discharge rate for some time if a breakdown occurred in an earlier time step. For consistency of the simulation, ceteris paribus, it is then desirable that the flow in the first time step of the breakdown is the same as in later time steps, also if the queue head is not moving upstream. This desire essentially extends the demand invariance principle of Lebacque and Khoshyaran (2005). Node models with capacity drop should adhere to this principle.

4.4.6   Comparison with literature

The capacity drop theory proposed in this section has significant advantages over the approach of Lu et al. (2009), who solve the entropy solution for unmodified LWR theory with a discontinuous flow-density relationship. Both feature similar traffic breakdowns; our emanation of traffic state (k_s, q_s) from the separating shock is directly comparable to the deceleration fans in their solution, as derived in Figure 2 in their paper – it is the same traffic state appearing under the same condition. A difference here is that Lu et al. (2009) take the limit $k_s \downarrow k_c$, which creates the infinite-shockwave-speed problem as witnessed by their numerical results, whereas we intentionally chose $k_s > k_c$ to avoid this. Indeed, the breakdown process with $k_s > k_c$ in our model can also be reproduced with unmodified LWR theory with a continuous non-concave FD, as illustrated by e.g. Daganzo (1999, pp. 86-88) who labels (k_s, q_s) a "coasting state".

The most important difference however lies in the discharge process. While our model produces the queue discharge state downstream of congestion, Lu et al. (2009) get back the free-flow capacity state there via acceleration shocks, as derived in Figures 4 and 5 in their paper. Thus, despite their discontinuous FD, their theory still inherits the absence of a capacity drop from unmodified LWR theory, i.e. the maximum flow out of a queue is not less than the maximum flow in free-flow.

We remark that the capacity drop theory proposed by Jin et al. (2015), where queue heads do not move upstream and shocks do not emanate waves, can be interpreted as an extreme case of the more general theory formulated here, with $\max Z' \uparrow 0, q_s \uparrow q_D$ so that, unlike reality, queue heads cannot move upstream and with

$$Q(k) = \begin{cases} K^{-1}(k) & \text{if } k \leq k_C \\ \dfrac{k_s - k}{k_s - k_C} q_C + \dfrac{k - k_C}{k_s - k_C} q_S & \text{if } k_C \leq k \leq k_S \\ K'^{-1}(k) & \text{if } k_S \leq k \end{cases} \qquad (4.27)$$

being a concave function, so that traffic state (k_s, q_s) cannot be emanated from a shock.

The node model with capacity drop proposed by Jin et al. (2015) is a special case of the model proposed here, with one incoming link, one outgoing link and a "dropped node capacity" equal to the minimum of the accelerating discharge rate of the incoming link and the merging discharge rate of the outgoing link.

Finally, our capacity drop theory can be seen as a simplified version of the recent proposal by Yuan et al. (2017a) who vary the queue discharge rate not only depending on whether the queue is a standing or a moving one, but also on the congestion density, thereby deviating from a fixed inverted-lambda FD. Our simplification is necessary for the link-level solution method that will be developed in the next section. Nevertheless, with a suitable choice of fixed inverted-lambda FDs, our theory exactly reproduces the lane-drop and on-ramp examples of Yuan et al. (2017a, p. 481).

4.5 Link model with capacity drop

In this section, we adapt the solution networks for the receiving and sending flow to implement the link dynamics with capacity drop proposed in the previous section. Rather than computing sending and receiving flows by solving the previous system of differential equations directly, which is cumbersome, we temporarily simplify the problem by not specifying how congestion dissolves. This results in algorithms that are almost as simple as the algorithms presented before in Subsection 4.3.3. In Subsection 4.5.1, we start with investigating the receiving flow; the sending flow is discussed in the Subsection 4.5.2. These two subsections result in algorithms listed in Subsection 4.5.3. Finally, consequences and mitigation of the initial simplification will be discussed in Subsection 4.5.4.

4.5.1 Computing receiving flows: solution network

The receiving flow $R(t) = N(x_0, t + \Delta t_{x_0}) - N(x_0, t)$ will be estimated with a solution network that mimics the above traffic flow theory for an inverted-lambda style FD. Like before, this solution network should consist of two types of paths: paths along the upstream link end enforcing that the link inflow does not exceed the free-flow capacity and paths originating from the downstream link end propagating congested traffic states upstream, but only to the extent such congestion exists at the downstream link end.

To restrict the inflow to the free-flow capacity, one path from (x_0, t) to $(x_0, t + \Delta t_{x_0})$ with cost $q_C \Delta t_{x_0}$, is sufficient. Next, we need to define paths from the downstream link end to $(x_0, t + \Delta t_{x_0})$. We create two types of such paths. Let T denote the first time the downstream link end is congested, which will follow from the output of the node model, to the extent congestion has not dissolved yet within the link. If such a T does not exist, no backward paths need to be considered at all. Otherwise, we firstly create paths corresponding to all wave speeds $v' \in Z'$ for which $t + \Delta t_{x_0} \geq T - L / v'$, representing all traffic states potentially emanated from the opposite link end. We secondly create additional paths from all $(x_L, \hat{t}) : \hat{t} < T \wedge L / (\hat{t} - t - \Delta t_{x_0}) \in V'(q_S)$ as if

$$\forall \hat{t} < T : N(x_L, \hat{t}) = N(x_L, T) + q_S \cdot (\hat{t} - T). \tag{4.28}$$

This second set of backward paths serves to emanate congested traffic state (k_S, q_S) from the separating shock if needed, i.e. we simplify the wave emanation from the shock as if these traffic states originated from the downstream link end as well. These paths are collectively enforced by the extra constraint

$$N(x_0, t + \Delta t_{x_0}) \leq N(x_L, T) + q_S \cdot (T - t - \Delta t_{x_0}) + k_S L \text{ if } t + \Delta t_{x_0} < T - L \max Z'. \tag{4.29}$$

Now, Theorem 4.3 below shows that this solution network is consistent with the Rankine-Hugoniot condition for shocks, allowing us to use the maximum backward information speed w from Equation (4.25) to truncate the boundary condition at $t + \Delta t_{x_0} + L/w$ yielding the corresponding time step size constraint $\Delta t_{x_0} \leq -L/w$.

Theorem 4.3. The above solution network solves $N(x_0, t + \Delta t_{x_0})$ **in a way consistent with shocks satisfying the Rankine-Hugoniot condition.**

Proof. This statement is only relevant when T exists, otherwise it is theoretically impossible for $(x_0, t + \Delta t_{x_0})$ to be congested, and the solution network contains no backward paths either. If T exists, there will always exist at least one backward path reaching $(x_0, t + \Delta t_{x_0})$. The solution network also contains a stationary path along x_0 representing the free-flow regime. Thus, the transition of the upstream link end from free-flow to congestion is implicitly handled by the minimum envelope of both regimes, which is consistent with a Rankine-Hugoniot separating shock. Since shocks within the congested regime are also handled by a minimum envelope, this extends to all shocks.   ∎

Corollaries 4.1 and 4.2 below give us the additional desirable properties that shocks do not emanate (k_s, q_s) if their inflow does not exceed the queue discharge rate and that receiving flows are non-negative.

Corollary 4.1. The above solution network does not simulate wave emanation from a shock if the flow into the shock does not exceed the discharge rate q_D.

Proof. Since Theorem 4.3 ensures the solution network contains no other shocks than Rankine-Hugoniot shocks, a necessary condition for a shock to start emanating (k_s, q_s) is that its inflow exceeds the queue discharge rate.   ∎

Corollary 4.2. The above solution network ensures non-negative receiving flows.

Proof. This statement is trivial if T does not exist. If T exists and is the same as in the previous time step, Theorem 4.3 proves the statement, since $\partial N / \partial t \geq 0$ in each regime and the Rankine-Hugoniot condition implies continuity of N at the separating shock. The time step size constraint $\Delta t_{x_0} \leq -L/w$ ensures that T is known before its existence matters. Corollary 4.1 ensures that the simulation of wave emanation from the separating shock cannot be constraining immediately after a previous queue has dissolved, so an increase of T cannot cause a negative jump in the permitted N at x_0 either.   ∎

Finally, Theorem 4.4 below proves that if there is no capacity drop, this solution network for the receiving flow produces the same result as the solution network derived in Section 4.3. Hence the solution network proposed here handles both cases with and without a capacity drop.

Theorem 4.4. The above solution network is equivalent to Algorithm 4.3 if $(k_D, q_D) = (k_C, q_C)$ **regardless of** (k_S, q_S) **.**

Proof. Firstly, the Theorem of Daganzo (2005a) permits excluding backward paths corresponding to wave speeds $v' \in [\min Z', 0)$ for which $t + \Delta t_{x_0} < T - L/v'$, without changing the solution. Theorem 4.1 is not affected by this. Secondly, Corollary 4.1 ensures that the extra constraint of Equation (4.29) is void, removing the dependency on (k_S, q_S). This completes the equivalence. ∎

4.5.2 Computing sending flows: solution network

Our next task is to specify a solution network to estimate the sending flow $S(t) = N(x_L, t + \Delta t_{x_L}) - N(x_L, t)$ mimicking the traffic flow theory for an inverted-lambda style FD. Paths originating from the upstream link end constrain $N(x_L, t + \Delta t_{x_L})$ as in Section 4.3. However, the constraint enforcing the outflow capacity in case of congestion becomes more complex with a capacity drop: here, the queue discharge rate should be substituted for the capacity, but this constraint should not be active when there is no congestion on the link so that the outflow can attain the free-flow capacity as well.

Let us therefore investigate when the outflow should be constrained to the queue discharge rate. The constraint must be applied after the last time θ the outflow was congested with a flow less than q_D. This constraint will last up till some time $\Theta > \theta$, which will, upon creating congestion in the node model, initially be set to ∞, consistent with our simplification that we ignore the dissolution of congestion. Later, in Subsection 4.5.4 below, we will replace this with a finite value. The constraint can be implemented as

$$N(x_L, t + \Delta t_{x_L}) \leq \begin{cases} N(x_L, \theta) + q_D \cdot (t + \Delta t_{x_L} - \theta) & \text{if } t + \Delta t_{x_L} < \Theta \\ \infty & \text{otherwise} \end{cases} \qquad (4.30)$$

which completes the description of the sending flow algorithm.

Theorem 4.5, Corollary 4.3 and Theorem 4.6 below show that this approach is consistent with Rankine-Hugoniot shocks, produces a sending flow not smaller than zero nor larger than the free-flow capacity and gives the same result as the solution network in Section 4.3 if there is no capacity drop.

Theorem 4.5. The solution network described above solves $N(x_L, t + \Delta t_{x_L})$ **in a way consistent with shocks satisfying the Rankine-Hugoniot condition.**

Proof. Shocks within the free-flow regime and between the free-flow regime and the queue discharge state are handled by a minimum envelope of all valid paths that could be constraining at $(x_L, t + \Delta t_{x_L})$. ∎

Corollary 4.3. The solution network described above ensures non-negative sending flows not exceeding the free-flow capacity.

Proof. Each traffic state emitted from the upstream link end, as well as the outflow capacity constraint, has $0 \leq \partial N / \partial t \leq q_C$. Theorem 4.5 ensures N is continuous at shocks as well. ∎

Theorem 4.6. The solution network described above is equivalent to Algorithm 4.2 if $(k_D, q_D) = (k_C, q_C)$**.**

Proof. The equivalence is straightforward for $t + \Delta t_{x_L} < \Theta$, when T exists. For other times, Algorithm 4.2 has an outflow capacity constraint which the algorithm described above lacks. However, Corollary 4.3 shows that this absence of the outflow capacity constraint cannot affect the solution. ∎

4.5.3   Computing sending and receiving flows: algorithms

Algorithms 4.4 and 4.5 implement the link model with support for the capacity drop as previously proposed. As shown above, for $(k_D, q_D) = (k_C, q_C)$ they yield solutions identical to the algorithms of the Section 4.3 without capacity drop. Note that the extended versions presented here use several new variables whose values persist over time steps: Θ, θ and T are persistent real numbers with initial values $\Theta := 0$ and $T := \infty$. Their meanings have been defined in Subsection 4.5.2, with the addition that we set T to ∞ as a placeholder when it would not exist otherwise. These variables are used by the node model as well, as we see later in Section 4.6.

Algorithm 4.4. Sending flow with optional capacity drop.

➢ $t_1 := t + \Delta t_{x_L} - L / \min Z$.

➢ $N(x_L, t + \Delta t_{x_L}) := N(x_0, t_1) - \kappa(\min Z) L$.   Apply constraint of point t_1.

➢ $t_2 := \lfloor t_1 \rfloor_{\Delta t_{x_0}}$.

➢   Loop:

     ➢ $t_2 := \min(t_2 + \Delta t_{x_0}, t + \Delta t_{x_L} - L / \max Z)$.   Find next t_2.

     ➢   If $t_1 = t_2$:

         ➢   Exit the loop.

     ➢ $q := \frac{N(x_0, t_2) - N(x_0, t_1)}{t_2 - t_1}$.

     ➢   If $t_1 + \frac{L}{\min V(q)} < t + \Delta t_{x_L} < t_2 + \frac{L}{\max V(q)}$:

         ➢ $N(x_L, t + \Delta t_{x_L}) \lesssim N(x_0, t_1) + q \cdot (t + \Delta t_{x_L} - t_1) - K(q) L$.   Apply constraint of interval (t_1, t_2).

     ➢ $N(x_L, t + \Delta t_{x_L}) \lesssim N(x_0, t_2) - \kappa\left(\frac{L}{t + \Delta t_{x_L} - t_2}\right) L$.   Apply constraint of point t_2.

     ➢ $t_1 := t_2$.   Move t_1 to t_2.

➢   If $t + \Delta t_{x_L} < \Theta$:

     ➢ $N(x_L, t + \Delta t_{x_L}) \lesssim N(x_L, \theta) + q_D \cdot (t + \Delta t_{x_L} - \theta)$.   Apply outflow discharge rate constraint.

➢ $S(t) := N(x_L, t + \Delta t_{x_L}) - N(x_L, t)$.   Set sending flow.

Algorithm 4.5. Receiving flow with optional capacity drop[*].

> $N\left(x_0, t + \Delta t_{x_0}\right) := N\left(x_0, t\right) + q_C \Delta t_{x_0}$. $\qquad$ Apply inflow capacity constraint.
> If $T < t + \Delta t_{x_0} + L / w$:
>> $t_1 := \max\left(T, t + \Delta t_{x_0} + L / \max Z'\right)$.
>> $N\left(x_0, t + \Delta t_{x_0}\right) \lesssim N\left(x_L, t_1\right) + q_S\left(t + \Delta t_{x_0} - t_1\right) + k_S L$. $\qquad$ Emanate $\left(k_S, q_S\right)$ up to point t_1 inclusive.
>> $t_2 := \lfloor t_1 \rfloor_{\Delta t_{x_L}}$.
>> Loop:
>>> $t_2 := \min\left(t_2 + \Delta t_{x_L}, t + \Delta t_{x_0} + L / \min Z'\right)$. $\qquad$ Find next t_2 .
>>> If $t_1 \geq t_2$:
>>>> Exit the loop.
>>> $q' := \frac{N(x_L, t_2) - N(x_L, t_1)}{t_2 - t_1}$.
>>> If $t_1 - \frac{L}{\max V'(q')} < t + \Delta t_{x_0} < t_2 - \frac{L}{\min V'(q')}$:
>>>> $N\left(x_0, t + \Delta t_{x_0}\right) \lesssim N\left(x_L, t_1\right) + q' \cdot \left(t + \Delta t_{x_0} - t_1\right) + K'(q') L$. Apply constraint of interval $\left(t_1, t_2\right)$.
>>> $N\left(x_0, t + \Delta t_{x_0}\right) \lesssim N\left(x_L, t_2\right) + \kappa'\left(\frac{L}{t_2 - t - \Delta t_{x_0}}\right) L$. $\qquad$ Apply constraint of point t_2 .
>>> $t_1 := t_2$. $\qquad$ Move t_1 to t_2 .
> $R(t) := N\left(x_0, t + \Delta t_{x_0}\right) - N\left(x_0, t\right)$. $\qquad$ Set receiving flow.

For brevity of notation, we use a compound assignment operator $a \lesssim b$ meaning $a := \min(a, b)$, a floor-to-multiple-of operator $\lfloor a \rfloor_b$ meaning $\lfloor a/b \rfloor \cdot b$ and a ceil-to-multiple-of operator $\lceil a \rceil_b$ meaning $\lceil a/b \rceil \cdot b$.

For Algorithm 4.1, the Courant-Friedrichs-Lewy (1928) condition requires that the time step sizes of nodes are chosen such that for each link,

$$\Delta t_{x_0} \leq -L / w \wedge \Delta t_{x_L} \leq L / \max Z .$$ (4.31)

The computational complexity of these algorithms is the same as in Section 4.3.

4.5.4 Dissolving congestion

As indicated before, we still need to address the dissolution of congestion. To this end, Figure 4.8 shows example output of the previous algorithms with sufficiently small time steps, along with the implied traffic states within the link. Although the onset of congestion correctly follows our traffic flow theory with capacity drop, we can identify three problems related to the resolution of congestion, that occur if the inflow later again exceeds the queue discharge rate. In order of decreasing severity, the problems are that:

1. for large t, the outflow q_D is inconsistent with the inflow q_C, violating our free-flow differential equation;

---

[*] Use $\left(k_C, q_C\right) = \left(k_D, q_D\right) = \left(k_S, q_S\right)$ and $w = \min Z'$ if there is no capacity drop.
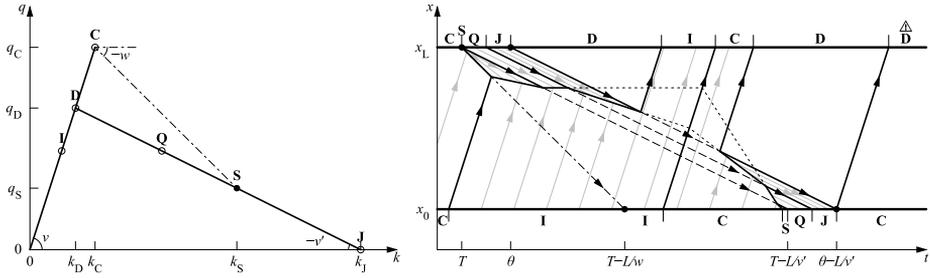
Figure 4.8. Problems related to dissolution of congestion. Letters indicate traffic states in the FD (left) and solution (right).

2. the low inflow q_I has dissolved the queue, but the queue re-appears later, violating our shock admissibility criterion with an incorrectly active separating shock;

3. the low inflow has dissolved traffic state Q within the queue, but this traffic state re-appears later, violating the shock admissibility criterion with an incorrect traffic state behind the separating shock.

The first problem is solved by setting Θ to its correct value rather than infinity. This also solves the second problem if T is adjusted accordingly. We will investigate how to do so below. It addresses the third problem as well, but does not solve it entirely if the congested branch of the FD is non-linear.

Essentially, the previous algorithms need to be supplemented with a mechanism that marks congested traffic states as dissolved, recovering the free-flow capacity. Hence, consider a point (x_0, t_0) on the upstream link end and a point (x_L, t_L) on the downstream link end that is preceded by congestion, possibly including the queue discharge state. We will now investigate when we can prove that this congestion will be cleared. Since a queue will always grow if the inflow exceeds the queue discharge rate, assume that the flow prior to (x_0, t_0) does not exceed q_D. Then, in the worst case, the link inflow will be q_C after t_0, such that (x_0, t_0) emanates an acceleration fan. The last traffic state in this fan not exceeding q_D will reach the downstream link end at $t_0 + L / \min V(q_D)$, and the corresponding wave will constrain $N(x_L, t_0 + L / \min V(q_D))$ to $N(x_0, t_0) + (q_D / \min V(q_D) - k_D)L$. Thus, we know the congestion will be dissolved if

$$N(x_0, t_0) + \left(\frac{q_D}{\min V(q_D)} - k_D \right) L \leq N(x_L, t_L) + q_D \cdot \left(t_0 + \frac{L}{\min V(q_D)} - t_L \right) \Leftrightarrow$$
$$N(x_0, t_0) - k_D L - q_D t_0 \leq N(x_L, t_L) - q_D t_L \tag{4.32}$$

Since this analysis is based on a valid forward path with speed $\min V(q_D)$, one does not need to verify that the flow prior to (x_0, t_0) is not larger than q_D. Moreover, if the head of a moving

jam present before (x_L, t_L) has reached the upstream link end in the time step prior to (x_0, t_0), Equation (4.32) becomes an equality so that congestion on the link is also treated as dissolved.

If congestion is dissolved, T can be increased to t_L, ensuring that the dissolved congestion does not constrain the future inflow. If the link is now free of congestion, T can be set to ∞ and Θ can be set to $t_0 + L / \min V(q_D)$, ensuring that both later inflow and outflow are unconstrained.

Algorithm 4.6 applies this theory, increasing T as congestion is dissolved and incrementing t_0 otherwise, until either of them lies in the future. It must be executed for every link at the

Algorithm 4.6. Marking congestion as dissolved.

➢ If $T < t$:
   ➢ Loop:
      ➢ $t_0 := \max\left(t_0, \left\lfloor T - \frac{L}{\min V(q_D)} \right\rfloor_{\Delta t_{x_0}} + \Delta t_{x_0}\right)$.   Ensure forward path reaches x_L after T.

      ➢ If $t_0 > t$:   Ensure value of $N(x_0, t_0)$ is already known.
         ➢ Exit the loop.
      ➢ If $N(x_0, t_0) - k_D L - q_D t_0 \leq N(x_L, T) - q_D T$:   Check whether (x_0, t_0) can dissolve cong. at (x_L, T).

         ➢ $\hat{T} := \min\left(\begin{array}{c} t_0 + \frac{L}{\min V(q_D)}, \\ \lfloor T \rfloor_{\Delta t_{x_L}} + \Delta t_{x_L} \end{array}\right)$.   See what more (x_0, t_0) could dissolve of downstream time step.

         ➢ $q' := \frac{N\left(x_L, \lfloor T \rfloor_{\Delta t_{x_L}} + \Delta t_{x_L}\right) - N\left(x_L, \lfloor T \rfloor_{\Delta t_{x_L}}\right)}{\Delta t_{x_L}}$.   Compute outflow in downstream time step.

         ➢ If $q' < q_D$:   If less than q_D, not all congestion before $(x_L, \hat{T})$ necessarily dissolves.

            ➢ $\hat{T} :\leq \lfloor T \rfloor_{\Delta t_{x_L}} + \frac{N\left(x_L, \lfloor T \rfloor_{\Delta t_{x_L}}\right) - q_D \lfloor T \rfloor_{\Delta t_{x_L}} - N(x_0, t_0) + k_D L + q_D t_0}{q_D - q'}$.   Thus decrease $\hat{T}$ appropriately[*].

         ➢ If $\hat{T} = \lfloor T \rfloor_{\Delta t_{x_L}} + \Delta t_{x_L}$:   If downstream time step dissolved completely.
            ➢ If $\hat{T} \geq \theta$:
               ➢ $T := \infty$.   All currently known congestion dissolved.
               ➢ $\Theta := t_0 + \frac{L}{\min V(q_D)}$. Recover outflow to free-flow once solving forward path reaches x_L.
               ➢ Exit the loop.
         ➢ Else:
            ➢ $t_0 := t_0 + \Delta t_{x_0}$.   Try dissolving remainder of downstream time step with next forward path.
         ➢ $T := \hat{T}$.   Mark dissolved congestion as such.
         ➢ If $T \geq t$:
            ➢ Exit the loop.
      ➢ Else:
         ➢ $t_0 := t_0 + \Delta t_{x_0}$.   Try dissolving congestion with next forward path instead.

---

[*] Decrease $\hat{T}$ until $N(x_0, t_0) - k_D L - q_D t_0 \leq N(x_L, \hat{T}) - q_D \hat{T}$, by solving the equality.

beginning of each time step $t : \Delta t_{x_0} \,|\, t \vee \Delta t_{x_L} \,|\, t$, i.e. before computing the sending flow or the receiving flow with the previous algorithms. t_0 is a persistent variable initialised at $-\infty$.

4.6   Node model with capacity drop

As an example of how to adapt a node model to support a capacity drop, we extend the Tampère et al. (2011) node model for unsignalised intersections in this section. The original version of this node model distributes the receiving flows of the outgoing links proportional to the capacities and turning fractions of the incoming links, thereby supporting any number of incoming and outgoing links, maximising flows to a user-equilibrium and conserving turning fractions. For simplicity, our extended node model will assume that the capacities used for distributing flows are the free-flow capacities.

To accommodate the requirements from Subsection 4.4.5, the node model is able to modify the sending and receiving flows: a high sending flow is reduced to the accelerating discharge rate if it cannot be transmitted in its entirety, and a high receiving flow is reduced to the merging discharge rate if it is smaller than the sum of the sending flows directed at it. The invariance principle is satisfied, as proven in Appendix 4.A.

Algorithm 4.7 is the extended algorithm, following the notation and structure of Tampère et al. (2011). C_i and D_i denote the free-flow capacities and accelerating discharge rates for incoming links integrated over the time step, E_j denotes the merging discharge rate for outgoing links. $\tilde{R}_j$ denotes the unallocated portion of receiving flow R_j, U_j indicates which incoming links are constrained by an outgoing link and J indicates which outgoing links still need to be solved. Calculated in step 2 of the algorithm, C_{ij} denote the oriented capacities, which are the products of the free-flow capacities of incoming links and the corresponding turning fractions. Step 4 finds the most constraining outgoing links $\hat{J} \subseteq J$ [*] and the corresponding reduction factor $\hat{a}$. Then, step 5 first eliminates demand-constrained incoming links ($S_i \le \hat{a} C_i$). A capacity drop must then occur on the remaining incoming links, so T_i is reduced[†] to t and their sending flows are capped at their accelerating discharge rates, initialising Θ_i and θ_i. If some sending flows are indeed reduced, step 5 is restarted as there may now exist additional demand-constrained incoming links. Eventually, for the remaining incoming links, the outflows must be smaller than the accelerating discharge rates. If the supply-constrained transition flows G_{ij} together exceed the merging discharge rate E_j of the outgoing link, the entire node model is restarted from step 3, with $R_j = E_j$ to incorporate the capacity drop on the entrance of the outgoing link as well, preserving only previous

---

[*] Unlike Tampère et al. (2011), we process multiple outgoing links with identical reduction factors simultaneously instead of sequentially, as the arbitrary order might otherwise influence which capacity drops are activated. Without capacity drops, it would not change any outcome, as identical reduction factors would stay the same until all are processed.

[†] **Correction.** The article in *Transportation Research Part B* erroneously described this action as $T_i := t$ instead of $T_i := \min(T_i, t)$.

Algorithm 4.7. Node model with optional capacity drops.

> Determine all S_i, S_{ij} and R_j.　　　　　*Step 1: Retrieve link constraints and initialise supplies and sets.*
> $\forall i:\ C_i := q_{C,i}\Delta t$, $D_i := q_{D,i}\Delta t$.
> $\forall j:\ E_j := q_{E,j}\Delta t$, $\tilde{R}_j := R_j$, $U_j := \{i\,|\,S_{ij} > 0\}$.
> $J := \{j\,|\,U_j \neq \varnothing\}$.

> $\forall i\,|\,S_i > 0:\ \forall j:\ C_{ij} := C_i S_{ij}/S_i$.　　　　*Step 2: Determine oriented capacities.*

> If $J = \varnothing$: Stop.　　　　　　　　　*Step 3: Stopping criterion.*

> $\forall j \in J:\ a_j := \tilde{R}_j / \sum_{i \in U_j} C_{ij}$.　　　*Step 4: Determine most restrictive constraints.*
> $\hat{a} = \min\limits_{j \in J} a_j$, $\hat{J} := \arg\min\limits_{j \in J} a_j$.

> > *Step 5: Determine flows of corresponding incoming link sets and recalculate receiving flows.*
> If $\exists i \in \bigcup_{j \in \hat{J}} U_j \,|\, S_i \leq \hat{a}C_i:\ \forall i \in \bigcup_{j \in \hat{J}} U_j \,|\, S_i \leq \hat{a}C_i:$
> > $G_i := S_i$.
> > $\forall j:\ G_{ij} := S_{ij}$.
> > $\forall j \in J:$
> > > $\tilde{R}_j := \tilde{R}_j - G_{ij}$, $U_j := U_j \setminus \{i\}$.
> > > If $U_j = \varnothing:\ J := J \setminus \{j\}$.
> Else:
> > $\forall i \in \bigcup_{j \in \hat{J}} U_j:\ T_i := \min(T_i, t)$, $\Theta_i := \infty$, $\theta_i := t + \Delta t$. Mark time step as cong. and intialise discharging.
> > If $\exists i \in \bigcup_{j \in \hat{J}} U_j \,|\, D_i < S_i:$　　　　Check for sending flows exceeding accelerating discharge rates.
> > > $\forall i \in \bigcup_{j \in \hat{J}} U_j \,|\, D_i < S_i:$　　　　Reduce them to accelerating discharge rates.
> > > > $\forall j:\ S_{ij} := S_{ij} D_i/S_i$.
> > > > $S_i := D_i$.
> > > Restart step 5.　　Re-check for demand-constrained incoming links since we reduced sending flows.
> > $\forall i \in \bigcup_{j \in \hat{J}} U_j:$
> > > $G_i := \hat{a}C_i$.
> > > $\forall j:\ G_{ij} := \hat{a}C_{ij}$.
> > If $\exists \hat{j} \in \hat{J} \,|\, E_{\hat{j}} < \sum_i G_{\hat{i}j}:$
> > > $\forall \hat{j} \in \hat{J} \,|\, E_j < \sum_i G_{\hat{i}j}:\ R_j := E_j$.
> > > $\forall j:\ \tilde{R}_j := R_j$, $U_j := \{i\,|\,S_{ij} > 0\}$.
> > > $J := \{j\,|\,U_j \neq \varnothing\}$.
> > > Go to step 3.
> > $\forall i \in \bigcup_{j \in \hat{J}} U_j:$
> > > $\forall j \in J:$
> > > > $\tilde{R}_j := \tilde{R}_j - G_{ij}$.
> > > > If $j \notin \hat{J}:$
> > > > > $U_j := U_j \setminus \bigcup_{j \in \hat{J}} U_j$.
> > > > > If $U_j = \varnothing:\ J := J \setminus \{j\}$.
> > $J := J \setminus \hat{J}$.
> Go to step 3.

reductions of sending and receiving flows. If this is not needed, the algorithm simply continues with the next iteration.

The flow maximisation property of the Tampère et al. (2011) model has a special benefit for our extension: it will prevent unnecessary activation of capacity drops. The choice of links forwhich to activate a capacity drop is not necessarily flow-maximising however: the order of activation will be dictated by the reduction factors of the outgoing links, with the most severe reduction first. As an example of non-flow-maximising behaviour, in case of a simple merge, the sending flow of both incoming links is reduced to their accelerating discharge rates simultaneously, while one such reduction might have sufficed.

4.7 Numerical examples

In this section, we investigate the qualitative properties of the proposed model in numerical examples. We start by demonstrating the elementary improvements of our model over the original LTM with triangular FDs in Subsection 4.7.1. After this synthetic example, we test the model in a real motorway corridor network in Subsection 4.7.2.

4.7.1 Elementary model features

To demonstrate the elementary features of our extended link and node models, we simulate a highway with two on-ramps and visualise the resulting traffic states along the highway. All links have identical FDs. The synthetic demand and FD, with annotated model results for the highway, are shown in Figure 4.9, including a comparison with a model with a triangular FD.

Let us first look at the extended LTM. While the main road's inflow is initially low, it quickly increases to the free-flow capacity, showing a forward acceleration fan. When a low flow from the 9 km on-ramp tries to merge, this generates a stop-and-go wave whose traffic state is emanated from the shock at its tail. The reason that there is no standing queue is that the sum of the accelerating discharge rate from the main road and the low flow from the parallel road do not exceed the free-flow capacity downstream of the merge. Nevertheless, the flow downstream of the merge still drops because the on-ramp flow is not equal to the capacity drop of the stop-and-go wave. A little less than half an hour into the simulation, the stop-and-go wave dissolves due to temporarily low inflow, so that when the inflow is high again later, it can continue without breaking down.

Around the same time, the flow from the ramp increases, causing a standing queue whose outflow equals the merging discharge rate of the link downstream of the merge. Once the queue on the ramp has been dissolved, the head of the queue on the main road starts to move upstream, showing a backward acceleration fan at its head. This is because now the sum of the sending flows does fit into the free-flow capacity downstream, and the node model has no memory of the previous congestion. The queue eventually dissolves, recovering the free-flow capacity of the main road.

Meanwhile, the upstream end of this queue has encountered an inflow higher than its discharge rate, causing it to emanate state (k_s, q_s) from the separating shock. As remarked in Subsection 4.4.6, this can also be interpreted as a deceleration fan. While this high inflow is temporarily interrupted, the queue partially dissolves, starting a new deceleration fan when the high inflow resumes.
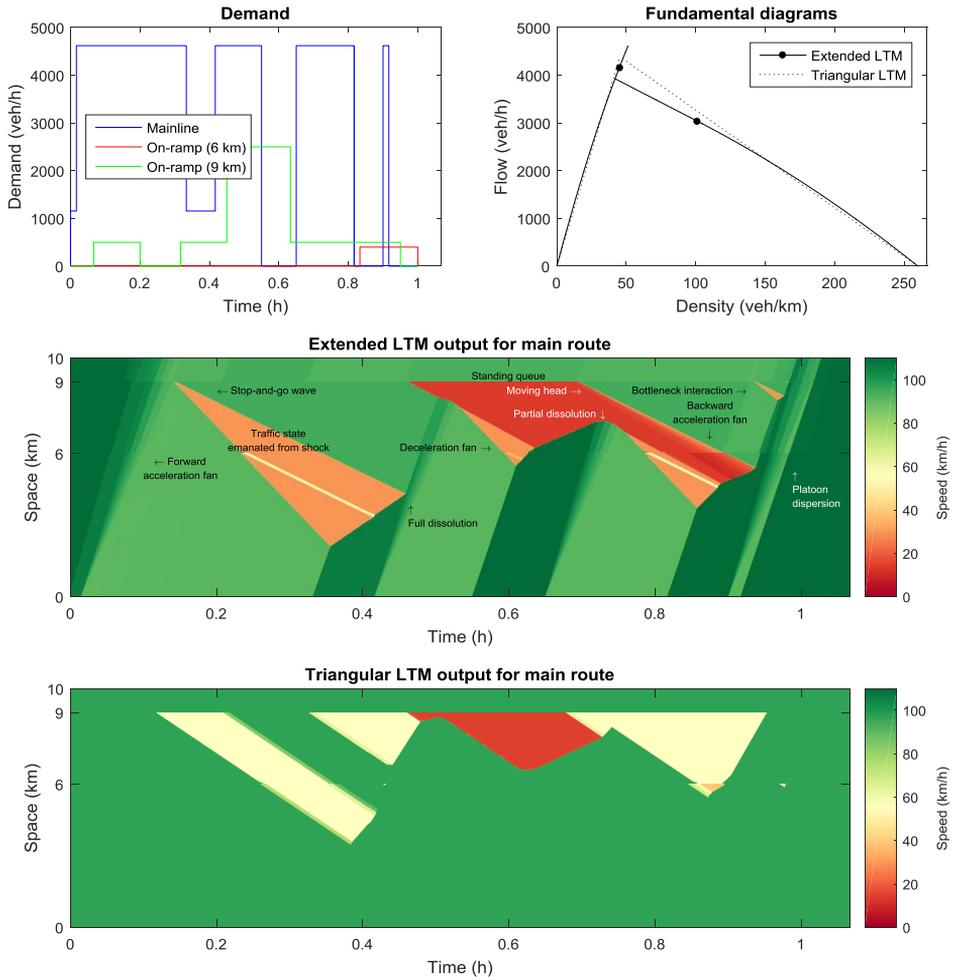
Figure 4.9. Simulation input and results for the example highway.

When the head of the queue passes the 6 km on-ramp, which now has a non-zero demand, the flow downstream of this on-ramp increases beyond the accelerating discharge rate of that link: the outflow of the queue remains the same, but the on-ramp can now add more flow. When this higher flow reaches the downstream on-ramp, traffic breaks down again. This means bottlenecks can interact with each other, which can produce an oscillatory traffic pattern even if the demand is constant. We will see more of this in the next subsection. Note that in general the interacting bottlenecks need not be on-ramps. They could include e.g. lane drops, off-ramps, any other geometric discontinuity or even (simplified) capacity funnels (Buckley and Yager, 1974).

Finally, near the end of the simulation, a short platoon with a flow equal to the free-flow capacity originates from upstream. Because the platoon disperses into an acceleration fan, it

does not break down when it reaches the 6 km on-ramp even though the on-ramp demand is still non-zero.

If we look at the results of the model with a triangular FD, we see that the acceleration fans are not present and that the speed in free-flow conditions does not decrease with increasing traffic density. Albeit standing queues can dissolve into moving queues if the merging demand drops, such moving queues are unable to grow, since unlike in our extended LTM, the inflow into a moving queue can never exceed its outflow. This also precludes the formation and growth of stop-and-go waves. Finally, the triangular LTM does not reproduce the alternating network outflow pattern of the extended LTM resulting from the capacity drops affecting the merges themselves and the highway upstream of the merges. Note that for both our extended LTM and the triangular LTM, all shocks are crisp, showing that the small numerical error of the original LTM is maintained in our extension.

4.7.2   Qualitative properties on a motorway corridor network

To investigate the qualitative behaviour of the model in a realistic setting, we simulate the Dutch A13 motorway corridor from the Kleinpolderplein interchange near Rotterdam to the Ypenburg interchange near The Hague. This motorway stretch has a length of 12 km, with five off-ramps and six on-ramps. The number of lanes varies between three and four, excluding ramp lanes. We model the corridor as a network of 23 links and 24 nodes as depicted in Figure 4.10. We base the capacities on the Dutch motorway capacity manual (Witteveen+Bos and TU Delft, 2011) and construct FDs with a capacity drop of 15% for accelerating and 10% for merging. Besides this fully extended model, we also conduct the same simulations in a partially extended model with a capacity drop for merging only (similar to Jin et al. (2015)) and a straight congested FD branch, as well as traditional simulations with triangular FDs.

We simulate three evening peaks in September 2012 with different patterns of congestion. In each simulation, the mainline demand is estimated from one-minute flow measurements of the loop detectors at the beginning of the corridor, and on-ramp demands and off-ramp split fractions are estimated from the differences between flow measurements of loop detectors downstream and upstream of the ramp. This estimation of demand is quite rough and hence both the extended model and the original model show large quantitative errors compared to reality. We nevertheless believe it suffices for the purpose of illustrating the qualitative properties of the model.

The results are plotted in Figure 4.11. Qualitatively, we see that our extended LTM replicates the variation of traffic speeds in free-flow conditions, as the used fundamental diagrams contain non-linear free-flow branches. Moreover, the capacity drop in the fully extended
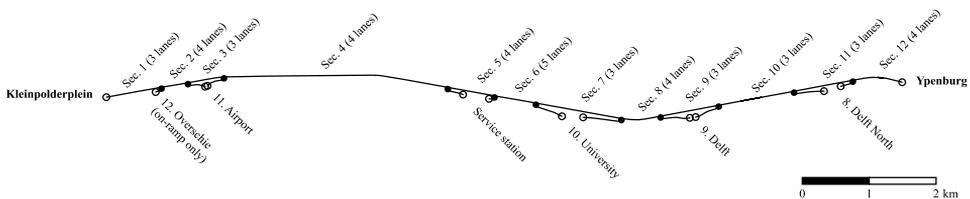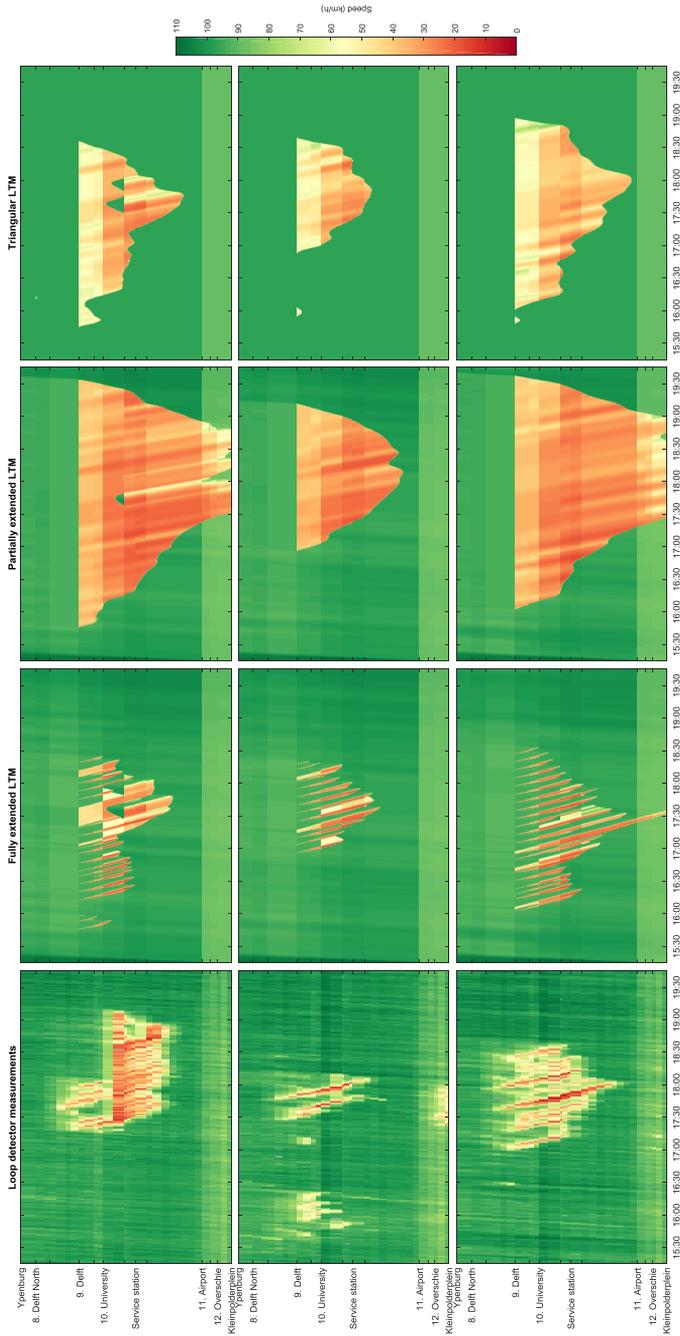

Figure 4.10. A13 corridor network.

Figure 4.11. A13 measurements and simulation results[*] for three evening peaks.

---

[*] **Correction.** The article in *Transportation Research Part B* reports slightly different numerical results for the fully extended LTM because of a round-off error in the calculation of $\hat{T}$ in Algorithm 4.6.

model allows it to reproduce both the onset and propagation of stop-and-go waves, resulting in the structure of the congestion pattern being much closer to reality than the reference model. The fully extended LTM is able to produce complex mixtures of both moving queues and standing queues, whose composition can be fine-tuned by adjusting the capacity drop. In the partially extended LTM on the other hand, the capacity drop for merging only increases the severity of the congestion, but does not create stop-and-go waves.

Interestingly, stop-and-go waves generated from an on-ramp can be beneficial for the downstream flow compared to having a standing queue at the same on-ramp, because it results in the on-ramp flow being added on top of the queue discharge rate rather than claiming a part of it. In other words, a moving queue effectively meters the inflow into the original bottleneck. The partially extended model, which lacks this metering effect, has lower flows downstream of the main bottleneck and thus more severe congestion than the fully extended model, despite the latter model having additional capacity drops for accelerating and equal capacity drops for merging.

Once the moving queue passes another on-ramp, this metering effect on the original bottleneck is reduced. This mechanism of interacting bottlenecks, also mentioned in Subsection 4.7.1, causes repetitive patterns of stop-and-go waves like in reality. The oscillations are a logical consequence of the presence of moving jams and relatively constant demand. The period of oscillation, which is dictated by the distance and wave speeds between the interacting bottlenecks, may be modified by changing the network definition, e.g. by adding explicit capacity funnels.

We also observe that a stop-and-go wave can trigger a standing queue when passing an upstream bottleneck. This again matches reality and it works even though the node model has no memory of capacity drops in previous time steps: the stop-and-go wave creates queues on links in front of the bottleneck, such that after it passed, the accelerating discharge rates of these links exceed the free-flow capacity of the motorway downstream of the bottleneck.

4.8 Conclusions

In this chapter, we extended the LTM, first to handle continuous concave non-triangular FDs, and later to include a capacity drop as well following an extension of LWR theory. For the former case, we showed based on Daganzo (2005a; 2005b) that the link model matches LWR theory up to the accuracy permitted by the discrete time steps with constant link in- and outflows, unlike previously proposed formulations. Next, based on an extension of LWR theory, we modified the solution algorithms of both the link model and the node model to include an optional capacity drop with an inverted-lambda style FD, paying special attention to both the transition from free-flow to congestion and vice versa. Here, the node model is an extension of Tampère et al. (2011), that does not memorise the capacity drops in previous time steps. When the inflow into the node is too high, it first attempts to create stop-and-go waves on the relevant incoming links to reduce the incoming flows to the queue discharge rates, and switches to standing queues, restricted to at most the queue discharge rates of the outgoing links, if the previous reduction would be insufficient. Overall, the result is a computationally-efficient first-order simulation model including optional capacity drops, applicable to general networks, able to produce acceleration fans or rarefaction waves and able to simulate both the onset and propagation of both standing and moving jams.

Both without and with the capacity drop, the numerical error is very small, leading to crisp shockwaves between traffic states. This is an advantage over cell-based models like the CTM and its previously proposed extensions, and also over second-order models with cell-based numerical schemes. Our numerical examples show that our extended model is the first first-order model able to produce traffic patterns featuring the creation, growth and dissolution of both standing queues and stop-and-go waves, consistent with a capacity drop. For further research, we recommend investigating to what extent the qualitative advantages of our extended LTM materialise as quantitative benefits. Since our results show considerable oscillatory interaction between bottlenecks, it would also be relevant to investigate how the model results are affected by details of the network definition, such as capacity funnels, and how this compares to higher-order traffic flow theory.

Acknowledgements

4.A Proof of node model invariance

In this appendix, we prove that the node model formulated in Section 4.6 produces the same transition flows in the first time step of a breakdown as in subsequent time steps, ceteris paribus. To this end, we must show that if we reduce a sending flow within the Tampère et al. (2011) algorithm, instead of setting the sending flow directly to this lower value before starting the algorithm, the algorithm produces the same transition flows regardless.

Therefore, let us define a relaxed version of the Tampère et al. (2011) algorithm, such that, at the beginning of step 4a[*], not all incoming links with insufficient demand $\{i \in U_j \mid S_i \le a_j C_i\}$ need to be solved as demand-constrained, but for some reason only a non-empty subset of them (because the final values of S_i might be smaller than the current values). We now show that this change does not affect the outcomes.

Lemma 4.1. The relaxed algorithm has a finite number of iterations.

Proof. In each iteration, either step 4a removes at least one element from at least one U_j (namely i from U_j), or step 4b removes at least one element from J (namely $\hat{j}$). Since these operations are not reversible and all sets are finite, the stop criterion $J = \varnothing$ is satisfied in a finite number of iterations.   ∎

---

[*] The equivalent position in our extended algorithm in Section 4.6 is the beginning of step 5. All further references to step numbers in this appendix also refer to Tampère et al. (2011).

Lemma 4.2. All demand-constrained incoming links are recognised as such by the relaxed algorithm.

Proof. In the relaxed algorithm, step 4a may not empty $U_{\hat{j}}$ and then $\hat{j} \in J \neq \varnothing$ in the next iteration of the algorithm. As proven in Appendix B.1 of Tampère et al. (2011), no a_j will decrease over iterations (this proof also holds for our relaxed algorithm). In later iterations, $a_{\hat{j}} = \min_{j \in J} a_j$ will therefore never be smaller than in previous iterations. This means that all incoming links $i \in U_{\hat{j}}$ satisfying $S_i \leq a_{\hat{j}} C_i$ in a previous iteration, but that were not solved as demand-constrained yet, will again satisfy $S_i \leq a_{\hat{j}} C_i$, regardless of $\hat{j}$ provided that $S_{ij} > 0$. Hence, if the algorithm arrives at step 4b, it is guaranteed that $\forall i \in U_{\hat{j}} \mid S_i > a_{\hat{j}} C_i$ and hence no demand-constrained incoming links are still a member of $U_{\hat{j}}$. If step 4b is never reached, $U_{\hat{j}}$ must be empty when the stop criterion is satisfied, which eventually happens due to Lemma 4.1. Hence all demand-constrained incoming links are removed from all U_j. ∎

Theorem 4.7. Each transition flow G_i produced by the relaxed algorithm gets the same value as in the original algorithm.

Proof. In each step 3 following step 4, some G_{ij} have been subtracted from the numerator of each a_j and some $C_{ij} \geq G_{ij}$ have been subtracted from the denominator, see Appendix B.1 of Tampère et al. (2011). Since both modifications are commutative, neither the order nor the number of iterations in which step 4 removes incoming links from $U_{\hat{j}}$ can impact the first $a_{\hat{j}}$ received by step 4b, but only the set of incoming links that are removed. Hence, since Lemma 4.2 guarantees that at least all demand-constrained incoming links have been removed from $U_{\hat{j}}$, the relaxed algorithm can only fail by removing too many incoming links from $U_{\hat{j}}$. Since, by Tampère et al. (2011), a removal cannot decrease $a_{\hat{j}}$, the first $a_{\hat{j}}$ in step 4b can only be larger than or equal to the correct value. However, $a_{\hat{j}}$ can only get too large after an element from $U_{\hat{j}}$ is incorrectly removed, yet the criterion for step 4a ensures that an element can only be removed from $U_{\hat{j}}$ incorrectly after $a_{\hat{j}}$ is too large. Hence $a_{\hat{j}}$ is correct, step 4b produces the correct flows $G_i = a_{\hat{j}} C_i$ for $i \in U_{\hat{j}}$. If there is a next iteration, the modifications of a_j for all remaining $j \in J$ in step 3 are therefore correct. Because all these modifications commute, the $a_{\hat{j}}$ in a next step 4b is again not impacted by the order of removal of incoming links from $U_{\hat{j}}$ by either steps 4a or 4b. Again Lemma 4.2 guarantees that all required removals from $U_{\hat{j}}$ have been carried out, and the criterion of step 4a again guarantees that not too many incoming links have been classified as demand-constrained and removed. So a next step 4b also produces correct results. By induction, all steps 4b together produce correct flows for all supply-constrained incoming links. By Lemma 4.2, the overall algorithm produces correct flows for all demand-constrained incoming links as well. ∎

Chapter 5

The Link Transmission Model with variable fundamental diagrams and initial conditions

Abstract

The Link Transmission Model is a macroscopic network traffic flow simulation tool based on Lighthill-Whitham-Richards theory. While its efficiency and accuracy are superior to the well-known Cell Transmission Model, applications of its current numerical formulations are limited by the inability to apply changes to the fundamental diagrams of links within a simulation and the need to start the simulation with an empty network. We resolve both limitations by developing methodology for initialising the discrete-time link model with a non-empty initial condition and for computing within-link densities during the simulation, which can then serve as an initial condition for continued simulation with a new fundamental diagram. Since the computation of within-link densities is algebraic, no new numerical errors are introduced. Optional support for multiple commodities, subcritical delays and platoon dispersion, are retained. The resulting model is demonstrated on a motorway corridor network with variable speed limits and dynamic lane management.

This chapter is a version of the following manuscript under review:

5.1 Introduction

Lighthill and Whitham (1955) and Richards (1956) were the first to model traffic flow using a differential equation combining conservation of vehicles with an equilibrium flow-density relationship referred to as the fundamental diagram (FD). The traditional cell-based numerical solution scheme to this Lighthill-Whitham-Richards (LWR) theory is known as the Cell Transmission Model (CTM) (Daganzo, 1994) and can simulate the propagation of traffic in a network of links and nodes using macroscopic variables (Daganzo, 1995a). While this solution scheme with links discretised into small homogeneous cells is intuitive, it has turned out to be neither the most efficient in terms of computation time nor the most accurate in terms of numerical error. Based on Newell's (1993) recipe using the cumulative numbers of vehicles as the main variable, Yperman et al. (2005) and Gentile and Papola (2009a; 2009b) developed the Link Transmission Model (LTM). This alternative numerical solution scheme for solving traffic propagation in a network does not discretise links into small cells, improving the efficiency and accuracy compared to the CTM and variants (Yperman et al., 2005; Yperman, 2007, pp. 46-49).

Since then, the LTM has accumulated improvements and extensions. Yperman et al. (2006) developed a multi-commodity version keeping track of heterogeneous traffic composition, to which Yperman (2007) added support for general nodes with multiple incoming and outgoing links. Gentile (2010) added general nodes to his formulation as well. These and other general node models have further been improved to satisfy the criteria for traffic flow at junctions formulated by Lebacque and Khoshyaran (2005) and Tampère et al. (2011), resulting in a family of node models described by Smits et al. (2015), and possibilities to simulate signalised junctions (Tampère et al., 2011) and ramp metering (Hajiahmadi et al., 2013).

Furthermore, while the formulations by Yperman et al. (2005; 2006) were limited to triangular FDs on links, Yperman (2007) proposed piecewise-linear FDs as well, and Gentile and Papola (2009a; 2009b) and Gentile (2010) proposed strictly concave FDs. Such FDs introduce subcritical delays, i.e. demand-dependent increases in link travel time before the capacity of the link is exceeded as illustrated in Figure 5.1. Due to an initially sharp front of a platoon dispersing into an acceleration fan, they also introduce platoon dispersion as illustrated in Figure 5.2 (Geroliminis and Skabardonis, 2005). Chapter 4 of this thesis improved the numerical accuracy in both proposals and allowed any concave FD. Chapter 4 furthermore extended the link and node models to support the capacity drop phenomenon with a mixture of moving jams and standing queues. For triangular FDs, other link model extensions have been proposed for supporting multiple vehicle types (Smits et al., 2011) or variable speed limits (Hajiahmadi et al., 2013).

Contrary to the above discrete-time approaches, Raadsen et al. (2016a) introduced an event-based continuous-time version of the Link Transmission Model, to which multi-commodity support and general concave FDs without capacity drop also have been added (Raadsen and Bliemer, 2018).

Meanwhile researchers have also been working on creating an iterative variant of the numerical solution scheme (Himpe et al., 2016), computing a user-equilibrium traffic assignment (Gentile et al., 2007; Gentile, 2015; Himpe and Tampère, 2016), investigating mathematical properties of the continuous-time network loading (Jin, 2015; Han et al., 2016) and user-equilibrium (Han et al., 2015) and utilising the LTM for optimisation purposes (Hajiahmadi et al., 2016; Van de Weg et al., 2016; Long et al., 2016).
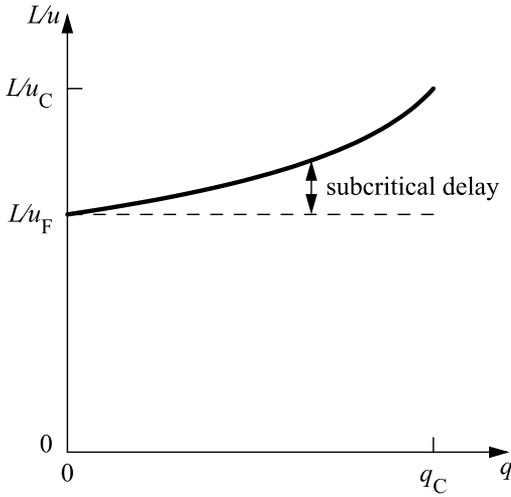
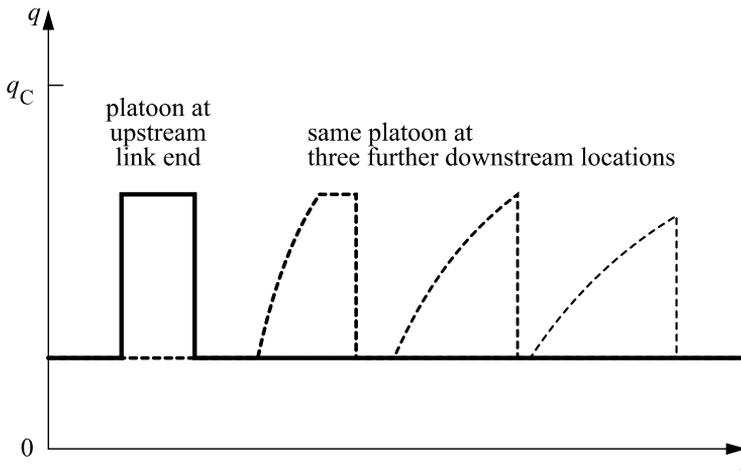**Figure 5.1. Link travel time L/u in stationary homogenous free-flow conditions as a function of flow q.**

Figure 5.2. Flow q versus time t at four locations showing platoon dispersion.

5.1.1   Problem statement

The above literature offers a rich set of features for macroscopically simulating the propagation of traffic in a network over time, also known as dynamic network loading. Nevertheless, except for the triangular-FD variable speed limits model by Hajiahmadi et al. (2013), all of these studies assume that the FDs of all links remain constant during the simulation, which can be problematic for certain applications. Firstly, it precludes time-varying link-based traffic control. Besides variable speed limits, this could also include e.g. temporary lane closures, lane reversal (Wolshon and Lambert, 2006) or operation of peak-hour lanes such as hard shoulder running (Sultan et al., 2008; Geistefeldt, 2012; Guerrieri and Mauro, 2016). Secondly, it precludes time-varying driving behaviour within a simulation. This may e.g. occur due to changes in precipitation, visibility and road surface conditions (Chung et al., 2006; Rakha et al., 2008; Dixit et al., 2012a; Kwon et al., 2013), changes to driver activation levels based on time-of-day (Dixit et al., 2012a) or due to evacuation conditions (Yuan et al., 2014), or changes to the average vehicle composition over time.

Another drawback of current LTM simulations is the need to start with an empty network. Only Jin (2015) considered more general initial conditions for triangular FDs, but did not develop a numerical solution scheme. This missing feature can also be restrictive, particularly for real-time applications. For example, non-empty initial conditions are used for rolling horizon optimisation of traffic control (Papageorgiou et al., 2003).

The solutions to these two limitations are related. To instantaneously change the FD of a link, it suffices to compute the density of traffic at all places within the link, and use those as the initial condition for a subsequent simulation with the new FD (see Figure 5.3). In the CTM this is trivial, since the state vector of the model consists entirely of the current cell densities. On the contrary, the LTM state vector does not contain within-link densities and spans more than one time instant. The LTM thus needs to be supplemented with two additional procedures: one to compute the within-link densities, and one to employ these as initial conditions. Hence, although the LTM is normally more efficient than the CTM, it will incur some additional computational cost each time within-link densities need to be used. A
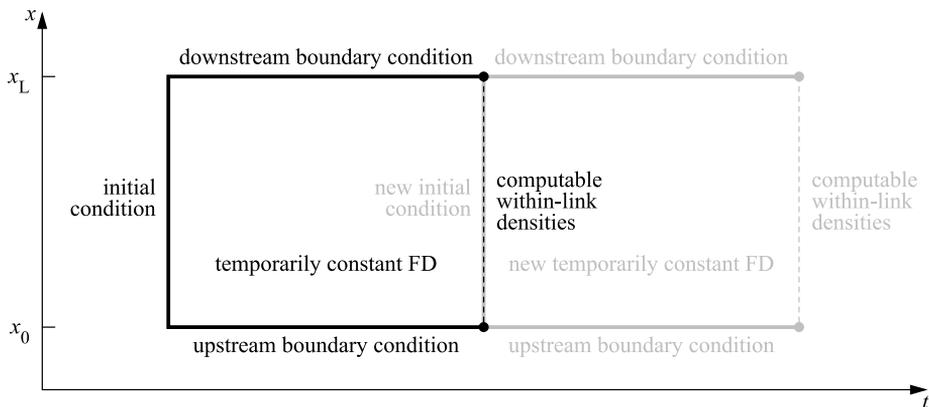

Figure 5.3. Space-time (x,t) illustration of an FD change for a link from x_0 to x_L.

convenient property of this setup is that one can even use the within-link densities at time t^* to decide whether and how the FD or within-link should change at the same time t^*. This enables real-time corrections to within-link densities or FD parameters (i.e. data fusion) as well as simulations with real-time traffic-responsive control measures.

5.1.2   Contribution of this chapter

The main contribution of this chapter is a novel link model for the discrete-time LTM that includes such methodologies for computing within-link densities and applying initial conditions, offering a single, computationally-efficient, LWR-theory-based simulation tool supporting subcritical delays and platoon dispersion, non-empty initial conditions, computation and modification of densities within links and changes to any FD parameters during the simulation.

As shown by Chapter 4, it is possible to develop constant-FD discrete-time LTM link models without numerical errors other than those introduced by the time discretisation of node flows. This suggests the possibility of using algebraic descriptions of within-link densities for particular choices of the FD shape, akin to how Raadsen et al. (2016a) use boundary flows without numerical errors. This chapter will exploit that possibility to avoid introducing any additional numerical errors, so that e.g. changing an FD into an identical FD does not affect the numerical results in any way.

This chapter is structured as follows. Section 5.2 describes the structure of the LTM and explains how our extension is embedded in this structure. Section 5.3 introduces the main methods needed to construct an LTM link model, relating to the traffic flow theory and the FD. Section 5.4 applies these methods to formulate our extended LTM link model. A practical demonstration of the extended LTM is provided in the simulation study in Section 5.5, including comparisons with the unextended LTM and the CTM. This is followed by a discussion of the extended model in Section 5.6 and our conclusions in Section 5.7.

5.2   Structure of the Link Transmission Model and its extension

The LTM consists of a link model describing traffic flow on homogenous road stretches, and a node model describing traffic flow at discontinuities, including intersections. This section describes the role of both models within the LTM and discusses how they are affected by our extension for variable FDs and initial conditions.

Please note that one can refer to Appendix 5.A for a complete overview of all notation that will be introduced in this section and the next two sections.

The link model of the LTM originally consists of two components. The first is the sending flow algorithm that computes the maximum number of vehicles $S(t)$ that can leave the link between now t and some point in the near future $t + \Delta t$, under the assumption that the downstream node does not impose any constraints. Similarly, the second is the receiving flow algorithm that computes the maximum number of vehicles $R(t)$ that can enter the link between now t and some point in the near future $t + \Delta t$, again under the assumption that in this case the upstream node does not impose any constraints. Both are done by comparing the maximum possible cumulative vehicle number N at the end of the time window $(t, t + \Delta t)$ with the known value of N at the start of the time window:

$$S(t) = N^{\max}\left(x_L, t + \Delta t_{x_L}\right) - N\left(x_L, t\right)$$
$$R(t) = N^{\max}\left(x_0, t + \Delta t_{x_0}\right) - N\left(x_0, t\right) \text{,}$$

(5.1)

where x_L denotes the downstream link end, x_0 denotes the upstream link end, and $N^{\max}$ denotes N assuming the downstream and upstream node impose no constraints on outflow and inflow respectively during $(t, t + \Delta t)$. Consistent with Yperman (2007, p. 25) and Chapter 4, each node can be updated with a different time window Δt. We use Δt_{x_0} and Δt_{x_L} to denote the time window sizes of the upstream and downstream node of a link respectively. To be clear, this means that every incoming link of the same node has the same Δt_{x_L} and every outgoing link of the same node has the same Δt_{x_0}. By the Courant-Friedrichs-Lewy (CFL) condition (Courant et al., 1928), the time window sizes of nodes must satisfy

$$\begin{cases} \Delta t_{x_0} \leq -\dfrac{L}{v'} \\ \Delta t_{x_L} \leq \dfrac{L}{u_F} \end{cases}$$

(5.2)

for all connected links, where L is the link length, u_F is the free speed, and $v' < 0$ is the wave speed in congestion.

The sending and receiving flows from Equation (5.1) serve as inputs to the node model of the LTM which, after resolving conflicts and applying node capacity constraints, determines the actual link inflows and outflows during such time windows, known as transition flows, forming new parts of the boundary conditions in the link models of the links connected to the node. This interaction between the link and node models is outlined in Algorithm 5.1.

The extension of the LTM to variable FDs and initial conditions, as proposed in this chapter, adds a third component to the link model, namely the ability to compute the densities within links. Furthermore, the sending and receiving flow algorithms are adapted to take the initial condition into account. Figure 5.4 illustrates the three link model components in the space-time plane: computation of within-link densities at time t, computation of the sending flow for a time window $(t, t + \Delta t_{x_L})$, and computation of the receiving flow for a time window

Algorithm 5.1. Calculating node flows for time window $(t, t + \Delta t)$, after Yperman et al. (2005).

> Compute sending flow $S_i(t)$ for each incoming link i using the link model.　　　　*Step 1.*
> Compute receiving flow $R_j(t)$ for each outgoing link j using the link model.　　　　*Step 2.*
> Determine turning fraction $S_{ij}(t) / S_i(t)$ for each turn ij.　　　　*Step 3.*
> Compute transition flow $G_{ij}(t)$ for each turn ij using the node model.　　　　*Step 4.*
> Extend downstream boundary condition of each incoming link i with $\sum_j G_{ij}(t)$.　　　　*Step 5.*
> Extend upstream boundary condition of each outgoing link j with $\sum_i G_{ij}(t)$.　　　　*Step 6.*
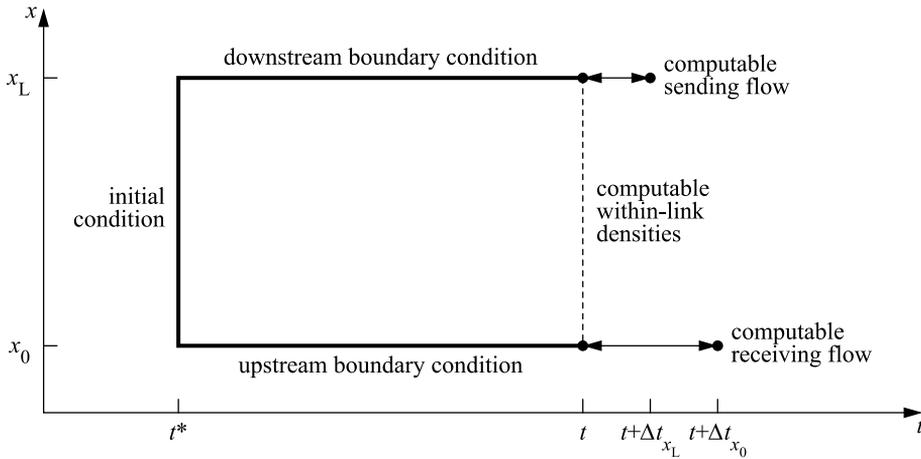
Figure 5.4. Illustration of the three link model components in the space-time (x,t) **plane.**

$\left(t, t+\Delta t_{x_0}\right)$. The figure also illustrates their inputs: an upstream boundary condition at location x_0, a downstream boundary condition at location x_L, and an initial condition at time t^*.

The node model on the other hand is not affected by our extension of the LTM. Despite the sending and receiving flows being calculated differently by the extended link model, the node model processes these sending and receiving flows in the same way. Any driving behaviour parameters of the node model, e.g. priorities or signal capacities, can be trivially changed during the simulation. We will therefore not discuss node models any further in this chapter. The disaggregation of traffic required for step 3 of Algorithm 5.1 can be modelled exogenously by means of pre-specified turning fractions or endogenously with a multi-commodity LTM. Appendix 5.B describes how to add support for multiple commodities to our extended LTM.

As mentioned in the introduction, the LTM can be used with either discrete or continuous time. In case of discrete time, the Δt of each node is its constant time step size. In case of continuous time, each Δt is variable and determined ad hoc. The use of continuous time requires the FD to be piecewise-linear or to be discretised as such (Bliemer and Raadsen, 2018). Since the FD shape we choose in Subsection 5.3.2 below is generally not piecewise-linear, we opt for the discrete-time approach in this chapter. For each time t^* one wishes to change the FD, we require that t^* is an integer multiple of both Δt_{x_0} and Δt_{x_L}, because the link model assumes that the FD does not change within any time step.

5.3   Link modelling prerequisites

In this section, we discuss the mathematical prerequisites for deriving the (extended) LTM link model, namely the used traffic flow theory and FD. The LWR traffic flow theory is introduced in Subsection 5.3.1. Our chosen FD shape is introduced in Subsection 5.3.2.

Finally, Subsection 5.3.3 explains the types of shocks and fans that can occur in the solution when applying the aforementioned LWR theory and FD shape, and the impact of shocks on solution methods.

5.3.1   Lighthill-Whitham-Richards traffic flow theory

As indicated in the introduction, we develop our model for the propagation of traffic on a link following LWR traffic flow theory. LWR theory traditionally is built on conservation of vehicles in the form of a scalar conservation law

$$\frac{\partial k(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0 \, , \tag{5.3}$$

where x indicates the position along the link, t indicates time, $k(x,t)$ describes the density of traffic and $q(x,t)$ describes the flow of traffic. From there, LWR theory proceeds by substituting a continuous FD of traffic flow $Q(k)$, mapping any density k to a corresponding flow q:

$$\frac{\partial k(x,t)}{\partial t} + \frac{dQ(k)}{dk}\frac{\partial k(x,t)}{\partial x} = 0, \quad q(x,t) = Q(k(x,t)) \, . \tag{5.4}$$

Using the Lax (1957) shock admissibility or entropy condition, it has a unique solution that can be constructed using the method of characteristics. While the FD is constant, each characteristic has a constant wave speed $dx/dt = dQ(k)/dk$ and carries a constant traffic state $(k, Q(k))$. We remark it is sufficient for $dQ(k)/dk$ to exist only almost everywhere, by treating any problem with $k \rightarrow Q(k)$ with sharp corners as a limiting case of the problem with a similar function with smooth corners. This results in an effectively multi-valued $dQ(k)/dk$ at previous jumps.

Differential Equation (5.4) can also be reformulated into the Hamilton-Jacobi equation

$$\frac{\partial N(x,t)}{\partial t} - Q\left(-\frac{\partial N(x,t)}{\partial x}\right) = 0, \quad q(x,t) = \frac{\partial N(x,t)}{\partial t}, \quad k(x,t) = -\frac{\partial N(x,t)}{\partial x} \, , \tag{5.5}$$

which preserves both the FD and conservation of vehicles (Newell, 1993). Its solution $N(x,t)$ is continuous, but is continuously-differentiable only almost everywhere, matching the solution of Equation (5.4). Along each characteristic, we have

$$\partial N = Q(k)\partial t - k\partial x \, . \tag{5.6}$$

Jumps in $q(x,t)$ or $k(x,t)$ are permitted along contact discontinuities, which have characteristics run in parallel on both sides, or along shocks, which absorb characteristics from one or both sides as time t progresses. The solution may contain rarefaction waves, also known as fans, when characteristics diverge from a point on an initial or boundary condition (Evans, 2002).

5.3.2 Smulders fundamental diagram

To solve the propagation of traffic along a link based on LWR theory, we need a specification of the FD. As suggested in the introduction, we will restrict ourselves to a specific shape to avoid introducing numerical errors within the link model itself. We introduce the FD shape here, and explain the avoidance of numerical errors in Subsection 5.4.1.

The FD shape we choose is the Smulders (1990, p. 117) FD. Its congested branch is linear in the flow-density plane, while its free-flow branch is linear in the speed-density plane, thus parabolic in the flow-density plane. It can be uniquely characterised by a free speed u_F, a critical speed $u_C \leq u_F$, a capacity q_C and a jam density k_J. An example diagram is shown in Figure 5.5. Note that setting $u_C = u_F$ yields the triangular FD as a special case of the Smulders FD, and hence is also supported.

One special feature we add to this FD is that we extend its congested branch with zero flow past the jam density:

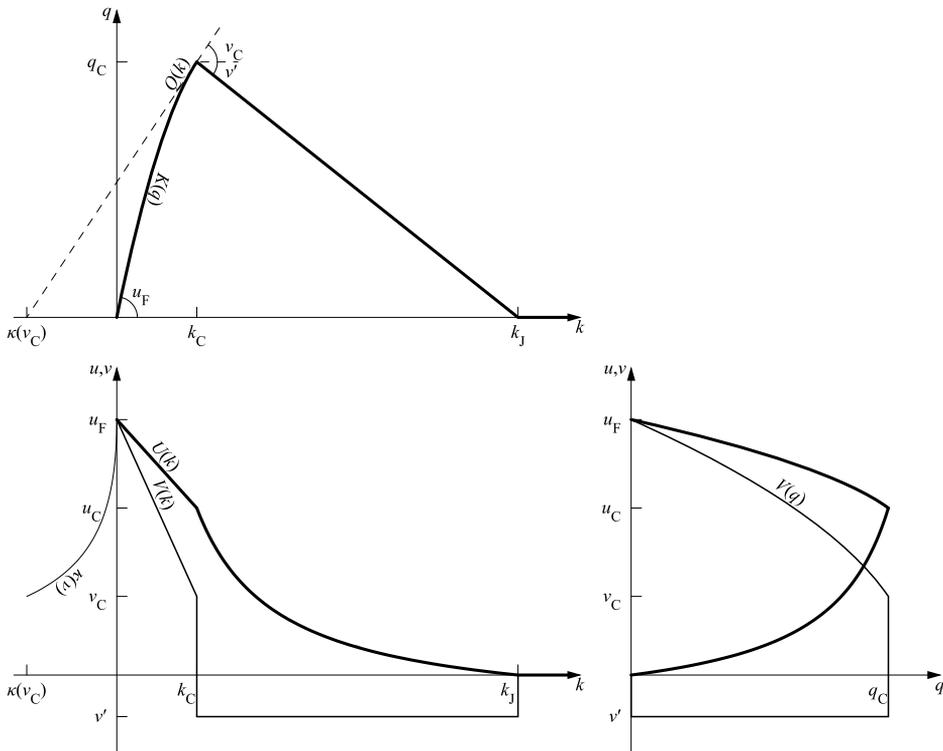$$\forall k \geq k_J : Q(k) = 0. \tag{5.7}$$

Figure 5.5. Smulders fundamental diagram (thick solid lines) and corresponding wave speeds (medium solid lines), including the extension beyond jam density.

Although densities exceeding jam density cannot occur in the solution if such densities are not present in the initial condition (Daganzo, 2005a, p. 192), this may be unavoidable when the FD changes during a simulation and the current traffic on the link is to be preserved, for example when a lane is closed while the link is congested.

The formulas for speed $U(k)$ and flow $Q(k)$ as functions of density k are

$$U(k) = \begin{cases} u_F - \dfrac{u_F - u_C}{k_C} k & \text{if } k \le k_C \\ v'\left(1 - \dfrac{k_J}{k}\right) & \text{if } k_C \le k \le k_J \\ 0 & \text{if } k_J \le k \end{cases}, \tag{5.8}$$

$$Q(k) = kU(k) = \begin{cases} u_F k - \dfrac{u_F - u_C}{k_C} k^2 & \text{if } k \le k_C \\ v'(k - k_J) & \text{if } k_C \le k \le k_J \\ 0 & \text{if } k_J \le k \end{cases}$$

where $k_C = q_C / u_C$ is the critical density. For free-flow traffic states ($k \le k_C$), the wave speed $V(k)$ as a function of density k is

$$V(k) = \frac{dQ(k)}{dk} = u_F - 2\frac{u_F - u_C}{k_C} k . \tag{5.9}$$

After some algebraic transformations of the formulas for $Q(k)$ and $V(k)$ for the free-flow branch, we find

$$K(q) = k \big|_{k:Q(k)=q \wedge k < k_C} = \begin{cases} \dfrac{q}{u_F} & \text{if } u_F = u_C \\ \dfrac{1}{2}\dfrac{k_C}{u_F - u_C}\left(u_F - \sqrt{u_F^2 - 4\dfrac{u_F - u_C}{k_C} q}\right) & \text{otherwise} \end{cases} \tag{5.10}$$

$$V(q) = V(K(q)) = \sqrt{u_F^2 - 4\frac{u_F - u_C}{k_C} q}$$

describing the density $K(q)$ and wave speed $V(q)$ as functions of flow q. We abbreviate the critical wave speed $v_C = V(k_C) = V(q_C) = 2u_C - u_F$. The intersection $\kappa(v) \le 0$ of a tangent line of wave speed $v \in [v_C, u_F]$ to the free-flow branch in the flow-density plane with the density axis is

$$\kappa(v) = K(q) - \frac{q}{v}\bigg|_{q:V(q)=v} = \begin{cases} 0 & \text{if } u_F = u_C \\ \dfrac{1}{4}\dfrac{k_C}{u_F - u_C}\left(2u_F - \dfrac{u_F^2}{v} - v\right) & \text{otherwise} \end{cases}. \tag{5.11}$$

5.3.3 Types of shocks and fans

Before proceeding to develop the LTM link model using the LWR theory and FD shape described above, it is helpful to understand what types of shocks and fans may occur in the analytical solutions of links. This subsection explains the possible types of shocks and fans, as

well as the impact of shocks on solution methods. Each type is illustrated with an example Riemann problem.

In most traffic flow applications, $k \to Q(k)$ is concave. With a concave FD in traffic flow, rarefaction waves occur when traffic accelerates (acceleration fans) whereas shocks occur when traffic decelerates (deceleration shocks). This is illustrated in Figure 5.6 with the solutions (densities) and characteristics (arrows) of example Riemann problems.

In Hamilton-Jacobi Equation (5.5), the so-called Hamiltonian $\partial N / \partial x \to -Q(-\partial N / \partial x)$ is then convex, which allows for direct solution of any point $N(x,t)$ from initial and boundary conditions by applying variational theory (Evans, 2002; Daganzo, 2005a). Whenever there are multiple candidate characteristics from the initial and boundary conditions passing through point (x,t), variational theory tells us that the correct solution $N(x,t)$ is the minimum of all values corresponding to those candidate characteristics. This Newell-Luke minimum principle (Newell, 1993; Luke, 1972) allows us to solve $N(x,t)$ without explicitly keeping track of any shocks. In this process, variational theory allows us to include candidate characteristics that do not carry the true traffic state at the initial or boundary condition, but some other traffic state, as such inclusion will not affect the minimum $N(x,t)$.

The FD from the previous subsection is however only concave for $k < k_J$. The extension of the congested branch beyond jam density gives rise to additional types of possible shocks and fans. Specifically, the solution can now also involve deceleration fans with constant traffic
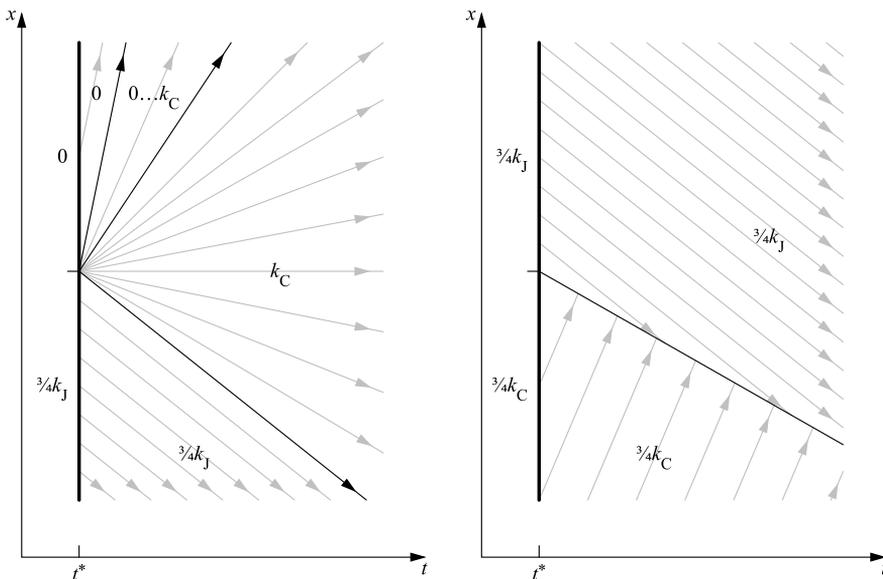

Figure 5.6. Example Riemann problems showing an acceleration fan (left) and a deceleration shock (right).
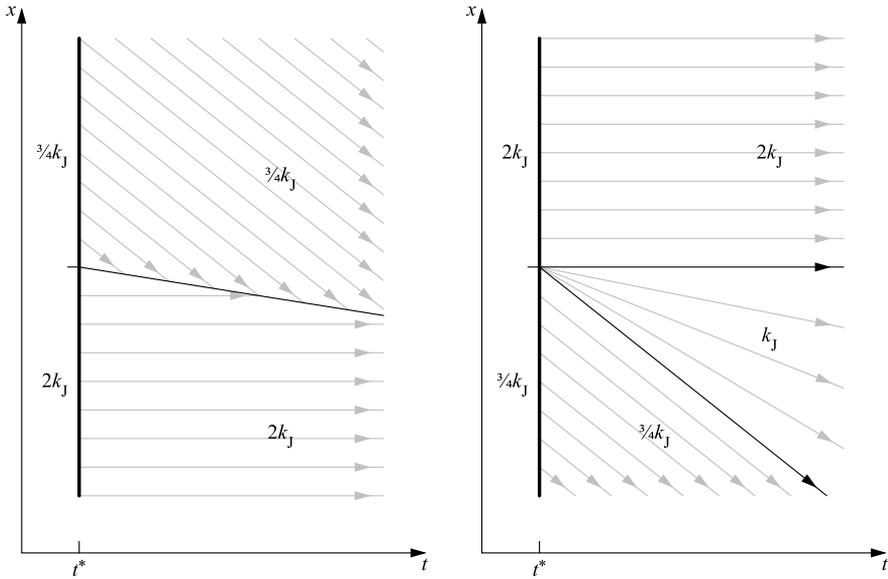
Figure 5.7. Example Riemann problems showing an acceleration shock (left) and a deceleration fan (right).
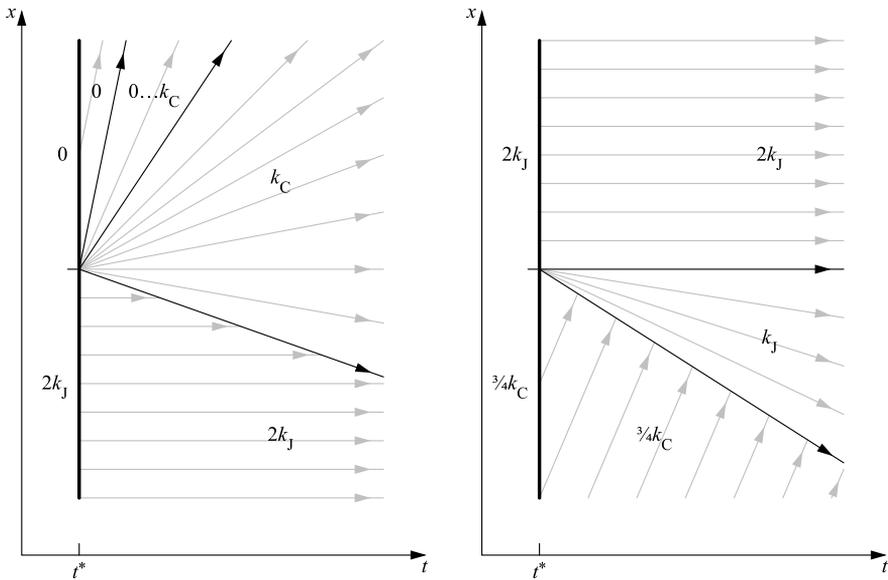

Figure 5.8. Example Riemann problems showing a combined acceleration shock and acceleration fan (left) and a combined deceleration shock and deceleration fan (right).

state $(k_J, 0)$ and acceleration shocks with shock speeds slower than the wave speed in congestion $v' = q_C / (k_C - k_J)$. Examples of such acceleration shocks and deceleration fans are illustrated in the Riemann problems of Figure 5.7. These possibilities are in addition to the acceleration fans and deceleration shocks that can normally occur in a solution with a concave FD. The solution may even involve combined shocks and fans, as illustrated in Figure 5.8.

The introduction of acceleration shocks has an important consequence for solution methods based on variational theory: we can no longer apply the Newell-Luke minimum principle to resolve any conflicts between candidate characteristics. Yet, we still have that all candidate characteristics pose some sort of constraints on N. Using logical reasoning about the traffic processes beneath a pair of conflicting candidate characteristics, we can still deduce whether the candidate characteristic yielding minimum N wins the conflict or whether the one yielding maximum N wins. Specifically, in case of a potential deceleration shock, both candidate characteristics form upper bounds on N, so that the minimum N wins. Likewise, in case of a potential acceleration shock, they form lower bounds on N, so that the maximum N wins. Overall, we thus can still construct the solution using variational theory, as long as we are aware of the nature of the comparison of candidate characteristics. The link model formulation in the next section takes advantage of this.

5.4   Link model formulation

With the LWR theory and FD specification from Section 5.3, we now develop the link model in this section. This section is structured as follows. Subsection 5.4.1 discusses the shape of the boundary and initial conditions. Using these shapes, Subsection 5.4.2 outlines the general solution method for all link model components. Then, Subsection 5.4.3 derives a procedure for computing the within-link densities, Subsection 5.4.4 derives a procedure for computing the receiving flow, and Subsection 5.4.5 derives a procedure for computing the sending flow.

5.4.1   Shapes of boundary and initial conditions

Since the LTM assumes link inflows and link outflows to be constant for the duration of some chosen time windows, the upstream and downstream boundary conditions of the link consist of piecewise-constant flows, with each piece formed by the solution of the node model of the LTM. The cumulative vehicle number N will thus be piecewise-linear in t. Note that if this does not hold for the algebraic network-level solution, the LTM results will be a numerical approximation. This approximation is imposed by the structure of the LTM and also occurs in other discrete-time traffic simulation models.

The shape of initial conditions is chosen to be restricted for a consistency reason. The initial density profile of a link must be piecewise-linear in x (that is, along the link), with permitted jumps. Thus, in combination with the Smulders FD, in general the initial cumulative vehicle number N will be piecewise-parabolic in x. The pieces need not have equal spatial length. Note that piecewise-constant, constant and zero density profiles are special cases of the piecewise-linear density profile.

possible shapes of within-link traffic state profiles that can be produced through simulation with the Smulders FD shape and piecewise-constant boundary flows. This consistency is proven in Lemma 5.1 and Theorem 5.1 below, provided that a minimum amount of time has

passed since the initial condition. We use $L = x_L - x_0$ to denote link length, i.e. the distance from the upstream end x_0 to the downstream end x_L.

Lemma 5.1. Along a characteristic with wave speed $\partial x / \partial t \in [v_C, u_F]$, it holds that $\partial N = -\partial x \kappa (\partial x / \partial t)$. This expression is parabolic in ∂x.

Proof. Rearranging Equation (5.6), we find

$$\partial N = Q(k) \partial t - k \partial x = \left(Q(k) \frac{\partial t}{\partial x} - k \right) \partial x = -\left(k - \frac{Q(k)}{\partial x / \partial t} \right) \partial x. \tag{5.12}$$

To be a characteristic, traffic state $(k, Q(k))$ must match with wave speed $\partial x / \partial t$. Assuming $(k, Q(k))$ is a free-flow traffic state ($k \leq k_C$) with corresponding wave speed $\partial x / \partial t \in [v_C, u_F]$, the factor between parentheses in Equation (5.12) equals the definition of $\kappa(v)$ in Equation (5.11). After substitution, we get

$$\partial N = -\partial x \kappa \left(\frac{\partial x}{\partial t} \right) = \frac{1}{4} \frac{u_F^{\ 2} k_C}{u_F - u_C} \partial t - \frac{1}{2} \frac{u_F k_C}{u_F - u_C} \partial x + \frac{1}{4} \frac{k_C}{u_F - u_C} \frac{\partial x^2}{\partial t} \tag{5.13}$$

which is parabolic in ∂x. ∎

Theorem 5.1. With N piecewise-parabolic in x at the initial condition, and N piecewise-linear in t at the upstream and downstream boundary conditions, L/v_C or more time later, N is again piecewise-parabolic in x.

Proof. All traffic states within the link at the target time must have been carried there by characteristics originating from the initial and boundary conditions. Characteristics emanated from the downstream boundary condition result in N being piecewise-linear in x, since N is piecewise-linear in t at the boundary and no rarefaction waves are possible. Characteristics emanated from the initial condition cannot result in traffic states with $k < k_C$ as those characteristics leave the downstream link end within L/v_C time. Because of the shape of the FD for $k \geq k_C$, rarefaction waves will carry a constant traffic state of either (k_C, q_C) or $(k_J, 0)$, resulting in N being linear in x. Other characteristics will simply translate the initial condition, which has N piecewise-parabolic in x. Finally, characteristics emanated from the upstream boundary condition will carry constant traffic states that result in N piecewise-linear in x, unless they are part of rarefaction waves that, based on Lemma 5.1, result in N parabolic in x (e.g. Figure 5.9 shows $k = -\partial N / \partial x$ linear in x). Combining all pieces from all characteristics completes the proof. ∎

Theorem 5.1 implies we can use the same procedure for applying initial conditions at the start of a simulation as we can for changing the FD, which involves computing and preserving the current within-link density profile. Thus, we can and will develop an exact procedure for this computation, so that no (new) numerical errors are introduced. It implies a minimum time of L/v_C between the last application of an initial condition to the link and the computation of
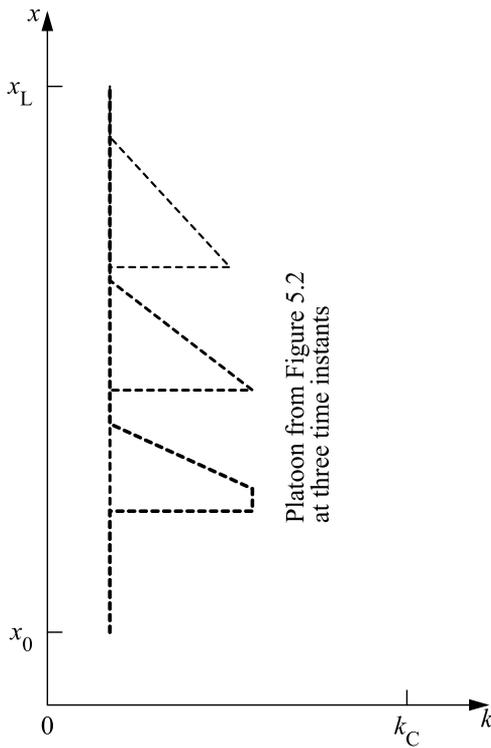
Figure 5.9. Density k **versus location** x **at various times for the platoon example from Figure 5.2.**

within-link densities. In practice this will pose a limit on e.g. how frequently the FD of a link can be changed during a simulation. If needed, this limitation can always be circumvented by splitting long links into shorter ones.

The above piecewise descriptions of the boundary and initial conditions are put in mathematical form in Table 5.1. Each boundary condition piece AB is specified by a start time t_A, an end time t_B, a start cumulative N_A, and an end cumulative N_B. Each initial condition piece AB is specified by an upstream location x_A, a downstream location x_B, an upstream cumulative N_A, a downstream cumulative N_B, and a second derivative of the cumulative over space f_{AB}. Because of Theorem 5.1, computed within-link density profiles can be described using the same piecewise formulation as the initial condition.

Table 5.1. Piecewise descriptions of the boundary and initial conditions.

	Downstream boundary condition	Initial condition	Upstream boundary condition
Space-time illustration			
Space-time domain	$\{x_L\} \times [t^*, t]$	$[x_0, x_L] \times \{t^*\}$	$\{x_0\} \times [t^*, t]$
Piece parameters	t_A, t_B, N_A, N_B	$x_A, x_B, N_A, N_B, f_{AB}$	t_A, t_B, N_A, N_B
Piece space-time illustration			
Piece space-time domain	$\{x_L\} \times [t_A, t_B]$	$[x_A, x_B] \times \{t^*\}$	$\{x_0\} \times [t_A, t_B]$
Piece cumulative N	$N(x_L, \tau) = N_{AB}(\tau)$ $= \frac{(t_B - \tau)N_A + (\tau - t_A)N_B}{t_B - t_A}$	$N(x, t^*) = N_{AB}(x)$ $= \frac{(x_B - x)N_A + (x - x_A)N_B}{x_B - x_A}$ $- f_{AB}\frac{(x_B - x)(x - x_A)}{2}$	$N(x_0, \tau) = N_{AB}(\tau)$ $= \frac{(t_B - \tau)N_A + (\tau - t_A)N_B}{t_B - t_A}$
Piece flow $q = \partial N / \partial t$	$q(x_L, \tau) = q_{AB} = \frac{N_B - N_A}{t_B - t_A}$	$q(x, t^*) = Q\big(k(x, t^*)\big)$	$q(x_0, \tau) = q_{AB} = \frac{N_B - N_A}{t_B - t_A}$
Piece density $k = -\partial N / \partial x$	$k(x_L, \tau) \overset{\text{assumed}}{=} k_J + \frac{q_{AB}}{v'}$	$k(x, t^*) = \frac{N_A - N_B}{x_B - x_A} + f_{AB}\frac{(x_B - x) - (x - x_A)}{2}$ $k_{A^+} = \lim_{x \downarrow x_A} k(x, t^*)$ $= \frac{N_A - N_B}{x_B - x_A} + f_{AB}\frac{x_B - x_A}{2}$ $k_{B^-} = \lim_{x \uparrow x_B} k(x, t^*)$ $= \frac{N_A - N_B}{x_B - x_A} - f_{AB}\frac{x_B - x_A}{2}$	$k(x_0, \tau) \overset{\text{assumed}}{=} K(q_{AB})$
Piece second spatial derivative $-\partial k / \partial x = \partial^2 N / \partial x^2$	$-\frac{\partial k}{\partial x}(x_L, \tau) = 0$	$-\frac{\partial k}{\partial x}(x, t^*) = f_{AB}$	$-\frac{\partial k}{\partial x}(x_0, \tau) = 0$

5.4.2   General solution method

Now that the shapes of the boundary and initial conditions and within-link density profiles have all been determined, we proceed with developing algorithms for the three link model components. These components are responsible for the computation of the within-link density

profile, the computation of constraints on future link inflow that determine the receiving flow, and the computation of constraints on future link outflow that determine the sending flow. All three components follow the same general solution procedure, which can be summarised as follows:

1. Iterate over the pieces of the boundary and initial conditions in counter-clockwise direction in the space-time plane in , starting with the downstream boundary.

2. Analyse how candidate characteristics emanated from the current piece or point affect the solution.

3. Apply the relevant candidate characteristics to the solution, resolving any conflicts with previous candidate characteristics in a pre-specified way.

If all candidate characteristics relevant to the solution are considered, all relevant real characteristics must be among the candidates. If the conflicts between candidate characteristics are also resolved correctly, the eventually resulting candidate characteristics must be real characteristics and our resulting solution must thus be correct. Because the FD from Subsection 5.3.2 is not entirely concave in the flow-density plane, we explicitly consider the nature of the candidate characteristics we are comparing as proposed in Subsection 5.3.3.

Note that the conflict resolution step looks for conflicts between two sets of candidate characteristics: the previous candidates and the new candidates. The output is a new, merged set of candidate characteristics with conflicts resolved. In this resolution process we not only remove multivaluedness of N in the candidate solution but also maintain continuity of N in the candidate solution by eliminating crossing candidate characteristics. This is permitted because characteristics cannot cross, and greatly simplifies the resolution process. Namely, in case of a conflict, the merging of both input sets can only introduce one new shock in the output set, with candidate characteristics from each one input set on one side of the new shock. Only if one input set fully dominates the other, the output set equals this input set and no new shock is introduced.

To ensure all candidate characteristics possibly affecting the link are taken into account, the traffic states on the downstream boundary are assumed to be on the congested branch of the FD and the traffic states of the upstream boundary are assumed to be on the free-flow branch of the FD. These assumptions are indicated in Table 5.1 and are based on the need for candidate characteristics to enter the link. The kind of candidate characteristics emanated from a piece of the initial condition depend on whether the density is lower than the critical density (Free-flow), between the critical density and jam density (Congested), or exceeding the jam density (Jammed). We will treat these cases separately. Therefore, ensure that all the pieces AB satisfy

$$\underbrace{\forall x \in (x_A, x_B): k(x, t^*) \leq k_C}_{\text{type Free-flow}} \quad \vee \quad \underbrace{\forall x \in (x_A, x_B): k_C \leq k(x, t^*) \leq k_J}_{\text{type Congested}} \quad \vee \quad \underbrace{\forall x \in (x_A, x_B): k_J \leq k(x, t^*)}_{\text{type Jammed}} \quad (5.14)$$

by splitting unclassifiable pieces into smaller pieces wherever the density passes k_C or k_J.

In the following three subsections, each component is discussed individually.

5.4.3   Computing the within-link density profile

We start in this subsection with the computation of within-link densities. Although the proof of Theorem 5.1 already hints at which candidate characteristics to take into account to compute the within-link densities, it does not yet offer formulas for them nor specifies how to resolve conflicts between characteristics that result in multi-valued N. We will now develop the method in detail, calculating the within-link density profile at (current) time t, assuming a time difference with the initial condition of at least L/v_C in accordance with Theorem 5.1. The output consists of an exact, piecewise description of $N(x,t)$ on $[x_0, x_L] \times \{t\}$, without any discretisation of space within the link, in the same piecewise format as the initial condition.

All candidate characteristics relevant to the computation of the within-link density profile are described in Table 5.2. Each row represents a class of relevant candidate characteristics emanating from different pieces and points of the boundary and initial conditions. Space-time visualisations of how these characteristics emanate from the boundary and initial conditions are provided in the first column, including the naming of points used later in the row.

The second column gives a mathematical description of the space-time origins of the candidate characteristics as well as their destinations at time t when the within-link density profile is calculated. These destination sets are intersected with $[x_0, x_L] \times \{t\}$ to eliminate candidate characteristics that have left the link before time t and hence do not affect the within-link densities at that time.
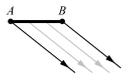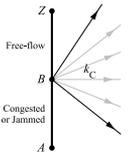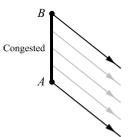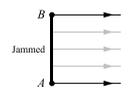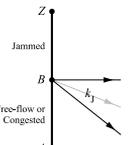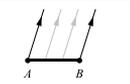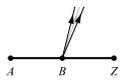
For the remaining candidate characteristics, the third column gives the formulas for the corresponding values of cumulative N at their destinations, obtained by integrating N over the candidate characteristic. The formula for its second spatial derivative $\partial^2 N / \partial x^2 = -\partial k / \partial x$ is also listed; due to Theorem 5.1 its value is constant. Together, the spatial endpoints of the destinations in the second column and the formulas in the third column allow construction of a candidate piece of the piecewise-parabolic solution.

Finally, the fourth column indicates how any conflicts with previously found candidate pieces are to be resolved, assuming candidate characteristics are considered in counter-clockwise order as described in Subsection 5.4.2:

- "N/A" means that no conflicts are possible;
- "min" means that the candidate characteristics yielding minimum N should be kept;
- "max" means that the candidate characteristics yielding maximum N should be kept.

If there are conflicts to be resolved, the "↑" symbol indicates that the new candidate characteristics corresponding to this row of the table must be on the upstream side of a newly created shock. This specifies the searching direction for the possible shock through the previous candidate pieces. The location of any shock between the new candidate piece and a previous candidate piece between their spatial endpoints is to be determined with the quadratic formula. If a shock is found, the previous piecewise solution is truncated to the shock location before the new (truncated) piece is added to the solution. Appendix 5.C.1 describes this process in more detail.

Table 5.2. Candidate characteristics affecting the within-link density profile.

Space-time illustration	Candidate characteristics beginnings and endings	Ending cumulative N and second spatial derivative $-\partial k/\partial x = \partial^2 N/\partial x^2$	Con-flict reso-lution
[1] Piece of downstream boundary condition:			
	$\{x_L\}\times[t_A,t_B]\to$ $\begin{bmatrix}\max\left(x_0,x_L+(t-t_A)v'\right),\\ x_L+(t-t_B)v'\end{bmatrix}\times\{t\}$	$N(x,t)=N\left(x_L,t+\frac{x_L-x}{v'}\right)+(x_L-x)k_J$ $-\frac{\partial k}{\partial x}(x,t)=0$	N/A
[2] Consecutive pieces of initial condition, type Congested or Jammed upstream of type Free-flow:			
	$\{x_B\}\times\{t^*\}\to$ $\begin{bmatrix}\max\left(x_0,x_B+(t-t^*)v'\right),\\ x_L\end{bmatrix}\times\{t\}$	$N(x,t)=N_B+(t-t^*)q_C-(x-x_B)k_C$ $-\frac{\partial k}{\partial x}(x,t)=0$	$\overline{\wedge}$ min
[3] Piece of initial condition, type Congested:			
	$[x_A,x_B]\times\{t^*\}\to$ $\begin{bmatrix}\max\left(x_0,x_A+(t-t^*)v'\right),\\ x_B+(t-t^*)v'\end{bmatrix}\times\{t\}$	$N(x,t)=N\left(x-(t-t^*)v',t^*\right)$ $+(t-t^*)v'k_J$ $-\frac{\partial k}{\partial x}(x,t)=f_{AB}$	N/A
[4] Piece of initial condition, type Jammed:			
	$[x_A,x_B]\times\{t^*\}\to$ $[x_A,x_B]\times\{t\}$	$N(x,t)=N(x,t^*)$ $-\frac{\partial k}{\partial x}(x,t)=f_{AB}$	$\overline{\wedge}$ max
[5] Consecutive pieces of initial condition, type Free-flow or Congested upstream of type Jammed:			
	$\{x_B\}\times\{t^*\}\to$ $\begin{bmatrix}\max\left(x_0,x_B+(t-t^*)v'\right),\\ x_B\end{bmatrix}\times\{t\}$	$N(x,t)=N_B+(x-x_B)k_J$ $-\frac{\partial k}{\partial x}(x,t)=0$	$\overline{\wedge}$ max
[6] Piece of upstream boundary condition:			
	$\{x_0\}\times[t_A,t_B]\to$ $\begin{bmatrix}x_0+(t-t_B)V(q_{AB}),\\ \min\left(x_0+(t-t_A)V(q_{AB}),x_L\right)\end{bmatrix}\times\{t\}$	$N(x,t)=\frac{(t_B-t)N_A+(t-t_A)N_B}{t_B-t_A}$ $-(x-x_0)K(q_{AB})$ $-\frac{\partial k}{\partial x}(x,t)=0$	$\overline{\wedge}$ min
[7] Consecutive pieces of upstream boundary condition:			
	$\{x_0\}\times\{t_B\}\to$ $\begin{bmatrix}x_0+(t-t_B)V(q_{BZ}),\\ \min\left(x_0+(t-t_B)V(q_{AB}),x_L\right)\end{bmatrix}\times\{t\}$	$N(x,t)=N_B-(x-x_0)\kappa\left(\frac{x-x_0}{t-t_B}\right)$ $-\frac{\partial k}{\partial x}(x,t)=\frac{1}{2}\frac{k_C}{u_F-u_C}(t-t_B)^{-1}$	$\overline{\wedge}$ min

We will now describe each row of Table 5.2:

[1] Each piece of the downstream boundary condition is assumed to be congested and hence generates candidate characteristics with wave speed v', representing a constant congested traffic state. Because all these characteristics run in parallel, no conflict occurs.

[2] Characteristics from pieces of the initial condition of type Free-flow can no longer affect the link interior at time t because they reach the downstream link end before that time. However, if such a piece has another piece of type Congested or Jammed upstream of it, the acceleration fan emanated from between these pieces can affect the solution. This fan carries the capacity state (k_c, q_c). Its upstream end is limited by wave speed v', while its downstream end can always reach the downstream link end. If the fan conflicts with any previous candidate characteristics, the fan represents an upper bound on the flow into the congestion ($k \geq k_c$) represented by the previous candidate characteristics. A conflict is thus resolved by taking minimum N, possibly resulting in a deceleration shock.

[3] A piece of the initial condition of type Congested emanates candidate characteristics with wave speed v', translating the traffic states from the initial condition and therefore preserving the second spatial derivative. They run parallel to previous candidate characteristics and hence no conflicts occur.

[4] A piece of the initial condition of type Jammed emanates candidate characteristics with zero speed. If these conflict with previous candidate characteristics, the previous candidates represent the outflow ($k \leq k_J$) from this jam. A conflict is thus resolved by taking maximum N, possibly resulting in an acceleration shock.

[5] If a piece of the initial condition of type Jammed has another piece of type Free-flow or Congested upstream of it, a deceleration fan is emanated from between the pieces. This fan carries the jammed state $(k_J, 0)$. Its wave speeds range from v' to zero. Like the previous row, a conflict is resolved by taking maximum N, possibly resulting in an acceleration shock.

[6] Each piece of the upstream boundary condition is assumed to be free-flow and hence generates candidate characteristics representing a constant free-flow traffic state, with the wave speed from Equation (5.10). They form an upper bound on flow into downstream traffic states in case of conflict, which is thus resolved by taking minimum N, possibly resulting in a deceleration shock.

[7] Between two consecutive pieces of the upstream boundary condition, an acceleration fan can be emanated. and Eq. determine the cumulative and second spatial derivative at its destination. Like the previous row, a conflict is resolved by taking minimum N, possibly resulting in a deceleration shock.

5.4.4 Computing the receiving flow

The receiving flow is the maximum flow that can enter the link at x_0. At the start of the simulation, the downstream boundary condition is unknown, thus it is clearly not possible to determine all link inflow constraints at once. We shall assume that the downstream boundary

condition is only known up to the current time t, which, by the CFL condition, suffices to compute any constraints on future cumulative inflow until $t - L/v'$. Equation (5.1) converts these cumulative inflow constraints into the receiving flow.
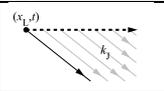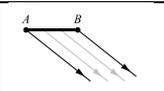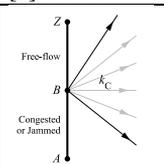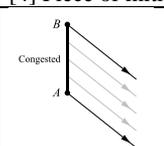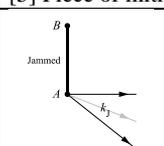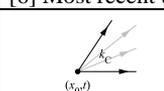
A main feature of the computation of the receiving flow in this subsection and the sending flow in the next subsection is that they consider the implications of the initial and boundary conditions for all future cumulative inflow and outflow simultaneously, instead of looking at only one time step at a time. This means that the bulk of the work in Algorithm 5.1 is done in steps 5 and 6 rather than steps 1 and 2, similar to Gentile and Papola (2009b) and Gentile (2010). This increases the computational efficiency compared to Section 4.3 because it avoids the need to iterate over all possible wave speeds every time step.

Similar to Subsection 5.4.3, Table 5.3 lists all candidate characteristics relevant for future inflow constraints. The first column provides a space-time visualisation of the characteristics. The second column indicates the space-time origins and destinations of the characteristics. The third column gives formulas for the destination cumulative N and its second time derivative $\partial^2 N / \partial t^2 = \partial q / \partial t$. Because the second time derivative is always constant, it is possible to compute an exact, piecewise-parabolic description of the cumulative inflow constraints. However, since the calculation of receiving flows only needs to know the cumulative inflow constraint at a finite set of times, it is more convenient to directly calculate the cumulative inflow constraint at these specific times only. The fourth column indicates whether to keep the minimum or maximum N in case of conflicts. The "$\rightarrowtail$" symbol specifies the searching direction for shocks, indicating that the new candidate characteristics corresponding to this row will be on the before side of any shock with previous candidate characteristics (see Appendix 5.C).

We will now describe each row of Table 5.3:

[1] Assuming the link outflow after current time t would be zero, the assumed future downstream boundary condition generates candidate characteristics with wave speed v' carrying traffic state $(k_J, 0)$. Because of satisfaction of the CFL condition, this assumption does not affect the relevant part of the solution ($t + \Delta t_{x_0} \leq t - L/v'$) and applying these candidate characteristics is thus entirely optional. We do include this row for two reasons. First, it makes it easier to update the cumulative inflow constraints when new pieces of the downstream boundary condition become known, without redoing the entire computation. This is discussed further below. Secondly, it ensures the cumulative inflow constraints always end in constant N (row [1] or row [5]), allowing the constraints applying to an infinite future to be more easily stored in finite computer memory.

[2] Each piece of the downstream boundary condition produces congested traffic states with wave speed v'. No conflict occurs.

[3] Characteristics from between pieces of the initial conditions of type Congested or Jammed and type Free-flow form a fan carrying capacity state (k_C, q_C), between wave speeds v' and v_C. In case of conflict with previous candidate characteristics, this fan represents inflow into previous congestion ($k \geq k_C$). A conflict is thus resolved by taking minimum N.

Table 5.3. Candidate characteristics affecting future link inflow.

Space-time illustration	Candidate characteristics beginnings and endings	Ending cumulative N and second time derivative $\partial q/\partial t = \partial^2 N/\partial t^2$	Conflict reso-lution
[1] Most recent downstream point:			
	$\{x_L\} \times [t,\infty) \rightarrow$ $\{x_0\} \times \left[t - \frac{L}{v'},\infty\right)$	$N(x_0,\tau) = N(x_L,t) + Lk_J$ $\frac{\partial q}{\partial \tau}(x_0,\tau) = 0$	N/A
[2] Piece of downstream boundary condition:			
	$\{x_L\} \times [t_A,t_B] \rightarrow$ $\{x_0\} \times \left[t_A - \frac{L}{v'}, \; t_B - \frac{L}{v'}\right]$	$N(x_0,\tau) = N\left(x_L, \tau + \frac{L}{v'}\right) + Lk_J$ $\frac{\partial q}{\partial \tau}(x_0,\tau) = 0$	N/A
[3] Consecutive pieces of initial condition, type Congested or Jammed upstream type Free-flow:			
	$\{x_B\} \times \{t^*\} \rightarrow$ $\{x_0\} \times \left[t^* - \frac{x_B - x_0}{v'},\infty\right)$	$N(x_0,\tau) = N_B + (\tau - t^*)q_C + (x_B - x_0)k_C$ $\frac{\partial q}{\partial \tau}(x_0,\tau) = 0$	$\rightarrowtail$ min
[4] Piece of initial condition, type Congested:			
	$[x_A, x_B] \times \{t^*\} \rightarrow$ $\{x_0\} \times \left[t^* - \frac{x_A - x_0}{v'}, \; t^* - \frac{x_B - x_0}{v'}\right]$	$N(x_0,\tau) = N\left(x_0 - (\tau - t^*)v', t^*\right) - (\tau - t^*)v'k_J$ $\frac{\partial q}{\partial \tau}(x_0,\tau) = v'^2 f_{AB}$	N/A
[5] Piece of initial condition, type Jammed:			
	$\{x_A\} \times \{t^*\} \rightarrow$ $\{x_0\} \times \left[t^* - \frac{x_A - x_0}{v'},\infty\right)$	$N(x_0,\tau) = N_A + (x_A - x_0)k_J$ $\frac{\partial q}{\partial \tau}(x_0,\tau) = 0$	$\rightarrowtail$ max
[6] Most recent upstream point:			
	$\{x_0\} \times \{t\} \rightarrow$ $\{x_0\} \times [t,\infty)$	$N(x_0,\tau) = N(x_0,t) + (\tau - t)q_C$ $\frac{\partial q}{\partial \tau}(x_0,\tau) = 0$	$\rightarrowtail$ min

[4] Each piece of the initial condition of type Congested produces congested traffic states with wave speed v'. No conflict occurs.

[5] The upstream end of a piece of the initial condition of type Jammed generates a fan of candidate characteristics carrying traffic state $(k_J, 0)$, between wave speeds v' and zero. In case of conflict, previous characteristics represent the outflow ($k \leq k_J$) from this jam fan. A conflict is thus resolved by taking maximum N.

[6] From the last known upstream point (x_0, t), one candidate characteristic carrying capacity state (k_C, q_C) restricts future inflow. In case of conflict, this represents inflow

into previous congestion ($k \geq k_c$). A conflict is thus resolved by taking minimum N. This row is equivalent to limiting the receiving flow to $q_c \Delta t_{x_0}$.

We now discuss how the cumulative inflow constraints resulting from this computation can be updated when new pieces of the downstream and upstream boundary conditions become known in steps 5 and 6 of Algorithm 5.1. For a new piece of the upstream boundary condition, this is as simple as applying row [6] again.

Due to the previous assumption of zero flow in row [1], a new piece of the downstream boundary condition can only relax the constraints on future inflow. However, the upper bounds posed by rows [3] and [6] still apply. Based on the mathematical principle

$$\min\left(a, \max\left(b, c\right)\right) = \max\left(\min\left(a, b\right), \min\left(a, c\right)\right), \tag{5.15}$$

we can update the previously calculated cumulative inflow constraints by taking the maximum N, provided that the minimum N operations associated with rows [3] and [6] are first applied to the updated candidate characteristics from rows [1] and [2]. This does not require iterating over the initial condition on every update of the downstream boundary condition: since all constraints based on rows [3] have identical flow q_c, one is the most constraining for every update, requiring only that one to be applied. The searching direction for the update is "↤".
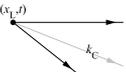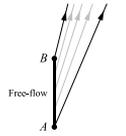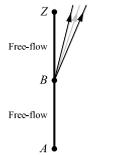
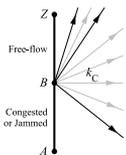### 5.4.5   Computing the sending flow

The sending flow is the maximum flow that can leave the link at x_L. At the start of the simulation, the upstream boundary condition is unknown, thus it is clearly not possible to determine all link outflow constraints at once. We shall assume that the upstream boundary condition is only known up to the current time t, which, by the CFL condition, suffices to compute any constraints on future cumulative outflow until $t + L/u_F$. Equation (5.1) converts these cumulative outflow constraints into the sending flow.

Similar to the last two subsections, Table 5.4 lists all candidate characteristics relevant for future outflow constraints. Unlike Table 5.3, the third column does not list the second time derivative. The reason for this is that it is not always constant; a piecewise-parabolic description of the cumulative outflow constraints thus cannot be constructed. Since the calculation of sending flows only needs to know the cumulative outflow constraint at a finite set of times, one directly calculates the cumulative outflow constraint at these specific times only. Because the calculation of outflow constraints only involves candidate characteristics associated with the concave free-flow branch of the FD, all conflicts are resolved by taking minimum N. The searching direction for shocks is always "↤", meaning that new candidate characteristics will be on the after side of a shock with previous candidate characteristics. An example is provided in Appendix 5.C.2.

We will now describe each row of Table 5.4:

[1] From the last known downstream point (x_L, t), one candidate characteristic carrying capacity state (k_c, q_c) restricts future outflow. This row is equivalent to limiting the sending flow to $q_c \Delta t_{x_L}$.

Table 5.4. Candidate characteristics affecting future link outflow.

Space-time illustration	Candidate characteristics beginnings and endings	Ending cumulative N	Conflict resolution
[1] Most recent downstream point:			
	$\{x_L\}\times\{t\} \rightarrow$ $\{x_L\}\times[t,\infty)$	$N(x_L,\tau)=N(x_L,t)+(\tau-t)q_C$	N/A
[2] Piece of initial condition, type Free-flow:			
	$[x_A,x_B]\times\{t^*\} \rightarrow$ $\{x_L\}\times\left[t^*+\frac{x_L-x_B}{V(k_{B^-})}, t^*+\frac{x_L-x_A}{V(k_{A^+})}\right]$	$N(x_L,\tau)=N(X(\tau),t^*)-(x_L-X(\tau))\kappa\left(\frac{x_L-X(\tau)}{\tau-t^*}\right)$ where $X(\tau)=\dfrac{x_L-(\tau-t^*)\left(u_F+\left(\frac{N_A-N_B}{x_B-x_A}+f_{AB}\frac{x_A+x_B-2x_0}{2}\right)\frac{dV(k)}{dk}\right)}{1-(\tau-t^*)f_{AB}\frac{dV(k)}{dk}}$ with $\frac{dV(k)}{dk}=-2\frac{u_F-u_C}{k_C}$	⇐ min
[3] Consecutive pieces of initial condition, type Free-flow upstream of type Free-flow:			
	$\{x_B\}\times\{t^*\} \rightarrow$ $\{x_L\}\times\left[t^*+\frac{x_L-x_B}{V(k_{B^+})}, t^*+\frac{x_L-x_B}{V(k_{B^-})}\right]$	$N(x_L,\tau)=N_B-(x_L-x_B)\kappa\left(\frac{x_L-x_B}{\tau-t^*}\right)$	⇐ min
[4] Consecutive pieces of initial condition, type Congested or Jammed upstream of type Free-flow:			
	$\{x_B\}\times\{t^*\} \rightarrow$ $\{x_L\}\times\left[t^*+\frac{x_L-x_B}{V(k_{B^+})},\infty\right)$	$N(x_L,\tau)=\begin{cases} N_B-(x_L-x_B)\kappa\left(\frac{x_L-x_B}{\tau-t^*}\right) & \text{if } \tau\leq t^*+\frac{x_L-x_B}{v_C} \\ N_B+(\tau-t^*)q_C-(x_L-x_B)k_C & \text{if } \tau\geq t^*+\frac{x_L-x_B}{v_C} \end{cases}$	⇐ min
[5] Piece of initial condition, type Free-flow, connected with piece of upstream boundary condition:			
	$\{x_0\}\times\{t^*\} \rightarrow$ $\{x_L\}\times\left[t^*+\frac{L}{V(k_{A^+})}, t^*+\frac{L}{V(q_{AZ})}\right]$	$N(x_L,\tau)=N_A-L\kappa\left(\frac{L}{\tau-t^*}\right)$	⇐ min
[6] Piece of upstream boundary condition:			
	$\{x_0\}\times[t_A,t_B] \rightarrow$ $\{x_L\}\times\left[t_A+\frac{L}{V(q_{AB})}, t_B+\frac{L}{V(q_{AB})}\right]$	$N(x_L,\tau)=\frac{(t_B-\tau)N_A+(\tau-t_A)N_B}{t_B-t_A}-LK\left(\frac{N_B-N_A}{t_B-t_A}\right)$	⇐ min
[7] Consecutive pieces of upstream boundary condition:			
	$\{x_0\}\times\{t_B\} \rightarrow$ $\{x_L\}\times\left[t_B+\frac{L}{V(q_{AB})}, t_B+\frac{L}{V(q_{BZ})}\right]$	$N(x_L,\tau)=N_B-L\kappa\left(\frac{L}{\tau-t_B}\right)$	⇐ min
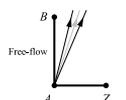
[2] A piece of the initial condition of type Free-flow emits candidate characteristics that either diverge ($\partial k / \partial x < 0$), converge ($\partial k / \partial x > 0$), or run in parallel ($\partial k / \partial x = 0$). Within the fan, the candidate characteristic ending at (x_L, τ) originates from the point $(X(\tau), t^*)$, where $X(\tau)$ is the solution of $x_L - X(\tau) = (\tau - t^*) V(k(X(\tau), t^*))$. If in case of convergence, the point of convergence is upstream of x_L, the link outflow is unaffected.

[3] A fan of candidate characteristics can be emanated from between two pieces of the initial condition of type Free-flow. If the density upstream of the point does not exceed the density downstream of the point, no fan occurs and the link outflow is unaffected.

[4] Similarly, a fan of candidate characteristics is emanated from between two pieces of the initial condition of type Congested or Jammed and type Free-flow. For the infinite part of the fan carrying constant capacity state (k_c, q_c), only the first point in time needs to be applied: row [1] will then ensure the capacity state is continued into later sending flows.

[5] A fan of candidate characteristics can be emanated from between a piece of the initial condition of type Free-flow and an adjacent piece of the upstream boundary condition. If the wave speed downstream of the point does not exceed the wave speed after the point, no fan occurs.

[6] Each piece of the upstream boundary condition generates candidate characteristic carrying a constant free-flow traffic state, with corresponding wave speed.

[7] A fan of candidate characteristics can be emanated from between two pieces of the upstream boundary condition. If the flow after the point does not exceed the flow before the point, no fan occurs.

When a new piece of the downstream boundary condition becomes available in step 5 of , row [1] can simply be applied again to update the cumulative outflow constraints. When a new piece of the upstream boundary condition becomes known in step 6 of Algorithm 5.1, one can update the constraints by applying either row [5] or [7], followed by row [6] for the new piece.

5.5   Simulation study

As a demonstration of the capabilities of the extended link model, we simulate the Dutch A13 motorway corridor from the Kleinpolderplein interchange near Rotterdam to the Ypenburg interchange near The Hague, with a length of 12 km, having five off-ramps and six on-ramps. Subsection 5.5.1 introduces the simulation setup and scenario, followed by visualisations of the results in Subsection 5.5.2 and model computation times in Subsection 5.5.3.

The raw input and output data of all simulations in this section are published by Van der Gun (2018).

5.5.1   Simulation setup and scenario

The link capacities in our simulations are based on the Dutch motorway capacity manual with a 2% truck percentage (Grontmij, 2015), reduced with 5% to correct for the absence of a capacity drop in our current model. For uncontrolled links, we assume a free speed of 110 km/h, a critical speed of 90 km/h and a wave speed in congestion of −20 km/h. The node model is taken from Tampère et al. (2011). In the simulations, traffic is disaggregated using destination-based commodities (see Appendix 5.B).

The same corridor was already simulated in Subsection 4.7.2 for various evening peaks in September 2012. In this chapter, we will focus on the evening peak of 11 September 2012. Around 15:45 that day, a crash occurred between the 10th off-ramp and on-ramp (University), just before the end of the peak hour lane. Despite the absence of injuries, the incident partly blocked the road for more than one hour until around 17:00, after which the released traffic triggers the recurrent bottleneck downstream of the 9th on-ramp (Delft). The flow and speed as measured on loop detectors just downstream of the incident location are shown in the top half of Figure 5.10.

During this evening peak, variable speed limits and dynamic lane management were applied, primarily through variable message signs (VMSs) mounted on gantries. The variable speed limits were partly set manually by the traffic management centre and partly automatically by the queue tail warning system (Dutch: filestaartbeveiliging). In both cases the purpose is to avoid unsafe situations. The opening and closing of lanes is always manual. Besides lane closures at the incident location, the entire peak hour lane was closed for the duration of the incident, presumably to keep a clear path for emergency services. Within the simulated time period there were also two temporary closures of parts of the peak hour lane and the closure of the peak hour lane at the end of the peak period. All displayed speed limits and lane closures are visualised in the bottom half of Figure 5.10. They give rise to a substantial number of FD changes throughout the simulation, which makes this scenario an interesting test case for the link model improvements developed in this chapter. In this demonstration we will simply impose the speed limits as they were displayed in reality rather than making them responsive to the congestion as it occurs in the simulation, although we emphasise the latter is also possible with our model, since it does not need to know any FDs ahead of time.

To account for the spatial detail in the traffic control measures of this scenario – each link can only have one FD at a time – we split links at VMS gantries. For lane drops, the merging area is explicitly modelled with the capacity contribution of the terminating lane equal to 20% of a normal lane. We also insert an extra link of 100 m length around the crash location, so that we can separately model the driving behaviour near the crash location. The resulting network graph consists of 48 links and 49 nodes and is depicted in Figure 5.11.

For changes in the availability of lanes, we scale the FD of a link in the flow-density plane with respect to the origin towards the capacity value corresponding to the open lanes. We use the upstream gantry to decide whether lanes are open. The variable speed limits yield the following transformation of the FD in the speed-density plane:

$$
\begin{pmatrix} k_C \\ k_J \\ u_F \\ u_C \end{pmatrix} \xrightarrow{\; VSL(\bar{u}) \;} \begin{pmatrix} \max\left(k_C, \frac{k_J}{1 - \bar{u}/v'}\right) \\ k_J \\ \bar{u} \\ \min\left(u_C, \bar{u}\right) \end{pmatrix}. \tag{5.16}
$$
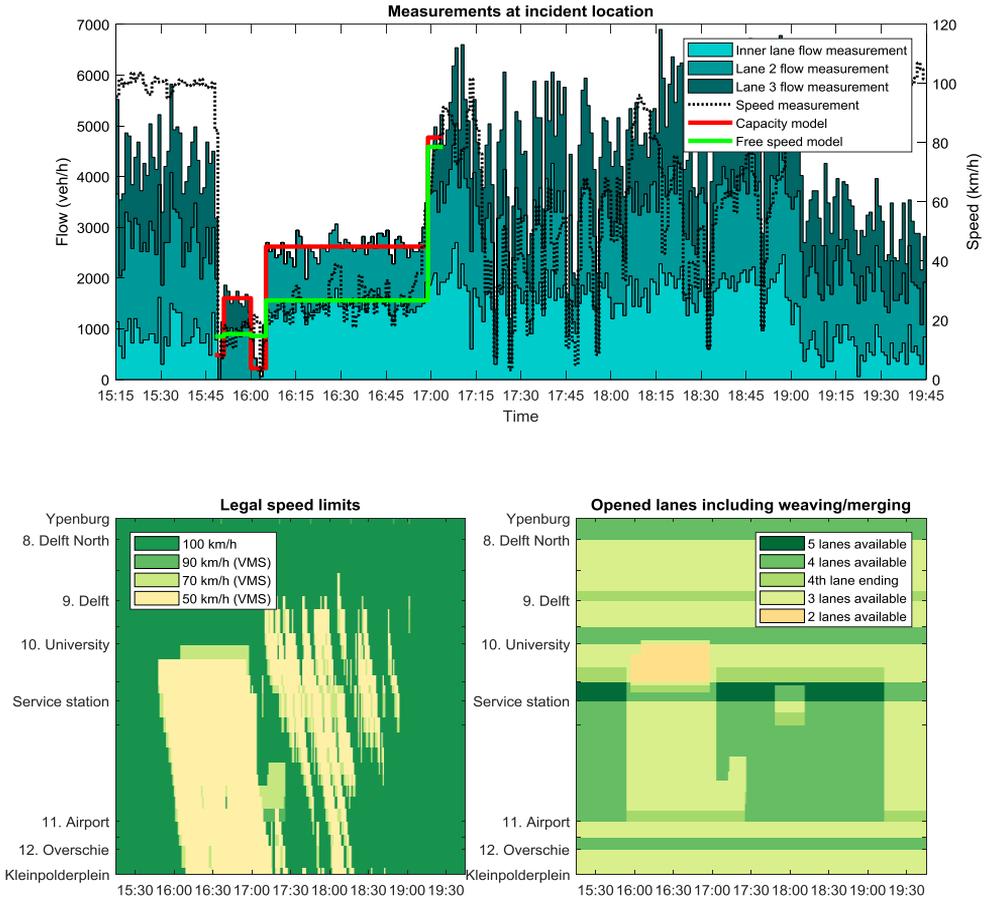
Figure 5.10. A13 flow at crash location and applied traffic control.
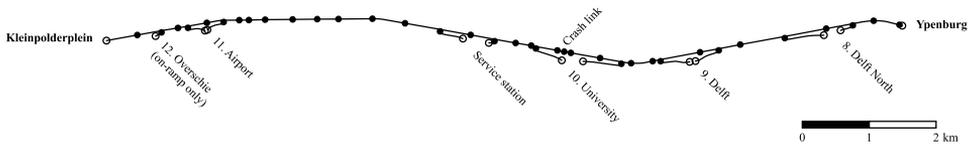
Figure 5.11. A13 corridor network.

Similar to Hegyi et al. (2005a; 2005b), for each density, this transformation reduces the possible speeds to adhere to the new speed limit $\bar{u}$, while maintaining the extended Smulders FD shape. (Because the speed limits are not used for homogenisation purposes, a possible increase in capacity as suggested by e.g. Cremer (1979) and Smulders (1990) is less important in this scenario.) To account for the relatively high non-compliance for VMS-displayed speed limits in the Netherlands, we add 20 km/h to the displayed speed limit in simulation. The

speed limits are also taken from the upstream gantry, unless there is an on-ramp between the link and that gantry and the downstream gantry displays a higher speed limit.

During the incident, the crash link uses a triangular FD with a low capacity and a low free speed set based on measurements. The used parameters are indicated in the top half of Figure 5.10. Since the wave speed in congestion is unchanged, the jam density is also low. While the crash site could alternatively be modelled more simply as a single node with a node flow constraint, similar to e.g. the ramp metering by Hajiahmadi et al. (2013), we opted for a link to test our extended link model with relatively extreme FD changes.

Our simulation study also includes non-empty initial conditions at the start of the simulation for the motorway links. To construct the initial conditions, we take the first flow measurement from each loop detector, transform those into densities for each link using the free-flow FD branch and linearly interpolate the densities between the loops.

We conduct three separate simulations of this evening peak:

- The first simulation uses the LTM with variable FDs and initial conditions as proposed in this chapter. Time step sizes differ per node based on the CFL conditions of connected links ($\Delta t \in \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{12}, \frac{1}{16}, \frac{1}{24}, \frac{1}{48}\}$ min).

- The second simulation uses the LTM with fixed FDs, i.e. no variable speed limits and an always-open peak hour lane, and without initial conditions. The capacity restriction at the crash site is modelled as a simple node flow constraint. Time step sizes again differ per node.

- The third simulation emulates the CTM solution scheme with variable FDs and initial conditions. It uses an equal time step size $\Delta t = \frac{1}{41}$ min for all nodes, based on the most constraining CFL condition. The number of cells is then maximised in each link, as this minimises the numerical error of the CTM (Van Wageningen-Kessels, 2013). The CTM solution scheme is emulated by simplifying the initial conditions to a constant density per cell, updating these after every time step using CTM rules, and starting the next time step from the updated initial conditions.

5.5.2   Visualisations of simulation results

The simulation results are visualised in Figure 5.12, which shows the flow, speed and density from the loop detector measurements along with the results from the three simulations. Note that since Dutch loop detectors only record flow and time mean speed, the space mean speed is only an estimate and hence the density, computed as $k = q/u$, may be rather inaccurate for high densities (Knoop et al., 2009). Because the LTM only computes traffic states at the boundaries of space-time rectangles as in Figure 5.3, the simulation results are further post-processed for the visualisation.

With the non-empty initial conditions, we see that the LTM no longer needs a warm-up period before providing realistic results. After the onset of the incident, the closing of the peak hour lane turns out to be important to reproduce the queue spillback – with fixed FDs, the queue density is too high and the spillback is too slow. The speed limits from the queue tail warning result in a more gradual speed reduction than just one shock, but, in accordance with Equation (5.16), do not affect traffic already driving slower than the new free speed.
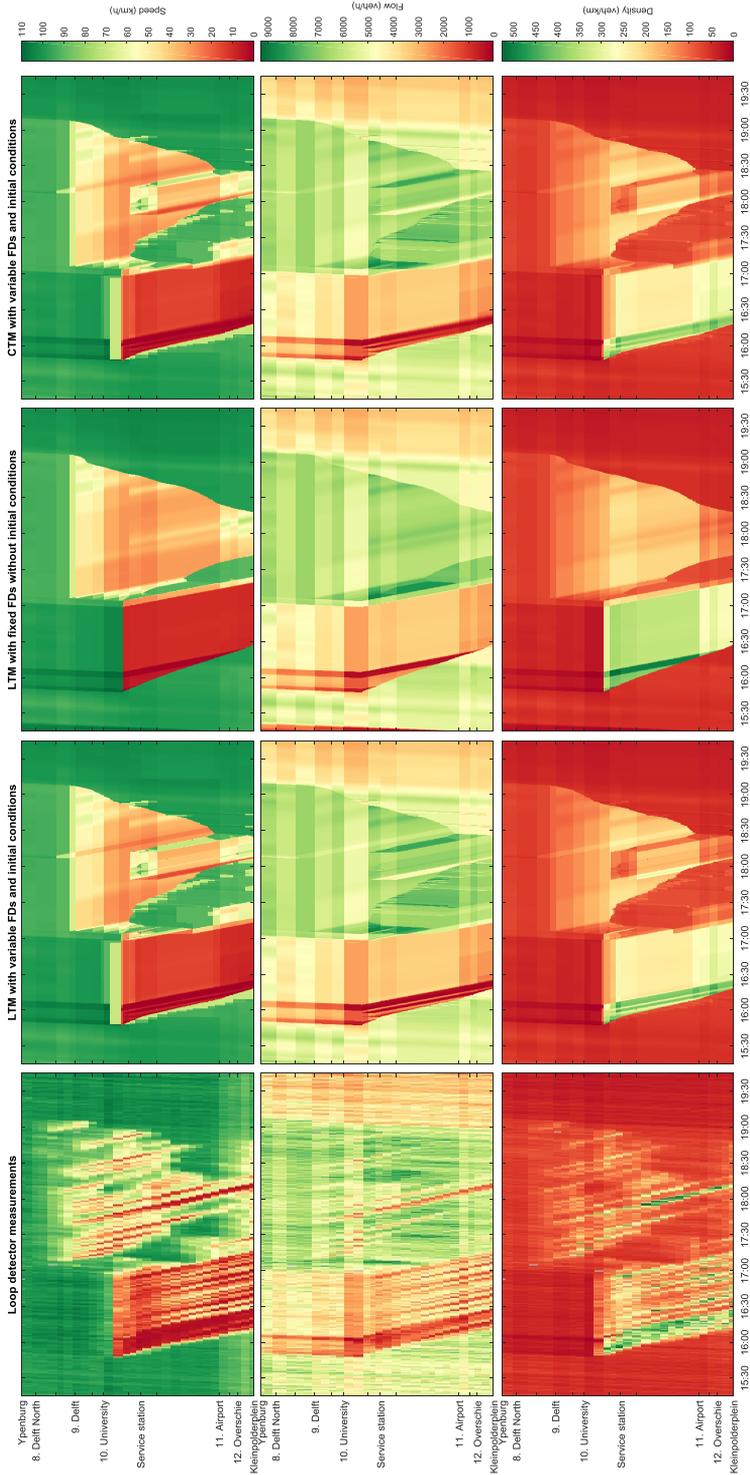
Figure 5.12. A13 measurements and simulation results for the extended LTM, traditional LTM, and extended CTM.
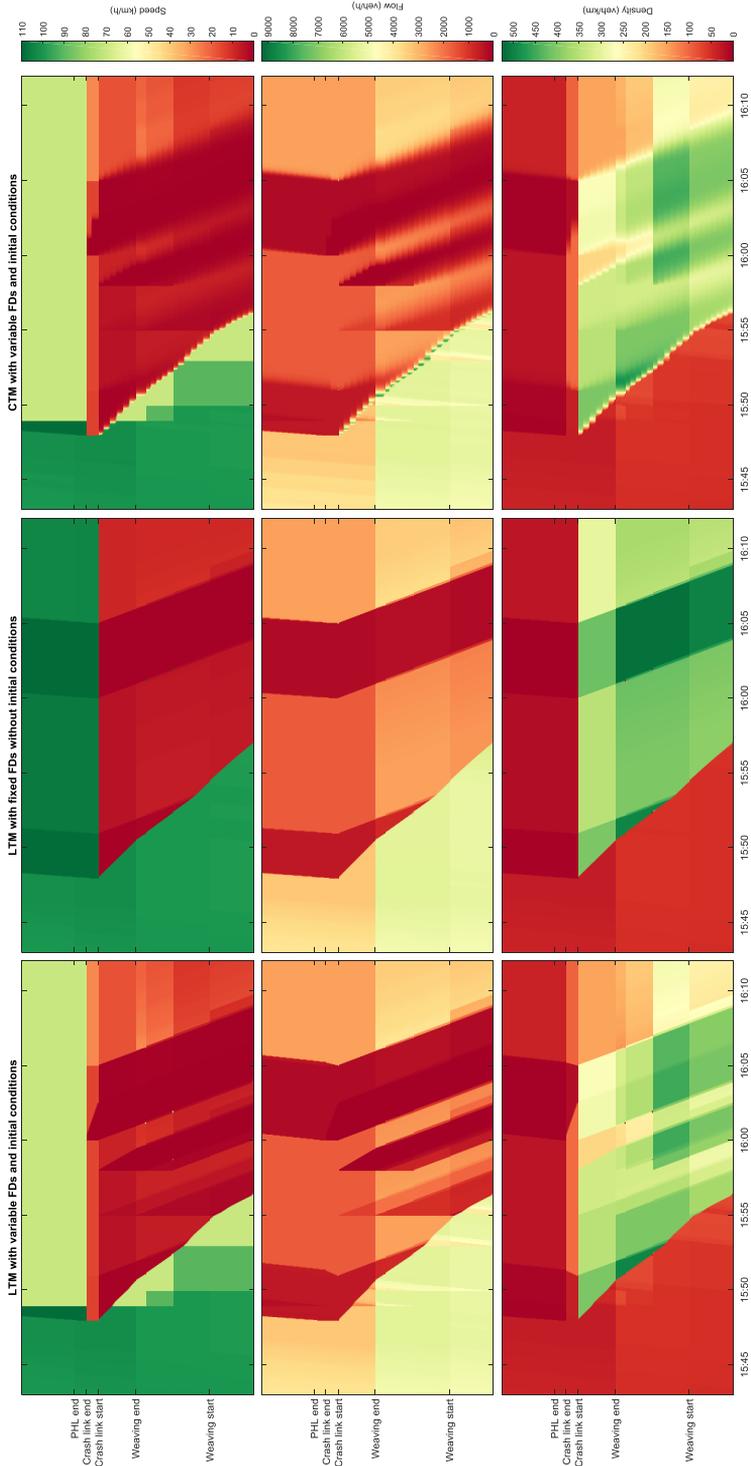
Figure 5.13. Congestion onset after the crash in simulation results for the extended LTM, traditional LTM, and extended CTM.

The CTM and LTM with variable FDs and initial conditions yield very similar results, but the CTM has greater numerical diffusion. This increased diffusion in the CTM is particularly evident in the less crisp contact discontinuities between traffic states in congestion.

Figure 5.13 zooms in on the onset of congestion and shows that the extended LTM correctly handles FD changes that result in the jam density being exceeded. The solution contains two acceleration shocks, one on the crash link starting at 16:00 due to a change in driving behaviour at the crash location and one on the link before that starting at 15:58 due to a lane closure, and no negative flows or speeds occur. At this zoom level, we also clearly see that the numerical diffusion of the CTM is much greater than that of the LTM.

After the incident is cleared, the re-opening of the closed lanes and removal of the speed limits are well-reproduced by the extended LTM link model. The growing oscillations in the congestion upstream of the incident and the stop-and-go waves originating from the recurrent bottleneck are not reproduced in the simulation, but these are well-known limitations of LWR theory to which the LTM and CTM calculate solutions. Despite this limitation, the shape of the congestion is reasonably well reproduced. One particularly interesting detail is that the most severe stop-and-go wave, reaching Kleinpolderplein around 18:10, does show up in the simulation – albeit with a lower density – due to an earlier lane drop making existing congestion more severe.

Overall, we conclude that our extended LTM works as expected based on LWR theory and our FD specifications, and that the simulation result indeed improves by taking the variable FDs and initial conditions into account. Compared to the CTM, this is achieved with considerably less numerical diffusion.

5.5.3   Model computation times

We finally compare the computation time of the simulations. In addition to the three simulations from the previous subsection, we include three more simulations where the LTMs and CTM use the same discretisation, i.e. an equal time step size $\Delta t = \frac{1}{41} \min$ for all nodes and only one CTM cell per link. This decreases the accuracy and computation time for the CTM, but further increases the accuracy and computation time for the LTMs.

All six simulations of 4.5 hours simulated time are executed 101 times in randomised order, using a single thread on a PC with an Intel Xeon E5-1650 3.20 GHz CPU and 32 GB RAM. Table 5.5 lists the median execution times per simulation type. The execution times include initialisation of the simulation model but exclude file input/output. The post-processing of results for visualisations like in the previous subsection is also excluded from this benchmark.

The results show that the (extended) LTM can be a very fast simulation tool. Despite the large number of FD changes in our simulated scenario, we see that their impact on the computation time of the LTM remains fairly small. The ability of the LTM to use different time step sizes per node greatly improves its performance in a network with some short links, while remaining very accurate. Our extended LTM is about four times as fast as a CTM for a comparable numerical accuracy like in Figure 5.12. Conversely, with an equal discretisation and similar computation time, our extended LTM has considerably better numerical accuracy than the CTM. Despite being important for CTM accuracy, the number of cells within links only has a minor impact on CTM efficiency because their computations are a lot simpler than general nodes.

Table 5.5. Computation times of the simulations.

	LTM with variable FDs and initial conditions	LTM with fixed FDs without initial conditions	CTM with variable FDs and initial conditions
Optimal discretisation (Subsection 5.5.2)	0.073 s	0.064 s	0.296 s
Equal discretisation (equal time steps, one CTM cell per link)	0.239 s	0.228 s	0.271 s

5.6   Discussion

In our link model formulation, all vehicles simultaneously on a link are assumed to adhere to the same FD, and thus simultaneously change their driving behaviour when a new FD is specified for the link. Some researchers might instead prefer a model that uses a multi-class LWR theory (Logghe and Immers, 2008), which was applied by Smits et al. (2011) and Hajiahmadi et al. (2013) in an LTM context. Such multi-class approaches allow e.g. assigning driving behaviour to vehicles upon the moment they enter the link, which they then retain while driving on the link, which Hajiahmadi et al. (2013) did for variable speed limits. This may be more realistic in some cases, e.g. when traffic rules for the link are displayed on gantries at the link entrance like in our case study, as drivers are unable to see rule changes after they entered the link. However, there are also important disadvantages to this alternative modelling approach. Firstly, it becomes much more difficult to develop a non-cell-based solution scheme, particularly when the model simultaneously has to support non-triangular FDs whose parameters can vary arbitrarily, and first-in-first-out behaviour no longer applies. Secondly, the assumption of driving behaviour being fixed after entering the link may itself be invalid, as drivers notice other drivers behaving differently, if drivers adapt to VMSs they see downstream, if information is provided in-vehicle, or if the FD change reflects a change in environmental conditions such as the weather.

For readers primarily interested in using our model to start their simulations with a non-empty network, we would like to remark that this uses as input an explicit initial distribution of density over space, and that this is not the only possible way to start an LTM simulation with a non-empty network. For example, when simulating two scenarios that only diverge after a certain point in time, it can be more convenient to simply make a copy of the model's computer memory at that point in simulation time, than to compute the within-link densities and use those to start the second simulation. Similarly, in our case study we could alternatively have used loop flow measurements from shortly before the studied time period to warm up the links. This however does rely on the initial data indeed consisting of point measurements over some prior time period (as opposed to spatial measurements at one moment in time) during which the FD is known, and if the FD changes during or prior to the simulation, one still needs to use our extended link model.

Another point of discussion is the chosen Smulders FD shape. While the choice of this specific shape is a limitation compared to other LTM formulations supporting more general FD shapes, our chosen free-flow branch shape does support subcritical delays and platoon dispersion. Although the shape of the congested branch is less sophisticated, we believe it suffices for most use cases, and its simplicity allows us to compute within-link densities

without having to wait for all slow congested waves from the initial condition to leave the link. The non-negativity of speeds and flows when densities exceed the jam density enhances the robustness of model applications.

One previous LTM feature that is nevertheless missing from our current FD shape is a capacity drop. Chapter 4 previously formulated such an extension to the LTM, using an inverted-lambda style FD, which was able to produce stop-and-go waves. As a generalisation of Algorithm 4.6, it is possible to divide the within-link space into an area where only free-flow traffic states and one selected "coasting state" with a density higher than the free-flow critical density, and a more downstream area where only traffic states with flows lower than the queue discharge rate can occur. This can be used to algebraically compute within-link densities, also in case of an initial condition. However, we are not aware of whether/how one can simultaneously satisfy Equation (5.7). The adaptation to include a capacity drop also makes the derivation and formulation of the extended LTM considerably different, more complex, and harder to understand from the perspective of traditional LWR theory. For these reasons, we chose to not include a capacity drop in this chapter. We do encourage future research into this topic because depending on its results, it could be very useful to combine a capacity drop and support for link-based traffic management in one simulation model.

As an alternative to our chosen Smulders FD in this chapter, we could have adopted an FD that is piecewise-linear in the flow-density plane in its entirety. Since this results in less smooth subcritical delay function than Figure 5.1, we believe this can only be beneficial if combined with an event-based continuous-time network solution scheme (Raadsen and Bliemer, 2018) instead of the less-complex and more-developed discrete-time scheme we used. An event-based adaptation of our work can however be an interesting avenue for future research that is likely feasible.

Some readers may wonder whether our extended link model can be embedded in the iterative LTM solution scheme (Himpe et al., 2016), allowing larger time steps than Equation (5.2) and warm-starting an optimisation or equilibration procedure (Himpe and Tampère, 2016). In this case, the most recent parts of the link boundary conditions that are needed for the computation of sending and receiving flows, are not yet known, but an estimate is available, allowing those estimates to be iteratively updated until sufficient convergence is reached. Hence, it is possible to make it work. First, at the start of a time step, one takes the discrete-time constraints on future cumulative inflows and outflows and makes a back-up copy in computer memory. Then, "update" the link boundary conditions as estimated, and compute the sending and receiving flows. As long as another iteration of the same time step is necessary, revert the constraints to their back-up copy and "update" them with the newer estimates instead. Once converged, continue with the next time step. Because the support for initial conditions in the sending and receiving flow computations is retained and the computation of within-link densities is not affected by the larger time steps, this makes all features of our extended link model available.

Finally, we emphasise the computational efficiency of our extended link model. Based on our remarks in Subsection 5.4.4 and Appendix 5.C.2, we conclude that the computation of sending and receiving flows requires the same or fewer iterations than previous discrete-time link model formulations in literature. Each computation of within-link density profiles and application of initial conditions only incurs a one-time computational cost per link, so the additional computation time for these new features should be limited for most applications, e.g. in case of a limited number of FD changes. The high numerical accuracy of the LTM is preserved. The high efficiency and accuracy have been validated in Section 5.5.

The computer memory requirements are also limited and independent of the duration of the simulation. For the calculation of within-link densities, only the most recent parts of the downstream boundary condition (last $-L/v'$ time) and the upstream boundary condition (last L/v_C time) need to be known. Older traffic states from the boundary conditions cannot affect the current within-link densities anymore and can thus be forgotten, so that longer simulations do not require more computer memory. Note that the initial condition must be remembered, since initial densities exceeding the jam density may remain present on the link indefinitely. The computations of the receiving and sending flows additionally need to store cumulative inflow and outflow constraints for a finite number of time points in the near future. These data can be handled efficiently using a (variable-size) circular buffer data structure instead of a traditional array.

5.7   Conclusions

With a derivation from LWR theory, we extended the link model of the discrete-time LTM with the capabilities of using non-empty initial conditions and computing within-link density profiles, using a Smulders FD with no negative flows or speeds beyond the jam density. When combined, they can be used to fuse data or to instantaneously change the FDs of links within an LTM simulation. These extra capabilities, which were previously only available using memory-less numerical schemes like the CTM, introduce no new numerical error to the LTM. Since each invocation of the new capabilities involves only a one-time computational cost penalty, the LTM remains suitable for the efficient multi-commodity simulation of large general networks, and outperforms the CTM.

We demonstrated the benefit of the abilities to change FDs and use initial conditions with a numerical example with variable speed limits and dynamic lane management. These and other possible applications of this robust general-purpose link model extension, e.g. involving traffic control, changes in environmental conditions and driver behaviour, and real-time forecasts significantly extend the potential application domain of the LTM and, by extension, efficient large-scale simulations based on LWR theory. We look forward to future research on such advanced applications.

Acknowledgement

5.A Notation

x	location
x_0	location of upstream link end
x_L	location of downstream link end
L	link length
t	time in general or current time
τ	time in general
t^*	time of most recent initial condition
Δt	time step size
Δt_{x_0}	time step size of upstream node
Δt_{x_L}	time step size of downstream node
$S(t),\ S_i(t)$	sending flow (of link i) in time step starting at time t
$S_{ij}(t)$	part of sending flow of link i destined for link j in time step starting at time t
$R(t),\ R_j(t)$	receiving flow (of link j) in time step starting at time t
$G_{ij}(t)$	transition flow from link i to link j in time step starting at time t
$k,\ k(x,t)$	density (at space-time point (x,t))
$q,\ q(x,t)$	flow (at space-time point (x,t))
$N,\ N(x,t)$	cumulative number of vehicles (at space-time point (x,t))
$N^{\max}(x,t)$	upper bound on cumulative number of vehicles at space-time point (x,t)
$Q(k)$	flow as function of density k
$U(k)$	speed as function of density k
$V(k),\ V(q)$	free-flow branch wave speed as function of density k or flow q
$K(q)$	free-flow branch density as function of flow q
$\kappa(v)$	flow-density plane tangent line density-axis intercept (non-positive) as function of free-flow branch wave speed v
u_F	free speed
u_C	critical speed, i.e. speed at capacity
k_C	critical density, i.e. density at capacity
v_C	critical wave speed, i.e. maximum wave speed at capacity
v'	wave speed in congestion (negative)
k_J	jam density
x_A	location of point A
t_A	time of point A
N_A	cumulative number of vehicles of point A
f_{AB}	second spatial derivative of cumulative number of vehicles on piece AB
$N_{AB}(x),\ N_{AB}(\tau)$	cumulative number of vehicles on piece AB as function of location x or time τ
q_{AB}	flow on piece AB
k_{A^+}	density immediately downstream of point A
k_{A^-}	density immediately upstream of point A

5.B Multi-commodity support

Since Yperman et al. (2006), the LTM can be used in a multi-commodity fashion, e.g. to disaggregate traffic by destination for route choice purposes. Yperman et al. (2006) disaggregate traffic (potentially) leaving a link by looking at the traffic composition when this traffic entered the link. We note that this is problematic when the network is not empty at the start of the simulation. There are two possible solutions.

The first solution is to construct a virtual link inflow history prior to the start of the simulation, which is then only used to disaggregate traffic on the link. This is the solution we use for the simulations in Section 5.5. Multiple virtual past time steps can be used to denote changes in initial traffic composition along the link.

Alternatively, this problem can be avoided by instead storing the traffic composition per link as an explicit first-in-first-out (FIFO) queue in computer memory, whose initial state can be provided along with the initial conditions for the traffic positioning on the link. Here, "queue" refers to the computer data structure, not to the traffic phenomenon. Each entry in this queue is an unordered collection of pairs of vehicle type and vehicle amount. Each time step, each node partially or completely consumes entries from the fronts of the queues of incoming links until it cannot accept more flow, and adds a new entry to the backs of the queues of each outgoing link that receives flow. For partially consumed entries, the consumption of vehicle types is proportional to their amounts.

A secondary benefit of such an explicit FIFO queue is that it is more efficient, since the computational cost of disaggregation no longer depends on the total number of commodities in the link or network, and it avoids storing a separate cumulative inflow history per commodity in memory. The node model algorithm can be modified to prefer strict FIFO over strict conservation of turning fractions. (Node model applications usually prefer strict conservation of the specified turning fractions, that is, when they change the outflow of a link, the composition of this outflow is not updated accordingly.) We refer to Subsection 3.3.2 for a further elaboration.

5.C Conflict resolution examples

In this appendix we explain the conflict resolution for Section 5.4 in more detail, with two examples: computing the within-link density profile in continuous coordinates (Part 5.C.1) and computing link outflow constraints in discrete coordinates (Part 5.C.2).

5.C.1 Conflict resolution in continuous coordinates

Suppose we are calculating the within-link density profile, for which the intermediate solution so far consists of four pieces EF, FG, GH, and HI ($x_E < x_F < x_G < x_H < x_I$, $N_E \geq N_F \geq N_G \geq N_H \geq N_I$). Using the second and third columns of Table 5.2, we construct some new candidate piece OP ($x_O < x_P$, $N_O \geq N_P$). Suppose the conflict resolution method indicated in the fourth column is "$\uparrow$ min". The upward "$\uparrow$" arrow means we start at low x, i.e. we first decide whether the upstream point O of the candidate piece should be added to the intermediate solution. This is the case if x_O lies in front of the previous upstream end of the intermediate solution ($x_O < x_E$) or if N_E is smaller (because of "min") than the previous intermediate solution evaluated at x_0.

If true, and if $x_P > x_E$, we need to find the shock between the new candidate characteristics and those of the previous intermediate solution. In this example we assume $x_F \le x_O < x_G$. Then, piece EF can be discarded immediately since only the new candidate piece can be upstream of the shock (because of "↑"). The next piece of the previous intermediate solution is FG. If $x_G \le x_P$ and $N_{OP}(x_G)$, the cumulative value of piece OP evaluated at x_G, is smaller (because of "min") than N_G, piece FG is also discarded. This is repeated until a next piece, e.g. piece HI, for which this condition no longer holds. Finally, solving

$$\begin{cases} N_{OP}(x) - N_{HI}(x) = 0 \\ \dfrac{d}{dx}\big(N_{OP}(x) - N_{HI}(x)\big) > 0 \end{cases} \tag{5.17}$$

for x using the quadratic formula ($>$ because "min") gives the location of the shock, to which both pieces are then truncated. The new candidate piece is now added to the intermediate solution.

Implementers are advised to be wary of the finite precision of floating-point number representations, to prevent rounding errors from resulting in extremely short pieces in the calculated solution that do not exist in the true algebraic solution.

5.C.2 Conflict resolution in discrete coordinates

Suppose we are calculating the link cumulative outflow constraints with discrete time steps, for which the intermediate solution so far is

$$\begin{cases} N^{\max}\left(x_L, t + \Delta t_{x_L}\right) = N_1 \\ N^{\max}\left(x_L, t + 2\Delta t_{x_L}\right) = N_2 \\ N^{\max}\left(x_L, t + 3\Delta t_{x_L}\right) = N_3 \\ N^{\max}\left(x_L, t + 4\Delta t_{x_L}\right) = N_4 \\ N^{\max}\left(x_L, t + n\Delta t_{x_L}\right) = N_5 + (n-5)q_C\Delta t_{x_L} \quad \forall n \in \{5,6,7,\ldots\} \end{cases} \tag{5.18}$$

for some N_1, N_2, N_3, N_4, N_5. Using the second and third columns of Table 5.4, we construct some new constraints

$$\begin{cases} \hat{N}^{\max}\left(x_L, t + 2\Delta t_{x_L}\right) = \hat{N}_2 \\ \hat{N}^{\max}\left(x_L, t + 3\Delta t_{x_L}\right) = \hat{N}_3 \\ \hat{N}^{\max}\left(x_L, t + 4\Delta t_{x_L}\right) = \hat{N}_4 \\ \hat{N}^{\max}\left(x_L, t + 5\Delta t_{x_L}\right) = \hat{N}_5 \end{cases} \tag{5.19}$$

with some $\hat{N}_2, \hat{N}_3, \hat{N}_4, \hat{N}_5$. The conflict resolution method indicated in the fourth column is "← min". The leftward "←" arrow means we start at high time, i.e. we first decide whether the constraint at time $t + 5\Delta t_{x_L}$ must be lowered in our intermediate solution (because of "min").

If true ($\hat{N}_5 < N_5$), we need to find the shock between the new candidate characteristics and those of the previous intermediate solution. We update the intermediate solution at $t + 5\Delta t_{x_L}$,

and discard the previous constraints for time $t + 6\Delta t_{x_L}$ and later since the shock must be before the new candidate characteristics (because of "←"). We next look at time $t + 4\Delta t_{x_L}$ and update it as well if $\hat{N}_4 < N_4$. This iteration continues until all new constraints are processed or until e.g. $t + 3\Delta t_{x_L}$ if $\hat{N}_3 \geq N_3$. The updated intermediate solution is now

$$\begin{cases} N^{\max}\left(x_L, t + \Delta t_{x_L}\right) = N_1 \\ N^{\max}\left(x_L, t + 2\Delta t_{x_L}\right) = N_2 \\ N^{\max}\left(x_L, t + 3\Delta t_{x_L}\right) = N_3 \\ N^{\max}\left(x_L, t + 4\Delta t_{x_L}\right) = \hat{N}_4 \\ N^{\max}\left(x_L, t + 5\Delta t_{x_L}\right) = \hat{N}_5 \end{cases} \tag{5.20}$$

Note that this searching direction allows terminating the iteration as soon as the shock is found; e.g. the computation of $\hat{N}_2$ can be skipped. This avoids the need to iterate over the entire new set of candidate characteristics when not all of them are relevant, increasing computational efficiency compared to Gentile and Papola (2009b) and Gentile (2010).

Chapter 6

Real-time bus-bridging scheduling and dynamic simulation for multimodal urban network disruptions

Abstract

Unplanned disruptions of multimodal transportation systems in urban areas can have large negative consequences and may require bus-bridging services to help stranded public transport passengers. Bus-bridging optimisation methods however assume that demand and travel times are known beforehand. We overcome this limitation by extending an existing static bus-bridging optimisation model into a novel real-time scheduling method and coupling it with a dynamic multimodal simulation model to test it in a hypothetical disruption of metro, tram and car traffic in Rotterdam, the Netherlands. We also show how various properties of the simulation model affect the assessment of the impacts of the disruption on both car traffic and public transport, including bus bridging. The results highlights the important role of alternative public transport options and the relevance of road travel times for the success of bus bridging, and lead to suggestions for improving the real-time bus-bridging scheduling.

This chapter is a version of the following manuscript in preparation:

6.1 Introduction

Metro systems often form a fundamental component of the multimodal transportation system of large urban areas, serving large volumes of passengers along the major urban corridors. Unplanned disruptions of the operation of the metro system, for example due to malfunctioning infrastructure, accidents, or other emergencies, can consequently have large negative impacts on the travel times experienced by many travellers. In such cases, authorities frequently have to resort to dispatching additional temporary bus services to help stranded passengers to get to their destinations despite the metro outage. More specifically, they typically employ some form of bus bridging, which involves running buses between the metro stations affected by the disruption.

Although bus-bridging services can be designed by hand, Kepaptsoglou and Karlaftis (2009) already recognised the benefit of designing the bus-bridging network with optimisation models selecting which bus routes to use and subsequently how buses are allocated to those routes. Along other improvements, Jin et al. (2016) later enhanced their idea by integrating both optimisation processes into a single model. Van der Hurk et al. (2016) proposed a formulation that allows the simultaneous selection of bus routes and frequencies, solving large-case problems in real-life situations. Gu et al. (2018) increased the bus-bridging flexibility by allowing a single bus vehicle to drive more than one route over time, i.e. allowing each bus vehicle to traverse any sequence of stations, and split the optimisation into a first stage to determine which buses drive which routes and a second stage to optimise the ordering of these routes.

A limitation of the above methods is that the bus-bridging passenger demand matrix and bus travel times between stations must be known beforehand. While that may be a reasonable assumption if the disruption only affects the metro system, this data is much more difficult to obtain beforehand in case of a less common multimodal disruption that also affects the road network. Firstly, the road congestion may then impact the travel times for the bus-bridging buses, so that they can no longer be (reliably) retrieved from e.g. historic data. Secondly, delays and blockages in the road network could also disrupt regular bus and tram lines, which makes it rather hard to estimate whether those existing bus and tram lines will serve part of the stranded passengers and reduce the demand for bus bridging, or instead generate new stranded passengers themselves and increase the bus-bridging demand. Despite these issues with applying a bus-bridging optimisation model, designing efficient bus bridging by hand is likely to become even more difficult in this case.

This lack of accurate information beforehand can however be circumvented by adding a feedback loop to the bus-bridging optimisation. For example, the travel times on the road network and the number of stranded passengers at the affected metro stations can be monitored continuously, and can be used to adjust the bus-bridging schedule periodically. In this chapter, we will explore the feasibility of this option by developing such a real-time approach, using an existing dynamic multimodal simulation model (Chapter 2) to represent a virtual reality in which travel times and stranded passengers are simulated and recorded, and an existing bus-bridging optimisation model (Gu et al., 2018) to periodically compute new bus-bridging schedules based on the latest recorded counts. We test the developed approach in a case study on a large-scale network of the city of Rotterdam, the Netherlands, consisting of a hypothetical emergency evacuation of an underground metro station that also obstructs cars and trams in the aboveground road network.

The contribution of this chapter is threefold. First, we contribute the formulation of the novel real-time bus-bridging scheduling method mentioned above. Second, we test the performance of this new method in a multimodal disruption in an urban setting, resulting in suggestions for further improvements to the bus bridging. Third, we analyse how various properties of the simulation model affect the ex-ante impact assessment of the disruption, including bus bridging, improving our understanding of what aspects of the simulation are most important.

This chapter is structured as follows. The section below first describes the methodology of both the multimodal simulation model and the bus-bridging optimisation model, followed by an explanation of how we set up the real-time scheduling method linking both models. The section thereafter describes the Rotterdam case and its simulation results, comparing various formulations of the simulation model. The final section discusses the implications of the results for both the proposed real-time bus-bridging scheduling method and the design of the simulation model used for evaluating the impacts of an urban public transport disruption.

6.2   Methodology

One of the main contributions of this chapter is the novel real-time scheduling method for bus bridging that accounts for dynamic passenger demand and travel times. The building blocks for the formulation and testing of this scheduling method consist of

- the static bus-bridging optimisation model with fixed passenger demand and bus travel times, as discussed by Gu et al. (2018), and

- the dynamic multimodal simulation model that includes impacts of disruptions and counter-measures, implementing the framework from Chapter 2.

Therefore, in Subsections 6.2.1 and 6.2.2, we first describe each of these models conceptually, while the interested reader is referred to the source paper/chapter for further details. Afterwards, in Subsection 6.2.3, we present the new framework in which these models are integrated and adapted to allow for the implementation and evaluation of real-time bus bridging.

6.2.1   Static bus-bridging optimisation model

A metro disruption often leads to link closure between two turnover stations, where track crossover is available. Metro lines then operate in short routing mode to transport passengers beyond the two turnover stations. Bus-bridging strategies are used to reconnect the affected metro link and transport the stranded passengers during metro disruptions with available bus resources. The bus-bridging service consists of two processes, namely

1. a dispatching process, where each bus drives one dispatching route from its depot to one of the affected metro stations, and

2. a bridging process, where each bus drives a sequence of bridging routes between metro stations to transport stranded passengers.

See Figure 6.1 for an example. Due to limited bus resources and high passenger demands, different bus-bridging strategies will certainly result in different efficiency. In this chapter, we use the flexible approach introduced by Gu et al. (2018), where each individual bus vehicle is
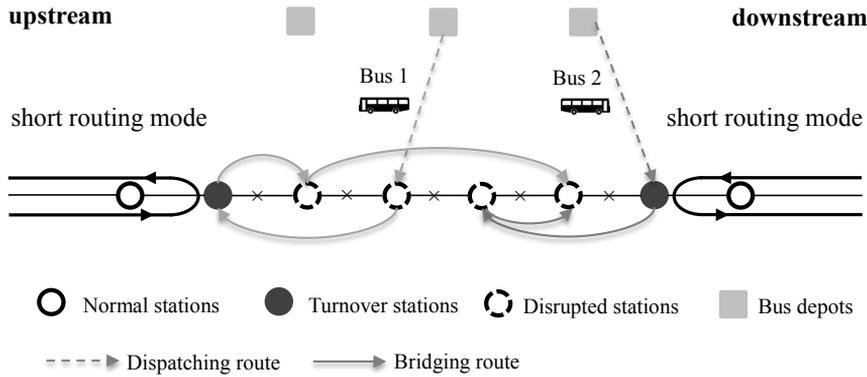
Figure 6.1. Example bus-bridging schedules proposed by the Gu et al. (2018) model.

assigned a tailored sequence of stations instead of being constrained to run the same routes repeatedly. As inputs, the method requires matrices of travel times between the stations and from the depots to the stations, and a matrix describing the origin-destination demand that is to be served. The approach consists of two stages:

1. a first stage optimising how often each bus drives each dispatching/bridging route, minimising the (longest) time needed to serve all passenger demand, and

2. a second stage optimising the order in which each bus drives each of its assigned bridging routes to serve the passenger demand as quickly as possible, minimising the average passenger delay.

Each stages optimises an integer linear programming model and is solved using the MIP solver in CPLEX. The resulting tailored bridging paths describe the stations that buses should visit in sequence once they are dispatched from the depot. As shown in Figure 6.1, they are often non-intuitive. Note that the approach can be easily extended to non-linear networks and more than two turnover stations, as we will do in our case study.

6.2.2   Dynamic multimodal simulation model

With respect to the simulation of the transportation system, this chapter implements the multimodal simulation framework for urban emergencies developed in Chapter 2, which combines an agent-based representation of travel demand with an advanced first-order macroscopic simulation of the propagation of traffic over time. Unlike e.g. Mezzo (Toledo et al., 2010), the advanced features of the traffic propagation do not depend on microscopic simulation.

The agent-based travel demand is based on activity-travel patterns of a synthetic population. Agents travel between their scheduled activities, where, in addition to choosing routes upon departure, are able to switch routes while travelling. Route choices are updated en-route every minute based on new information about the state of the transportation system, for both car drivers and public transport users. This information consists of instantaneous road link travel

times, which are averaged over 5 minutes to remove high-resolution oscillations (which are unlikely to affect route choice), and public transport schedules with expected delays, cancellations, and bus-bridging services. It is possible to use equilibrium route choice, but we do not adopt that approach here, where we focus on adaptive rerouting. The choice model is a probit model based on the expected travel time of the route and number of public transport boardings and alightings along the route. Each household is assigned a set of probit error terms at the start so that their members have consistent route choice preferences over the course of the simulation. Prior to the simulation, for all modes, we generate route choice sets from each link in the network to each destination, so that agents do not continually have to search for new routes during the simulation.

Public transport vehicles can be boarded and alighted by traveller agents along their routes. This enables us to simulate public transport usage, including bus bridging, in full detail at the level of individual runs, stops, and travellers, accounting for any incurred delays. In the model, buses and trams are part of the road traffic and are consequently hindered by and part of road congestion, unless they have separate infrastructure. While the model structure allows to make the number of boarding and alighting passengers affect dwell times, we omit this effect because we lack data on this relationship for the study area.

The propagation of traffic over time in the network is modelled by the an extension of the Link Transmission Model (Yperman et al., 2006), consisting of a link model and a node model. Nodes can have restrictions on what turns from incoming links to outgoing links are permitted.

The link model is a numerical solution scheme to kinematic wave theory (Lighthill and Whitham, 1955; Richards, 1956), solving traffic flow over homogenous road sections. For car traffic links, we use an extension of the original link model that supports Smulders (1990, p. 117) fundamental diagrams where the speed in subcritical traffic linearly decreases with increasing traffic density, which on motorways is further extended to include a capacity drop as well (Chapter 4). Railway capacities are assumed to suffice to serve all trains and metros and their signalling systems are not modelled.

To solve traffic flows at junctions and other discontinuities, we use the Flötteröd and Rohde (2011) node model without internal node supply constraints, i.e. accounting for merging and diverging conflicts but not for crossing conflicts. For signalised intersections, we solve the absence of crossing conflicts by separately applying capacity constraints on incoming links of signalised junctions, based on their green fraction. Upstream of signalised junctions, we model multiple separate incoming links for different approaches with separately controlled traffic lights, each with corresponding turn restrictions for the junction. This accounts for the lack of first-in-first-out behaviour upstream of signalised junctions and allows to set different green fractions per approach. We then first run a simulation with all traffic lights disabled, measure peak hourly flows on all approach links, use those to calculate a fixed-time signalling scheme with the Webster (1958) method for the critical conflict group, add all other conflict groups to the scheme, and store the resulting green fractions to be used in simulations with traffic lights enabled. While it is possible with this method to explicitly simulate red and green phases, we choose to only simulate the average effect on capacity using the green fractions, as otherwise any misspecification of the coordination between nearby nodes would cause a substantial capacity reduction in the model. This simplification removes the delay at signalised junctions in undersaturated conditions, but this is to some extent compensated for by the Smulders fundamental diagrams of the links which in our study have a substantially lower critical speed than their free speed.

To track individual agents in the macroscopic flows of the Link Transmission Model, a multi-commodity formulation is used. Based on the considerations of Chapter 3, we use one commodity per agent, use queue data structures on links to store traffic composition using a strict weak order, transmitting traffic over nodes with the incremental Flötteröd and Rohde (2011) node model and excluding the possibility of unrealistic gridlock due to agents being partially on one link and partially on the next link. To limit numerical errors in the simulation, we use a maximum time step size of 2 seconds for nodes along public transport lines, and 3 seconds for other nodes inside the study area, compared to 60 seconds for other nodes outside the study area.

6.2.3   Real-time bus-bridging scheduling method

Our proposed real-time scheduling method of bus-bridging services is based on repeated online application of the static bus-bridging optimisation described in Subsection 6.2.1. Each such application is based on the then-available information about the state of the transportation system, as derived from the dynamic multimodal simulation described in Subsection 6.2.2. More precisely, our real-time method assumes estimates of the following information to be available:

- current passenger counts waiting for bus-bridging services at each station;
- historical destination station fractions for each origin station;
- current travel times between stations and from depots to stations.

The passenger counts can be obtained by monitoring the stations. Since the destinations of these waiting passengers are difficult to monitor or predict in real-time (note that they could of course be retrieved from the model, but that such information cannot be reliably retrieved in practice), we use destination fractions obtained from studying metro passenger flows on an undisrupted day. These historical fractions can vary dependent on the time-of-day. Multiplication of the destination fractions with the passenger counts results in an origin-destination matrix that can be used as input to the static bus-bridging optimisation method described in Subsection 6.2.1. Finally, in order to ensure that each origin-destination pair gets served at least once, we add a value of 1 to each off-diagonal cell of this matrix.

Initially, all buses will be located at depots. Thus, in the first static optimisation round, the starting locations will correspond to the physical depots. This optimisation round results in a schedule for each bus. In each later static optimisation round, buses must first complete their previously-scheduled runs, so they effectively depart from the final stations in their previous schedule instead of from the physical depots. Additionally, the buses may not all be immediately available, since not each bus reaches the final station of their previous schedule at the same time. Thus, for each later optimisation round, starting locations are defined by the final station of the previous round and the remaining time until it completes the previous schedule. Combined, the final station and the remaining time can be treated as a virtual depot in the optimisation model. The travel time from such a virtual depot to any station equals the remaining time plus the travel time from the final station of the previous schedule to the first station of the next schedule (if these differ).

With this structure, each optimisation round effectively plans the evacuation of currently waiting stranded passengers to their destinations, and it is possible to repeatedly add more rounds to evacuate more people as time progresses, using newer input information. Each round, the completion time of the evacuation is minimised. In theory, each new optimisation
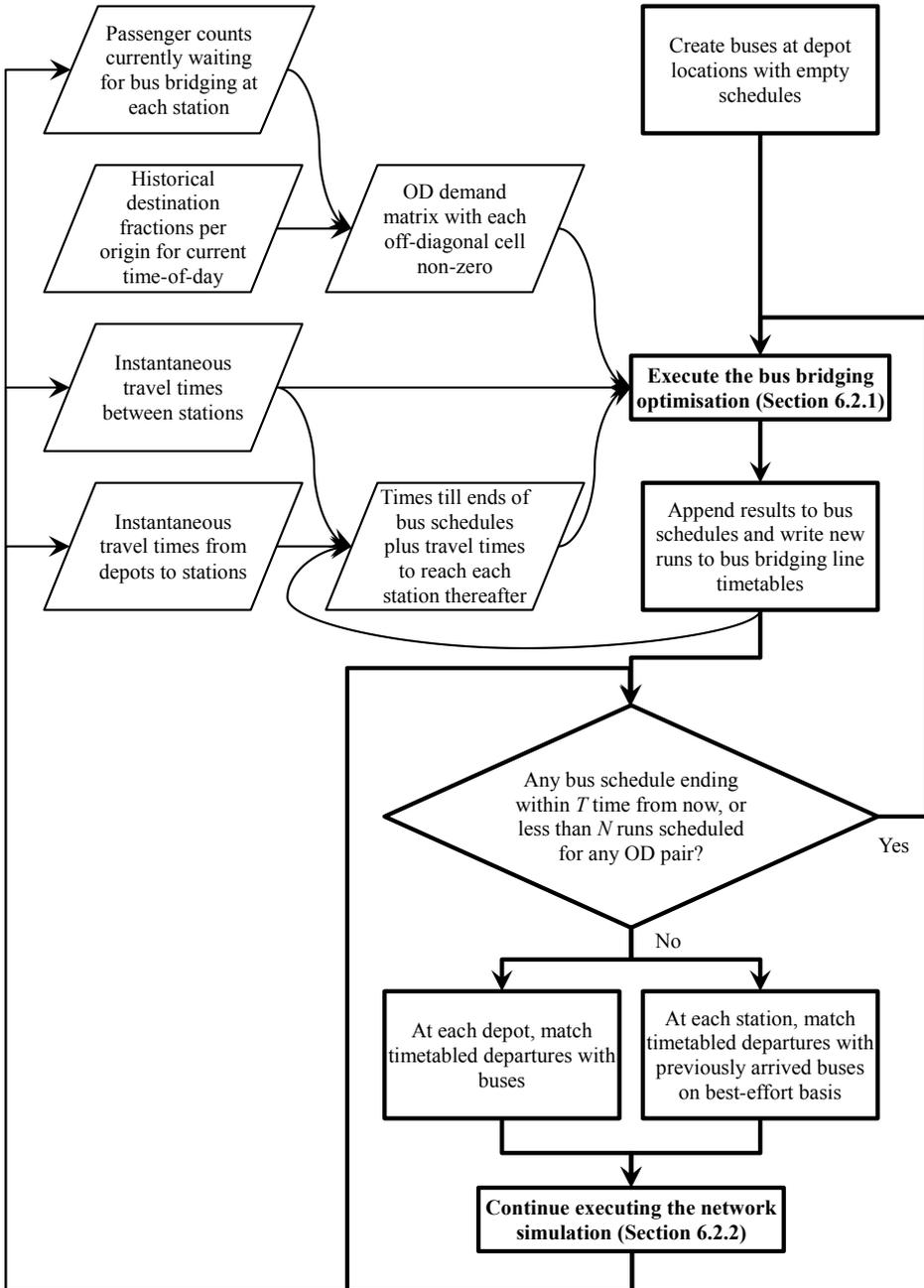
Figure 6.2. Flow chart of the real-time scheduling of bus-bridging services.

round can be postponed until one of the buses completes its schedule. However, in practice, it will be desireable to inform passengers in advance when buses are going to run, e.g. via announcements at stations and other traveller information systems. This leads to a trade-off between optimising as late as possible with the most recent and thus most accurate information on the state of the system, and optimising earlier so that the bus services can be announced to travellers in advance and passengers can plan their travel accordingly. To facilitate this planning by travellers, the scheduling method can no longer modify or cancel a bus-bridging run after it has been scheduled. We propose two parameters for configuring this trade-off, determining how far into the future the bus-bridging runs are scheduled:

- an amount of time $T \geq 0$ that each bus-bridging run is announced prior to departure;
- a minimum number of bus-bridging runs $N \geq 0$ for each origin-destination station pair that is scheduled to depart in the future.

T and N must be large enough so that travellers heading towards an affected station know a bus-bridging connection will be available once they get there, but small enough so that bus-bridging travel times do not vary too much between scheduling and execution of runs, as that would result in inefficient or delayed execution of the scheduled runs. In the case study of this chapter, we use $T = 30 \, \text{min}$ and $N = 1$, which means that all bus-bridging runs are announced to travellers at least 30 minutes before departure, and passengers always know at least one next departure time for each origin-destination pair.

Figure 6.2 summarises how this bus-bridging scheduling is integrated with the network simulation, or when applied in practice, with the real transportation system. Note that besides the bus-bridging scheduling described above, this flow chart also includes a level of operational control that allows bus vehicles to swap roles. For example, consider a situation where two buses are scheduled to depart from the same station at different times. If due to some delay, the bus that was scheduled to arrive at that station first, actually arrives later than the second bus, then the second bus may continue executing the schedule of the first bus and the first bus can take over the remaining schedule of the second bus, improving the punctuality of the departures. This is done by converting the bus schedules into timetables for each origin-destination pair immediately after they are optimised, and then dynamically assigning any arriving buses to the scheduled departures at each station. Together, this concludes our contribution of the formulation of our novel real-time bus-bridging scheduling method.

6.3   Case description

In this section, we describe the transportation network, the synthetic population and travel demand, the disruption scenario used for testing our methodology, and some configuration details of the simulation model. The research results are presented thereafter in Section 6.4.

6.3.1   Network

Our case study uses a network of Rotterdam, the Netherlands, which we based on the 2015 network of the static transportation model used by the Municipality of Rotterdam. The network is depicted in full in Figure 6.3 and zoomed in on the study area in Figure 6.4. As modes, our case study supports car driver, car passenger, train, metro, tram, bus, bus bridging, and slow (referring to cycling and walking). Counting all modes, the network has 167 centroids, 6752 junctions, 28924 links, 50057 turns, and 867 public transport stations/stops.
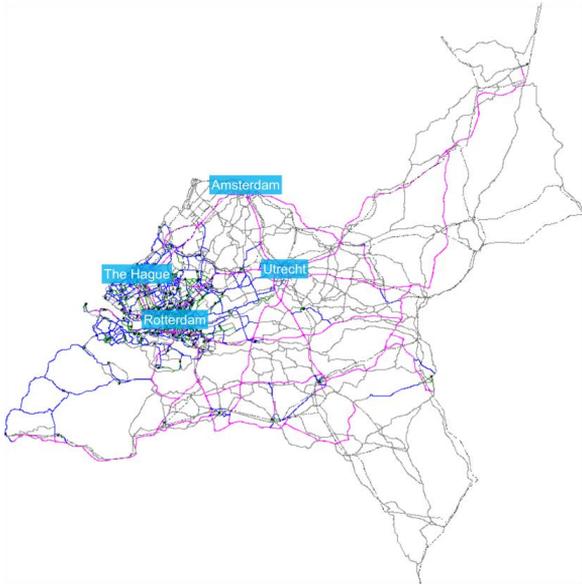
Figure 6.3. Overview of the full network, with motorways (purple), distributor roads (blue), access roads (green), pedestrian links (grey), railways (dashed). For reference, the network as shown spans approximately 200 km by 200 km.
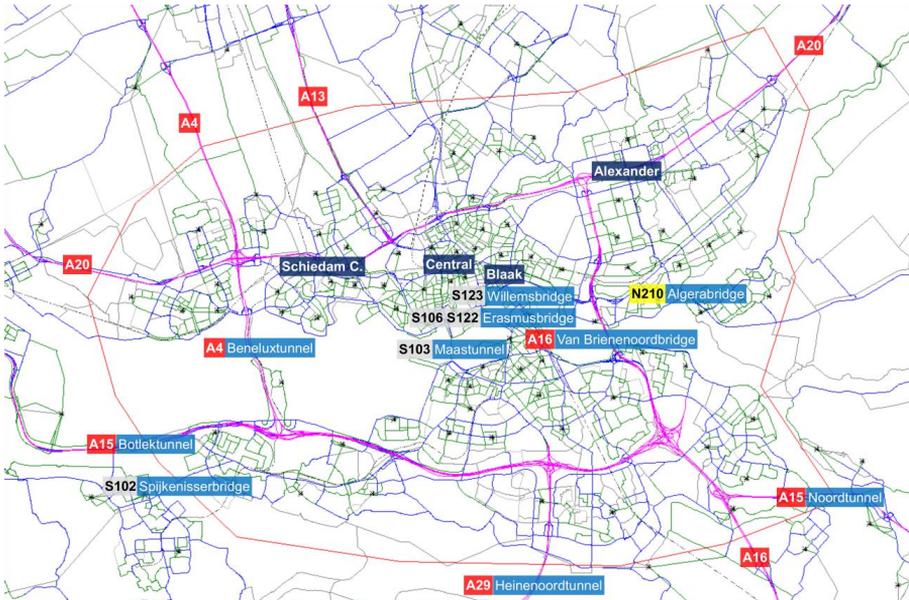
Figure 6.4. Closer view of the study area (red border) within the network, including motorway numbers (red) and names of major bridges and tunnels (light blue) and names of Intercity/express train stations (dark blue). For reference, the network as shown spans approximately 30 km (east to west) by 20 km (north to south).

For car traffic, we distinguish three main road types. The corresponding capacities are listed in Table 6.1, which for motorways vary depending on whether traffic is in free-flow, discharging at a fixed bottleneck, or discharging from a moving jam. Link free speeds are set based on the municipal model. The critical speed corresponding to the free-flow capacity is set 10 km/h less than the free speed, with a maximum of 90 km/h. The jam density is always set to 150 pcu/km/ln. The driving speed inside stop-and-go waves caused by the capacity drop on motorways is set to 10 km/h. Busways and railways for public transport (3232 km) and sidewalks for slow traffic (12728 km) are also modelled as links, but with arbitrary large capacities as modelling congestion on them does not fall within the scope of this research.

Table 6.2 gives an overview of the public transport modes in the model. Public transport vehicle capacities are based on seats and standing places in typical vehicles used in and around Rotterdam, thus ignoring variations per line and time-of-day. For bus, metro and tram, complete timetables are imported from the municipal model; for train, the incomplete timetables were extended before the morning peak and after the afternoon peak using the inter-peak schedule. The resulting public transport network is plotted in Figure 6.5. The slow mode speed inside the study area is set to 5 km/h, but is set to 18 km/h outside this area to compensate for the omission of urban public transport services in the model beyond the study area.

Table 6.1. Road link type capacities including capacity drops.

Road type	Total road length	Free-flow capacity	Standing queue discharge rate	Moving jam discharge rate
Motorway	2734 km	2200 pcu/h/ln	2100 pcu/h/ln	2000 pcu/h/ln
Distributor road	3801 km	1800 pcu/h/ln	1800 pcu/h/ln	1800 pcu/h/ln
Access road[*]	2215 km	1600 pcu/h/ln	1600 pcu/h/ln	1600 pcu/h/ln

Table 6.2. Public transport mode statistics and parameters.

Public transport mode	Total infra length	Stations/ stops	Lines	Dialy runs	Vehicle capacity	Maximum access/egress time	Passenger car equivalent
Train (HSL)	138 km	3	2	38	720 psg/veh	180 min	N/A
Train (Intercity)	2436 km	63	61	1787	939 psg/veh	180 min	N/A
Train (Sprinter)	2355 km	88	86	2692	716 psg/veh	180 min	N/A
Metro	157 km	65	25	1217	271 psg/veh	60 min	N/A
Tram	155 km	220	32	1741	182 psg/veh	15 min	4 pcu/veh
Bus	999 km	536	229	6317	98 psg/veh	10 min	2 pcu/veh
Bus bridging	N/A	N/A	N/A	N/A	81 psg/veh	60 min	2 pcu/veh

---

[*] Capacities of distributor roads are used for access road links connecting to signalised junctions.
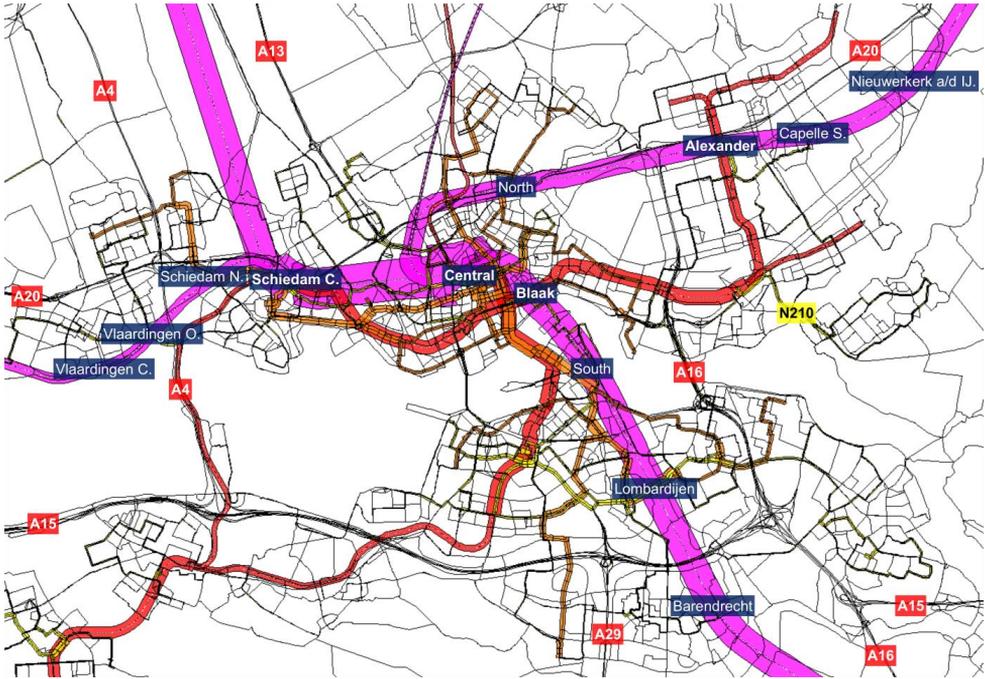
Figure 6.5. Study area public transport services, consisting of train (purple), metro (red), tram (orange), and bus (yellow), including train station names (dark blue, bold for Intercity/express) and major road numbers. Bar widths are proportional to the product of daily runs and vehicle capacity.

6.3.2   Population and travel demand

We use the existing Albatross model (Arentze and Timmermans, 2004; Arentze and Timmermans, 2008), which has been calibrated for the Netherlands using annual census data and travel diaries, to generate a synthetic population of households with activity-travel patterns for the household members, which they are instructed to execute in our simulation model. Albatross uses 4-digit postal codes to identify locations of activities; we have created and numbered the zones in our network such that each such location can be directly mapped to a centroid, with postal codes further away from our study area merged into aggregate centroids. Albatross uses four modes for trips: car driver, car passenger, public transport and slow. We use this for the mode and route choices in our model as follows:

1. If the Albatross mode is car driver, only car routes are available in our route choice model.
2. If the Albatross mode is car passenger, we look at whether another member of the same household has a car driver trip in its schedule at the same time for the same origin-destination pair. If so, the car passenger trip is linked to the car driver trip. Otherwise, the car passenger trip is instead treated as a public transport trip.
3. If the Albatross mode is public transport, both public transport routes and the slow mode route are available in our route choice model.

4. If the Albatross mode is slow, the slow mode route is chosen in our model.

Because the Albatross data is sampled with ratio 1:7 for a weekday in the model's base year 2004 and our case study is based on a 2015 network, we use the following calibration factors to determine weights for each household:

- Per centroid, we define factors for household count targets based on the 2015 population data from Statistics Netherlands. We define these factors separately for the household types single-person households, two-person households and households with children, unless the total number of households living in the centroid is less than 30.
- We define factors for the total number of trip ends for the modes car driver and public transport, with targets taken from the municipal model.
- We separately define factors for the number of trip ends of the modes car driver and public transport within the study area, again with targets derived from the municipal model.
- Using the fastest paths between origins and destinations, based on free speeds and ignoring waiting times, we define factors for the modes car driver and public transport based on the amount of travellers on particular sets of links near the edge of the study area, namely:
  - car traffic on A4/A13 Midden-Delfland and A20 Vlaardingen-West – Maassluis;
  - car traffic on A15 Noordtunnel, A16 Drechttunnel, A29 Heidenoordtunnel, and N915 Brug over de Noord;
  - car traffic on A15 Botlektunnel/-bridge and S102 Spijkenisserbridge;
  - car traffic on A20 Nieuwerkerk a/d IJssel – Moordrecht;
  - train users Railwaybridge Dordrecht;
  - train users Rotterdam-Alexander – Gouda;
  - train/metro users Schiedam Centre – Delft-South, Rotterdam Central – Schiphol, and Meijersplein – Berkel en Rodenrijs;
  - metro/bus users Spijkenisserbridge and Zalmplaat – Spijkenisse.

The values of all factors, and thus the household weights, were solved by iterative proportional fitting (Kruithof, 1937). After determining the weights, all households not living
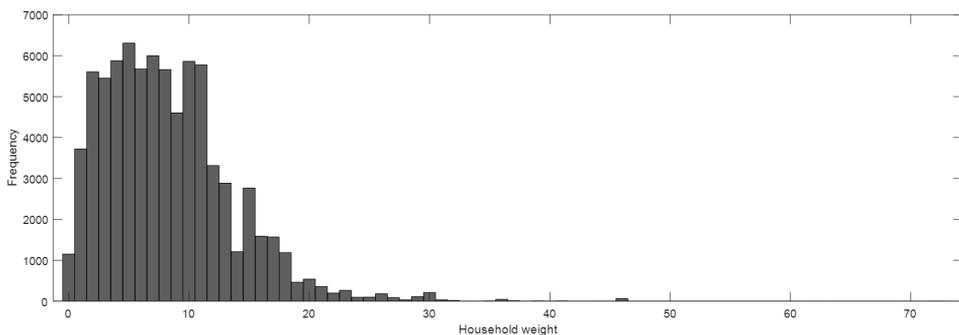

Figure 6.6. Calibrated household weights after rounding, excluding discarded households.

in the study area, not travelling to the area and not expected to travel through the area by either car or public transport are discarded. Since households with high weights may cause unrealistic local-momentous peaks in the travel demand, the weights of the remaining households are randomly rounded to integers, and households with weights greater than one are replaced with multiple copies with weight one, with trip departure times spaced 2 minutes apart and independent route choice preferences (thus smoothing the otherwise local-momentous peak). This results in 648,823 households. A histogram of the rounded weights is shown in Figure 6.6.

6.3.3   Disruption

The hypothetical disruption that we analyse in our case study involves an evacuation of the underground metro station Leuvehaven, indicated in Figure 6.7. Because of a limited number of switches in the metro tunnel, metro services are not only suspended for this station but also the next two stations south of it. Due to the need for emergency services at the site, we also assume that aboveground, the Schiedamsedijk road and its intersection with the S100 city centre ring road and the northern end of the Erasmusbridge will be unavailable as well, affecting car and tram traffic. The disruption lasts four hours from 16:00 to 20:00, thus coinciding with the afternoon peak hours.

The consequences of this disruption for public transport operation are depiced in Figure 6.8. Metro services have to turn around at stations Beurs in the north and Maashaven in the south, leaving stations Leuvehaven, Wilhelminaplein and Rijnhaven unserved. Tram services have to turn around at tram stops Beurs in the north, Vasteland in the west, and Wilhelminaplein in the southeast, leaving tram stop Leuvehaven unserved.
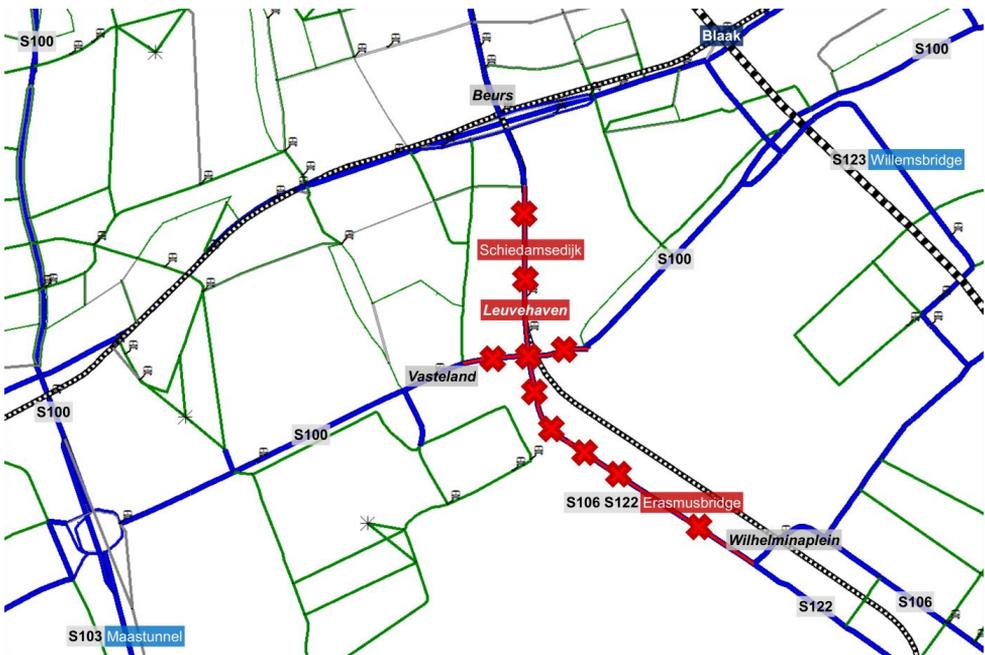
Figure 6.7. Road closures (red, crosses) during the disruption.

Excluding station/stop Leuvehaven, we therefore have five stops that are to be served by bus-bridging services: Beurs, Vasteland, Wilhelminaplein, Rijnhaven, and Maashaven. These buses are deployed from two bus depots: Kleiweg in the north and Sluisjesdijk in the south. We assume 10 buses are available, equally distributed over both depots. We chose bus-bridging routes between all stops and depots routing around the disaster site and using either the Maastunnel or the Willemsbridge instead of the unavailable Erasmusbridge. Buses depart from the depots at 16:30, half an hour after the start of the disruption. In the bus-bridging optimisation, the seated capacity of 44 psg/veh is used instead of the total capacity of 81 psg/veh, thus designing for a higher quality of service and incorporating a capacity-margin in case more people arrive between optimisation and bus departure or in case of errors in the estimation of destination fractions.
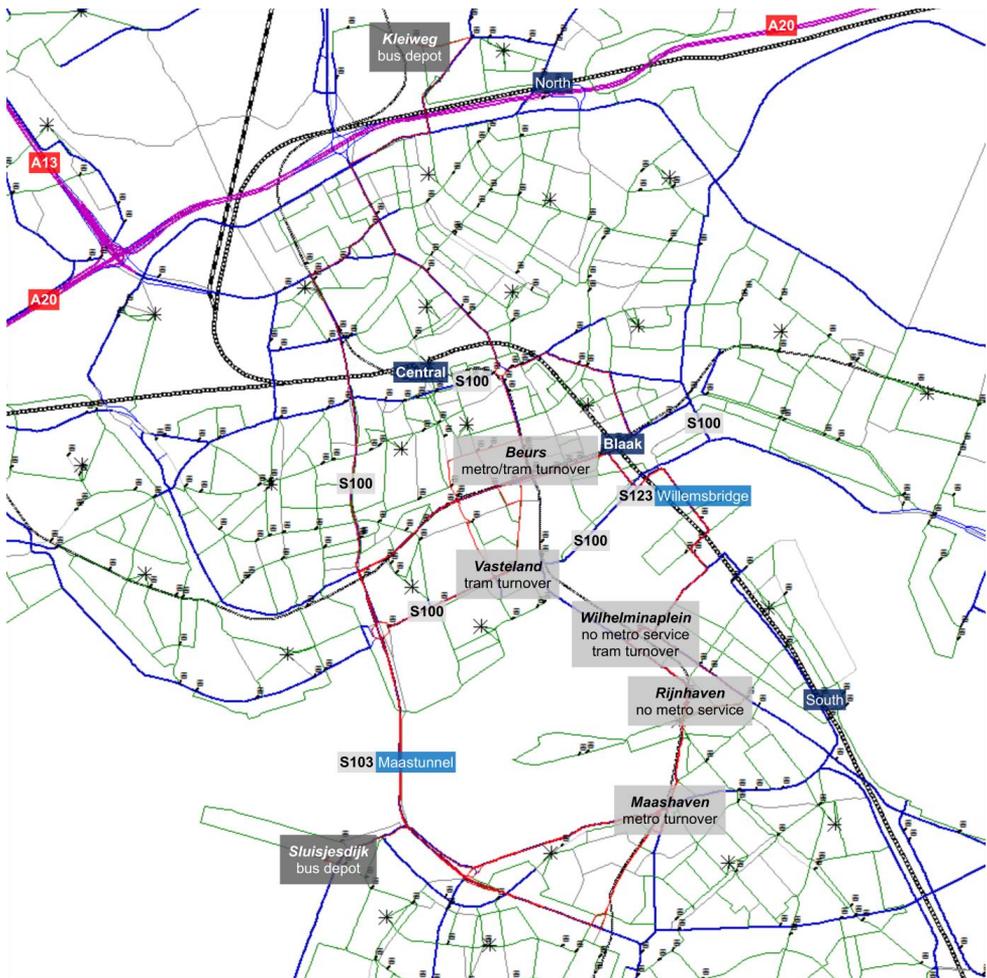
Figure 6.8. Bus routes (red lines), stops (light grey boxes) and depots (dark grey boxes) for bus bridging. Train stations (dark blue) are indicated but are not bus-bridging stops.

6.3.4   Simulation configuration

While an overview of the simulation methodology was already presented in Subsection 6.2.2, (case-specific) configuration details of the route set generation and the node model were not yet described. These descriptions are provided below.

The route choice set generation is as follows. For pedestrians, only shortest paths are generated (Dijkstra, 1959). For car drivers, routes are generated with the accelerated Monte Carlo method (Fiorenzo-Catalano and Van der Zijpp, 2001), using 300 iterations and a maximum dispersion coefficient of 0.12 minutes (0.002 hour), plus one extra shortest path search excluding links unavailable during the emergency scenario to ensure all travellers always have at least one option. For public transport users, we use a branch-and-bound algorithm to generate routes similar to Subsection 2.3.1, with additional constraints on transfers (walking distance between consecutively used lines must be minimal; no other modes are permitted between two train legs), a tolerance constraint of 1.05, a larger maximum number of transfers (7), and a larger maximum access/egress time dependent on the public transport mode. This is repeated for different times of day using different expected travel and waiting times based on the timetables in each period. To guarantee modal diversity in the public transport choice sets, the route set generation is also run separately with all public transport modes, only train, only train and metro and only train and tram. The public transport route set generation is also repeated excluding the disrupted segments during the emergency, both without and with bus-bridging services present (in the latter case, no bus-bridging waiting time is considered for route generation). The boarding penalty is set to 5 minutes.

For the node model, priority coefficients are based on the capacities of incoming links (Tampère et al., 2011), but we let access roads yield to distributor roads by multiplying the priority coefficients of access roads with 0.25. At signalised junctions, the signal capacities are used instead. For calculating signal schemes, we use a uniform loss time of 4 seconds, a maximum cycle time of 150 seconds, and a minimum green time of 5 seconds. When available, separate approach links for buses and trams are included, which are assumed to have priority and hence are not subject to traffic lights.

6.4   Results

We will now discuss the simulation results. First, Subsection 6.4.1 will analyse the results of our main model, as presented until now, including evaluation of the performance of the new bus-bridging scheduling method. Then, Subsection 6.4.2 will compare the results for different variations of the main model, analysing how various properties of the simulation model affect an ex-ante impact assessment of the disruption. Conclusions and implications follow in Section 6.5.

6.4.1   Main model

We first look at the flows and speeds in the road network, of which the averages between 16:00 and 20:00 are plotted in Figure 6.9 for both the disruption and a normal day. We see that the disruption in particular worsens the congestion in the northern side of Rotterdam, particularly urban roads leading towards the southbound directions of the Willemsbridge, Maastunnel and Van Brienenoordbridge. Around 16:30, queues in front of the remaining bridges and tunnels start building up, initially primarily towards the city centre but later also
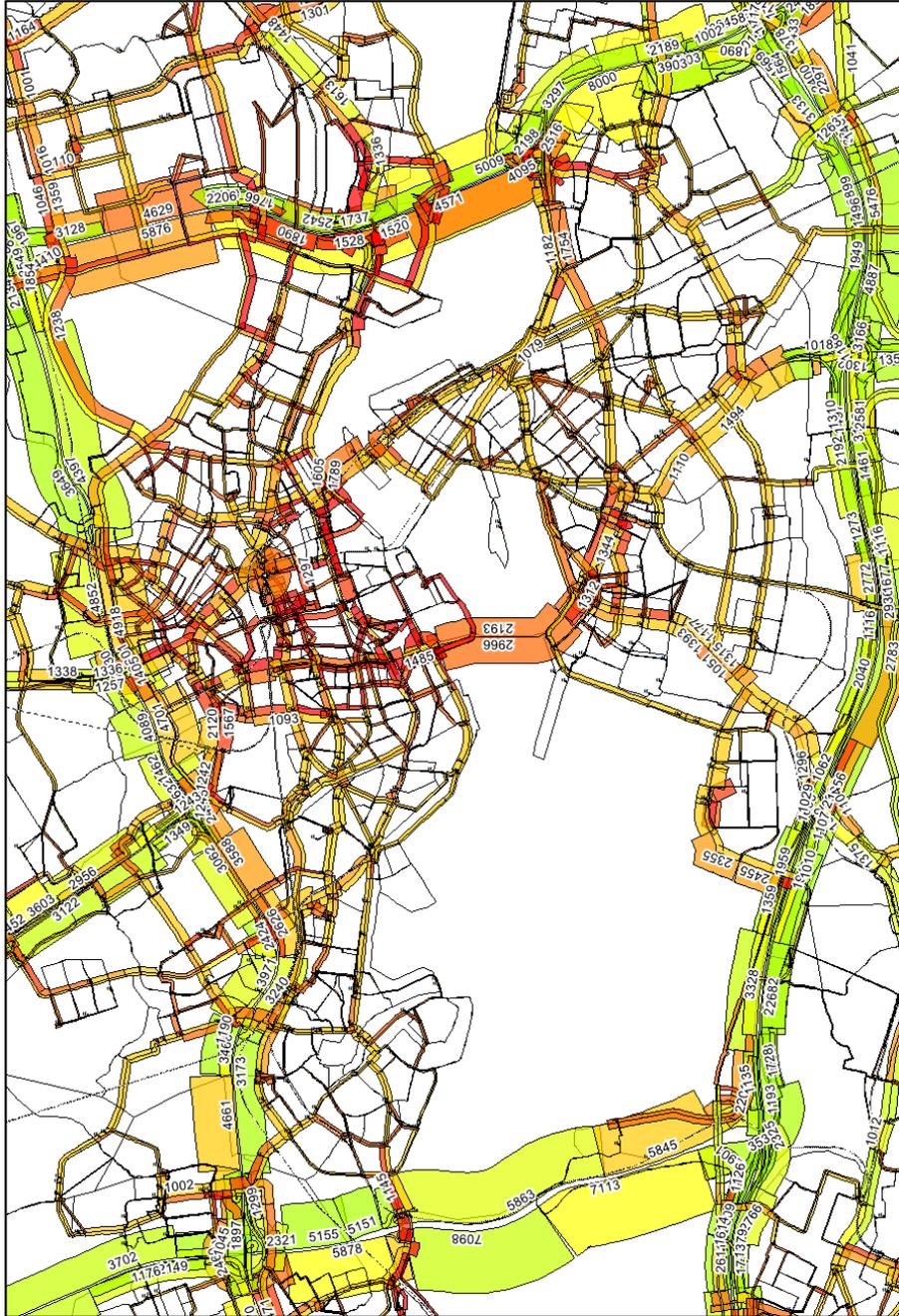
Figure 6.9. Average vehicular flows during the disruption (top) and the same period on a normal day (bottom).

Figure 6.10. Average public transport passenger flows (psg/h) during the disruption (top) and the same period on a
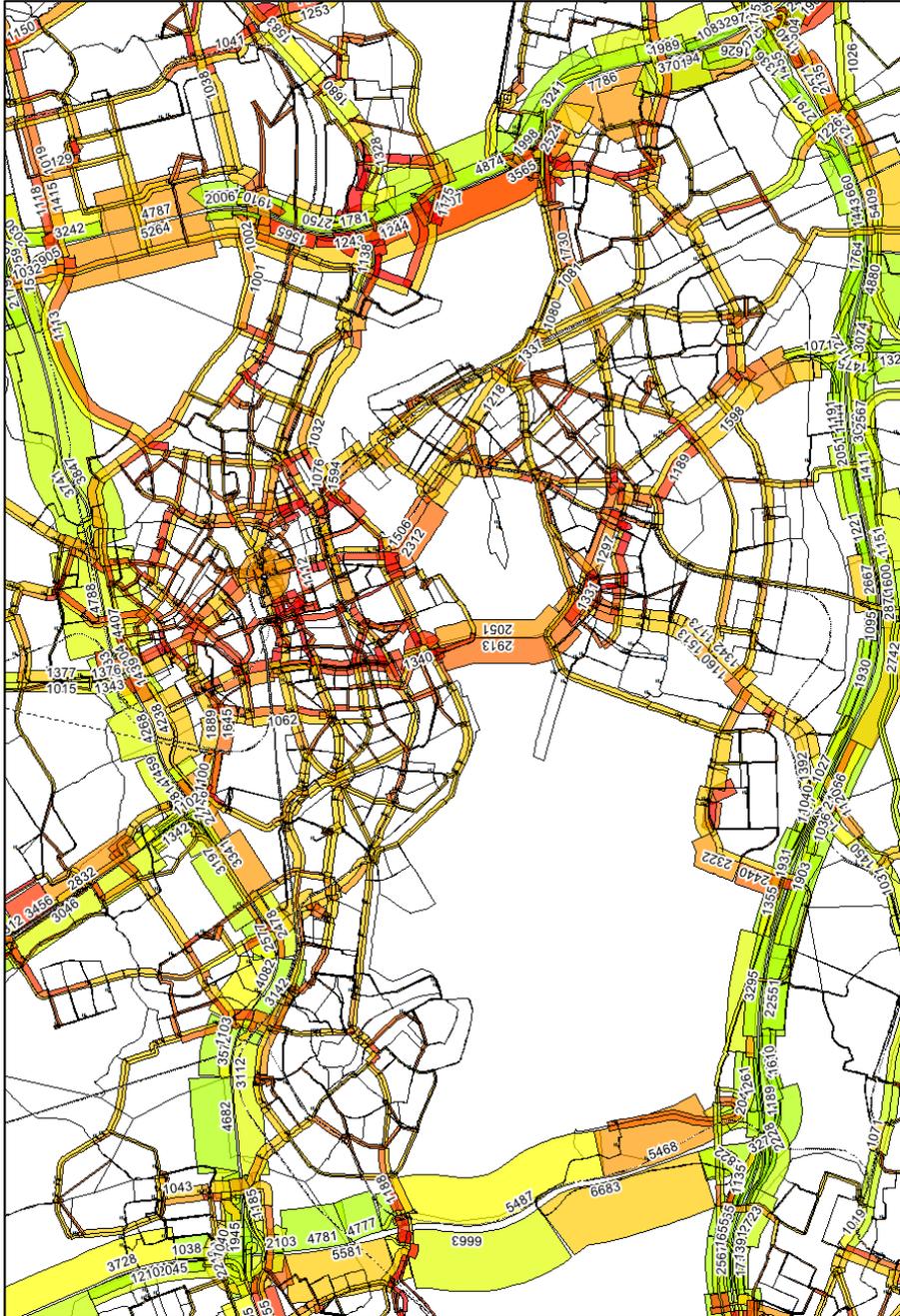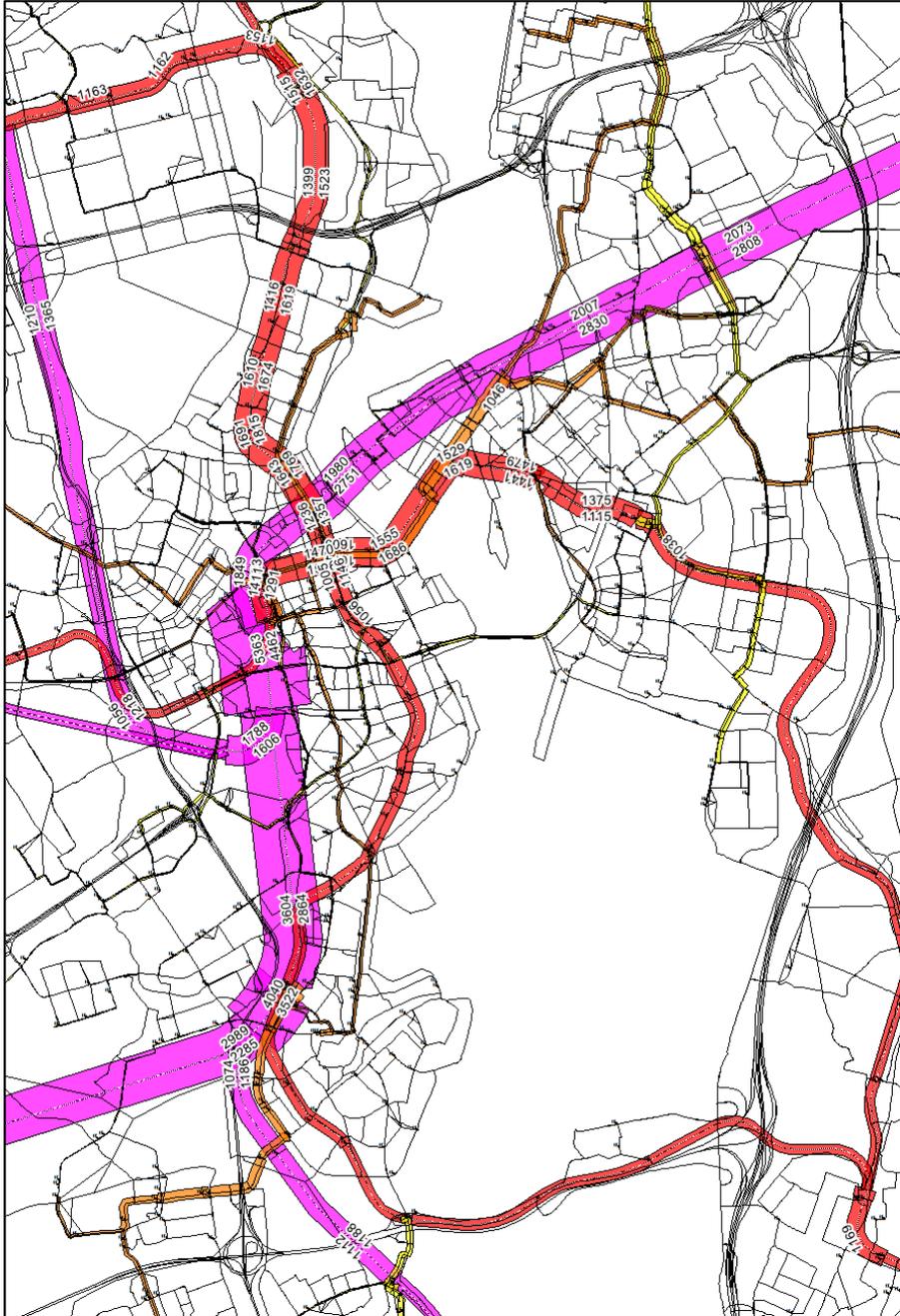
in other directions. In particular, the Willemsbridge queue spills back over the eastbound distributor parallel to the S100 city centre ring in the north. Like a normal day, the congestion peaks between 18:00 and 18:30, but the Maastunnel and Willemsbridge queues do not fully dissolve until around 20:30, i.e. after the re-opening of the Erasmusbridge at 20:00, whereas on a normal day those queues have fully dissolved by 19:30. During the disruption, all alternatives to the Erasmusbridge serve more traffic than usual. In the south of Rotterdam, we also observe congestion towards the north, but it is not much more extended than normally.

The average flows within the public transport system for the same time periods are plotted in Figure 6.10. The disrupted metro and tram lines serve much fewer passengers in the line segments that remain operational. A sizable increase in ridership is observed for the train, which apparently serves as an important alternative mode: considerably more people board at Central or Blaak and alight at South or Lombardijen, and vice versa. Usage of the train tunnel near the Willemsbridge peaks between 17:00 and 17:30, at which point it is serving about 2300 psg/h more than normally, and the high ridership numbers last longer. The metro lines through the Beneluxtunnel also form an important alternative route serving extra passengers, primarily travellers with origins or destinations in the southwest, with peak usage between 17:30 and 18:00 being about 650 psg/h more than normally. Closer to the disruption site, we also observe a considerable increase in usage of the regular bus line through the Maastunnel, primarily northbound, which begins very quickly after the start of the emergency.

Usage of the bus-bridging services is very minor compared to those other alternatives and is therefore hardly visible in Figure 6.10. The total number of passengers is listed in Table 6.3. The most important relation is from Beurs to Wilhelminaplein with 422 passengers, where they mostly continue travelling southeast with the truncated tram lines. The opposite direction is the next most important relation with 97 passengers. The bus-bridging lines between the metro turnover stations Maashaven and Beurs follow with 77 passengers northbound and 46 passengers southbound. The passenger counts are very low relative to the number of bus runs shown in Table 6.4, which, due to the low demand compared to the bus vehicle's capacity, does not show a concentration of services on particular origins or destinations. An explanation can be found in the long travel times of the bus-bridging services, shown in Table 6.5. This especially applies to bus routes via the Maastunnel, which, unlike the Willemsbridge, does not have bus lanes. Looking at the cells in Table 6.5, it would often have been faster to drive via an intermediate station over the Willemsbridge than directly through the Maastunnel.

We can also analyse the relative performance of the transportation system in the disaster scenario by looking at aggregate network statistics. Here, we use three types of indicators:

- output, or total number of passengers boarded and alighted, with each public transport mode;
- production, or total distance travelled, with each mode, and;
- accumulation, or total time spent, within and waiting for each mode.

The difference in these indicators compared to a normal undisrupted day is displayed in Figure 6.11. We see that the decrease in metro and tram usage is indeed met primarily with an increase in train usage, and with small increases in bus and bus-bridging usage. The increased congestion leads to more time spent in car traffic. The distance travelled by car also increases, likely due to detours to avoid the increased congestion. We also see here that the distance travelled and time spent walking also increase a lot. Some of this is likely a result of the

Table 6.3. Total bus-bridging passenger counts on all runs.

	To					
From		Beurs	Vasteland	Wilhelminaplein	Rijnhaven	Maashaven	All
	Beurs		3	422	0	46	**471**
	Vasteland	9		19	10	16	**54**
	Wilhelminaplein	97	15		0	8	**120**
	Rijnhaven	23	9	0		27	**59**
	Maashaven	77	35	17	35		**164**
	All	**206**	**62**	**458**	**45**	**97**	**868**

Table 6.4. Number of bus-bridging runs.

	To					
From		Beurs	Vasteland	Wilhelminaplein	Rijnhaven	Maashaven	All
	Beurs		9	13	8	9	**39**
	Vasteland	11		9	9	8	**37**
	Wilhelminaplein	8	10		14	10	**42**
	Rijnhaven	9	10	10		15	**44**
	Maashaven	10	7	9	13		**39**
	All	**38**	**36**	**41**	**44**	**42**	**201**

Table 6.5. Average realised bus-bridging travel time over all runs (minutes:seconds).

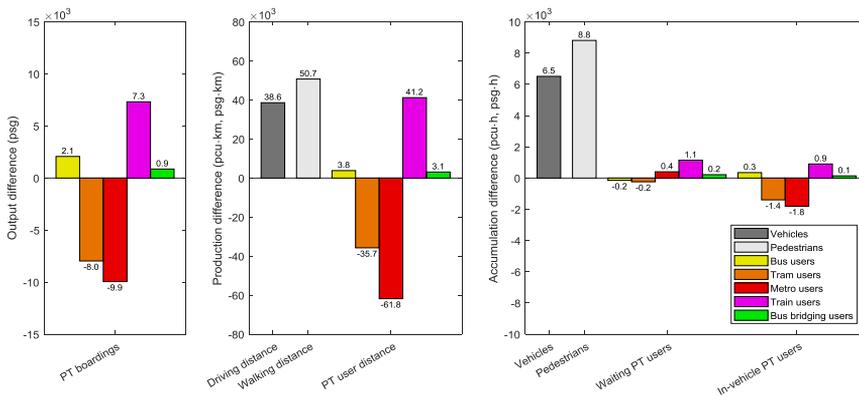|	To					
From		Beurs	Vasteland	Wilhelminaplein	Rijnhaven	Maashaven	All
	Beurs		2:23	8:14	33:09	33:04	**17:43**
	Vasteland	4:54		10:09	12:59	13:57	**10:06**
	Wilhelminaplein	7:51	5:58		13:28	2:17	**4:03**
	Rijnhaven	13:28	8:58	0:52		1:29	**5:30**
	Maashaven	18:46	12:03	2:24	1:25		**8:00**
	All	**11:12**	**7:05**	**5:35**	**9:40**	**10:49**	**8:54**

Figure 6.11. Differences of emergency scenario compared to a normal day.

increase in train usage increasing access/egress on foot, but when visualising network flows for pedestrians (not shown here), we also see significant increases in the pedestrian flows over the Erasmusbridge and through the Maastunnel, who are apparently satisfied by neither bus-bridging services nor other public transport services.

6.4.2   Model sensitivity analyses

To further analyse (the sensitivity of) the above results, in this subsection we perform sensitivity analyses towards model properties. These properties represent specific behavioural, control and simulation features that are included in the main model, but could also be excluded. Thus these sensitivity analyses show the importance of accounting for these features. Note that these model variations affect all simulations, including those of the undisrupted situations used as reference. Because the model variations affect the entire network over all time, the comparisons will focus on the aggregate indicators displayed in Figure 6.11. In case sizable differences at specific locations and times are observed, these are mentioned as well.

No capacity drop

The first simulation model property to investigate is removal of the capacity drop indicated in Table 6.1, i.e. fixing the motorway capacity to 2100 pcu/h/ln regardless of whether there is a moving jam, a standing queue, or no congestion. Despite the slight reduction in the motorway capacity in free-flow conditions and slight increase in moving-jam conditions, the resulting aggregate indicators in Figure 6.12, when compared to Figure 6.11, show that removal of the capacity drop leads to a smaller increase in time spent in car traffic. This is an expected effect: without capacity drops, any additional congestion in the emergency scenario will cause less additional delays. Nevertheless, the overall traffic pattern in the network looks very similar. There is a slightly larger increase in bus ridership, but overall the impacts on public transport are also minor.

No subcritical delays

The next model property to test is the subcritical delays in the road network, i.e. without the speed on road links linearly decreasing as a function of traffic density in subcritical traffic conditions. This increase of the critical speed to the free speed lowers travel time when road capacity is not exceeded. The resulting indicators are shown in Figure 6.13. We see a considerably larger increase in time spent in car traffic, which is a logical consequence of a stronger increase in road travel times compared to a normal day. An explanation is that the time spent in the emergency scenario is dominated by supercritical conditions, which differ more strongly from the subcritical conditions more common on a normal day if there are no subcritical delays. Within public transport, tram usage decreases a bit more and bus usage increases a bit more, but no other clear differences can be seen. Overall, the differences in public transport usage between the models are thus minor, likely because near the disruption site, most travel time is spent in congestion, whereas this model feature primarily affects travel times in the free-flow regime.

No traffic lights

Another model property is inclusion of traffic lights. Crossing conflicts at junctions are hence here ignored to test this feature. We do preserve the separate incoming links for separate
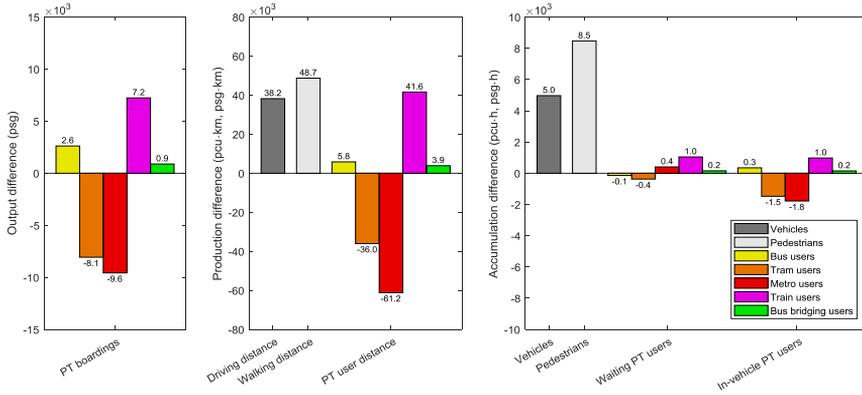
Figure 6.12. Differences of the emergency scenario compared to a normal day in the model without capacity drop.
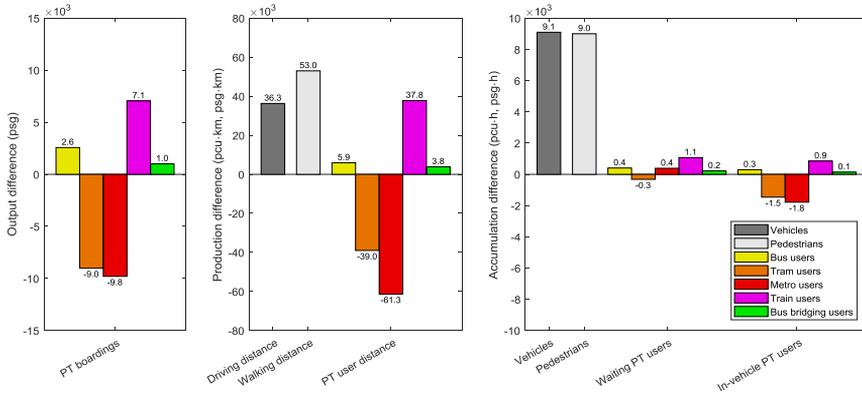

Figure 6.13. Differences of the emergency scenario compared to a normal day in the model without subcritical delays.
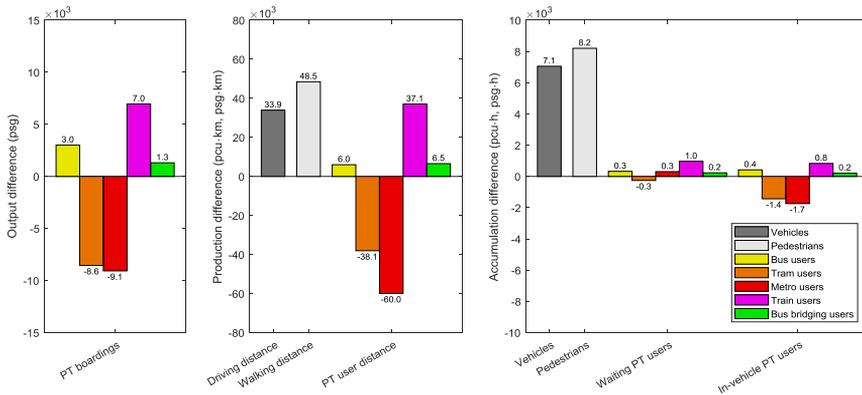

Figure 6.14. Differences of the emergency scenario compared to a normal day in the model without traffic lights.

approaches to the previously signalised intersections. The results are in Figure 6.14, showing that without modelling the traffic lights, the disruption seems worse compared to a normal day. An explanation is that the emergency scenario is dominated by merging conflicts towards the remaining bridges and tunnels, which are accounted for in both models, while crossing conflicts at signalised intersections are more constraining in the normal day scenario, which are now omitted. Therefore, removal of traffic lights may overestimate the differences between both scenarios.

Looking at the network flows, we note that the Maastunnel has a higher throughput during the disruption in this model variant, due to lack of spillback from signalised intersections into the tunnel. (Better signal controllers might avoid this in the main model.) The southbound direction even nearly constantly operates at its capacity of 3600 veh/h. The higher throughput appears to result in lower travel times that increase the use of both regular bus services and bus-bridging services, which is visible in Figure 6.14. Indeed, the bus-bridging travel time from Beurs to Maashaven through the Maastunnel is now only 23:15 minutes, and 9:19 minutes vice versa (compared to 33:04 minutes and 18:46 minutes when traffic lights are modelled).

No en-route choice for cars

In the main model, car drivers can update their route choice based on the prevailing travel times every minute while en-route. In the following, we analyse what happens when this capability is removed and car drivers remain on the route initially selected upon departure. To test this feature, we have to allow people to drive over the closed links if they chose the route before the links were closed. It is clear from Figure 6.15 that excluding this model property is problematic. Visualisation of network flows over time reveals the cause: the emergency scenario now results in complete gridlock, originating in the city centre of Rotterdam and spreading out from there, as car drivers do not choose routes avoiding the affected links and area.

No car traffic

Finally, we look at how the results are impacted if we omit car traffic from the model, thus not assigning car users to the network and only assigning buses and trams. Naturally, this results in substantial improvements in road travel times, and according to Figure 6.16 this leads to less pedestrians and train users and more bus users and bus-bridging users in comparison with the main model. As the average bus-bridging travel time of just 4:22 minutes allows fast and frequent services, the number of bus-bridging users spectacularly increases to 4587, even exceeding the extra number of train users in this model variant. Particularly the service between the metro turnover stations through the previously congested Maastunnel is now very popular, with 1437 passengers from Beurs to Maashaven and 933 vice versa. The increased use of bus-bridging services is also clearly visible in Figure 6.17.

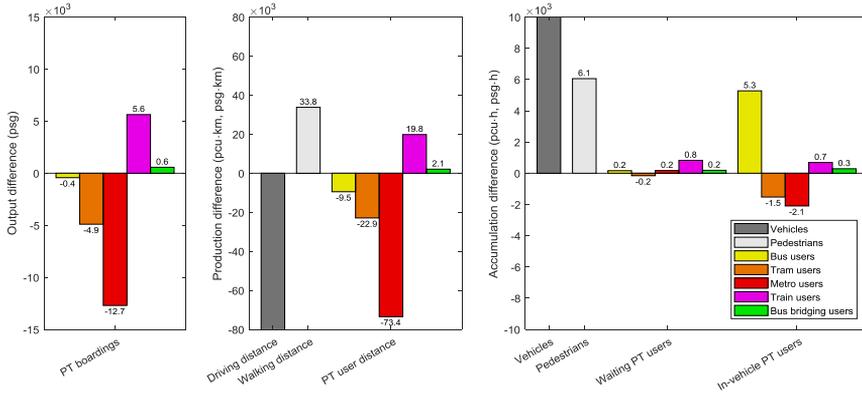**Figure 6.15. Differences of the emergency scenario compared to a normal day in the model without en-route choice for cars.**

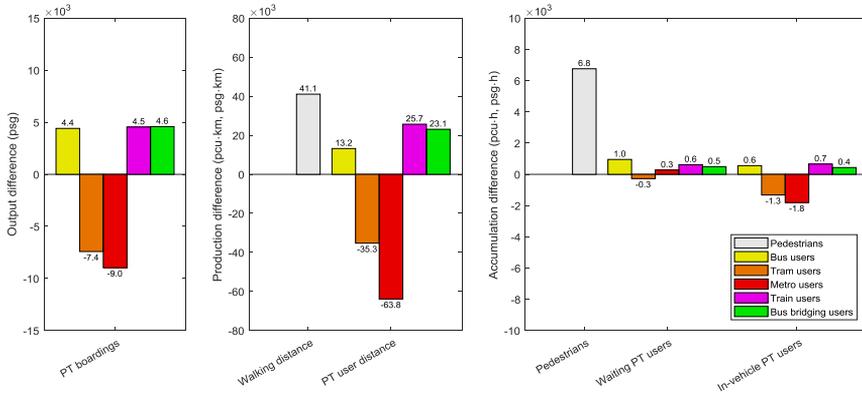**Figure 6.16. Differences of the emergency scenario compared to a normal day in the model without car traffic.**

Figure 6.17. Average public transport passenger flows (veh/h) during the disruption in the model without car traffic. Bus bridging is indicated in green.

6.5 Discussion and conclusion

In this chapter, we extended a static bus-bridging optimisation model into a novel real-time scheduling method; analysed the impacts of a major disruption of metro and tram operations with bus bridging in the city of Rotterdam, the Netherlands; and tested how these impacts relate to specific properties of the agent-based dynamic multimodal simulation model.

This case study shows that in Rotterdam, existing alternative routes for public transport users play a critical role. In all our model variants, regular bus and train services combined served more stranded passengers than the bus-bridging services. In that sense, providing bus-bridging services between the affected stations is not of primary importance (but can be considered as complementary), and one may prefer to focus on efficiently handling the extra passengers on alternative existing lines and modes.

In terms of simulation model features, the capacity drop on motorways and the difference between free speed and critical speed had substantial impacts on the extra delays for car users compared to a normal day. The differences are attributable to the additional congestion in the emergency scenario and the fact that these model features have different consequences for free-flow traffic and congested traffic. Their impacts on public transport users were however limited. Larger public transport impacts were found for the modelling of signalised intersections, so we can recommend detailed modelling of those with realistic signal controllers, ideally including their network coordination so that signal phases can be simulated explicitly. The relevance for public transport stems from the resulting travel times, which affects the extent to which road-bound public transport, including bus bridging, can compete against alternative rail-bound public transport for the stranded passengers. This is also clear from the simulation results without car traffic, where bus and bus-bridging services are consequently able to attract many more stranded passengers. En-route choice for car traffic was also shown to be of critical importance, as lack thereof created unrealistic gridlock during the disruption.

For other networks, this means that it is very important to assess the availability and capacity of alternative public transport routes, and whether/how they can cooperate with bus-bridging services. Alternative public transport routes are likely to be more effective in most non-radial urban public transport networks, especially if the scale of the disruption is local enough for walking to form a significant part of an alternative route. On the other hand, as identified by Pender et al. (2013), bus bridging may be more important in suburban and inter-city transportation where viable parallel systems tend to be less common. As to the simulation model features, the road travel times are of key importance. They affect the efficiency of bus bridging itself, and if bus bridging competes with alternative public transport routes, especially when rail-based, they affect the distribution of passengers over both systems. Which model features are most important for these travel times can be expected to be somewhat network-dependent, e.g. if bus(-bridging) services run on motorways, the capacity drop may be more important and signalised intersections may be less important.

Analysis of the results also suggests that significant improvements to the proposed real-time bus-bridging scheduling are possible, such as varying the routes of origin-destination pairs based on current travel times to avoid traffic jams. Offering bus-bridging services with intermediate stops instead of only direct origin-destination services can be expected to increase bus-bridging efficiency, but requires a reformulation of the static optimisation problem. It is also wise to somehow detect changes in the destination fractions compared to an undisrupted situation, as the case shows that bus-bridging demand is rather elastic as

stranded passengers also divert to other regular public transport and/or walking. To this end, it is also worth investigating how bus bridging can be scheduled to efficiently work together with alternative public transport routes. Finally, for future research we recommend comparing the impacts of different settings of parameters that control the time horizon of the real-time bus-bridging scheduling.

Acknowledgements

Chapter 7

Conclusions and recommendations

This thesis has identified requirements for, developed, and tested simulation models and components that are able to analyse how emergencies affect transportation and able to evaluate candidate transportation management plans. In this final chapter, the main findings are synthesised, its conclusions and implications are discussed, and some directions for future research are recommended.

7.1 Summary and main findings

The main body of this thesis can roughly be divided into three parts: development of the simulation framework (Chapter 2), improvements to the simulation model (Chapters 3-5), and application for optimisation purposes (Chapter 6). The main findings and conclusions are synthesised for each part.

7.1.1 Simulation framework

Chapter 2 developed a general methodological framework for simulating transportation networks during emergencies. To do so, it first identified seven main requirements, namely the need for the transportation model

- to be dynamic,
- to describe the (effects of) relevant choice behaviour,
- to predict or otherwise allow specification of initial conditions of the emergency,
- to include interactions between individuals, e.g. members of the same household,
- to include multiple modes of transportation and their interaction,
- to include travellers who are not directly affected and their behaviour, and
- to include emergency services.

The methodology consists of two main components. First, it includes an activity-based choice model, rooted in activity-travel patterns on a normal day, to describe choice behaviour at the level of agents, i.e. individual travellers or groups of them. To model the escalation of traveller's responses to the emergency, they can have normal behaviour, adaptation behaviour, or evacuation behaviour. Second, it includes a macroscopic or mesoscopic dynamic network loading model that can include multiple modes and emergency services.

Both can be subject to any (effects of) control measures undertaken by authorities in an attempt to obtain a better value of their goal function (Chapter 1). Together, this resulted in a simulation model that is suitable for computationally efficient simulation of a wide range of emergencies.

7.1.2   Model improvements

Chapter 3 considered the problem of coupling agent-based choice models with macroscopic or mesoscopic dynamic network loading in more detail. In case the possibility of en-route choice needs to be preserved, of which emergency simulation is an example, this leads to an interesting problem where the traffic composition needs to be disaggregated into commodities that represent the individual agents in the aggregated traffic flow. It is found that consistency issues can easily arise in either the traffic flow evolution, the propagation of agents, or both. Nevertheless, it is possible to address the problem with suitable model design choices, for example with the Link Transmission Model with small time steps, a strict weak ordering of vehicle parts, an incremental node model, and a relaxed first-in-first-out constraint.

In Chapters 4-5, the discrete-time Link Transmission Model has been extended to enhance its level of detail, resulting in three main findings:

- The Link Transmission Model can be extended from triangular fundamental diagrams to continuous concave fundamental diagrams. Unlike previous attempts in literature, this can be done without deviating from Lighthill-Whitham-Richards traffic flow theory. This extension for example allows the gradual reduction of speed on a road as it gets loaded up to capacity, and the dispersion of platoons.

- Using an extension of Lighthill-Whitham-Richards theory supporting inverted-lambda style fundamental diagrams, the Link Transmission Model can be further extended to also include a capacity drop once congestion has formed. This allows the production of stop-and-go waves in the first-order model, resulting in traffic patterns combining moving jams and standing queues.

- Based on the Smulders fundamental diagram shape, the Link Transmission Model can be extended to support non-empty initial conditions and computation of within-link densities. Together, this allows arbitrary changes to the driving behaviour on links during the simulation, for example in response to traffic management or changing environmental conditions.

None of these extensions affect the numerical accuracy of the Link Transmission Model, so that in all cases shockwaves between traffic states remain crisp and the model remains suitable for the simulation of large and general traffic networks as needed for our emergency simulations.

7.1.3   Optimisation application

Chapter 6 demonstrates the feasibility of using the framework to simulate multimodal transportation disruptions in large urban areas, including signalised intersections that were previously not modelled but do affect the simulation results. It also showcases how real-time optimisation during the emergency, in this case for bus bridging, can be embedded in the simulation framework as a control component. This serves as an example of how emergency plans with tactical decisions can be evaluated and how such simulations produce suggestions

for improved tactics, which can then be tested again. Furthermore, the above extensions of the Link Transmission Model are indeed found to have an impact on car traffic.

7.2   Research conclusions and implications

First, we can draw implications particularly relevant for model-based emergency management. The case study applications show that specific traffic flow interactions on the network, that have hitherto often been ignored, can have substantial impacts and hence should be explicitly modelled. These interactions relate to different traffic flows (of outbound, inbound, and background traffic), emergency services, and different transport modes.

In case of the evacuation in Chapter 2, interaction of outbound traffic, inbound traffic and background traffic causes congestion to spill back such that the roads towards the evacuation area become congested. This implies that such possible interactions between different types of traffic are important to take into account, as they can have a significant impact on the congestion pattern.

Additionally, the impact of emergency services heading towards the disaster site is to be considered, which is further supported by the case of the traffic accident on the motorway in Chapter 5, where traffic managers have to close the peak-hour lane upstream of the incident to ensure it can be used by emergency services, accelerating the spillback of congestion. Both of these cases in Chapters 2 and 5 show that it is advisable for practitioners to consider how queues build up over time, whether this affects emergency services heading towards the emergency, and what can be done to avoid this.

Chapters 2 and 6 show that it is also important to look for possible interactions between modes: congested road traffic may considerably slow down public transport, which in turn shifts demand to other public transport lines that have less delay, potentially leading to less efficient overall use of the public transport system capacity. In Chapter 6 road travel times were shown to be critical to the success of urban bus bridging in multimodal disruptions. While Chapter 2 suggests that encouraging car owners to evacuate by public transport may speed up evacuations, such efforts are less likely to succeed if the public transport suffers from the same delays as car traffic.

Secondly, we can draw implications particularly relevant for (emergency) modelling. The findings of Chapter 3 have implications for developers of agent-based models using first-order traffic flow modelling. The chapter shows that it is possible to develop accurate, computationally-efficient, first-order models propagating agents according to macroscopic traffic flow rules, but that inconsistencies can easily arise if the model is not carefully designed. Model developers are therefore advised to pay special attention to avoid systematic errors, limit numerical diffusion, prevent agents from traversing links faster than the free speed, carefully specify the ordering of vehicles on a link, choose an appropriate node model form, and prevent gridlock due to agents being unable to change routes. Agent-based en-route choice was indeed shown to prevent gridlock in Chapter 6.

Thirdly, we can draw implications for those who wish to adopt the simulation models and components developed and tested in this thesis. Although this thesis offers a flexible simulation framework for emergencies, it can still be a significant amount of work to set up such a simulation model, as can be seen in Chapter 2 and Chapter 6. Notably, the framework depends on an activity-based choice model for normal-day travel behaviour, and setting up a

dynamic traffic assignment model for an urban region is also a significant amount of work, even more so if the normal-day simulation needs to be calibrated and validated. While these two case studies in this thesis are more of an exploratory nature and hence lack serious calibration and validation efforts, this is problematic when much more accurate and detailed results are needed for testing and improving real emergency plans in practice. Since both activity-based choice models and dynamic traffic assignment models are not yet commonly used on large urban areas, practitioners willing to apply the framework offered in this thesis may need to significantly invest in their normal-day transportation and traffic models first. When such investments have already been made, extension of the normal-day models into an emergency simulation framework is relatively easy, and the results can be easily compared both between the models and with data collected in the real world.

7.3   Directions for future research

For future research, there are many possibilities in terms of practical applications of the modelling framework beyond the three case studies in this thesis. For example, we did not yet compare different control strategies, or optimise a disaster plan based on those results, as suggested in Chapter 1. One may obtain valuable practical knowledge from such future research, such as insights into what kind of control measures tend to be effective, in what combinations, how much can be gained from applying them, and to what extent this depends on specific characteristics of the emergency situation. For such model applications, it may also be interesting to run the model multiple times in a Monte Carlo setting to get an estimate of the certainty of the outcomes.

Aside from model applications, there are also various opportunities to improve and extend the methodology itself. These are discussed below.

For the choice modelling within our emergency simulation framework, one important recommendation for future research is to further develop these choice models mentioned in this thesis and calibrate them for various emergencies. The case studies in Chapters 2 and 6 used relatively simple choice models based on normal-day activity-travel patterns, but there is room for more advanced implementations. For example, the scheduling of trips required for evacuations can be further detailed and the adjustment of activities in response to disruptions could be included as well. Evidently, this will be challenging due to the infrequent occurrence of emergencies and the consequent lack of detailed empirical data that can be used for calibration of choice models. One possible way to overcome this is a stated-choice experiment, where the respondents are asked what choices they would make over time during an array of hypothetical emergencies for which all relevant data such as travel times are known. But even then, it is a challenge to estimate how much of that information would be available to the decision-maker in reality, and with what accuracy.

Another recommendation is to further improve multimodal equilibrium computation for agent-based choice models. In Chapter 2, achieving a reasonable equilibrium for a normal day took a large number of iterations. In Chapter 6, which had a much larger network and included signalised intersections, this would have been even more difficult. Nevertheless, computation of the normal-day equilibrium allows producing more accurate initial conditions for the emergency and permits the choice model to separate normal choice behaviour and adaptation behaviour in response to the emergency, which allows testing and comparing scenarios with varying degrees of public awareness of the emergency. Additionally, it potentially results in a model that can be used to quickly assess the impact of traffic

management on the normal-day equilibrium of a larger area, reducing the gap between transportation models used for strategic planning and the more detailed traffic models used for real-time control.

Finally, it can be recommended to further analyse the validity of the Link Transmission Model extensions. Particularly in case of the capacity drop, the model indeed produced stop-and-go waves in Chapter 4 where they would also occur in reality, but their severity and spacing may differ between simulation and reality. Difficulties here are that in reality, the capacity of a bottleneck is stochastic, capacity funnels may delay breakdowns, acceleration and deceleration rates are finite, and responses to link-based traffic management measures as in Chapter 5 are not instantaneous. While the Link Transmission Model extensions are expected to improve the response of the simulation model to traffic management measures, one should thus be careful when analysing the quantitative impacts of very targeted local measures. It is possible that higher-order models are needed for some applications, in which case there is a challenge in making this compatible with the guidelines for agent-based applications of Chapter 3. Multi-class traffic flow models could also enhance the accuracy of the traffic simulation, for example in multimodal emergencies, but again make it difficult to satisfy Chapter 3.

References

Abdelgawad, H., Abdulhai, B., Wahba, M. (2010). Multiobjective Optimization for Multimodal Evacuation. *Transportation Research Record*, 2196, 21-33.

Alvarez-Icaza, L., Islas, G.J. (2013). Hysteretic Cell Transmission Model. 16th International IEEE Annual Conference on Intelligent Transportation Systems, pp. 578-583. The Hague, the Netherlands.

Arentze, T.A., Timmermans, H.J.P. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B*, 38(7), 613-633.

Arentze, T.A., Timmermans, H.J.P. (2008). ALBATROSS: overview of the model, application and experiences. 2nd TRB Conference on Innovations in Travel Modeling. Portland, Oregon.

Baxter, P.J., Ancia, A. (2002). *Human Health and Vulnerability in the Nyiragongo Volcano Crisis, Democratic Republic of Congo, 2002.* World Health Organisation.

Bekhor, S., Kheifits, L., Sorani, M. (2014). Stability analysis of activity-based models: case study of the Tel Aviv transportation model. *European Journal of Transport and Infrastructure Research*, 14(4), 311-331.

Bish, D.R., Sherali, H.D., Hobeika, A.G. (2014). Optimal evacuation planning using staging and routing. *Journal of the Operational Research Society*, 65(1), 1-17.

Bliemer, M.C.J. (2007). Dynamic Queuing and Spillback in Analytical Multiclass Dynamic Network Loading Model. *Transportation Research Record*, 2029, 14-21.

Bliemer, M.C.J., Raadsen, M.P.H. (2018). Continuous-time general link transmission model with simplified fanning, part I: Theory and link model formulation. *Transportation Research Part B*.

Bliemer, M.C.J., Raadsen, M.P.H., Bell, M.G.H. (2016). Discrete and continuous time formulations of the link transmission model class. 6th International Symposium on Dynamic Traffic Assignment. Sydney, Australia.

Bowman, J.L. (2009). Historical development of activity-based models: theory and practice. *Traffic Engineering and Control*, 50, 59-62 & 314-318.

Brachman, M.L., Dragicevic, S. (2014). A spatially explicit network science model for emergency evacuation in an urban context. *Computers, Environment and Urban Systems*, 44, 15-26.

Buckley, D.J., Yager, S. (1974). Capacity Funnels near On-Ramps. In: Buckley, D.J. (ed.), *Transportation & Traffic Theory: Proceedings of the Sixth International Symposium on Transportation and Traffic Flow Theory*. Elsevier, pp. 105-123.

Burghout, W., Koutsopoulos, H.N., Andreasson, I. (2006). A Discrete-Event Mesoscopic Traffic Simulation Model for Hybrid Traffic Simulation. 9th IEEE Intelligent Transportation Systems Conference, pp. 1102-1107. Toronto, Ontario.

Carey, M., Watling, D. (2012). Dynamic traffic assignment approximating the kinematic wave model: System optimum, marginal costs, externalities and tolls. *Transportation Research Part B*, 46(5), 634-648.

Cetin, N., Burri, A., Nagel, K. (2003). Parallel Queue Model Approach to Traffic Microsimulations. 82nd Transportation Research Board Annual Meeting. Washington, D.C.

Charypar, D., Axhausen, W.H., Nagel, K. (2007). Event-Driven Queue-Based Traffic Flow Microsimulation. *Transportation Research Record*, 2003, 35-40.

Chiba, Y., Shaw, R., Banba, M. (2017). Japan's Experiences of Catastrophic Mountain Disasters in Wakayama. In: Banba, M., Shaw, R. (eds.), *Land Use Management in Disaster Risk Reduction*. 1st edn. Springer Japan, pp. 215-235.

Chiu, Y., Zhou, L., Song, H. (2010). Development and calibration of the Anisotropic Mesoscopic Simulation model for uninterrupted flow facilities. *Transportation Research Part B*, 44(1), 152-174.

Chung, E., Ohtani, O., Warita, H., Kuwahara, M., Morita, H. (2006). Does Weather Affect Highway Capacity? In *Proceedings of the 5th International Symposium on Highway Capacity and Quality of Service*. Japan Society of Traffic Engineers, pp. 139-146.

Chung, K., Rudjanakanoknad, J., Cassidy, M.J. (2007). Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B*, 41(1), 82-95.

Chung, Y. (2012). Assessment of non-recurrent congestion caused by precipitation using archived weather and traffic flow data. *Transport Policy*, 19, 167-173.

Cominetti, R., Correa, J. (2001). Common-Lines and Passenger Assignment in Congested Transit Networks. *Transportation Science*, 35(3), 250-267.

Courant, R., Friedrichs, K., Lewy, H. (1928). Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, 100, 32-74.

Cremer, M. (1979). *Der Verkehrsfluß auf Schnellstraßen: Modelle, Überwachung, Regelung*. Springer-Verlag.

Daganzo, C.F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B*, 28(4), 269-287.

Daganzo, C.F. (1995a). The cell transmission model, part II: Network traffic. *Transportation Research Part B*, 29(2), 79-93.

Daganzo, C.F. (1995b). A finite difference approximation of the kinematic wave model of traffic flow. *Transportation Research Part B*, 29(4), 261-276.

Daganzo, C.F. (1999). The Lagged Cell-Transmission Model. In: Ceder, A. (ed.), *Transportation and Traffic Theory: Proceedings of the 14th International Symposium on Transportation and Traffic Theory*. Pergamon, Oxford, pp. 81-104.

Daganzo, C.F. (2005a). A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B*, 39(2), 187-196.

Daganzo, C.F. (2005b). A variational formulation of kinematic waves: Solution methods. *Transportation Research Part B*, 39(10), 934-950.

Deka, D., Carnegie, J.A. (2010). Analyzing Evacuation Behavior of Transportation-Disadvantaged Populations in Northern New Jersey. 89th Transportation Research Board Annual Meeting. Washington, D.C.

Dijkstra, E.W. (1959). A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, 1(1), 269-271.

Dixit, V.V., Gayah, V.V., Radwan, E. (2012a). Comparison of Driver Behavior by Time of Day and Wet Pavement Conditions. *Journal of Transportation Engineering*, 138(8), 1023-1029.

Dixit, V.V., Wilmot, C., Wolshon, B. (2012b). Modeling Risk Attitudes in Evacuation Departure Choices. *Transportation Research Record*, 2312, 159-163.

Dombroski, M., Fischhoff, B., Fischbeck, P. (2006). Predicting Emergency Evacuation and Sheltering Behavior: A Structured Analytical Approach. *Risk Analysis*, 26(6), 1675-1688.

Edie, L.C. (1961). Car-Following and Steady-State Theory for Noncongested Traffic. *Operations Research*, 9(1), 66-76.

Evans, L.C. (2002). Nonlinear First-Order PDE. In *Partial Differential Equations*. American Mathematical Society, pp. 94-174.

Fiorenzo-Catalano, S., Van der Zijpp, N. (2001). A Forecasting Model for Inland Navigation Based on Route Enumeration. 29th European Transport Conference. Cambridge, United Kingdom.

Flötteröd, G., Lämmel, G. (2015). Bidirectional pedestrian fundamental diagram. *Transportation Research Part B*, 71, 194-212.

Flötteröd, G., Rohde, J. (2011). Operational macroscopic modeling of complex urban road intersections. *Transportation Research Part B*, 45(6), 903-922.

Fonseca, D.J., Lou, Y., Moynihan, G.P., Gurupackiam, S. (2013). Incident Occurrence Modeling during Hurricane Evacuation Events: The Case of Alabama's I-65 Corridor. *Modelling and Simulation in Engineering*, 2013.

Friedrich, M., Hofsäß, I., Wekeck, S. (2001). Timetable-Based Transit Assignment Using Branch and Bound Techniques. *Transportation Research Record*, 1752, 100-107.

Fu, H., Wilmot, C.G. (2004). A Sequential Logit Dynamic Travel Demand Model For Hurricane Evacuation. *Transportation Research Record*, 1882(1), 19-26.

Gao, S., Frejinger, E., Ben-Akiva, M. (2008). Adaptive Route Choice Models in Stochastic Time-Dependent Networks. *Transportation Research Record*, 2085, 136-143.

Geistefeldt, J. (2012). Operational Experience with Temporary Hard Shoulder Running in Germany. *Transportation Research Record*, 2278, 67-73.

Gentile, G. (2010). The General Link Transmission Model for Dynamic Network Loading and a Comparison with the DUE algorithm. In: Tampère, C.M.J., Viti, F., Immers, L.H. (eds.), *New Developments in Transport Planning: Advances in Dynamic Traffic Assignment*. Edward Elgar, pp. 153-178.

Gentile, G. (2015). Using the General Link Transmission Model in a Dynamic Traffic Assignment to simulate congestion on urban networks. *Transportation Research Procedia*, 5, 66-81.

Gentile, G., Papola, N. (2009a). Dynamic Traffic Assignment with Nonseparable Link Cost Functions and Queue Spillovers. In: Cascetta, E., *Transportation Systems Analysis: Models and Applications*. 2nd edn. Springer, pp. 464-480.

Gentile, G., Papola, N. (2009b). The Simplified Theory of Kinematic Waves Based on Cumulative Flows: Application to Macroscopic Link Performance Models. In: Cascetta, E., *Transportation Systems Analysis: Models and Applications*. 2nd edn. Springer, pp. 497-510.

Gentile, G., Meschini, L., Papola, N. (2007). Spillback congestion in dynamic traffic assignment: A macroscopic flow model with time-varying bottlenecks. *Transportation Research Part B*, 41(10), 1114-1138.

Geroliminis, N., Skabardonis, A. (2005). Prediction of Arrival Profiles and Queue Lengths Along Signalized Arterials by Using a Markov Decision Process. *Transportation Research Record*, 1934, 116-124.

Godunov, S. (1959). A difference scheme for numerical computation of discontinuous solutions of equations of fluid dynamics. *Matematischeskii Sbornik*, 47(89), 271-306.

Grontmij (2015). *Capaciteitswaarden Infrastructuur Autosnelwegen*. 4th edn. Dienst Water, Verkeer en Leefomgeving.

Gu, W., Yu, J., Ji, Y., Van der Gun, J.P.T., Pel, A.J., Zhang, H.M., Van Arem, B. (2018). Optimizing Tailored Bus Bridging Paths. 97th Transportation Research Board Annual Meeting. Washington, D.C.

Guerrieri, M., Mauro, R. (2016). Capacity and safety analysis of hard-shoulder running (HSR). A motorway case study. *Transportation Research Part A*, 92, 162-183.

Hajiahmadi, M., Corthout, R., Tampère, C.M.J., De Schutter, B., Hellendoorn, H. (2013). Variable Speed Limit Control Based on Extended Link Transmission Model. *Transportation Research Record*, 2390, 11-19.

Hajiahmadi, M., Van de Weg, G.S., Tampère, C.M.J., Corthout, R., Hegyi, A., De Schutter, B., Hellendoorn, H. (2016). Integrated Predictive Control of Freeway Networks Using the Extended Link Transmission Model. *IEEE Transactions on Intelligent Transportation Systems*, 17(1), 65-78.

Hamdar, S.H. (2004). *Towards Modeling Driver Behavior Under Extreme Conditions*. University of Maryland.

Han, K., Friesz, T.L., Szeto, W.Y., Liu, H. (2015). Elastic demand dynamic network user equilibrium: Formulation, existence and computation. *Transportation Research Part B*, 81, 183-209.

Han, K., Piccoli, B., Szeto, W.Y. (2016). Continuous-time link-based kinematic wave model: formulation, solution existence, and well-posedness. *Transportmetrica B*, 4(3), 187-222.

Hara, Y., Kuwahara, M. (2015). Traffic Monitoring immediately after a major natural disaster as revealed by probe data - A case in Ishinomaki after the Great East Japan Earthquake. *Transportation Research Part A*, 75, 1-15.

Hegyi, A., De Schutter, B., Hellendoorn, J. (2005a). Optimal Coordination of Variable Speed Limits to Suppress Shock Waves. *IEEE Transactions on Intelligent Transportation Systems*, 6(1), 102-112.

Hegyi, A., De Schutter, B., Hellendoorn, J. (2005b). Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C*, 13(3), 185-209.

Hegyi, A., Hoogendoorn, S.P., Schreuder, M., Stoelhorst, H., Viti, F. (2008). SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory. 11th International IEEE Conference on Intelligent Transportation Systems, pp. 827-832. Beijing, China.

Helbing, D., Balietti, S. (2012). Agent-Based Modelling. In: Helbing, D., *Social Self-Organization, Agent-Based Simulations and Experiments to Study Emergent Social Behavior*. Springer, pp. 25-70.

Himpe, W., Tampère, C.M.J. (2016). A dynamic user equilibrium algorithm that exploits warm starting capabilities of the iterative link transmission model. 6th International Symposium on Dynamic Traffic Assignment. Sydney, Australia.

Himpe, W., Corthout, R., Tampère, C.M.J. (2016). An efficient iterative link transmission model. *Transportation Research Part B*, 92(B), 170-190.

Hoogendoorn, R.G. (2012). *Empirical Research and Modeling of Longitudinal Driving Behavior Under Adverse Conditions*. TRAIL Thesis Series.

Hoogendoorn, S.P., Bovy, P.H.L. (2001). State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 215(4), 283-303.

Huang, S.K., Lindell, M.K., Prater, C.S., Wu, H.C., Siebeneck, L.K. (2012). Household Evacuation Decision Making in Response to Hurricane Ike. *Natural Hazards Review*, 13, 283-296.

Huibregtse, O. (2013). *Robust Model-Based Optimization of Evacuation Guidance*. TRAIL Thesis Series.

Illenberger, J., Flötteröd, G., Nagel, K. (2007). Enhancing MATSim with capabilities of within-day re-planning. Intelligent Transportation Systems Conference, pp. 94-99. Seattle, Washington.

Jayakrishan, R., Mahmassani, H.S., Hu, T. (1994). An evaluation tool for advanced traffic information and management systems in urban networks. *Transportation Research Part C*, 2(3), 129-147.

Jin, J.G., Teo, K.M., Odoni, A.R. (2016). Optimizing Bus Bridging Services in Response to Disruptions of Urban Transit Rail Networks. *Transportation Science*, 50(3), 790-804.

Jin, W. (2015). Continuous formulations and analytical properties of the link transmission model. *Transportation Research Part B*, 74, 88-103.

Jin, W., Gan, Q., Lebacque, J. (2015). A kinematic wave theory of capacity drop. *Transportation Research Part B*, 81(1), 316-329.

Joh, C.H., Timmermans, H.J.P., Arentze, T.A. (2006). Measuring and Predicting Adaptation Behavior in Multi-Dimensional Activity-Travel Patterns. *Transportmetrica*, 2(2), 153-173.

Kepaptsoglou, K., Karlaftis, M.G. (2009). The bus bridging problem in metro operations: conceptual framework, models and algorithms. *Public Transport*, 1(4), 275-297.

Kitamura, R., Fujii, S. (1998). Two Computational Process Models of Activity-Travel Choice. In: Gärling, T, Laitila, T, Westin, K (eds.), *Theoretical Foundations of Travel Choice Modeling*. Pergamon, pp. 251-279.

Knapen, L., Bellemans, T., Usman, M., Janssens, D., Wets, G. (2014). Within day rescheduling microsimulation combined with macrosimulated traffic. *Transportation Research Part C*, 45, 99-118.

Knoop, V.L. (2009). *Road Incidents and Network Dynamics: Effects on driving behaviour and traffic congestion*. TRAIL Thesis Series.

Knoop, V.L., Hoogendoorn, S.P., Van Zuylen, H.J. (2009). Empirical Differences Between Time Mean Speed and Space Mean Speed. In: Appert-Rolland, C., Chevoir, F., Gondret, P., Lassarre, S., Lebacque, J.P., Schreckenberg, M. (eds.), *Traffic and Granular Flow '07*. Springer-Verlag, pp. 351-356.

Knoop, V.L., Hoogendoorn, S.P., Van Zuylen, H.J. (2010). Rerouting behaviour of travellers under exceptional traffic conditions - an empirical analysis of route choice. *Procedia Engineering*, 3, 113-128.

Kolen, B. (2013). *Certainty of uncertainty in evacuation for threat driven response: Principles of adaptive evacuation management for flood risk planning in the Netherlands*. Radboud University Nijmegen.

Koshi, M., Iwasaki, M., Okhura, I. (1983). Some findings and an overview on vehicular flow characteristics. In: Hurdle, V.F., Hauer, E., Steuart, G.N. (eds.), *Proceedings of the Eighth International Symposium on Transportation and Traffic Theory*. University of Toronto Press, pp. 403-426.

Koshimura, S., Katada, T., Mofjeld, H.O., Kawata, Y. (2006). A method for estimating casualties due to tsunami inundation flow. *Natural Hazards*, 39(2), 265-274.

Kruithof, J. (1937). Telefoonverkeersrekening. *De Ingenieur*, 52(8), E15-E25.

Kurauchi, F., Bell, M.G.H., Schmöcker, J.D. (2003). Capacity Constrained Transit Assignment with Common Lines. *Journal of Mathematical Modelling and Algorithms*, 2(4), 309-327.

Kwon, T.J., Fu, L., Jiang, C. (2013). Effect of Winter Weather and Road Surface Conditions on Macroscopic Traffic Parameters. *Transportation Research Record*, 2329, 54-62.

Lämmel, G., Rieser, M., Nagel, K., Taubenböck, H., Strunz, G., Goseberg, N., . . . Birkmann, J. (2008). Emergency Preparedness in the Case of a Tsunami - Evacuation Analysis and Traffic Optimization for

the Indonesian City of Padang. In: Klingsch, W.W.F., Rogsch, F., Schadschneider, A., Schreckenberg, M. (eds.), *Pedestrian and Evacuation Dynamics*. Springer, pp. 171-182.

Laval, J.A., Leclercq, L. (2013). The Hamilton-Jacobi partial differential equation and the three representations of traffic flow. *Transportation Research Part B*, 52, 17-30.

Lax, P.D. (1957). Hyperbolic Systems of Conservation Laws II. *Communications on Pure and Applied Mathematics*, 10(4), 537-566.

Leach, J. (1994). *Survival Psychology*. Macmillan Press.

Lebacque, J.P., Khoshyaran, M.M. (2005). First-Order Macroscopic Traffic Flow Models: Intersection Modeling, Network Modeling. In: Mahmassani, H.S. (ed.), *Transportation and Traffic Theory: Flow, Dynamics and Human Interaction - Proceedings of the 16th International Symposium on Transportation and Traffic Theory*. Elsevier, pp. 365-386.

Lighthill, M.J., Whitham, G.B. (1955). On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London A*, 229(1178), 317-345.

Lin, D.Y., Eluru, N., Waller, S.T., Bhat, C.R. (2009). Evacuation Planning Using the Integrated System of Activity-Based Modeling and Dynamic Traffic Assignment. *Transportation Research Record*, 2132, 69-77.

Lindell, M.K. (2008). EMBLEM2: An empirically based large scale evacuation time estimate model. *Transportation Research Part A*, 42(1), 140-154.

Lindell, M.K., Prater, C.S. (2007). Critical Behavioral Assumptions in Evacuation Time Estimate Analysis for Private Vehicles: Examples from Hurricane Research and Planning. *Journal of Urban Planning and Development*, 133(1), 18-29.

Litman, T. (2006). Lessons From Katrina and Rita: What Major Disasters Can Teach Transportation Planners. *Journal of Transportation Engineering*, 132, 11-18.

Logghe, S., Immers, L.H. (2008). Multi-class kinematic wave theory of traffic flow. *Transportation Research Part B*, 42, 523-541.

Long, J., Chen, J., Szeto, W.Y., Shi, Q. (2016). Link-based system optimum dynamic traffic assignment problems with environmental objectives. *Transportation Research Part D*, in press.

Lu, Y., Wong, S.C., Zhang, M., Shu, C. (2009). The Entropy Solutions for the Lighthill-Whitham-Richards Traffic Flow Model with a Discontinuous Flow-Density Relationship. *Transportation Science*, 43(4), 511-530.

Luke, J.C. (1972). Mathematical Models for Landform Evolution. *Journal of Geophysical Research*, 77(14), 2460-2464.

Maassen, K. (2012). *Optimizing and Simulating Evacuation in Urban Areas*. Universität Duisburg-Essen.

Mahut, M. (2000). *A Discrete Flow Model For Dynamic Network Loading*. University of Montreal.

Mazaré, P., Dehwah, A.H., Claudel, C.G., Bayen, A.M. (2011). Analytical and grid-free solutions to the Lighthill-Whitham-Richards traffic flow model. *Transportation Research Part B*, 45, 1727-1748.

McNutt, S.R. (1996). Seismic Monitoring and Eruption Forecasting of Volcanoes: A Review of the State-of-the-Art and Case Histories. In: Scarpa, R., Tilling, R.I. (eds.), *Monitoring and Mitigation of Volcano Hazards*. Springer, pp. 99-146.

Meschini, L., Gentile, G. (2009). Simulating car-pedestrian interactions during mass events with DTA models: the case of Vancouver Winter Olympic Games. 37th European Transport Conference. Noordwijkerhout, the Netherlands.

Michoud, C., Bazin, S., Blikra, L.H., Derron, M., Jaboyedoff, M. (2013). Experiences from site-specific landslide early warning systems. *Natural Hazards and Earth System Sciences*, 13, 2659-2673.

Monamy, T., Haj-Salem, H., Lebacque, J. (2012). A Macroscopic Node Model Related to Capacity Drop. *Procedia Social and Behavioral Sciences*, 54, 1388-1396.

Muralidharan, A., Horowitz, R., Varaiya, P. (2012). Model Predictive Control of a Freeway Network with Capacity Drops. In *Proceedings of the ASME 5th Annual Dynamic Systems and Control Conference joint with the JSME 11th Motion and Vibration Conference*. American Society of Mechanical Engineers, pp. 303-312.

Murray-Tuite, P.M., Mahmassani, H.S. (2003). Model of Household Trip Chain Sequencing in an Emergency Evacuation. *Transportation Research Record*, 1831, 21-29.

Murray-Tuite, P.M., Mahmassani, H.S. (2004). Transportation Network Evacuation Planning with Household Activity Interactions. *Transportation Research Record*, 1894, 150-159.

Murray-Tuite, P.M., Wolshon, B. (2013). Evacuation transportation modeling: An overview of research, development and practice. *Transportation Research Part C*, 27, 25-45.

Nagel, K., Schreckenberg, M. (1992). A cellular automaton model for freeway traffic. *Journal de Physique*, 2(12), 2221-2229.

Nagel, K., Stretz, P., Pieck, M., Leckey, S., Donnelly, R., Barrett, C.L. (2008). *TRANSIMS traffic flow characteristics*. arXiv.

Newell, G.F. (1993). A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B*, 27(4), 281-287.

Nie, Y.M. (2011). A cell-based Merchant-Nemhauser model for the system optimum dynamic traffic assignment problem. *Transportation Research Part B*, 45(2), 329-342.

Noh, H., Chiu, Y., Zheng, H., Hickman, M., Mirchandani, P. (2009). Approach to Modeling Demand and Supply for a Short-Notice Evacuation. *Transportation Research Record*, 2091, 91-99.

Ortúzar, J.D., Willumsen, L.G. (2011). *Modelling Transport*. 4th edn. John Wiley & Sons.

Papageorgiou, M. (1990). Dynamic modelling, assignment and route guidance in traffic networks. *Transportation Research Part B*, 24(6), 471-495.

Papageorgiou, M. (1998). Some remarks on macroscopic traffic flow modelling. *Transportation Research Part A*, 32(5), 323-329.

Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., Wang, Y. (2003). Review of Road Traffic Control Strategies. *Proceedings of the IEEE*, 91(12), 2043-2067.

Pel, A.J. (2011). *Transportation Modelling for Regional Evacuations*. TRAIL Thesis Series.

Pel, A.J., Bliemer, M.C.J., Hoogendoorn, S.P. (2012). A review on travel behavior modelling in dynamic traffic simulation models for evacuations. *Transportation*, 39, 97-123.

Pender, B., Currie, G., Delbosc, A., Shiwakoti, N. (2013). Disruption Recovery in Passenger Railways: International Survey. *Transportation Research Record*, 2353, 22-32.

Petzäll, K., Petzäll, J., Jansson, J., Nordström, G. (2011). Time saved with high speed driving of ambulances. *Accident Analysis and Prevention*, 43(3), 818-822.

Polyak, B.T. (1990). A new method of stochastic approximation type. *Avtomatika i Telemakhanika*, 51(7), 937-946.

Qian, Z.S., Zhang, H.M. (2013). A Hybrid Route Choice Model for Dynamic Traffic Assignment. *Network and Spatial Economics*, 13(2), 183-203.

Raadsen, M.P.H., Bliemer, M.C.J. (2018). Continuous-time general link transmission model with simplified fanning, Part II: Event-based algorithm for networks. *Transportation Research Part B*.

Raadsen, M.P.H., Bliemer, M.C.J., Bell, M.G.H. (2014). An efficient event-based algorithm for solving first order dynamic network loading problems. 5th International Symposium on Dynamic Traffic Assignment. Salerno, Italy.

Raadsen, M.P.H., Bliemer, M.C.J., Bell, M.G.H. (2016a). An efficient and exact event-based algorithm for solving simplified first order dynamic network loading problems in continuous time. *Transportation Research Part B*, 92(B), 191-210.

Rakha, H., Farzaneh, M., Arafeh, M., Sterzin, E. (2008). Inclement Weather Impacts on Freeway Traffic Stream Behavior. 87th Transportation Research Board Annual Meeting. Washington, D.C.

Raney, B., Nagel, K. (2006). An improved framework for large-scale multi-agent simulations of travel behavior. In: Rietveld, P., Jourquin, B., Westin, K. (eds.), *Towards better performing European Transportation Systems*. Routledge, pp. 305-347.

Regnier, E. (2008). Public Evacuation Decisions and Hurricane Track Uncertainty. *Management Science*, 54(1), 16-28.

Richards, P.I. (1956). Shock Waves on the Highway. *Operations Research*, 4(1), 42-51.

Robinson, R.M., Khattak, A. (2010). Route Change Decision Making by Hurricane Evacuees Facing Congestion. *Transportation Research Record*, 2196, 168-175.

Robinson, R.M., Khattak, A., Sokolowski, J.A., Foytik, P., Wang, X. (2009). What is the Role of Traffic Incidents in Hampton Roads Hurricane Evacuations? 88th Transportation Research Board Annual Meeting. Washington, D.C.

Roncoli, C., Papageorgiou, M., Papamichail, I. (2015). Traffic flow optimisation in presence of vehicle automation and communication systems - Part I: A first-order multi-lane model for motorway traffic. *Transportation Research Part C*, 57, 241-259.

Rosenthal, U., 't Hart, P., Van Duin, M., Boin, A., Kroon, M., Otten, M., Overdijk, W. (2013). *Complexity in Urban Crisis Management: Amsterdam's response to the Bijlmer air disaster*. 2013 edn. Earthscan.

Sadri, A.F., Ukkusuri, S.V., Murray-Tuite, P.M., Gladwin, H. (2014). How to Evacuate: Model for Understanding the Routing Strategies during Hurricane Evacuation. *Journal of Transportation Engineering*, 140, 61-69.

Schreiter, T., Smits, E., Van Lint, J.W.C., Hoogendoorn, S.P. (2010). The Cell Transmission Model with Capacity Drop. 11th International TRAIL Congress. Delft, the Netherlands.

Shiwakoti, N., Liu, Z., Hopkins, T., Young, W. (2013). An Overview on Multimodal Emergency Evacuation in an Urban Network. 36th Australasian Transport Research Forum. Brisbane, Australia.

Smits, E., Bliemer, M.C.J., Van Arem, B. (2011). Dynamic Network Loading of Multiple User-Classes with the Link Transmission Model. 2nd International Conference on Models and Technologies for Intelligent Transportation Systems. Leuven, Belgium.

Smits, E., Bliemer, M.C.J., Pel, A.J., Van Arem, B. (2015). A family of macroscopic node models. *Transportation Researsch Part B*, 74, 20-39.

Smulders, S. (1990). Control of freeway traffic flow by variable speed signs. *Transportation Research Part B*, 24B(2), 111-132.

Srivastava, A., Geroliminis, N. (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C*, 30, 161-177.

Sultan, B., Meekums, R., Brown, M. (2008). The Impact of Active Traffic Management on Motorway Operation. IET Road Transport Information and Control Conference and ITS United Kingdom Members´ Conference. Manchester, United Kingdom.

Szeto, W.Y. (2008). Enhanced Lagged Cell-Transmission Model for Dynamic Traffic Assignment. *Transportation Research Record*, 2085, 76-85.

Tahmasseby, S. (2009). *Reliability in Urban Public Transport Network Assessment and Design*. TRAIL Thesis Series.

Tampère, C.M.J., Corthout, R., Cattrysse, D., Immers, L.H. (2011). A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B*, 45(1), 289-309.

Taylor, N.B. (2003). The CONTRAM dynamic traffic assignment model. *Network and Spatial Economics*, 3(3), 297-322.

Teng, H., Kwigizile, V., Xie, G., Kaseko, M., Gibby, A.R. (2010). The Impacts of Emergency Vehicle Signal Preemption on Urban Traffic Speed. *Journal of the Transportation Research Forum*, 49(1), 69-79.

Toledo, T., Cats, O., Burghout, W., Koutsopoulos, H.N. (2010). Mesoscopic simulation for transit operations. *Transportation Research Part C*, 18(6), 896-908.

Torné, J.M., Soriguera, F., Geroliminis, N. (2014). Coordinated active traffic management freeway strategies using the capacity-lagged cell transmission model. Transportation Research Board Annual Meeting, 93. Washington, D.C.

Trainor, J.E., Murray-Tuite, P.M., Edara, P., Fallah-Fini, S., Triantis, K. (2013). Interdisciplinary Approach to Evacuation Modeling. *Natural Hazards Review*, 14(3), 151-162.

Tu, H., Tamminga, G., Drolenga, H., De Wit, J., Van der Berg, W. (2010). Evacuation Plan of the City of Almere: Simulating the Impact of Driving Behavior on Evacuation Clearance Time. *Procedia Engineering*, 3, 67-75.

Urbanik, T. (2000). Evacuation time estimates for nuclear power plants. *Journal of Hazardous Materials*, 75(2), 165-180.

Van de Weg, G.S., Keyvan-Ekbatani, M., Hegyi, A., Hoogendoorn, S.P. (2016). Urban network throughput optimization via model predictive control using the link transmission model. Transportation Research Board Annual Meeting, 95. Washington, D.C.

Van der Gun, J.P.T. (2018). A13 motorway simulation study data for the Link Transmission Model with variable fundamental diagrams and initial conditions. *4TU.ResearchData*, doi:10.4121/uuid:c7dd3a61-1bed-4858-b294-93acd960a645.

Van der Hurk, E., Koutsopoulos, H.N., Wilson, N., Kroon, L.G., Maróti, G. (2016). Shuttle Planning for Link Closures in Urban Public Transport Networks. *Transportation Science*, 50(3), 947-965.

Van Driel, C.J.G., Van Arem, B. (2010). The Impact of a Congestion Assitant on Traffic Flow Efficiency and Safety in Congested Traffic Caused by a Lane Drop. *Journal of Intelligent Transportation Systems*, 14(4), 197-208.

Van Wageningen-Kessels, F.L.M. (2013). Numerical methods for mixed-class models. In *Multi-class continuum traffic flow models: Analysis and simulation methods*. TRAIL Thesis Series, pp. 159-189.

Van Wageningen-Kessels, F.L.M., Van Lint, J.W.C., Hoogendoorn, S.P., Vuik, C. (2010). Lagrangian Formulation of Multiclass Kinematic Wave Model. *Transportation Research Record*, 2199, 29-36.

Van Wageningen-Kessels, F.L.M., Van Lint, J.W.C., Vuik, C., Hoogendoorn, S.P. (2015). Genealogy of traffic flow models. *EURO Journal on Transportation and Logistics*, 4(4), 445-473.

Van Wageningen-Kessels, F.L.M., Yuan, Y., Hoogendoorn, S.P., Van Lint, J.W.C. (2013). Discontinuities in the Lagrangian formulation of the kinematic wave model. *Transportation Research Part C*, 34, 148-161.

Van Zuilekom, K., Van Maarseveen, M., Van der Droef, M. (2005). A Decision Support System for Preventive Evacuation of People. In: Van Oosterom, P., Zlatanova, S., Fendel, E.M. (eds.), *Geo-information for Disaster Management*. Springer, pp. 229-253.

Vorst, H.C.M. (2010). Evacuation Models and Disaster Psychology. *Procedia Engineering*, 3, 15-21.

Wardrop, J.G. (1952). Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institution of Civil Engineers*, 1(3), 325-362.

Webster, F.V. (1958). *Traffic signal settings.* London, U.K.: Road Research Laboratory.

Witteveen+Bos and TU Delft (2011). *Capaciteitswaarden Infrastructuur Autosnelwegen*. 3rd edn. Dienst Verkeer en Scheepvaart.

Wolshon, B., Lambert, L. (2006). Reversible Lane Systems: Synthesis of Practice. *Journal of Transportation Engineering*, 132(12), 933-944.

Wolshon, B., Marchive, E. (2007). Emergency Planning in the Urban-Wildland Interface: Subdivision-Level Analysis of Wildfire Evacuations. *Journal of Urban Planning and Development*, 133, 73-81.

Yin, W., Murray-Tuite, P.M., Ukkusuri, S.V., Gladwin, H. (2014). An agent-based modeling system for travel demand simulation for hurricane evacuation. *Transportation Research Part C*, 42, 44-59.

Younis, J., Anquetin, S., Thielen, J. (2008). The benefit of high-resolution operational weather forecasts for flash flood warning. *Hydrology and Earth System Sciences*, 12, 1039-1051.

Yperman, I. (2007). *The Link Transmission Model for Dynamic Network Loading*. KU Leuven.

Yperman, I., Logghe, S., Immers, L.H. (2005). The Link Transmission Model: An efficient implementation of the kinematic wave theory in traffic networks. 10th EURO Working Group on Transportation Meeting, pp. 122-127. Poznan, Poland.

Yperman, I., Logghe, S., Tampère, C.M.J., Immers, L.H. (2006). The Multi-Commodity Link Transmission Model for Dynamic Network Loading. 85th Transportation Research Board Annual Meeting. Washington, D.C.

Yuan, K., Knoop, V.L., Hoogendoorn, S.P. (2017a). A kinematic wave model in Lagrangian coordinates incorporating capacity drop: Application to homogeneous road stretches and discontinuities. *Physica A*, 465, 472-485.

Yuan, K., Knoop, V.L., Leclercq, L., Hoogendoorn, S.P. (2017b). Capacity drop: a comparison between stop-and-go wave and standing queue at lane-drop bottleneck. *Transportmetrica B*, 5(2), 145-158.

Yuan, Y., Pel, A.J., Hoogendoorn, S.P. (2014). The transition between normal and emergency driving behaviour during evacuation and its implications for traffic flow operations and traffic management. 17th IEEE International Conference on Intelligent Transportation Systems, pp. 2700-2705. Qingdao, China.

Zhao, Y., Sadek, A.W. (2013). Large-Scale Agent-Based Models for Transportation Network Management under Unplanned Events. In: Janssens, D., Yasar, A., Knapen, L. (eds.), *Data Science and Simulation in Transportation Research*. IGI Global, pp. 206-231.

Summary

Multimodal Transportation Simulation for Emergencies
using the Link Transmission Model

The finite capacity of transportation systems of urban regions is evidenced by the congestion and delays incurred by travellers by car or public transport. This situation worsens when major emergencies affect the transportation system. These emergencies are characterised by a strong decrease in transportation supply and/or a strong increase in transportation demand. This thesis looks into computer simulation models that authorities can use to assess the transportation consequences of emergencies and to test plans to manage the associated transportation problems.

An urban transportation system may be affected by a wide range of emergency conditions having these characteristics. Extreme weather, infrastructure damage, and public transport failures reduce or eliminate the capacity of parts of the transportation system, thus reducing transportation supply. Traffic accidents, train and aeroplane crashes, and building fires and collapses require dispatching of emergency services, thus generating additional transportation demand. Floods, tsunamis, hurricanes, sediment disasters, wildfires, volcanic eruptions, chemical spills, nuclear disasters and other industrial accidents are even more severe as they generate additional transportation demand in the form of evacuation traffic. More than one of those three effects can occur in a single emergency.

These emergencies pose management problems to authorities such as the public administration, the police, fire brigades, emergency medical services, traffic control centres, and public transport companies. This problem can be seen as an optimisation problem, consisting of decision variables, an objective, and a set of constraints. Here, the decision variables form the emergency plan, ranging from strategic and tactical decisions such as issuing travel advice, deploying roadblocks, rescheduling public transport, to operational decisions such as configuring traffic lights, ramp metering, or variable speed limits. The corresponding value of the objective can be obtained by simulating the transportation system in a computer simulation model, enabling both analysis of the impacts of the emergencies on urban areas and evaluation of candidate management plans. The goal of the research in this thesis is to identify requirements for, develop, and test (components of) this simulation model.

Simulation framework

A transportation simulation model that is used for such applications needs several features to capture the transport-related characteristics of emergencies. Namely, the simulation methodology must

- be dynamic,
- describe the (effects of) relevant choice behaviour,
- predict or otherwise allow specification of initial conditions of the emergency,
- include interactions between individuals, e.g. members of the same household,
- include multiple modes of transportation and their interaction,
- include travellers who are not directly affected and their behaviour, and
- include emergency services.

A simulation framework consisting of two main components is proposed. The first is an activity-based choice model, rooted in activity-travel patterns on a normal day, describing choice behaviour at the level of individual travellers or groups of them. The escalation of traveller's responses to the emergency is modelled with normal behaviour, adaptation behaviour, and evacuation behaviour. The second component is a macroscopic or mesoscopic dynamic network loading model used to simulate traffic propagation, with the ability to include multiple modes and emergency services. The macroscopic or mesoscopic nature allows the propagation of car traffic to be detailed, simple to calibrate, and computationally efficient. Both components can be subject to (effects of) control measures undertaken by authorities.

The simulation framework is tested by simulating a multimodal evacuation of the city of Delft, the Netherlands. Simulation of this emergency is shown to be computationally efficient. Its results highlight the importance of interaction between inbound, outbound, and background traffic, with consequences for incoming emergency services, the interaction between modes, with road congestion delaying public transport, and the potentially non-system-optimal nature of user choices.

Model improvements

Several aspects of this simulation framework and its model components can be further improved. Among these is the coupling between the agent-based choice model and the macroscopic or mesoscopic dynamic network loading model. As simulations of emergency conditions need the possibility of en-route choice behaviour, a problem arises where the traffic composition is to be disaggregated as commodity flows that represent individual agents in the aggregated traffic flow, risking consistency issues in either the traffic flow evolution, the propagation of agents, or both. These can be avoided with suitable model design choices like using an accurate link model with small time steps, a strict weak ordering of vehicle parts, an incremental node model, and a relaxed first-in-first-out constraint.

Other improvements are developed for the discrete-time Link Transmission Model, a macroscopic dynamic network loading model based on Lighthill-Whitham-Richards traffic flow theory that is suitable for use within the simulation framework. First, the fundamental diagram, describing the possible combinations of traffic flow and traffic density at any point along a road at any time, is extended from a triangular shape to a general continuous concave function. This for example enables the gradual reduction of speed on a road as it gets loaded

up to capacity, i.e. subcritical delays, and the dispersion of platoons. Second, support is added for inverted-lambda style fundamental diagrams, creating a capacity drop once congestion has formed. This also enables the production of stop-and-go waves, resulting in complex traffic patterns combining moving jams and standing queues. As shown in numerical examples, including a simulation of the Dutch A13 motorway corridor, both extensions meanwhile preserve the small numerical error and crisp shockwaves of the original Link Transmission Model.

By specifically choosing the Smulders fundamental diagram shape for the Link Transmission Model, additional functionality is unlocked. New algorithms create the possibility for roads to be non-empty at the start of the simulation, and the possibility to algebraically compute traffic density as a function of space within links. Neither of them introduces new numerical errors. In combination, these algorithms furthermore allow arbitrary changes to driving behaviour on links during the simulation, for example in response to traffic management or changing environmental conditions. The model is demonstrated on the A13 motorway corridor in a scenario with a traffic accident, variable speed limits, and dynamic lane management. This application shows that the extension works as expected based on Lighthill-Whitham-Richards theory and that the inclusion of variable fundamental diagrams and initial conditions indeed improves the result.

Optimisation application

Simulation of a large-scale metro, tram, and car traffic disruption in the city of Rotterdam, the Netherlands, is used to demonstrate how tactical decisions can be effectuated, specifically for providing bus bridging. Based on an existing static bus-bridging optimisation method, a new real-time bus-bridging scheduling method is developed and evaluated in the context of this emergency scenario. The results, which include a comparison of several variations of the simulation model, indicate that the availability of alternative public transport routes and the travel times on the road network are both important for how public transport is used during the emergency, and can result in very minor usage of the provided bus-bridging services. The inclusion of car traffic in the model, and en-route choices for car traffic, are critical. As to the modelling of traffic propagation, particularly the modelling of signalised intersections has an impact on public transport use in this specific case, but this might vary depending on the nature of the studied network. For the bus-bridging scheduling, consideration of multiple possible routes and possible intermediate stops can be expected to increase its success.

Conclusions and recommendations

A simulation framework for emergencies has been formulated and (components of) the simulation model have been demonstrated in three different case studies, and implications have been identified for model-based emergency management and for (emergency) modelling itself. For practitioners who want to apply the developed emergency simulation framework, it can be recommended to invest into normal-day transportation and traffic models with activity-based choice models and dynamic traffic assignment models, as this makes extension to emergency simulation relatively easy.

For future research, more applications of the modelling framework are recommended. It can for example be used to compare and optimise control strategies in different emergency situations. There are also opportunities to improve and extend the methodology. First, the

choice models included in the simulation framework can be further developed and calibrated for various emergencies. Secondly, the computation of normal-day agent-based multimodal equilibria can be improved. Lastly, one can further analyse the validity of the Link Transmission Model extensions.

Samenvatting

Multimodale Vervoerssimulatie voor Noodgevallen

gebruikmakend van het Linktransmissiemodel

De eindige capaciteit van vervoerssystemen van stedelijke gebieden manifesteert zich in de congestie en vertragingen die reizigers per auto of openbaar vervoer oplopen. Deze situatie verslechtert wanneer ernstige noodgevallen het vervoerssysteem treffen. Deze noodgevallen worden gekarakteriseerd door een sterke afname in het vervoersaanbod en/of een sterke toename in de vervoersvraag. Dit proefschrift onderzoekt computersimulatiemodellen die autoriteiten kunnen gebruiken om de vervoersgevolgen van noodgevallen te beoordelen en plannen te testen om de bijbehorende vervoersproblemen te beteugelen.

Een stedelijke vervoerssysteem kan worden getroffen door een breed scala aan noodsituaties. Extreem weer, schade aan infrastructuur en uitval van openbaar vervoer verminderen of elimineren de capaciteit van delen van het vervoersysteem, en verminderen daarmee het vervoersaanbod. Verkeersongevallen, trein- en vliegtuigcrashes en gebouwbranden en -instortingen vereisen de inzet van hulpdiensten, en genereren daarmee extra vervoersvraag. Overstromingen, tsunamis, orkanen, sedimentrampen, bosbranden, vulkaanuitbarstingen, chemische lekkages, kernrampen en andere industriële ongevallen zijn nog erger aangezien ze extra vervoersvraag genereren in de vorm van evacuatieverkeer. Meer dan één van deze drie effecten kan in dezelfde noodsituatie optreden.

Deze noodgevallen vormen beheersproblemen voor autoriteiten zoals het openbaar bestuur, de politie, de brandweer, ambulancediensten, verkeerscentrales en openbaarvervoerbedrijven. Dit probleem kan worden beschouwd als een optimalisatieprobleem, bestaande uit beslissingsvariabelen, een doelstelling en een reeks randvoorwaarden. De beslissingsvariabelen vormen hier het noodplan, variërend van strategische en tactische besluiten als het verstrekken van reisadvies, het inzetten van wegversperringen, het herplannen van openbaar vervoer, tot operationele besluiten als het configureren van verkeerslichten, toeritdoseringen of variabele snelheidslimieten. De bijbehorende waarde van de doelstelling kan worden verkregen door simulatie van het vervoerssysteem in een computersimulatiemodel, waardoor zowel analyse van de impacten van noodgevallen op stedelijke gebieden als evaluatie van voorgestelde beheersplannen mogelijk wordt. Het doel

van het onderzoek in dit proefschrift is het vaststellen van vereisten voor, het ontwikkelen van en het testen van (componenten van) dit simulatiemodel.

Simulatieraamwerk

Een vervoerssimulatiemodel dat wordt gebruikt voor zulke toepassingen heeft verschillende functionaliteiten nodig om de vervoersgerelateerde kenmerken van noodsituaties te vangen. De simulatiemethodologie moet namelijk

- dynamisch zijn,
- (effecten van) het relevante keuzegedrag beschrijven,
- de initiële toestand van de noodsituatie voorspellen of specificatie daarvan toestaan,
- interacties tussen individuën bevatten, bijv. leden van hetzelfde huishouden,
- meerdere vervoerswijzen en hun interactie bevatten,
- reizigers en hun gedrag bevatten die niet direct getroffen worden, en
- hulpdiensten bevatten.

Een simulatieraamwerk bestaande uit twee hoofdcomponenten wordt voorgesteld. De eerste is een activiteiten-gebaseerd keuzemodel, gegrond in activiteiten-reispatronen op een normale dag, dat keuzegedrag op het niveauvan individuele reizigers of groepen van hen beschrijft. De escalatie van reizigersreacties op de noodsituatie wordt gemodelleerd met normaal gedrag, aanpassingsgedrag en evacuatiegedrag. De tweede component is een macroscopisch of mesoscopisch dynamisch netwerkbeladingsmodel dat wordt gebruikt om verkeerspropagatie te simuleren, met de mogelijkheid om meerdere vervoerswijzen en hulpdiensten op te nemen. De macroscopische of mesoscopische aard staat toe dat dat de propagatie van autoverkeer gedetailleerd, eenvoudig te kalibreren en rekenkundig efficiënt is. Beide componenten kunnen onderhevig zijn aan (effecten van) beheersmaatregelen genomen door autoriteiten.

Het simulatiemodel wordt getest door een multimodale evacuatie van de stad Delft, Nederland te simuleren. Getoond wordt dat de simulatie van deze noodsituatie rekenkundig efficiënt is. De resultaten markeren het belang van interactie tussen inkomend, uitgaand en achtergrondverkeer, met gevolgen voor inkomende hulpdiensten, de interactie tussen vervoerswijzen, met vertraagd openbaar vervoer door verkeerscongestie, en de mogelijk niet-systeem-optimale aard van gebruikerskeuzen.

Modelverbeteringen

Verscheidene aspecten van dit simulatieraamwerk en zijn modelcomponenten kunnen verder worden verbeterd. Hieronder valt de koppeling tussen het agent-gebaseerde keuzemodel en het macroscopische of mesoscopische dynamisch netwerkbeladingsmodel. Aangezien simulaties van noodsituaties de mogelijkheid van keuzegedrag onderweg benodigen, ontstaat er een probleem waarbij de verkeerssamenstelling moet worden gedesaggregeerd als goederenstromen die individuele agents in de geaggregeerde verkeersstroom vertegenwoordigen, met risico's voor de consistentie van of de evolutie van de verkeersstroom, of de propagatie van agents, of beide. Deze kunnen worden vermeden met geschikte modelontwerpkeuzes zoals het gebruik van een nauwkeurig linkmodel met kleine tijdstappen, een strikte partiële ordening van voertuigdelen, een incrementeel knoopmodel en een versoepelde eerst-in-eerst-uitbeperking.

Andere verbeteringen worden ontwikkeld voor het discretetijd-Linktransmissiemodel, een macroscopisch dynamisch netwerkbeladingsmodel gebaseerd op Lighthill-Whitham-Richardsverkeersstroomtheorie dat geschikt is voor gebruik binnen het simulatieraamwerk. Ten eerste wordt het fundamenteel diagram, dat de mogelijke combinaties van verkeersintensiteit en -dichtheid op elk punt langs de weg op elk tijdstip beschrijft, uitgebreid van een driehoekige vorm naar een algemene continue concave functie. Dit maakt bijvoorbeeld mogelijk dat de snelheid op een weg geleidelijk afneemt terwijl hij tot capaciteit wordt beladen, d.w.z. subkritische vertragingen, en de dispersie van pelotons. Ten tweede wordt ondersteuning toegevoegd voor fundamenteel diagrammen in omgekeerdelambdastijl, een capaciteitsval creërend zodra congestie is gevormd. Dit maakt ook de productie van filegolven mogelijk, resulterend in complexe verkeerspatronen waarin zowel bewegende opstoppingen als staande files. Zoals getoond in numerieke voorbeelden, waaronder een simulatie van de Nederlandse A13-snelwegcorridor, behouden beide uitbreidingen de kleine numerieke fout en scherpe schokgolven van het originele Linktransmissiemodel.

Door voor het Linktransmissiemodel specifiek de fundamenteeldiagramvorm van Smulders te kiezen, komt aanvullende functionaliteit beschikbaar. Nieuwe algoritmes creëren de mogelijkheid voor wegen om niet leeg te zijn bij aanvang van de simulatie, en de mogelijkheid om verkeersdichtheid als functie van de plaats binnen links algebraïsch te berekenen. Geen van beide introduceert nieuwe numerieke fouten. In combinatie staan deze algoritmes tijdens de simulatie bovendien willekeurige wijzigingen toe in rijgedrag op links, bijvoorbeeld in reactie op verkeersmanagement of wijzigende omgevingscondities. Het model wordt gedemonstreerd op de A13-snelwegcorridor in een scenario met een verkeersongeval, variabele snelheidslimieten en dynamisch rijstrookbeheer. Deze toepassing laat zien dat de uitbreiding werkt zoals verwacht op basis van Lighthill-Whitham-Richardstheorie en dat het inbegrip van variabele fundamenteel diagrammen en beginvoorwaarden het resultaat inderdaad verbetert.

Optimalisatietoepassing

Simulatie van een grootschalige verstoring van metro-, tram- en autoverkeer in de stad Rotterdam, Nederland wordt gebruikt om te demonstreren hoe tactische besluiten kunnen worden bewerkstelligd, specifiek voor het leveren van busbruggen. Gebaseerd op een bestaande statische busbruggenoptimalisatiemethode wordt een nieuwe realtimebusbruggenplanningsmethode ontwikkeld en geëvalueerd in de context van deze noodsituatie. De resultaten, waaronder een vergelijking van verscheidene variaties van het simulatiemodel, wijzen erop dat de beschikbaarheid van alternatieve openbaarvervoerroutes en de reistijden in het wegennet beide belangrijk zijn voor hoe openbaar vervoer gebruikt wordt tijdens de noodsituatie, en kunnen resulteren in erg laag gebruik van de geleverde busbrugdiensten. Het inbegrip van autoverkeer in het model en keuzes onderweg voor autoverkeer zijn essentieel. Wat betreft de modellering van verkeerspropagatie heeft met name de modellering van verkeerslichtgeregelde kruispunten een impact op openbaarvervoergebruik in deze specifieke casus, maar dit zou kunnen variëren afhankelijk van de aard van het bestudeerde netwerk. Voor de planning van busbruggen kan worden verwacht dat beschouwing van meerdere mogelijke routes en mogelijke tussenstops het succes vergroten.

Conclusies en aanbevelingen

Een simulatieraamwerk voor noodgevallen is geformuleerd en (componenten van) het simulatiemodel zijn gedemonstreerd in drie verschillende casestudies, en implicaties zijn vastgesteld voor modelgebaseerd crisismanagement en voor (noodsituatie)modellering zelf. Voor beroepsbeoefenaren die het ontwikkelde noodsituatiesimulatieraamwerk willen toepassen kan worden aanbevolen om te investeren in verkeers- en vervoersmodellen voor normale dagen met activiteiten-gebaseerde keuzemodellen en dynamische verkeerstoedelingsmodellen, aangezien dit uitbreiding naar simulatie van noodsituaties relatief eenvoudig maakt.

Voor toekomstig onderzoek worden meer toepassingen van het modelraamwerk aanbevolen. Het kan bijvoorbeeld worden gebruikt om regelstrategieën in verschillende noodsituaties te vergelijken en optimaliseren. Er zijn ook mogelijkheden om de methodologie te verbeteren en uit te breiden. Ten eerste kunnen de keuzemodellen binnen het simulatieraamwerk verder worden uitgebreid en gekalibreerd voor diverse noodgevallen. Ten tweede kan de berekening van agent-gebaseerde multimodale evenwichten op normale dagen worden worden verbeterd. Tot slot kan men de validiteit van de Linktransmissiemodel-uitbreidingen verder analyseren.

About the author

Jeroen Peter Tjebbo van der Gun was born in 1990. In 2008, he completed the Dutch Gymnasium at Corderius College in Amersfoort. In the same year he started the BSc programme *Civiele Techniek* at the faculty of Civil Engineering and Geosciences of Delft University of Technology, completing his Propedeuse with distinction in 2009. As a minor, he chose *Applied Mathematics*. His Bachelor thesis focused on the microscopic simulation of rubbernecking congestion near traffic incidents. He completed his Bachelor's degree with distinction as well as the *Honours Programme Bachelor* in 2011.

After his Bachelor, he started the MSc programme *Civil Engineering* at the same faculty, choosing the *Transport & Planning* track. During this study, he was a member of the Internet and Automation Committee of study association Dispuut Verkeer. As part of an internship at Significance B.V. in The Hague, he wrote his Master thesis on multimodal mode and route choice modelling in travel demand models. He completed his Master's degree with distinction in 2013.

He worked as a PhD researcher at the Transport & Planning department of the faculty from 2013 to 2018, resulting in this PhD thesis. From 2015 to 2017, he was a member of the PhD Council of research school TRAIL. As of 2018, he works at the department as a researcher in the NWO project *Spatial and Transport Impacts of Automated Driving*, part of the *Smart Urban Regions of the Future* programme. Since 2017, he also works one day a week a as a traffic engineering software developer at Fileradar B.V. in Delft.

List of publications

Journal articles in preparation or under peer review

Van der Gun, J.P.T., Pel, A.J., Gu, W., Van Arem, B. (in preparation). Real-time bus-bridging scheduling and dynamic simulation for multimodal urban network disruptions.

Van der Gun, J.P.T., Pel, A.J., Van Arem, B. (under review). The Link Transmission Model with variable fundamental diagrams and initial conditions. Submitted to *Transportmetrica B: Transport Dynamics*. Abingdon, England: Taylor & Francis. ISSN 2168-0566.

Peer-reviewed journal publications

Van der Gun, J.P.T., Pel, A.J., Van Arem, B. (2017). Extending the Link Transmission Model with non-triangular fundamental diagrams and capacity drops. *Transportation Research Part B: Methodological*, vol. 98, pp. 154–178. Amsterdam, the Netherlands: Elsevier. ISSN 0191-2615. doi:10.1016/j.trb.2016.12.011

Van der Gun, J.P.T., Pel, A.J., Van Arem, B. (2016). A general activity-based methodology for simulating multimodal transportation networks during emergencies. *European Journal of Transport and Infrastructure Research (EJTIR)*, vol. 16(3), pp. 490–511. Delft, the Netherlands: Delft University of Technology. ISSN 1567-7141.

Peer-reviewed conference contributions

Gu, W., Yu, J., Ji, Y., Van der Gun, J.P.T., Pel, A.J., Zhang, M.H., Van Arem, B. (2018). Optimising Tailored Bus Bridging Paths. Paper no. 18-05145 presented at the 97th Transportation Research Board Annual Meeting (TRB 2018). Washington, D.C., 7–11 January.

Van der Gun, J.P.T., Pel, A.J., Van Arem, B. (2016). Extending the Link Transmission Model with general concave fundamental diagrams and capacity drops. Extended abstract presented at the 6th International Symposium on Dynamic Traffic Assignment (DTA 2016). Sydney, Australia, 28–30 June.

Van der Gun, J.P.T., Pel, A.J., Van Arem, B. (2016). Propagating agents with macroscopic dynamic network loading: challenges and possible solutions. *Procedia Computer Science*, vol. 83, pp. 914–920. Amsterdam, the Netherlands: Elsevier. ISSN 1877-0509. Presented at the 5th International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications (ABMTRANS 2016). Madrid, Spain, 24 May. doi:10.1016/j.procs.2016.04.185

Van der Gun, J.P.T., Pel, A.J., Van Arem, B. (2015). A General Activity-based Methodology for Simulating Multimodal Transportation Networks during Emergencies. Paper no. 11 presented at the 3rd International Conference on Evacuation Modeling and Management (ICEM 2015). Tainan, Taiwan, 1–3 June.

Van der Gun, J.P.T., Van Nes, R., Van Arem, B. (2014). Flexible Public Transport Modelling for Large Urban Areas. Extended abstract presented at the 3rd Symposium of the European Association for Research in Transportation (hEART 2014). Leeds, United Kingdom, 10–12 September.

Non-peer-reviewed seminar contribution

Van der Gun, J.P.T., Pel, A.J., Van Arem, B. (2014). An Activity-Based Multimodal Model Structure to assess Transportation Management Strategies for Urban Emergencies. Paper in: Tu, H., Pel, A.J. (Eds.). *Proceedings of Emergency Transport Management*, pp. 79–83. Delft, the Netherlands: TRAIL Research School. ISBN 978-90-78271-08-6. Presented at the Joint Chinese-Dutch Seminar on Transportation Management and Travel Behaviour for Urban Emergencies: Past, Present, and Future Research. Shanghai, China, 25–27 June.

Dataset

Van der Gun, J.P.T. (2018). A13 motorway simulation study data for the Link Transmission Model with variable fundamental diagrams and initial conditions. *4TU.ResearchData*. doi:10.4121/uuid:c7dd3a61-1bed-4858-b294-93acd960a645

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 200 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Gun, J.P.T. van der, *Multimodal Transportation Simulation for Emergencies using the Link Transmission Model,* T2018/3, May 2018.

Riessen, B. van, *Optimal Transportation Plans and Portfolios for Synchromodal Container Networks,* T2018/2, March 2018.

Saeedi, H., *Network-Level Analysis of the Market and Performance of Intermodal Freight Transport,* T2018/1, March 2018.

Ypsilantis, P., *The Design, Planning and Execution of Sustainable Intermodal Port-hinterland Transport Networks,* T2017/14, December 2017.

Han, Y., *Fast Model Predictive Control Approaches for Road Traffic Control,* T2017/13, December 2017.

Wang, P., *Train Trajectory Optimization Methods for Energy-Efficient Railway Operations,* T2017/12, December 2017.

Weg, G.S. van de, *Efficient Algorithms for Network-wide Road Traffic Control*, T2017/11, October 2017.

He, D., *Energy Saving for Belt Conveyors by Speed Control,* T2017/10, July 2017.

Bešinović, N., *Integrated Capacity Assessment and Timetabling Models for Dense Railway Networks,* T2017/9, July 2017.

Chen, G., *Surface Wear Reduction of Bulk Solids Handling Equipment Using Bionic Design,* T2017/8, June 2017.

Kurapati, S., *Situation Awareness for Socio Technical Systems: A simulation gaming study in intermodal transport operations,* T2017/7, June 2017.

Jamshidnejad, A., *Efficient Predictive Model-Based and Fuzzy Control for Green Urban Mobility,* T2017/6, June 2017.

Araghi, Y., *Consumer Heterogeneity, Transport and the Environment*, T2017/5, May 2017.

Kasraian Moghaddam, D., *Transport Networks, Land Use and Travel Behaviour: A long term investigation,* T2017/4, May 2017.

Smits, E.-S., *Strategic Network Modelling for Passenger Transport Pricing,* T2017/3, May 2017.

Tasseron, G., *Bottom-Up Information Provision in Urban Parking: An in-depth analysis of impacts on parking dynamics,* T2017/2, March 2017.

Halim, R.A., *Strategic Modeling of Global Container Transport Networks: Exploring the future of port-hinterland and maritime container transport networks,* T2017/1, March 2017.

Olde Keizer, M.C.A., *Condition-Based Maintenance for Complex Systems: Coordinating maintenance and logistics planning for the process industries,* T2016/26, December 2016.

Zheng, H., *Coordination of Waterborn AGVs,* T2016/25, December 2016.

Yuan, K., *Capacity Drop on Freeways: Traffic dynamics, theory and Modeling,* T2016/24, December 2016.

Li, S., *Coordinated Planning of Inland Vessels for Large Seaports,* T2016/23, December 2016.

Berg, M. van den, *The Influence of Herding on Departure Choice in Case of Evacuation: Design and analysis of a serious gaming experimental set-up,* T2016/22, December 2016.

Luo, R., *Multi-Agent Control of Urban Transportation Networks and of Hybrid Systems with Limited Information Sharing,* T2016/21, November 2016.

Campanella, M., *Microscopic Modelling of Walking Behavior,* T2016/20, November 2016.

Horst, M. van der, *Coordination in Hinterland Chains: An institutional analysis of port-related transport,* T2016/19, November 2016.

Beukenkamp, W., *Securing Safety: Resilience time as a hidden critical factor,* T2016/18, October 2016.

Mingardo, G., *Articles on Parking Policy,* T2016/17, October 2016.

Duives, D.C., *Analysis and Modelling of Pedestrian Movement Dynamics at Large-scale Events*, T2016/16, October 2016.

Wan Ahmad, W.N.K., *Contextual Factors of Sustainable Supply Chain Management Practices in the Oil and Gas Industry,* T2016/15, September 2016.

TRAIL

Summary

Emergencies disrupting urban transportation systems cause management problems for authorities. This thesis develops simulation methods that permit analysis thereof and evaluation of candidate management plans, tested in three case studies. It formulates a methodological framework using agent-based choice models and multimodal macroscopic dynamic network loading models, and develops extensions of the Link Transmission Model to deal with more complex and variable fundamental diagrams and initially non-empty roads.

About the Author

Jeroen Peter Tjebbo van der Gun received his MSc degree in Civil Engineering from Delft University of Technology in 2013. He performed his PhD research at the Transport & Planning department of the same university from 2013 to 2018.

ERASMUS UNIVERSITY ROTTERDAM

Radboud University Nijmegen

TUDelft Delft University of Technology

UNIVERSITY OF TWENTE.

TU/e Technische Universiteit Eindhoven University of Technology

university of groningen faculty of economics and business